De problem-solving-bundel (113 pagina's samenvatting theorie en voorbeelden/ oefeningen)

olympia

augustus 2012

inleiding

Er zijn verschillende olympiades waar de Vlaamse top mag aan deelnemen: BxMO, IMO
en later misschien nog RMM en EGMO .

De vragen op die olympiades vragen vooral wat ervaring om te kunnen concurreren met de andere landen.

In deze PDF worden alle voorbeelden gegeven die in $99,99$ procent van de gevallen voorkomen, zodat men weet wat er in een combinatoriek-, algebra-, getaltheorie- of meetkundevraag gebruikt kan worden.

Soms gebruikt men natuurlijk de verschillende onderwerpen gemixt, wat de juiste benodigdheden niet altijd triviaal te vinden maakt.

Het is dus niet enkel theorie, maar vaak ook mogelijke soorten van vragen die als voorbeeld geven die via Olympia becommentarieerd wordt (verbetering, hints),
indien een onopgeloste vraag door de lezer wordt ingezonden (aan te raden).

Veel succes en plezier met de mooie problem-solving die te ontdekken valt.

werking

De bedoeling is dat volgende onderwerpen worden doorlopen door de theorie door te nemen en dan de oefeningen te proberen.
Het is belangrijk dat men de voorbeelden en theorie goed snapt en de tijd durft te nemen om lang genoeg te proberen de oefeningen op te lossen voor de echte ervaring.
Via de link kan men de vraag terugvinden op Olympia met de oplossing en indien zonder, kan men zelf de oplossing indienen, die dan ook gecontroleerd , waarna enkele opmerkingen, hints als hulp worden gegeven.
Indien een onderwerp onduidelijk is, kan men een link naar completer bestand vinden over dat onderwerp met meer voorbeelden
(soms ook al theorie die verder in het hoofdstuk stond of een competitievraag die ook in de vragenlijst stond, die dan natuurlijk niet zelf meer moet worden ingezonden als ze er nog niet opgelost was).
Het kan gebeuren dat zo'n bijlage in het Engels is en heel algemeen/moeilijk en delen bevat die minder interessant zijn.

De volgorde van de onderwerpen apart kan men kiezen en makkelijk terug vinden:

combinatoriek
algebra en analyse
getaltheorie
meetkunde

**Bij raad, opmerkingen of vragen ivm de bundel,
kan men dit via olympia@problem-solving.be melden.**

1 problem-solving algemeen

Problem-solving is iets meer dan gewoon uitrekenenen.

Merk trouwens op dat een vraag met identieke oplossing niet steeds even makkelijk is:

Bewijs dat 2^k een veelvoud geeft dat enkel bestaat uit enen en tweeen.
Men kan hier direct een inductie op het aantal factoren 2 uitvoeren.

Bewijs dat 2^k een veelvoud geeft dat geen nullen bevat.
Wanneer men dit wil bewijzen voor een getal als 2^{45} , zit er niet direct een logica achter omdat er heel veel mogelijkheden zijn.

Soms kan een extra voorwaarde de vraag dus zelfs vergemakkelijken en moet men zich niet laten afschrikken door lange vragen.

We leggen kort uit hoe men een problem-solving-vraag kan oplossen a.h.v. een gekend bestand en eenvoudig IMOLL-voorbeeld.

Voorbeeld 1.1. *n is een natuurlijk getal.*
Een tango wordt gedanst op de getallenlijn.
Iedere danser maakt een stap naar links of rechts (van x naar $x+1$ of $x-1$)
De 2 personen starten in 0 en maakt dan n stappen gelijktijdig op het ritme.
Er zijn 3 regels waar ze zich moeten aan houden:
(1) *Iedere danser moet eindigen in 0.*
(2) *Als ze op dezelfde plaats staan, geven ze elkaar een kus (niet bij start en einde)*
(3) *Iedere danser moest even veel linkse punten als rechtse punten verdienen.*
Hierbij scoort men het aantal linkse punten gelijk aan het aantal stappen dat men al naar rechts had genomen bij een stap naar links.
Analoog als men naar rechts stapt, krijgt men het aantal stappen die al naar links genomen had, als rechtse punten erbij.

(*a Voor welke waarden van n is een tango mogelijk?*

(*b*) *Bewijs dat iedere tango minimum 1 kus bevat.*

stap 1: het begrijpen

Deze lange, saaie vraag mag al enkele keren gelezen worden om dan a en b apart te interpreteren. Werken met kleine waarden kan handig zijn om de derde voorwaarde goed te begrijpen.

stap 1: het begrijpen

stap 2: ideeen krijgen

Merk alvast op dat bij moeilijkere vragen men meerdere ideeen zal moeten durven afgaan en van stap 3 of 4 naar stap 2 kan moeten terugkeren.

Bij (a) elimineren we eenvoudig dat n oneven is en merken op dat 2 niet werkt, maar 4 wel bij de kleine waarden.
We kunnen L, R schrijven als de linkse en rechtse punten resp. die gescoord worden en merken op dat we de danspassen van 1 persoon mogen observeren.
Als we $n = 2A$ schrijven, maken we A stappen naar links en A naar rechts.
Een unie van 2 dansen is hier een nieuwe geldige tango, de viervouden zullen dus werken.
We zullen dus nog enkel moeten bewijzen wat het kleinste getal is van de vorm $4t+2$ dat werkt of bewijzen dat A altijd even moet zijn.

Bij (b) weten we uit het ongerijmde dat beide dansers verschillend starten bij een tegenvoorbeeld. De linkse speler zal dus altijd aan de linkerkant van de andere danser blijven tot het einde, als ze elkaar niet kruisen (pariteit toont aan dat ze niet met een verschil van 1 plaats van elkaar kunnen staan).

stap 3: ideeen uitwerken

(A) Na wat voorbeelden, zien we dat de $L + R = A^2$, dit bewijst dat $2|A^2$ zodat de vraag opgelost wordt.
We moeten dit dus nog aantonen:
Als we de i^{de} stap naar rechts, als j^{de} stap nomen, betekent dit dat er $j - i$ keer al naar links

gestapt werd.
$R = \sum j_r - i_r$ *waarbij* $\sum j_r$ *een som is van* A *elementen van* 1 *tot* $n = 2A$ *, maar* $\sum i_r$ *is gekend als* $1 + 2 + \cdots + A = \frac{A(A+1)}{2}$.
Analoog voor de stappen naar links: $L = \sum j_l - i_l$.
Nu is $\sum j_r + \sum j_l$ *gekend, omdat dit de stappen van* 1 *tot* n *waren.*
Hieruit volgt $A^2 = \frac{2A(2A+1)}{2} - 2\frac{A(A+1)}{2} = L + R = 2L$, *we hebben dus dat* $2|A$ *zodat* $4|n$.
Iedere viervoud werkt, laat de danser volgend patroon herhalen: $links, rechts, rechts, links$.
Hij scoort dan $2(k-1), 2k$ *punten naar links wanneer hij voor de* k^{de} *keer ons patroon danst en* 2 *keer* $2k-1$ *punten naar rechts.*
Dit toont aan dat er op zijn minst 1 *mogelijkheid is voor een tango als* $4|n$.
Antwoord: (a) *is enkel mogelijk als* $4|n$.
(B) *We gaan verder met onze veronderstelling uit het ongerijmde dat er geen kus is:*
De linkse speler 1 *zijn score voor links zal lager liggen dan de rechtse speler* 2*' linkse score.*
Dit omdat bij de j^{de} *stap* $(n > j > 0)$*, het aantal stappen naar rechts lager lag bij speler* 1 *dan bij speler* 2 *omdat speler* 2 *rechts van speler* 1 *is.*
Dat betekent dat de i^{de} *stap naar links eerder was bij speler* 1 *dan bij speler* 2 *en speler* 1 *bij die stap minder punten scoorde.*
Op het einde geeft speler 1 *minder linkse punten* P_{l1} *dan speler*2 *linkse punten had* $(P_{l2}$.
Analoog scoort speler 1 *meer punten naar rechts of* $P_{r1} > P_{r2}$
Omdat de linkse en rechtse scores gelijk zijn, geldt $P_{r2} = P_{l2}$ *voor speler* 2*, zodat* $P_{r1} > P_{r2} = P_{l2} > P_{l1}$ *waardoor* $P_{r1} > P_{l1}$*, contradictie omdat speler* 1 *ook de derde regel volgde.*
Uit de contradictie weten we dat onze veronderstelling fout was en de b-vraag toch klopte.

stap 4: controle

We controleren onze oplossing aandachtig en kijken of er geen fouten ingekropen zijn, tot slot zorgen we dat de oplossing ergens duidelijk staat en geen belangrijke stappen vergeten op te schrijven waren.

Bij functievergelijkingen bevoorbeeld moet de oplossing duidelijk gecontroleerd staan (zonder de benaming triviaal).

Polya's Problem Solving Techniques

In 1945 George Polya published the book *How To Solve It* which quickly became his most prized publication. It sold over one million copies and has been translated into 17 languages. In this book he identifies four basic principles of problem solving.

Polya's First Principle: Understand the problem

This seems so obvious that it is often not even mentioned, yet studens are often stymied in their efforts to solve problems simply because they don't understand it fully, or even in part. Polya taught teachers to ask students questions such as:

- Do you understand all the words used in stating the problem?
- What are you asked to find or show?
- Can you restate the problem in your own words?
- Can you think of a picture or diagram that might help you understand the problem?
- Is there enough information to enable you to find a solution?

Polya's Second Principle: Devise a plan

Polya mentions that there are many reasonable ways to solve problems. The skill at choosing an appropriate strategy is best learned by solving many problems. You will find choosing a strategy increasingly easy. A partial list of strategies is included:

- Guess and check
- Make an orderly list
- Eliminate possibilities
- Use symmetry
- Consider special cases
- Use direct reasoning
- Solve an equation
- Look for a pattern
- Draw a picture
- Solve a simpler problem
- Use a model
- Work backwards
- Use a formula
- Be ingenious

Polya's Third Principle: Carry out the plan

This step is usually easier than devising the plan. In general, all you need is care and patience, given that you have the necessary skills. Persist with the plan that you have chosen. If it continues not to work discard it and choose another. Don't be misled, this is how mathematics is done, even by professionals.

Polya's Fourth Principle: Look back

Polya mentions that much can be gained by taking the time to reflect and look back at what you have done, what worked, and what didn't. Doing this will enable you to predict what strategy to use to solve future problems.

So starting on the next page, here is a summary, in the master's own words, on strategies for attacking problems in mathematics class. This is taken from the book, *How To Solve It*, by George Polya, 2nd ed., Princeton University Press, 1957, ISBN 0-691-08097-6.

1. UNDERSTAND THE PROBLEM

   - **First.** You have to *understand* the problem.
   - What is the unknown? What are the data? What is the condition?
   - Is it possible to satisfy the condition? Is the condition sufficient to determine the unknown? Or is it insufficient? Or redundant? Or contradictory?
   - Draw a figure. Introduce suitable notation.
   - Separate the various parts of the condition. Can you write them down?

2. DEVISING A PLAN

   - **Second.** Find the connection between the data and the unknown. You may be obliged to consider auxiliary problems if an immediate connection cannot be found. You should obtain eventually a *plan* of the solution.
   - Have you seen it before? Or have you seen the same problem in a slightly different form?
   - *Do you know a related problem?* Do you know a theorem that could be useful?
   - *Look at the unknown!* Try to think of a familiar problem having the same or a similar unknown.
   - *Here is a problem related to yours and solved before. Could you use it?* Could you use its result? Could you use its method? Should you introduce some auxiliary element in order to make its use possible?
   - Could you restate the problem? Could you restate it still differently? Go back to definitions.
   - If you cannot solve the proposed problem, try to solve first some related problem. Could you imagine a more accessible related problem? A more general problem? A more special problem? An analogous problem? Could you solve a part of the problem? Keep only a part of the condition, drop the other part; how far is the unknown then determined, how can it vary? Could you derive something useful from the data? Could you think of other data appropriate to determine the unknown? Could you change the unknown or data, or both if necessary, so that the new unknown and the new data are nearer to each other?
   - Did you use all the data? Did you use the whole condition? Have you taken into account all essential notions involved in the problem?

3. CARRYING OUT THE PLAN

- **Third.** *Carry out* your plan.
- Carrying out your plan of the solution, *check each step*. Can you see clearly that the step is correct? Can you prove that it is correct?

4. LOOKING BACK

- **Fourth.** *Examine* the solution obtained.
- Can you *check the result*? Can you check the argument?
- Can you derive the solution differently? Can you see it at a glance?
- Can you use the result, or the method, for some other problem?

2 combinatoriek + algemene problem-solving

2.1 basis

dubbeltellen

Men kan bepaalde eigenschappen combinatorisch bekijken om eigenschappen elegant te bewijzen.

* Het is belangrijk dat men dus de klassieke formules om te tellen kent:

k elementen in volgorde plaatsen met keuze uit n elementen kan op $\frac{n!}{(n-k)!}$ manieren,

indien elementen meerdere keren mogen voorkomen, hebben we n^k manieren om rijen te vormen van k elementen

indien de volgorde niet belangrijk is, hebben we $\binom{n}{k} = \frac{n!}{(n-k)!k!}$ manieren om k elementen te selecteren uit n waarden

het aantal permutaties van een set $\{a_1, a_2, \cdots, a_s\}$ waarbij a_i k_i keer voorkomt en er in totaal n elementen zijn, is gelijk aan $\frac{n!}{k_1!k_2!\cdots k_s!}$

Deze linken op 2 manieren aan een zelfde probleem, kan leiden naar een contradictie of een ongelijkheid.

extremaalprincipe

Men bekijkt het kleinste of grootste element van een verzameling en door naar bepaalde eigenschappen te kijken of bewerkingen uit te voeren,

zien we dat er een groter/ kleinere waarde is, zodat ons extremum fout is, waardoor er ∞ veel elementen waarden zijn OF de vraag onmogelijk is.

We kunnen een oneindige afdaling doen bvb. om te zien dat er geen enkele waarde is die voldoet.

identiteiten

$a^3 + b^3 + c^3 - 3abc = (a + b + c)(a^2 + b^2 + c^2 - ab - ac - bc)$

$n^4 + 4x^4 = (n^2 + 2x^2 - 2xn)(n^2 + 2x^2 + 2xn)$ (identiteit van Sophie-Germaine)

$\binom{m}{n} \equiv \prod_{i=0}^{i=k} \binom{m_i}{n_i} \pmod{p}$ met

$m = m_k p^k + m_{k-1} p^{k-1} + \cdots + p_0$ en $n = n_k p^k + \cdots + n_0$ (indentiteit/stelling van Lucas)

$ap + bq$ met $ggd(p, q) = 1$ en $a, b \in \mathbb{N}$ kan alle waarden groter dan $pq - q - p$ aannemen, $pq - p - q$ is de grootste waarde die niet zo te schrijven is. (postzegelidenteit)

$n! \sim \sqrt{2\pi n} \cdot (\frac{n}{e})^n$ (stelling/ identiteit van Stirling) [vrij juist voor grote waarden]

$\binom{2p-1}{p-1} \equiv 1 \pmod{p^3}$ en $\binom{ap}{bp} \equiv \binom{a}{b} \pmod{p^2}$ (Wolstenholme's)

Het principe van inclusie exclusie (PIE)
Zij $A1, A2, \cdots, A_n$ eindige verzamelingen. Dan geldt
$|A_1 \cup A_2 \cup ... \cup A_n| =$
$|A_1| + |A_2| + \cdots + |A_n|$
$-|A_1 \cap A_2| - |A_1 \cap A_3| - \cdots - |A_{n-1} \cap A_n|$
$+|A_1 \cap A_2 \cap A_3| + |A_1 \cap A_2 \cap A_4| + \cdots + |A_{n-2} \cap A_{n-1} \cap A_n|$
$\cdots$
$+(-1)^{n+1}|A_1 \cap A_2 \cap \cdots \cap A_n|$

Voorbeeld 2.1. *Bepaal alle oplossingen voor* $x^5 + 2y^5 = 4z^5$ *in* $\mathbb{Z}$.
Bewijs. We zien dat $(0, 0, 0)$ een oplossing is en er niet exact 1 of 2 elementen 0 kan zijn.
We bekijken nu een andere oplossing (x, y, z)
We merken op dat x even is, maar dan is $4|4z^5 - x^5$ zodat $4|2y^5$ en dus $2|y$.
Vervolgens geldt dat $32|LL$ en dus moet het rechterlid dit ook zijn, zodat $2|z$.
$(\frac{x}{2}, \frac{y}{2}, \frac{z}{2})$ is dus ook een oplossing, analoog wordt $(\frac{x}{2^n}, \frac{y}{2^n}, \frac{z}{2^n})$ ook een oplossing.
De 3 getallen worden dus steeds kleiner met oneindige afdaling, maar dan kunnen ze niet geheel blijven waaruit we concluderen dat er geen andere oplossng was.

□

1. Zij n een oneven natuurlijk getal groter dan 1 en $c_1, c_2, \ldots, c_n$ gehele getallen.

   Voor iedere permutatie $a = (a_1, a_2, \ldots, a_n)$ van $\{1, 2, \ldots, n\}$ definiren we $S(a) = \sum_{i=1}^{n} c_i a_i$.

   Bewijs dat er - voor iedere a - een permutatie $b \neq a$ van $\{1, 2, \ldots, n\}$ bestaat zodat $n!$ een deler is van $S(a) - S(b)$.

   link

2. We definieren $p(n)$ als het aantal manieren om n te schrijven als de som van positieve getallen $\in \mathbb{N}$. Bewijs dat $p(n) - p(n-1)$ gelijk is aan het aantal manieren om n te schrijven als een som van positieve getallen $\in \{2, 3, 4, 5 \cdots\}$.

   link

3. Definieer

   $$a_n = \frac{n^2 + 1}{\sqrt{n^4 + 4}}$$

   voor $n = 1, 2, 3, \ldots$ en stel $b_n = a_1 a_2 \ldots a_n$. Toon aan dat

   $$b_n = \frac{\sqrt{2n^2 + 2}}{\sqrt{n^2 + 2n + 2}}$$

   en dat

   $$\frac{1}{(n+1)^3} \leq \frac{b_n}{\sqrt{2}} - \frac{n}{n+1} \leq \frac{1}{n^3}.$$

   link

4. Zij T de verzameling van alle geordende drietallen (p, q, r) van natuurlijke getallen. Vind alle functies $f : T \to \mathbb{R}$ zodat

   $$f(p,q,r) = \begin{cases} 0 \text{ als } pqr = 0, \\ 1 + \frac{1}{6}(f(p+1, q-1, r) + f(p-1, q+1, r) \\ +f(p-1, q, r+1) + f(p+1, q, r-1); \\ +f(p, q+1, r-1) + f(p, q-1, r+1)) \text{ als } pqr \neq 0. \end{cases}$$

   link

5. Beschouw de verzameling $S = \{1, 2, \cdots, 280\}$. Bepaal het kleinste positieve gehele getal n met de eigenschap dat in elke uit n elementen bestaande deelverzameling van S er 5 elementen te vinden zijn die paarsgewijs onderling ondeelbaar zijn. link

6. $n, k \in \mathbb{N}$ en S is een verzameling van n punten in een vlak.

   Dit op zo'n wijze dat er geen 3 punten collineair zijn en ieder punt A minimum k punten heeft die op een zelfde afstand van A liggen.

   Bewijs dat $k < 0.5 + \sqrt{2n}$

   link

2.2 bedekkingen

Soms wordt in een vraag de mogelijkheid om iets te bedekken gevraagd.

Door een kleuring te gebruiken (enkele vakken in groepen verdelen) en eigenschappen zoals de pariteit te bekijken per object,
probeert men te bewijzen dat het al dan niet kan.

Voorbeeld 2.2. *Bewijs dat een $m*n$rechthoek enke volledig bedekt kan worden met $d*1$-blokken als $d|m$ en/ of $d|n$.*

Bewijs. Het is triviaal dat $d|mn$ moet gelden.

Bekijk het rechthoek als de unie van roosterpunten, zodat $(1,1),(1,n),(m,n),(m,1)$ de hoekpunten zijn.

Kleur (i,j) met een kleur $t \equiv i+j \pmod{d}$ zodat $t \in \{1,2,\cdot,d\}$ zit.

Het is duidelijk dat ieder $d*1$ blokje nu ieder kleur exact 1 keer bedekt.

Er moet dus gelden dat ieder kleur even vaak voorkomt, wat niet zo is:

Zij $m = kd+p$ en $n = dl+q$, dan bevat het $m*dl$ bord ieder kleur even vaak, alsook het $kd*q$-bord.

Het overige $p*q$ bord kan onmogelijk ieder kleur even vaak hebben.
Omdat $d|pq$ moeten $ggd(d,q)$ en $ggd(d,p) > 1$ zodat het volgende geldt:

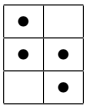
Als $p+q \leq d$ komt het kleur d er niet voor of slechts 1 keer terwijl $pq > d$.
Als $p+q > d$ komt het kleur d er $q+p+1-d$ keer voor, terwijl kleur $d-1$ er $q+p+2-d$ keer voorkomt.
Contradictie en dus komt niet ieder kleur even vaak voor en is er geen bedekking mogelijk.

□

1. We hebben een $9 * 9$schaakbord waarvan er 46 vakjes gekleurd worden in't rood, bewijs dat er een $2 * 2$vierkant kan worden gevonden waarvan er minimum 3 rood zijn. link

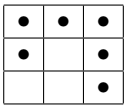
2. We verdelen een vierkant in enkele rechthoeken die heel het vierkant bedekken zonder overlapping. Bewijs dat als iedere lijn // met een zijde van het vierkant het binnenste van een rechthoek snijdt, niet alle rechthoeken aan de omtrek van het vierkant grenzen. link

3. Op een $10 * 10$ bord wordt een gebied van $4n$ vakjes gekleurd. Er geldt dat het kan bedekt worden met n $2 * 2$-vierkantjes. Bij een andere bedekking, kan men er n van volgende tetramino's plaatsen:



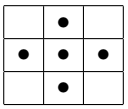
   Wat is de minimale waarde die n kan aannemen? Kan je alle mogelijke waarden van n vinden?link

4. Bepaal alle $m \times n$ rechthoeken die kunnen bedekt worden met "haken" zoals aangegeven in de figuur, die bestaan uit 6 eenheidsvierkanten (zie figuur met de 6 gevulde vierkantjes). Rotaties en spiegelingen van haken is toegelaten. De rechthoek dient bedekt te zijn zonder gaten noch overlappingen en geen enkel stuk van een haak mag buiten de rechthoek vallen.



   link

5. Bepaal het maximaal aantal kruisjes



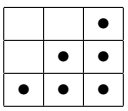
   dat men kan plaatsen in een $10 * 11$ rechthoek. link

6. We hebben een $999 * 999$ bord waarop we een stuk plaatsen die op de volgende manier beweegt: het kan naar een ander vlakje gaan van het bord als de 2 vierkanten een zijde gemeenschappelijk hebben en iedere stap staat loodrecht op de vorige (men stapt dus nooit in 1 keer over een $1 * 3$rechthoek). Men wil een zo lang mogelijke cyclus maken,waarbij men ieder vlakje slechts 1 keer bewandelt en eiindigt bij het eerste vlakje (een gesloten kring dus). Hoeveel vakken kan die cyclus maximaal bevatten?

   link

7. Een trapvormige doos met 3 treden van breedte 2 is gemaakt uit 12 eenheidskubusjes. (zoals de 6 bolletjes, maar met dikte 2)
   Bepaal alle natuurlijke getallen n waarvoor het mogelijk is om een kubus met zijde n te maken als je slechts beschikt over dergelijke bouwstenen.

   link



8. We hebben een $2011 * 2011$-tafel die we bedekken met een eindig aantal $52 * 52$ doeken .
   In ieder van de 2011^2 vakjes schrijven we het aantal doeken die weer op plaatsten.
   Wat is het grootste aantal gelijke cijfers $\neq 0$ dat we kunnen hebben? link

2.3 inductie

Bij inductie wordt een uitdrukking bewezen voor alle natuurlijke getallen vanaf k, dit door het te bewijzen voor k. (inductiebasis IB)
Vervolgens als het geldt voor n, bewijst men dat het ook geldt voor $n+1$.
Bij volledige inductie bewijst men bij de tweede stap dat de vraag geldt voor $n+1$ als het waar was $\forall i \in \{k, k+1 \cdots, n\}$

Voorbeeld 2.3. *(kleine stelling van Fermat) Er geldt dat* $n^p \equiv n \pmod{p}$ *als* p *priem is* $\forall n \in \mathbb{N}$.

Bewijs. Voor $n = 0$ en 1 is de vraag de triviaal. (IB)
Als de vraag geldt voor n, bewijzen we dat $(n+1)^p \equiv n+1 \pmod{p}$.
Lemma: Er geldt dat $p|\binom{p}{i}$ als $0 < i < p$ omdat $i!, (p-i)!$ geen factoren p bevatten.
$(n+1)^p = n^p + 1 + \sum_{i=1}^{p-1} \binom{p}{i} n^i \equiv n^p + 1 \equiv n+1 \pmod{p}$ door de inductiebasis en ons lemma.
Met inductie geldt de stelling van Fermat nu voor alle getallen $n \in \mathbb{N}$.

□

vb.'en

1. $n \in \mathbb{N}$, bepaal alle getallen x waarvoor geldt dat

$$0 = 1 + x + \frac{x(x+1)}{2!} + \cdots + \frac{x(x+1)\cdots(x+n)}{(n+1)!}$$

link

2. k is een vast natuurlijk getal.
Een winkelketen wil zoveel mogelijk sombrero's verkopen.
Iedere klant kan 2 anderen een sombrero doen kopen na zijn aankoop (het telt niet als de klant door iemand anders al overhaald was).
Iedere klant die zo (direct of via keten) minimum k personen een sombrero liet kopen, wint een DVD. Bewijs dat wanneer men n sombrero's verkocht, men maximaal $\frac{n}{k+2}$ DVD's geeft moeten weggeven.

link

3. $n > 0$ is een natuurlijk getal .
Op een balans willen we gewichten van $2^0, 2^1, \cdots 2^{n-1}$ kilo plaatsen zodat ieder gewicht elk op zijn beurt wordt geplaatst op zo'n wijze dat de rechtste schaal nooit zwaarder weegt dan de linkse.
Hoeveel manieren zijn er hiervoor?

link

4. Bewijs dat $\forall n \in \mathbb{N}_0$ de verzameling $\{2, 3 \cdots, 3n+1\}$ verdeeld kan worden in n triplets die de zijden van een stompe driehoek zijn.

link

5. Voor elk positief geheel getal n wordt S(n) als volgt gedefinieerd: S(n) is het grootste positieve gehele getal zodanig, dat voor elk natuurlijk getal $k \leq S(n)$ het getal n^2 te schrijven is als de som van k kwadraten van positieve gehele getallen. (a) Bewijs dat $S(n) \leq n^2 - 14$ voor alle $n > 4$. (b) Bepaal een getal n waarvoor geldt $S(n) = n^2 - 14$. (c) Bewijs dat er oneindig veel positieve gehele getallen zijn waarvoor geldt $S(n) = n^2 - 14$.

link

6. Zijn $A_1, \ldots, A_n$ twee aan twee verschillende deelverzamelingen van $\{1, 2, \ldots, n\}$. Bewijs dat er een element $x \in \{1, 2, \ldots, n\}$ is zodat de verzamelingen $A_i \setminus \{x\}$ allemaal verschillend zijn.

link

7. In het gecoordinatiseerde vlak beschouwt men een eindige verzameling roosterpunten V .
Is het mogelijk alle punten van V met 1 van beide kleuren, rood of wit, te kleuren zo, dat aan de volgende voorwaarde is voldaan: voor elke rechte D evenwijdig aan 1 van de coordinaatassen is de absolute waarde van het verschil tussen het aantal rode punten en het aantal witte punten dat op D ligt kleiner dan of gelijk aan 1.

link

8. Bepaal alle gehele getallen $n > 0$ waarvoor er natuurlijke getallen $a_1, a_2 \cdots a_n$ bestaan zodat

$$\frac{1}{2^{a_1}} + \frac{1}{2^{a_2}} + \cdots + \frac{1}{2^{a_n}} = \frac{1}{3^{a_1}} + \frac{2}{3^{a_2}} + \cdots + \frac{n}{3^{a_n}}.$$

link

2.4 invariantie en contradictie

Wanneer men wil bewijzen dat iets niet kan bij een combinatorische vraag, zijn er enkele manieren die vaak werken:

I Men zegt vanuit het ongerijmde dat de vraag wel kan opgelost worden en door de eigenschappen van de oplossing te bekijken,
bekomt men een contradictie waardoor er geen oplossing kon zijn (het ongerijmde was fout)

II Men bekijkt een eigenschap die invariant is in de vraag, waarbij die eigenschap bij de start en het einde verschillend is, waaruit volgt dat we het einde nooit kunnen bereiken.

III Men gebruikt een eigenschap die monotoon is bij iedere stap met een minimaal verschil, wanneer de eigenschap begrensd is, zijn er slechts een eindig aantal oplossingen.

opmerking Natuurlijk moet je opletten als je blijft proberen te bewijzen dat het niet werkt, dat er niet gewoon wel een oplossing was.

Voorbeeld 2.4. *We hebben de getallen van* 1 *tot* 2012^9 *in een pot gestoken.*
Iedere keer als we 2 *getallen* x *en* y *eruit halen, worden ze vervangen door* $(x-1006)(y-1006)+1006$ *en steken dit ene getal terug in de pot.*
Welk getal kunnen we vinden als er slechts 1 *getal meer in de pot zit?*

Bewijs. Merk op dat het getal 1006 en x vervangen wordt door 1006 en dit getal invariant blijft (in de pot terug wordt gestoken).
Dit getal blijft dus in de pot en zal het laatste getal 1006 zijn.

□

We beginnen eens extra vroeg aan de oefeningen:

1. 5 Lege emmers met een inhoud van 2 liter staan op de hoekpunten van een regelmatige vijfhoek. Assepoester en haar boze stiefmoeder voeren afwisselend een stap uit.
   De stiefmoeder mag bij haar stap telkens 1 liter water verdelen over de 5 emmers zoals ze wil.
   Daarna mag Assepoester 2 emmers die naast elkaar staan kiezen en ledigen.
   Indien een emmer kan overlopen door de stiefmoeder wint ze, in het andere geval Assepoester, wie heeft een winnende strategie? link

2. Men heeft 6 dozen B_1 tot B_6 met in elke doos 1 munt. Men voert hierop handelingen uit, 2 soorten zijn toegestaan:

   -handeling type 1: men neemt uit doos B_i 1 munt en legt er 2 in doos $B_(i+1)$ (dit kan voor $i \in \{1,2,3,4,5\}$) -handeling type 2: men neemt uit doos B_i1 munt en wisselt dan de inhoud van dozen $B_(i+1)$ en $B_(i+2)$ (dit kan voor$i \in \{1,2,3,4\}$)

   Is het mogelijk na een eindig aantal handelingen de dozen B1 tot B5 leeg te krijgen en in doos B6 $2010^{2010^{2010}}$ munten.

   Opmerking: in deze internationale competitie is $a^{b^c} = a^{(b^c)}$ ter verduidelijking.link

1. Zij $r \geq 2$ een vast natuurlijk getal, en zij F een oneindige familie van verzamelingen, allemaal van grootte r, en geen twee ervan zijn disjunct. Bewijs dat er een verzameling bestaat van grootte $r-1$ die iedere verzameling uit F snijdt.link

2. We hebben n dubbelkleurige stenen die we plaatsen in een rij. Bij de start wijzen ze allen met hun witte kant naar boven en in iedere stap kiezen we een witte steen met 2 buren , waarna we de buren omdraaien en de witte steen wegnemen. (de stenen zijn zwart-wit bvb) Bewijs dat we $n-2$ beurten lang kunnen spelen aesa 3 $\not| n-1$. link

3. Een verzameling A van gehele getallen noemen we somvol als elk element $a \in A$ de som is van twee (niet noodzakelijk verschillende) elementen $b, c \in A$. Een verzameling A van gehele getallen noemen we nulsomvrij als 0 het enige gehele getal is dat niet te schrijven is als de som van de elementen van een eindige, niet-lege deelverzameling van A. (waarbij ieder element van A max 1 keer mag voorkomen) Bestaat er een somvolle, nulsomvrije verzameling van gehele getallen?link

4. Aan elk hoekpunt van een regelmatige vijfhoek wordt een geheel getal toegevoegd zo, dat de som van deze vijf getallen strikt positief is.
   Indien aan drie opeenvolgende hoekpunten respectievelijk de getallen x, y en z toegevoegd zijn waarbij $y < 0$, dan is de volgende bewerking toegelaten: het drietal (x, y, z) wordt vervangen door het drietal $(x+y, -y, y+z)$.
   Deze bewerking wordt hernomen zolang ten minste 1 van de 5 getallen strikt negatief is.
   Bepaal of deze procedure noodzakelijkerwijs stopt na een eindig aantal van deze bewerkingen.link

5. Een stapel van n kiezelsteentjes wordt in een vertikale kolom geplaatst. Deze configuratie is volgens de volgende regels opgesteld. Een kiezelsteentje kan verplaatst worden als het bovenaan een kolom ligt die minstens twee kiezelsteentjes meer bevat dan de kolom rechts ervan (als daar geen kiezelsteentjes liggen, mag je dit beschouwen als een kolom met 0 kiezelsteentjes). In iedere fase kies je een kiezelsteentje dat beweegbaar is, en verplaats je het naar de top van de kolom rechts ervan. Als er geen kiezelsteentjes meer verplaatst kunnen worden, noemen we dat een finale configuratie. Voor iedere n, toon aan dat, ongeacht welke keuzes er gemaakt worden in iedere fase, de finale configuratie uniek is. Omschrijf deze configuratie in termen van n. link

6. We hebben een eindig aantal punten in het vlak waarvan er geen 3 collineair zijn, die we met S noteren. Een windmolen is een proces dat begint met een rechte l die door 1 punt $P \in S$ gaat. De lijn draait et de klok meer om het draaipunt P to er voor't eerst een ander pnt van S op deze lijn ligt, dat nieuwe punt wordt het nieuwe draaipunt. We zeggen dat Q een klap van de molen krijgt. De lijn draait nu met de klok mee om Q en de windmolen draait zo oneindig door. Laat zien dat we punt P van S en een lijn l door P kunnen kiezen zodat er een windmolen ontstaat waarbij elk punt van S ∞ veel klappen van de molen krijgt. link

7. $\forall n \in \mathbb{N}$ wordt $h(n)$ voor $n > 1$ berekend als volgt (r is het laatste cijfer van n):
   als $r = 0$ dan is $h(n) = \frac{n}{10}$
   als $r > 0$ wordt R het grootste getal rechts met enkel digits $\geq r$ en L de rest.
   $h(n) = L\overbrace{R-1}\overbrace{R-1}$ (het getal door L te laten volgen door 2 keer $R-1$)
   Bewijs dat er een k bestaat zodat $h^k(n) = 1$.link

2.5 duivenhokprincipe (DVH-principe)

Zijn $n, k \in \mathbb{N}_0$.
Als men n duiven verdeelt over k duivenhokken, dan bestaat er een duivenhok dat minstens $\lfloor \frac{n-1}{k} \rfloor + 1$ duiven bevat.

Voorbeeld 2.5. *Binnen een cirkel met straal 16 liggen 650 gegeven punten.*
Definieer een ring als het vlakdeel dat begrepen is tussen twee concentrische cirkels met stralen 2 en 3 respectievelijk.
Bewijs dat men een ring kan plaatsen zodat minstens 10 van de 650 punten bedekt worden door deze ring.

Bewijs. Maak de cirkel met straal 16 nog iets groter tot een straal van 16.
Teken rond ieder van de 650 punten een ring.
De som van de oppervlakten van de ringen is $650 * 5\pi = 3250\pi$ en deze liggen allen in de cirkel met oppervlakte 361π.
Toevallig is $3250 = 9 * 361 + 1$ zodat er wegens 't DVH-principe een punt is dat in 10 ringen ligt.
Wanneer men nu een ring legt met centrum dat punt, waren er min. 10 centra van die 650 punten op een afstand tussen 2 en 3 zodat ze op onze geplaatste cirkel liggen.
Hiermee is het gevraagde bewezen.

□

Terug een oefenreeks met een Nederlandstalige bijlage bij problemen

1. Bewijs dat iedere deelverzameling met 55 elementen uit $\{1, 2, \cdots, 100\}$ minimum 2 elementen heeft met verschil 9.link

2. (i) 15 stoelen worden geplaatst rond een cirkelvormige tafel met bijhorende naamkaartjes voor de 15 gasten. De gasten merken deze kaartjes echter niet op, en het blijkt dat niemand op de juiste plaats zit. Toon aan dat de tafel zodanig gedraaid kan worden dat er tegelijkertijd twee mensen wel op de juiste plaats zitten. (ii) Geef een voorbeeld van een schikking waarbij 1 iemand op de juiste plaats zit en elke rotatie van de tafel ervoor zorgt dat er maximum 1 op de juiste plaats zit.

   link

3. Gegeven zijn tien verschillende positieve gehele getallen van twee cijfers.
   Bewijs dat men ze kan verdelen over twee verzamelingen zo, dat de som van de getallen in de ene verzameling gelijk is aan de som van de getallen in de andere verzameling.link

4. Op een $n*n$-bord worden alle getallen uit $\{1, 2, \cdots, n^2\}$ geplaatst op een apart vakje. Bewijs dat er 2 vakjes zijn die met een zijde aan elkaar grenzen zodat het verschil tussen hun waarden $\geq n$ is. link

5. Bart en Ria spelen het volgende spel. Bart schrijft n verschillende natuurlijke getallen op een papier, met n een natuurlijk getal. Ria mag enkele van deze getallen wegstrepen (ze mag er ook geen enkel wegstrepen, maar ze mag ze zeker niet allemaal wegstrepen). Daarna mag Ria voor elk van de overblijvende getallen een + of een zetten en de som bepalen van de getallen die ze op deze manier bekomt. Als deze som deelbaar is door 2003 dan wint Ria. In het andere geval wint Bart.
   Voor welke waarden van n heeft Bart een winnende strategie, en voor welke waarden van n heeft Ria een winnende strategie? link

6. Gegeven is een verzameling M van 1985 verschillende gehele getallen groter dan nul. Geen van die getallen heeft een priemdeler groter dan 26.
   Bewijs dat M vier getallen bevat waarvan het product de vierde macht is van een geheel getal. link

7. 6 Vragen werden gesteld op een IMO, zodat ieder paar van problemen door meer dan 40 Niemand kon alle vragen.
   Bewijs dat er minimum 2 deelnemers 5 vragen oplosten. Geldt dit ook als we ≥ 0.4 hebben?
   link

8. $x_1, x_2, \ldots, x_n$ zijn reeele getallen zodat $x_1^2 + x_2^2 + \ldots + x_n^2 = 1$. Bewijs dat, $\forall k \in \mathbb{N}$ zodat $k \geq 2$, er gehele getallen $a_1, a_2, \cdots a_n$ bestaan, niet allen gelijk aan 0, zodat $|a_i|$ $lek1$ voor alle i en zodat
   $$|a_1x_1 + a_2x_2 + \ldots + a_nx_n| \leq \frac{(k-1)\sqrt{n}}{k^n - 1}.$$
   link

9. 21 meisjes en 21 jongens deden mee in een wiskunde-competitie. Het bleek dat iedere deelnemer maximum zes problemen had opgelost, en voor iedere 2 deelnemers bestaande uit een jongen en een meisje, was er minstens n probleem dat zowel door de jongen als het meisje was opgelost. Toon aan dat er een probleem was dat opgelost werd door minstens drie jongens en drie meisjes.

   link

10. Zij p een oneven priemgetal en n een natuurlijk getal. In het cordinatenvlak liggen er acht verschillende punten met gehele cordinaten op een cirkel met diameterlengte p^n. Bewijs dat er een driehoek bestaat met zijn hoekpunten onder de gegeven punten zodanig dat de kwadraten van de lengtes van zijn zijden natuurlijke getallen zijn die deelbaar zijn door p^{n+1}.

    link

2.6 winnende strategieen

Bij een vraag moeten we bewijzen dat iemand een winnende strategie heeft bij een spel,
dit kan door een kleuring of modulorekenen of andere elegante eigenschappen die worden uitgebuit.

stelling van Zermelo

Deze stelling zegt dat ieder spel tussen 2 personen waar toeval niet in meespeelt, geen gelijkspel mogelijk is en de spelers elk op hun beurt een zet doen:
1 van de 2 spelers geeft dan een winnende strategie.

Voorbeeld 2.6. *(QED-competitie)*
ALbert en Philip bestellen een zak met 2011 *frieten. Albert start met het eten van enkele frieten en eet om zijn beurt met Philip. Ze vorken* $1, 2, 5$ *frietjes op per keer. Degene die de laatste friet opeet, betaalt de rekening. Bewijs dat Philip zijn portemonnee kan laten zitten.*

Bewijs. Als Albert 1 frietje neemt, neemt Philip er 2.
Nam Albert er 2 of 5 neemt Philip er 1. Op die manier is er na Philip's beurt steeds een 3voud opgegeten, zodat hij niet de 2011^{de} friet at.

□

vb.'en

1. Het $Y2K$ spel wordt gespeeld op een 1×2000 rooster als volgt: twee spelers schrijven elk op beurt een S of een O op een leeg vakje. De eerste speler die eerst drie opeenvolgende vakjes kan maken met SOS wint het spel. Als alle vakjes gevuld zijn zonder SOS te produceren, dan eindigt het in een gelijkspel. Als de ene speler begint, bewijs dan dat de andere speler een winnende strategie heeft.link

2. Arne en Bart spelen op een 8×8 schaakbord volgend spel: beginnend met Arne kleuren ze om de beurt een nog niet gekleurd veld, Arne in het rood en Bart in het blauw. De winnaar is degene die als eerste een 2×2 vierkant helemaal in zijn eigen kleur kan kleuren. Bewijs dat Bart altijd kan voorkomen dat Arne wint. link

3. We beschouwen een polynoom $f(x) = x^4 + a_3x^3 + a_2x^2 + a_1x + a_0$.

   Albert Einstein en Homer Simpson spelen een spel waarbij ze om hun beurt 1 v.d. coefficienten $a_0, a_1, \ldots, a_4$ een waarde geven in de vorm t^n waarbij t op voorhand bepaald is en $n \in \mathbb{N}$. Albert start.
   Na 5 zetten is het spel gedaan, wanneer alle coefficienten ingevuld zijn.
   Als $f(x)$ een reele wortel geeft, wint Homer, indien het 4 complexe wortels zijn , wint Albert.
   Voor welke waarden van t geeft Albert een winnende strategie? link

4. Een rij van 2009 kaarten, die elk een gouden en een zwarte zijde hebben, ligt op tafel. Bij het begin liggen alle kaarten met hun gouden zijde naar boven. Twee spelers, spelen een spel waarbij afwisselend 50 opeenvolgende kaarten worden gekozen waarvan de meest linkse kaart met goud boven lag en draait hierbij die 50 kaarten om. Bepaal of dit spel altijd eindigt en wie er een winnende strategie heeft/ altijd kan winnen? link

5. We beschouwen een polynoom $f(x) = x^{2012} + a_{2011}x^{2011} + \cdots + a_1x + a_0$.
   Albert Einstein en Homer Simpson spelen een spel waarbij ze om hun beurt 1 v.d. coefficienten $a_0, a_1, \ldots, a_{2011}$ een waarde geven. Albert start.
   Na 2012 zetten is het spel gedaan, wanneer alle coefficienten ingevuld zijn.
   Homer wil dat $f(x)$ deelbaar is door $m(x)$ en Einstein wint als dit niet zo is.
   (a) Wie kan winnen als $m(x) = x - 2012$?
   (b) Wat als $m(x) = x^2 + 1$? link

6. Het liegebeestspel wordt gespeeld door 2 spelers A en B.
   Bij de start kiest A natuurlijke getallen x, N met $1 \leq x \leq N$ en zegt enkel de waarde N aan B.
   Speler B mag nu enkel vragen stellen door een set S te geven aan A en vragen of x in S zit. (hij mag meerdere keren de zelfde verzameling geven)
   A antwoordt met ja of nee , maar mag liegen op zo'n wijze dat tussen iedere $k+1$ opeenvolgende vragen, hij minstens 1 keer eerlijk antwoordde.
   B mag zoveel vragen (eindig natuurlijk) stellen en moet dan een set X geven met n gehele getallen. Als $x \in X$ wint B en anders verliest hij.
   Bewijs dat
   1 Als $n \geq 2^k$, B een winnende strategie heeft.
   2 Als k groot genoeg is, er is een natuurlijke $n \geq 1.99^k$ zodat B geen winnende strategie heeft. link

2.7 meetkunde binnen de combinatoriek

Het gebeurt vaak dat er interessante vraagjes over een meetkundige constructie plaats vindt op een grote olympiade.

Men kijkt naar specifieke eigenschappen die vaak logisch zijn en simpel te bewijzen zijn in een lemma.

Een logische stelling is de volgende

Stelling 2.7. *(tapijtenstelling)*
Wanneer enkele tapijten die samen de oppervlakte van de kamer hebben gelegd worden, is de oppervlakte die dubbel gelegd werd, gelijk aan de oppervlakte die niet bedekt werd.
Indien iets $m > 2$ keer belegd werd, moet je wel $m - 1$ keer de oppervlakte rekenen.

Enkele goede voorbeelden zijn belangrijk om het principe te verstaan:

1. $ABCD$ is een vierhoek zodat M, N de middens zijn van $[AD], [BC]$.
   BM, CM, AN, DN snijden de vierhoeken in 7 delen, het deel dat grenst aan $[AB], [CD]$ is II, III en de vierhoek in't midden heeft opp. I.
   Bewijs dat $I = II + III$. link

2. In het vlak liggen 100 punten, geen drie op n lijn. Beschouw alle mogelijke driehoeken met drie van deze punten als hoekpunten.
   Bewijs dat ten hoogste 70 procent van deze driehoeken scherphoekig is. link

3. Zij $n \geq 5$ een natuurlijk getal. Vind het grootste natuurlijk getal k zodat er een veelhoek bestaat met n hoekpunten (convex of niet, maar niet zelf-snijdend!) die k interne 90° hoeken heeft. link

4. Beschouw een vlak met Carthesiaans cordinatenstelsel. Voor ieder punt met gehele cordinaten, teken een schijf rond dit punt (met dat punt als midden dus) en straal 1/1000. (i) Bewijs dat er een gelijkzijdige driehoek bestaat met zijn drie hoekpunten in het inwendige van verschillende schijven. (ii) Toon aan dat iedere gelijkzijdige driehoek met zijn drie hoekpunten in het inwendige van verschillende schijven een zijde heeft met lengte meer dan 96. link

5. Zij $n \geq 3$ een natuurlijk getal. Zij $C_1, C_2, \ldots, C_n$ eenheidscirkels in het vlak, met middens $O_1, O_2, \ldots, O_n$ respectievelijk. Als geen enkele rechte meer dan twee van de cirkels snijdt, bewijs dan dat $\sum_{1 \leq i < j \leq n} \frac{1}{O_i O_j} \leq \frac{(n-1)\pi}{4}$. link

6. In een vlak hebben we n rechthoeken met parallelle zijden. De zijden van verschillende rechthoeken liggen op verschillende rechten. De grenzen van de rechthoeken snijden het vlak in verschillende (samenhangende) regio's. Men zegt dat een regio aangenaam is als die minstens n van de hoekpunten van de originele rechthoeken op zijn grenzen heeft. Bewijs dat de som van het aantal hoekpunten van alle aangename regio's minder is dan $40n$ (niet-convexe regio's zijn toegelaten, de rest van het vlak buiten je figuur telt ook als regio). link

7. Tien gangsters staan op een vlakke ondergrond, en de onderlinge afstanden tussen hen zijn allemaal verschillend.
   Om klokslag twaalf schiet iedere gangster de dichtstbijzijnde andere gangster neer. Wat is het minimumaantal vermoorde gangsters?
   Kan je trouwens ook als combinatoriekvraag bewijzen dat het maximum niet 10 is? link

8. Dertien vogeltjes strijken neer op een plat vlak zodanig dat er van elk vijftal vogeltjes minstens 4 op dezelfde cirkel zitten. Bewijs dat er een cirkel bestaat waar minstens 6 vogeltjes op zitten. Kan men dit resultaat nog scherper stellen? link

9. Ieder paar overstaande zijden van een convexe zeshoek heeft de volgende eigenschap: de afstand tussen hun middens is gelijk aan $\frac{\sqrt{3}}{2}$ keer de som van hun lengtes. Toon aan dat alle hoeken van de zeshoek gelijk zijn. link

10. Beschouw een convexe veelhoek P. Aan elke zijde b van P associren we de oppervlakte van de grootste driehoek met b als zijde, die volledig binnen P ligt. Bewijs dat de som van deze oppervlakten minstens dubbel zo groot is als de oppervlakte van P . link

2.8 grafentheorie

Een graaf bestaat uit een verzameling V van knopen of vertices, en uit een eindige verzameling E van zijden of edges. Alle zijden zijn paren van knopen. Twee knopen X, Y zijn verbonden als er zijde $\{X, Y\}$ (XY) is.
Hierbij geven we wat uitleg over de terminologie:

- Een propere graaf is een een graaf waarbij iedere zijde verbonden is met een andere zijde en geen enkel punt met zichzelf verbonden is.
- Een gerichte graaf is een graaf waarbij de zijden gericht zijn.
- Een complete graaf K_n is een propere graaf met n punten waarbij iedere 2 punten verbonden zijn met 1 zijde.
- Een k-partiete graaf is een graaf waarvan de set met de punten kan worden verdeeld in k disjuncte deelgrafen waarbij in iedere deelgraaf er geen enkele zijde is die 2 punten uit die deelverzameling verbindt.

  In een bipartite (letterlijk: "2-delige") graaf is de verzameling van knopen V opgesplits in twee delen: V_1 en V_2. De enige toegelaten zijden gaan verbinden knopen van V_1 met knopen van V_2. Er zijn dus geen zijden met twee knopen inV_1, noch met twee knopen in V_2.
- De graad van een knoop k (notatie $d(k)$ of $deg(k)$) is het aantal keren dat k een eindpunt is van een zijde, er geldt uiteraard dat $\sum d(x) = 2|E|$
- een pad is een opeenvolging van verschillende punten die met elkaar verbonden zijn via zijden, indien het start- en eindpunt a en b zijn, is de afstand $d(a, b)$ gelijk aan het aatal zijden bij het kortste pad van a naar b.
- indien bij een pad het eind- en beginpunt hetzelfde is, noemen we het een cyclus
- een graaf is verbonden als voor ieder paar punten er een pad is, die de 2 punten verbindt
- een verbonden graaf zonder cycels heet een boom, een boom heeft exact $n - 1$ zijden en minimum 2 punten met afstand 1.
- Een Hamiltoncykel is een cykel die ieder punt exact 1 keer bevat
- een Eulerpad is een pad waarbij alle zijden tot het pad behoren
  Een Eulercycel is een cycel die alle zijden aandoet.
- een planaire graaf is een graaf waarbij de zijden met lijnen kunnen worden getekend die elkaar niet snijden, zo'n graaf met n punten heeft maximaal $3n - 6$ punten

stellingen van Euler

Een simpele, samenhangende graaf bevat een Euler-cykel dan en slechts dan als de graad van alle knopen even is.

Een verbonden, planaire graaf met v knopen, e zijden en f gebieden (gebeid= deel van het vlak omringd door zijden), houdt zich aan de regel $v + f = e + 2$, dat geldt dan ook voor de convexe veelvlakken, waarbij f voor het aantal vlakken staat.

stelling van Turan

Als een simpele graaf met $n = t(p-1) + r$ punten met $0 \leq r < p-1$ meer dan

$$\frac{(p-2)n^2 - r(p-1-r)}{2(p-1)}$$

zijden heeft, bestaat er een K_p die een subgraaf is.

stelling van Hall Neem een bipartiete graaf met X,Y de subgrafen die inwendig geen zijden bevatten:

indien voor elke deelverzameling S van X, de elementen van S samen met minimum $|S|$ elementen van Y verbonden zijn, dan is er een matching die alle elementen van X matcht.

Hierbij is $|X| \geq |Y|$ en bedoelen we met een matching dat ieder punt van X met een uniek punt van Y verbonden is met een zijde (een punt in Y is maximaal 1 keer verbonden).

stelling van Kuratowskil

Een graaf is planair aesa het K_5 en $k_{3,3}$ niet bevat.

stelling van Ramsey

Deze stelling zegt dat een complete graaf met minimum $R_(n_1, \cdots, n_c)$ waarvan de zijden in c kleuren $1, 2 \cdots c$ gekleurd worden, dan bestaat er een complete subgraaf met n_i punten waarvan alle zijden in kleur i gekleurd zijn.

Gekende voorbeelden zijn $R(3,3) = 6, R(3,3,3) = 17$ en in het algemeen is de bovengrens ; $R(\underbrace{3,3,\cdots,3}_{k}) \leq \lfloor k!e \rfloor + 1$

Zij die dit te kort uitgelegd vonden of er meer over willen weten, kunnen deze bijlage doornemen .

1. Bewijs dat in een groep van 6 personen er 3 personen te vinden zijn die elkaar niet kennen of elkaar wel kennen.

   link

2. Bewijs dat als we een graaf beschouwen die een boom is, er een punt is die tot alle langste paden behoort.

   link

3. De volgende handeling is toegestaan op een eindige graaf: kies een willekeurige cykel van lengte 4 (als er zijn),
   en kies een willekeurig boog in die cykel, en verwijder die van de graaf.
   Voor een vast natuurlijk getal $n \geq 4$, vind het minimum aantal bogen van een graaf dat verwijderd kan worden door herhaaldelijk deze handeling uit te voeren op de complete graaf met n knopen.

   link

4. Zij n een even natuurlijk getal.
   Toon aan dat er een permutatie $x_1, x_2, \ldots, x_n$ van $1, 2, \ldots, n$ bestaat
   zodat voor iedere $1 \leq i \leq n$ het getal x_{i+1} is n van de getallen $2x_i, 2x_i - 1, 2x_i - n, 2x_i - n - 1$ (met $x_{n+1} = x_1$).

   link

5. Bekijk $n \geq 2$ punten op de omtrek van een cirkel met straal 1. Zij q het aantal lijnstukken met de eindpunten in die n punten, met lengte groter dan $\sqrt{2}$. Bewijs dat $3q \leq n^2$.

   link

6. De leden van een internationaal genootschap komen uit 6 verschillende landen. De ledenlijst bevat 1978 namen, genummerd van 1 tot en met 1978. Bewijs dat er ten minste 1 lid is wiens nummer gelijk is aan de som van de nummers van twee van zijn landgenoten of tweemaal zo groot als het nummer van 1 van zijn landgenoten.

   link

7. Voor een eindige graaf G, stel $f(G)$ gelijk aan het aantal driehoeken en $g(G)$ het aantal viervlakken gevormd door de bogen van de graaf G. Vind de kleinste constante c zodat

   $$g(G)^3 \leq c \cdot f(G)^4$$

   voor elke graaf G.

   link

8. Op een planeet hebben we 2^N landen met $N \geq 4$. Ieder land heeft een vlag met N stroken die naast elkaar liggen. Geen 2 landen hebben er eenzelfde vlag.
   Een verzameling van N vlaggen is divers als we ze kunnen leggen tot een $N * N$vierkant zodat alle N velden/stroken op de hoofddiagonaal dezelfde kleur hebben. Vind het kleinste aantal vlaggen dat we nodig hebben, zodat we er steeds N kunnen vinden die een diverse verzameling kunnen vormen. link

9. $P_1, \cdots P_s$ zijn rekenkundige rijen van gehele getallen. De volgende condities gelden: a) ieder geheel getal behoort tot minstens 1 rekenkundige rij b) iedere rij bevat minimum 1 getal dat geen enkele andere rij bevat Met n noteren we het kgv van de verschillen van de s rijen en schrijf $n = p_1^{a_1} \cdots p_k^{a_k}$ als priemfactorontbinding. Bewijs dat $s \geq 1 + \sum_{i=1}^{k} a_i(p_i - 1)$.

   link

2.9 veeltermen en complexe getallen gebruiken

1. wortels en ontbinding:

   Een niet-constante veelterm van graad n heeft exact n complexe nulpunten (niet noodzakelijk verschillend).

   Indien $P(\alpha_i) = 0$ voor i als $n+1$ waardenis, dus allen een wortel een wortel zijn van een veelterm met graad P, dan is die veelterm de nulveelterm.

2. complexe wortels:

   $\forall n \in \mathbb{N}$ is $\omega_n = e^{\frac{2\pi i}{n}}$ een eenheidswortel, omdat $\omega_n^n = 1$ en $0 = 1 + \omega_n + \cdots + \omega_n^{n-1} = \frac{\omega_n^n - 1}{\omega_n - 1}$

3. rational root theorem:

   Zij $f(X) = a_n X^n + a_{n-1} X^{n-1} + \cdots + a_1 X + a_0$ een veelterm met gehele cofficinten en als geldt dat de breuk $\frac{p}{q}$ een wortel van f is met $ggd(p,q) = 1$, dan geldt dat $p|a_0$ en $q|a_n$.

4. delingsalgoritme:

   Zij $f(X)$ een veelterm met domein R met $R = \mathbb{Q}, \mathbb{R}$of $\mathbb{C}$ dan zijn er veeltermen $p(X), q(X), r(X) \in R[X]$ zodat $f(X) = p(X)q(X) + r(X)$ met de graad van r kleiner dan die van f, hierbij zijn q, r uniek in functie van p.

5. Vieta :

   $f(X) = a_n X^n + a_{n-1} X^{n-1} + \cdots + a_1 X + a_0$ heeft de n nulpunten z_1 tot z_n, dan geldt dat $\frac{(-1)^i a_{n-i}}{a_n} = \sum_{sym} z_{j1} z_{j2} \cdots z_{ji}$

6. veeltermvergelijkingen:

   Voor alle veeltermen met gehele cofficinten geldt dat $a - b|P(a) - P(b)$ als a, b verschillende gehele getallen zijn.

Een veelterm construeren, kan dus veel informatie geven en eist wat oefenen.

Enkele voorbeelden, indien er nog wat problemen zijn, kan volgende veeltermen en irreducibiliteitsbijlage helpen door meet voorbeelden te zien.

1. Zij $P \in \mathbb{Z}[\mathbb{X}]$ een niet-constante veelterm van graad n.

   Bewijs dat er maximaal $n+2$ gehele a-waarden bestaan zodat $P(a)^2 = 1$.

   link

2. Er geldt dat de veelterm $P(x) = X^n + a_{n-1}X^{n-1} + \cdots + a_1X + a_0 \in \mathbb{Z}[X]$ de n wortels $a_{n-1}, a_{n-2} \cdots a_1$ geeft. Vind alle mogelijkheden voor $P(x)$. link

3. Zij n, k natuurlijke getallen zodat de veelterm $x^{2k} - x^k + 1 | x^{2n} + x^n + 1$. Bewijs dat $x^{2k} + x^k + 1 | x^{2n} + x^n + 1$ ook geldt. link

4. Zij a_1, a_2, a_3, a_4, a_5 reele getallen zodat $\forall k \in \{1, 2, 3, 4, 5\}$ geldt dat
   $\frac{a_1}{k^2+1} + \frac{a_2}{k^2+2} + \frac{a_3}{k^2+3} + \frac{a_4}{k^2+4} + \frac{a_5}{k^2+5} = \frac{1}{k^2}$
   Vind dan de waarde van $\frac{a_1}{37} + \frac{a_2}{38} + \frac{a_3}{39} + \frac{a_4}{40} + \frac{a_5}{41}$?

   link

5. Zij $P(x)$ een veelterm van graad $n > 1$ met gehele cofficinten en k is strikt natuurlijk. We beschouwen de veelterm $Q(x) = P(P(\cdots P(x))\cdots)$ waarin P k keer voorkwam.
   Bewijs dat er maximaal n gehele getallen bestaan waarvoor geldt dat $Q(t) = t$.

   link

6. Bewijs dat de verzameling van alle reele getallen x die voldoen aan de ongelijkheid

   $$\sum_{k=1}^{k=70} \frac{k}{x-k} \geq 1.25$$

   de vereniging is van een aantal disjuncte intervallen, waarbij de som van de lengtes van die intervallen gelijk is aan 1988.

   link

7. Zij a, b, c, d, e, f 6 strikt natuurlijke getal waarvoor geldt dat $abc + def$ alsook $ab + bc + ca - (df + de + ef)$ deelbaar zijn door $S = a + b + c + d + e + f$. Bewijs dat S meer dan 2 delers heeft. link

8. Een woord is een eindig rijtje letters uit een of ander alfabet.
   Een woord heet repeterend als het bestaat uit twee of meer dezelfde woorden die achter elkaar geplakt zijn (zo zijn *ababab* en *abcabc* repeterend, maar *ababa* en *aabb* niet).
   Bewijs dat als een woord de eigenschap heeft dat elke verwisseling van twee aangren- zende letters zorgt dat het woord repeterend wordt, dan alle letters van het woord hetzelfde moeten zijn.
   (Merk op dat je twee aangrenzende letters die hetzelfde zijn, ook mag verwisselen, waarbij het woord dus onveranderd blijft.) link

9. Bewijs dat er een convexe 1990-hoek bestaat met de volgende eigenschappen: (1) alle hoeken zijn gelijk; (2) de lengten van de zijden zijn een permutatie van de getallen $1^2, 2^2, 3^2 \cdots 1990^2$

   link

2.10 irreducibiliteit

defenitie

Een veelterm $f(x) \in R[x]$ is reducibel als er niet-constante veeltermen $g, h \in R[X]$ bestaan zodat $f(x) = g(x)h(x)$, in het andere geval is f irreducibel.

lemma van Gauss f is irredubel over $\mathbb{Z}$ aesa ze irreducibel is over $\mathbb{Q}$.

criterium van Eisenstein

Zij $f(x) = f_n x^n + f_{n-1}x^{n-1} + \cdots + f_1 x + f_0$ een veelterm over $\mathbb{Z}$.
Als er een priemgetal p bestaat zodat $p|f_k$ $\forall k \in \{0, 1, \cdots, n-1\}$, $p \not| f_n$ en $p^2 \not| f_0$,
dan is $f(x)$ irreducibel over $\mathbb{Z}$.

uitbegreid criterium van Eisenstein

Zij $f(x) = f_n x^n + f_{n-1}x^{n-1} + \cdots + f_1 x + f_0$ een veelterm over $\mathbb{Z}$.
Als er een priemgetal p bestaat zodat $p|f_k$ $\forall k \in \{0, 1, \cdots, t\}$, $p \not| f_{t+1}$ en $p^2 \not| f_0$,
dan geeft $f(x)$ een irreducibele deler van graad $\geq t + 1$ over $\mathbb{Z}$.

criterium van Eisenstein voor veeltermen in 2 variabelen

Zij $f(x) = \sum_{i=0}^{i=n} P_i(x)y^i$ en veronderstel dat $Q(x)|P_i(x)$ voor $i \in \{0, 1, 2, 3, 4 \cdots, n-1\}$,$Q(x)$ / $|P_n(x)$ en $Q(x)^2 \not| P_0(x)$, hierbij zijn P_i, Q veeltermen met gehele coefficienten die constant mogen zijn.
Hierbij moest $Q(x)$ een priem/ irreducibele veelterm zijn.
Dan is f irreducibel.

reductie mod p

Zij $f(x) = f_n x^n + f_{n-1}x^{n-1} + \cdots + f_1 x + f_0$ een veelterm over $\mathbb{Z}$ en p een priemgetal. Als $g(x) = g_n x^n + g_{n-1}x^{n-1} + \cdots + g_1 x + g_0$ de veelterm is met $p|g_i - f_i$ en $g_i \in \{0, 1, \cdots, p-1\}$. Als $p \not| f_n$ en $g(x)$ is irreducibel, is $f(x)$ dat ook.

1. Bewijs dat $f(X) = X^n + 5X^{n-1} + 3$ voor $n > 1$ irreducibel is in de gehele getallen. link

2. Bewijs dat volgende polynoom onontbindbaar is in $\mathbb{Z}[x, y]$

$$x^{200}y^5 + x^{51}y^{100} + x^{106} - 4x^{100}y^5 + x^{100} - 2y^{100} - 2x^6 + 4y^5 - 2$$

   link

3. Zij $a_1, a_2 \cdots a_n$ verschillende gehele getallen en beschouw de veelterm $P(x) = (x - a_1)(x - a_2) \cdots (x - a_n) - 1$ Toon aan dat $P(x)$ irreducibel is over $\mathbb{Z}$.

   Bewijs dat ook $\prod(x - a_i)^2 - 1$ irreducibel is over $\mathbb{Z}$. link

4. Als geldt dat $a \in \mathbb{Q}$, bewijs dat dan geldt dat

$$X^{2^n}(X + a)^{2^n} + 1$$

   irreducibel is in $\mathbb{Q}[X]$ voor iedere natuurlijk getal n. link

2.11 genererende functies

Een genererende functie is een veelterm genoteerd als $g_a(z) = \sum_{a \in A} z^a$ waarbij A een deelverzameling is van $\mathbb{Z}$.
$f(X) = a_0 + a_1 X + a_2 X^2 + \cdots$ noemen we de genererende functie van $(a_n)_n$
De technieken die gebruikt worden:

- Een gelijkheid van rijen bewijzen, kan dan door te bewijzen door te tonen dat de genererende functies gelijk zijn.
- Het aantal manieren waarop iets kan, wordt dan makkelijker bepaald als de som en product van verschillende functies.
- De som $s_n = a_1 + a_2 + \cdots + a_n$ kunnen we in de genererende functie genereren als $\frac{f(X)}{1-X}$
- snake-oilmethod: de som van een bepaalde som berekenen we door de sommen in een rij te steken en de genererende functie van die rij te bepalen, hiervoor proberen we een juiste monische macht te plaatsen, waarna we gekende formules kunnen toepassen

We geven hier de belangrijke genererende functies/Taylorreeksen:

1. $\frac{1}{(1-x)^{k+1}} = \sum_{n \geq 0} \binom{n+k}{n} x^n$
2. $(1+x)^a = \sum_{k \geq 0} \binom{a}{k} x^k$
3. $e^x = \sum_{n \geq 0} \frac{x^n}{n!}$
4. $ln\frac{1}{1-x} = \sum_{n \geq 1} \frac{x^n}{n!}$
5. $sinx = \sum_{n \geq 0} (-1)^n \frac{x^{2n+1}}{(2n+1)!}$
6. $cosx = \sum_{n \geq 0} (-1)^n \frac{x^{2n}}{(2n)!}$
7. $Bgtanx = \sum_{n \geq 0} (-1)^n \frac{x^{2n+1}}{(2n+1)}$

Merk op dat een functie zoals

$$sinh(x) = \frac{e^x - e^{-x}}{2} = 0.5[\sum_{n \geq 0} \frac{x^n}{n!} + \sum_{n \geq 0} \frac{(-x)^n}{n!}] = \sum_{n \geq 0} \frac{x^{2n+1}}{(2n+1)!}$$

Analoog is

$$cosh(x) = \frac{e^x + e^{-x}}{2} = \sum_{n \geq 0} \frac{x^{2n}}{(2n)!}$$

Voor velen is dit heel erg nieuw en kan een uitgebreid bestand helpen.

Zoals altijd kan een voorbeeld iets abstracts verduidelijken en hier is dat zeker nodig:

Voorbeeld 2.8. *Er geldt dat $n \in \mathbb{N}$, we hebben 2 verschillende rijen van n positieve getallen $a_1, a_2, \ldots, a_n$ en $b_1, b_2, \ldots, b_n$ zodat de sommen $a_1 + a_2,\ a_1 + a_3,\ \ldots,\ a_{n-1} + a_n$ en $b_1 + b_2,\ b_1 + b_3,\ \ldots,\ b_{n-1} + b_n$ dezelfde zijn op een permutatie na. Bewijs dat n een macht van2 is.*

Bewijs. We schrijven de genererende functies als $F(x) = x^{a_1} + x^{a_2} + \cdots + x^{a_n}$ en $G(x) = x^{b_1} + x^{b_2} + \cdots + x^{b_n}$.
We zien dat

$$F^2(x) - G^2(x) = \left(\sum_{i=1}^{n} x^{2a_i} + 2\sum_{1 \leqslant i \leqslant j \leqslant n} x^{a_i + a_j}\right) - \left(\sum_{i=1}^{n} x^{2b_i} + 2\sum_{1 \leqslant i \leqslant j \leqslant n} x^{b_i + b_j}\right)$$

$$= F(x^2) - G(x^2).$$

Omdat $F(1) = G(1) = n$, hebben we dat 1 een nulpunt is van zekere graad $k, (k \geq 1)$ v.d. veelterm $F(x) - G(x)$.
Dus $F(x) - G(x) = (x-1)^k H(x)$, zodat

$$F(x) + G(x) = \frac{F^2(x) - G^2(x)}{F(x) - G(x)} = \frac{F(x^2) - G(x^2)}{F(x) - G(x)} = \frac{(x^2-1)^k H(x^2)}{(x-1)^k H(x)} = (x+1)^k \frac{H(x^2)}{H(x)}$$

We zien voor $x = 1$ nu dat

$$2n = F(1) + G(1) = (1+1)^k \frac{H(x^2)}{H(x)} = 2^k,$$

zoda $n = 2^{k-1}$.

□

1. Bewijs dat voor elk natuurlijk getal n geldt dat $\sum_{k=1}^{n} \binom{n+k-1}{2k-1} = F_{2n}$

   link

2. Zij p een oneven priemgetal. Bepaal het aantal deelverzamelingen A van de verzameling $\{1, 2, \cdots 2p\}$ waarvoor geldt: (1) A bevat precies p elementen; (2) de som van alle elementen in A is deelbaar door p.

   link

3. De veeltermen $a(x), b(x), c(x), d(x)$ voldoen aan $a(x^5) + xb(x^5) + x^2c(x^5) = (1 + x + x^2 + x^3 + x^4)d(x)$. Toon aan dat $a(1) = 0$.

   link

4. $(F_n)_{n\geq 1}$ is de Fibonaccirij met $F_1 = F_2 = 1, F_{n+2} = F_{n+1} + F_n (n \geq 1)$, en $P(x)$ is een veelterm van graad 990 die voldoet aan $P(k) = F_k,$ voor $k = 992, ..., 1982$. Bewijs dat $P(1983) = F_{1983} - 1$.

   link

5. We noemen een rij $a_0, a_1, \cdots, a_n$ van rele getallen m-gebalanceerd voor een natuurlijke $m \geq 1$ als de sommen
   $a_0 + a_m + \cdots$
   $\sum_{i=0}^{i=[n/m]} a_{im+1}$
   $\cdots$
   $\sum_{i=1}^{i=[n/m]+1} a_{im-1}$
   allen gelijk zijn ([] staat hier voor de entierfunctie)
   Zij $a_0, a_1 \cdots a_{49}$ een gebalanceerde rij voor iedere $m \in \{3, 5, 7, 11, 13, 17\}$.
   Bewijs dat $a_0 = \cdots = a_{49} = 0$.

   link

6. Zij $m, n \geq 2$ natuurlijke getallen en $a_1, a_2, \ldots, a_n$ gehele getallen, waarvan er geen enkele een veelvoud is van m^{n-1}. Toon aan dat er gehele getallen $e_1, e_2, \ldots, e_n$ bestaan, niet allemaal gelijk aan 0, met $|e_i| < m$ voor alle i, zodat $e_1a_1 + e_2a_2 + \cdots + e_na_n$ een veelvoud is van m^n.

   link

7. Zij n een positief geheel getal groter dan 1. Voorts zijn er n in een cirkel geplaatste lampen $L_0, \cdots, L_{n-1}$. Op elk moment is iedere lamp AAN of UIT. Een reeks stappen $S_0, S1 \cdots$ wordt uitgevoerd. Stap S_j heeft alleen invloed op de toestand van L_j (de toestand van alle andere lampen blijft onveranderd) en wel als volgt: (1) als L_{j-1} AAN is, dan verandert S_j de toestand van L_j van AAN naar UIT of van UIT naar AAN; (2) als L_{j-1} UIT is, dan verandert S_j de toestand van L_j niet. De lampen zijn mod(n) genummerd, dat wil zeggen $L_n = L_0$ etcetera. In het begin zijn alle lampen AAN. Bewijs dat (a) er een positief getal $M(n)$ bestaat zodanig, dat na $M(n)$ stappen alle lampen weer AAN zijn; (b) als $n = 2^k, k \in N$, alle lampen weer AAN zijn na $n^2 - 1$ stappen; (c) als$n = 2^k + 1, k \in N$ alle lampen weer AAN zijn na $n^2 - n - 1$ stappen.

   link

8. n, k zijn natuurlijke getallen zodat $k \geq n$ en $2|k - n$. We hebben $2n$ lampen geordend van 1 tot $2n$. In het begin zijn alle lampen uit. Bij iedere stap doen we een lamp branden of doven we een brandende lamp. (we wisselen de status van 1 lamp) Het aantal manieren bestaande uit k stappen om alle lampen van 1 tot n te doen branden en de andere gedoofd te laten, noemen we N. (de lampen van $n+1$ tot $2n$ mogen aan zijn geweest, maar zijn bij het einde uit) Het aantal manieren die behoren tot N, maar waarbij de lampen $n + 1$ tot $2n$ allen gedurende de k stappen gedoofd bleven, noemen we M. Bepaal $\frac{N}{M}$.

   link

uitzonderlijke stellingen

Enkel bij een vraag 3 of 6 kan eens een heel uitzonderlijke stelling voorkomen die nodig is om de vraag op te lossen. Hier enkele voorbeelden (maar bij een IMO-training of dergelijke zelden nodig natuurlijk).

Vandermonde determinant

Als we een matrix hebben van de vorm $V_n = \begin{vmatrix} 1 & x_1 & x_1^2 & \cdots & x_1^{n-2} & x_1^{n-1} \\ 1 & x_2 & x_2^2 & \cdots & x_2^{n-2} & x_2^{n-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 1 & x_n & x_n^2 & \cdots & x_n^{n-2} & x_n^{n-1} \end{vmatrix}$

Dan geldt dat $Det(V_n) = \prod_{1 \leq i < j \leq n} (x_j - x_i)$

basis galoistheorie

$z = a + b\sqrt{d}$ geeft geconjugeerde $z' = a - b\sqrt{d}$ en norm $N(z) = zz' = a^2 - db^2$. $(a, b \in \mathbb{Z})$

i is de imaginaire eenheid en $\omega = \frac{-1+\sqrt{3}i}{2}$.

Ieder element van $\mathbb{Z}[i], \mathbb{Z}[\omega]$ is uniek te schrijven als het product van priemgetallen.

$\mathbb{Z}[i]$ geeft als priemgetallen de waarden x waarvoor geldt dat $N(x)$ priem is of $|x| \equiv 3 \pmod 4$ en als eenheden $\pm 1, \pm i$.

$\mathbb{Z}[\omega]$ geeft als priemgetallen de waarden x waarvoor geldt dat $N(x)$ priem is of $|x| \equiv 2 \pmod 3$ en als eenheden $\pm 1, \pm\omega, \pm(1 + \omega)$.

Indien dit te kort, is kan men kijken op 't IMOmath bestand hierover

CNS

$f(z_1, z_2, \cdots, z_k)$ is een polynoom is $F[z_1, z_2, \cdots, z_k]$.
De graad van f, $degf = \sum t_i$ $(t_i \geq 0)$
Veronderstel dat de coefficient van $\prod z_i^{t_i}$ in $f \neq 0$.
Als $S_1, s_2, \cdots, S_k$ deelverzamelingen zijn van F zodat $|S_i| > t_i$ $\forall i \in \{1, 2 \cdots, k\}$,
dan bestaan er $s_i \in S_i$ zodat $f(s_1, s_2, \cdots, s_k) \neq 0$

Voorbeeld 2.9. *We kijken naar het roostervierkant met punten uit* $\{1, 2, 3, 4, 5\}^2$ *in het vlak van* $\mathbb{R}^2$. *Wat is het minimum aantal cirkels dat we moeten tekenen om alle* 25 *punten te bedekken?*

Bewijs. Ten eerste moet men een configuratie met 5 cirkels vinden.
Stel uit het ongerijmde dat het werkt met 4 cirkels.
We defenieren $P(x, y) := \prod_{i=1}^{4} \left((x - a_i)^2 + (y - b_i)^2 - r_i^2 \right)$, als het product van de 4 voorstellingen van de cirkels.
Er geldt dat $\deg(P) = 8 = \big(|S| - 1\big) + \big(|S| - 1\big)$ en de coefficient van $x^4 y^4$ is niet gelijk aan 0. $(t_1 = t_2 = 4)$
Stel $S_1 = S_2 = \{1, 2, 3, 4, 5\}$, dus $|S_1| = |S_2| > t_1 = t_2$ zodat CNS zegt dat er s_1, s_2 bestonden zodat $P(s_1, s_2) \neq 0$, contradictie.

□

voor zij die het te abstract vonden, maar meer over CNS willen weten.

1. Bepaal alle oplossingen in natuurlijke getallen van m, n die voldoen aan

$$m + n - \frac{3mn}{m+n} = \frac{2011}{3}.$$

link

2. De rij $a_0, a_1, a_2, \ldots$ wordt gedefinieerd als $a_0 = 2,\ a_{k+1} = 2a_k^2 - 1$ met $k \geq 0$. Bewijs dat als een oneven priemgetal $p|a_n$, dat dan $2^{n+3}|p^2 - 1$. link

3. $n \in \mathbb{N}$ en we beschouwen de verzameling

$$S = \{(x, y, z) \mid x, y, z \in \{0, 1, \ldots, n\}, x + y + z > 0\}.$$

Bepaal het minimum aantal vlakken die alle punten van S bevat, maar niet het punt $(0, 0, 0)$

link

Diverse combinatoriekvragen

De meeste soorten van combinatoriekvragen konden opgelost worden met 1 van vorige methodes. Het blijft echter creatief denken en daarom is oefenen nog steeds belangrijker dan de uitzonderlijke stellingen.

1. Zij $a_0, a_1, a_2, \ldots$ een stijgende rij van natuurlijke getallen zodat ieder natuurlijk getal op een unieke manier kan uitgedrukt worden in de vorm $a_i + 2a_j + 4a_k$ met i, j, k niet noodzakelijk verschillend.
   Bepaal a_{1998}. link

2. Gegeven zijn 9 punten in de ruimte met al hun verbindingslijnen, waarbij geen vier punten in een vlak liggen.
   Een aantal verbindingslijnen wordt blauw gekleurd, een aantal rood en de overige blijven ongekleurd.
   Bepaal de kleinste waarde van $n \in N$ met de eigenschap dat als precies n verbindingslijnen gekleurd zijn,
   de verzameling van gekleurde verbindingslijnen een driehoek bevat waarvan de zijden dezelfde kleur hebben.

   link

3. In een eindige rij, geldt er dat iedere opeenvolgende a elementen een strikt negatieve som heeft en iedere opeenvolgende b elementen een strikt positieve som.
   Hoeveel elementen kan de rij maximaal hebben? link

4. In een groep van 120 personen zijn er sommige koppels bevriend.
   Een zwak kwartet is een verzameling van vier personen waarvan er precies 1 koppel bevriend is.
   Wat is het maximum aantal zwakke kwartetten? link

5. Een buitenaards ras heeft drie geslachten: male, female en emale.
   Een getrouwd tripel bestaat uit drie personen, van elk geslacht 1, die allemaal van elkaar houden.
   Een wezen mag hoogstens tot n getrouwd tripel behoren. We gaan er verder van uit dat houden van een wederzijds gevoel is.
   Het ras zendt een expeditie uit om een andere planeet te koloniseren. Er gaan n males, n females en n emales mee. Elk lid van de expeditie houdt van minstens k personen van elk van de andere twee geslachten. De bedoeling is om zoveel mogelijk getrouwde tripels te vormen.
   (a) Toon aan dat als n even is en $k = \frac{n}{2}$, het zelfs niet altijd mogelijk is om 1 getrouwd tripel te vormen.
   (b) Toon aan dat als $k \geq \frac{3n}{4}$, het altijd mogelijk is om n disjuncte getrouwde tripels te vormen en zo alle expeditieleden te trouwen. link

6. n Jongens $a_1, a_2, \cdots a_n$ en n meisjes $m_1, m_2, \cdots, m_n$ zijn op een feestje. Jongens schudden geen handen met elkaar, net zoals de meisjes geen andere vrouwtjes een hand gaven. a_i gaf nooit een hand aan m_i $\forall i \in \{1, 2 \cdots, n\}$. We willen de $2n$ gasten verdelen in t groepen zodat: 1. In iedere groep is het aantal jongens gelijk aan 't aantal meisjes. 2. In iedere groep schudden geen 2 personen een hand aan elkaar.

   m is het aantal koppels (a_i, m_j) die elkaar een hand gaven. Bewijs dat het mogelijk is de groepn te verdelen met $t \leq max(2, \frac{2m}{n} + 1)$. link

Algemene Combinatoriekvragen

3 algebra

3.1 ongelijkheden

Stelling 3.1. *(AM-GM-HM) Voor $n \in \mathbb{N}$ $a_1, ..., a_n > 0$ geldt:*

$$\frac{a_1 + a_2 + \cdots + a_n}{n} \geq \sqrt[n]{a_1 a_2 \cdots a_n} \geq \frac{n}{\frac{1}{a_1} + \cdots + \frac{1}{a_n}}.$$

de triviale benodigdheden

* Kwadraten zijn positief

* ontbindingen

* maximum/minimum beschouwen van de variabelen

* zorgen dat de gelijkheidsgevallen niet verdwenen zijn en in die gevallen nog gelijkheid blijft gelden

(men mag dus de vraag niet te veel vereenvoudigen dat de laatste stappen niet waar meer zijn)

Voorbeeld 3.2. *(IMC 1999) Gegeven reële getallen $x_1, ..., x_n > -1$, met $x_1^3 + x_2^3 + ... + x_n^3 = 0$. Bewijs dat*

$$x_1 + ... + x_n \leq \frac{n}{3}.$$

Oplossing. Merk op dat

$$x^3 - \frac{3}{4}x + \frac{1}{4} = (x+1)\left(x - \frac{1}{2}\right)^2$$

Voor $x = x_i$ is $(x+1)\left(x - \frac{1}{2}\right)^2 \geq 0$, aangezien $x_i > -1$. Tellen we dit nu op voor alle x_i dan komt er

$$(x_1^3 + ... + x_n^3) - \frac{3}{4}(x_1 + ... + x_n) + \frac{n}{4} \geq 0$$

$$\frac{3}{4}(x_1 + ... + x_n) \leq \frac{n}{4}$$

Na deling door $\frac{3}{4}$ geeft dit het te bewijzen. □

1. Bewijs dat $\forall a,b,c \in \mathbb{R}_0^+$ er geldt dat $\frac{a}{bc} + \frac{b}{ac} + \frac{c}{ab} \geq \frac{2}{a} + \frac{2}{b} - \frac{2}{c}$ en zeg wanneer er gelijkheid geldt.

   voorbeeld

2. Bepaal alle drietallen (x,y,z) die voldoen aan $(x+y)^2 = z$ $(x+z)^2 = y$ $(y+z)^2 = x$

   voorbeeld

3. ABC is een driehoek met $P,Q,R \in [BC],[AC],[AB]$. TB:

$$\min\{[AQR],[BRP],[CQP]\} \leq \frac{1}{4} \cdot [ABC]$$

   link

4. Gegeven dat x,y,z positieve rele getallen zijn die voldoen aan $xyz = 32$,
   vind de maximumwaarde van

$$x^2 + 4xy + 4y^2 + 2z^2.$$

   link

5. Vindt alle positieve getallen zodat $a+b = ab$ en $\frac{a}{b^2+4} + \frac{b}{a^2+4} \geq 0.5$.

   klik

6. (*)

   $a,b,c > 0$ zodat $\frac{1}{a} + \frac{1}{b} + \frac{1}{c} = a+b+c$. Bewijs dat $\frac{1}{(2a+b+c)^2} + \frac{1}{(a+2b+c)^2} + \frac{1}{(a+b+2c)^2} \leq \frac{3}{16}$

   klik

7. (*)

   Zij a, b, c rele getallen zodat $ab+bc+ca \leq 3abc$. Bewijs dat $\sqrt{\frac{a^2+b^2}{a+b}} + \sqrt{\frac{b^2+c^2}{b+c}} + \sqrt{\frac{c^2+a^2}{c+a}} + 3 \leq \sqrt{2}\left(\sqrt{a+b} + \sqrt{b+c} + \sqrt{c+a}\right)$

   klik

8. (*)

   Zij $n \geq 3$ en $a_2, a_3, \cdots, a_n > 0$ zodat

$$\prod_{i=2}^{i=n} a_i = 1.$$

   TB:

$$\prod_{i=2}^{i=n} (1+a_i)^i > n^n.$$

   klik

9. (enkel passend)

   Bestaat er een polynoom $P(x)$ met 2012 reeele nulpunten zodat

$$P(a)^3 + P(b)^3 + P(c)^3 \geq 3P(a)P(b)P(c)$$

   geldt $\forall a,b,c, \in \mathbb{R} | a+b+c = 0$?

   http://olympia.problem-solving.be/node/2018

Stelling 3.3. *(Cauchy-Schwarz [CS]) Voor* $a_1, \cdots, a_n, b_1, \cdots, b_n \in \mathbb{R}$ *geldt:*

$$(a_1^2 + a_2^2 + ... + a_n^2) \cdot (b_1^2 + b_2^2 + ... + b_n^2) \geq (a_1 b_1 + ... + a_n b_n)^2.$$

Gelijkheid treedt op als en slechts als $\rho \begin{pmatrix} a_1 & a_2 & \cdots & a_n \\ b_1 & b_2 & \cdots & b_n \end{pmatrix} \leq 1.$
(de waarde $\frac{a_i}{b_i}$ *constant is voor alle i)*
Dit kan uiteraard worden vervormd in andere vormen zoals:
(Cauchy-Schwarz in Engelvorm) Voor $a_1, ..., a_n \in \mathbb{R}$ *en* $b_1, ..., b_n > 0$ *geldt:*

$$\left(\frac{a_1^2}{b_1} + \frac{a_2^2}{b_2} + \cdots + \frac{a_n^2}{b_n} \right) \geq \frac{(a_1 + a_2 + \cdots + a_n)^2}{b_1 + b_2 + ... + b_n}.$$

hints:
* homogeniseren; zorgen dat de graad van iedere veelterm gelijk is om passende voorwaarden te mogen stellen en de gewone stellingen te kunnen toepassen
* substituties: vervangen van de variabelen door een combinatie van nieuwe variabelen om de ongelijkheid te vereenvoudigen

Voorbeeld 3.4. $a, b, c > 0$
TB:

$$(a^2 + 2)(b^2 + 2)(c^2 + 2) \geq 3(a + b + c)^2$$

Bewijs. Er zijn 2 variabelen in $[0, 1], [1, \infty]$ wegens het duivenhokprincipe. Stel dat dit a, b zijn. Dan is $(a^2 + 2)(b^2 + 2) \geq 3(a^2 + b^2 + 1)$ omdat $(a^2 - 1)(b^2 - 1) \geq 0$.
$(a^2 + b^2 + 1)(1 + 1 + c^2) \geq (a + b + c)^2$ vervolledigt het bewijs.

□

oefenen

1. Vind alle oplossingen $x, y, z \in \mathbb{R}^+ : \frac{1}{x} + \frac{1}{y} + \frac{1}{z} = 2010$ $x + y + z = \frac{3}{670}$

   klik

2. $P_1, P_2, \ldots, P_n$ is een permutatie van $\{1, \ldots, n\}$. Bewijs dat $\sum_{i=1}^{n-1} \frac{1}{P_i + P_{i+1}} > \frac{n-1}{n+2}$. klik

3. De strikt positieve rele getallen p, q, r voldoen aan $p+q+r = 1$. Bewijs dat $7(pq+qr+rp) \leq 2 + 9pqr$.klik

4. Zij $a, b, c > 0$ met $abc = 1$. Bewijs dat

$$\frac{1}{a^3(b+c)} + \frac{1}{b^3(a+c)} + \frac{1}{c^3(b+a)} \geq 1.5$$

   klik

5. Zij a, b, c drie positieve rele getallen. Bewijs dat

$$\left(\frac{a}{b} + \frac{b}{c} + \frac{c}{a}\right)^2 \geq (a+b+c)\left(\frac{1}{a} + \frac{1}{b} + \frac{1}{c}\right).$$

   klik

6. Zij $x, y, z \in \mathbb{R}_0^+$ met $xyz \geq 1$. Toon aan dat

$$\frac{x^5 - x^2}{x^5 + y^2 + z^2} + \frac{y^5 - y^2}{x^2 + y^5 + z^2} + \frac{z^5 - z^2}{x^2 + y^2 + z^5} \geq 0.$$

   klik hier

7. (*)

   Bewijs dat voor alle positieve reele getallen a, b, c geldt dat

$$\frac{a}{\sqrt{a^2 + 8bc}} + \frac{b}{\sqrt{b^2 + 8ca}} + \frac{c}{\sqrt{c^2 + 8ab}} \geq 1.$$

   klik hier

8. (*)

   $x, y, z \in \mathbb{R}^+$ zodat $x + y + z = xy + yz + zx$ Bewijs dat $\frac{1}{x^2+y+1} + \frac{1}{y^2+z+1} + \frac{1}{z^2+x+1} \leq 1$ en zeg wanneer er gelijkheid geldt.

   klik

9. (*)

   $a, b, c \in \mathbb{R}$ $\frac{1}{a^2+1} + \frac{1}{b^2+1} + \frac{1}{c^2+1} = 2$ TB: $ab + bc + ac \leq 1.5$

   klik

Stelling 3.5. *(Holder)*
Zij $P_1, p_2, \cdots, p_k \in \mathbb{R}^+ > 0$ en neem k rijen van n positieve, rele getallen met a_{ij} het element van de i^{de} rij en j^{de} kolom en $s = \frac{1}{\frac{1}{p_1}+\cdots+\frac{1}{p_k}}$ Dan geldt er dat

$$\prod_{i=1}^{k} \sqrt[p_i]{a_{i1}^{p_i} + \cdots + a_{i1}^{p_i}} \geq \sqrt[s]{\sum_{j=1}^{n} a_{1j}^s a_{2j}^s \cdots a_{kj}^s}$$

Stelling 3.6. *(Minkowski) Zij $p > r \in \mathbb{R}^+ > 0$ en neem k rijen van n positieve, rele getallen met a_{ij} het element van de i^{de} rij en j^{de} kolom. Dan geldt er dat*

$$\sqrt[r]{\sum_{j=1}^{n}(\sum_{i=1}^{k} a_{ij}^p)^{\frac{r}{p}}} \geq \sqrt[p]{\sum_{i=1}^{k}(\sum_{j=1}^{n} a_{ij}^r)^{\frac{p}{r}}}$$

Stelling 3.7. *(gelijkheid bij gelijkheid)*
Wanneer men wil bewijzen dat het extremum optreedt, wanneer alle termen gelijk zijn, is het voldoende uit het ongerijmde een contradictie te bekomen.

Voorbeeld 3.8. *$a_1 + a_2 + \cdots + a_n = 1$ Vind het minimum van $\sum_{i=1}^{i=n} \sqrt{a_i^2 + \frac{1}{a_i^2}}$?*
Bewijs. Stel dat het minimum optreedt bij $(a, b, x_3, x_4, \cdots, x_n)$ in te vullen voor $(a_1, a_2, \cdots, a_n)$ met $a \neq b$
(merk op dat we kunnen permuteren en dus vanaf er 2 getallen niet gelijk zijn)
Het is voldoende te tonen dat $(a, b) \to (\frac{a+b}{2}, \frac{a+b}{2})$ zorgt dat de uitdrukking kleiner wordt.
$\sqrt{a^2 + \frac{1}{a^2}} + \sqrt{a^2 + \frac{1}{a^2}} > 2\sqrt{(\frac{a+b}{2})^2 + \frac{1}{(\frac{a+b}{2})^2}}$
kwadraten en $AM - GM, CS$ toont aan dat de ongelijkheid strikt geldt als $a \neq b$.
Er geldt dus dat als het minimum optreedt, alle elementen gelijk zijn; $a_i = \frac{1}{n}$ en dus wordt het minimum $\sqrt{n^4 + 1}$.
***
merk op dat deze vraag ook een direct gevolg is van Minkowski . □

Voor zij dit niet zo direct zien, kan men zich tot de moeilijke stellingen-bijlage wenden.

niks beter dan zelf aan de slag te gaan

1. Vind alle $\alpha, \beta > 0$ zodat geldt $\forall x_1, x_2, \cdots, x_n, y_1, \cdots, y_n > 0$ dat $(\sum x_i^\alpha)^{\frac{1}{\alpha}} (\sum y_i^\beta)^{\frac{1}{\beta}} \geq \sum x_i y_i$

   klik

2. Als $a, b, c > 0$ en $ab + bc + ca = 1$, bewijs dan dat

$$\sqrt[3]{\frac{1}{a} + 6b} + \sqrt[3]{\frac{1}{b} + 6c} + \sqrt[3]{\frac{1}{c} + 6a} \leq \frac{1}{abc}.$$

   klik

Stelling 3.9. *(Gewogen Jensen) Zij I een interval, $a_i \in I$, $k_i \in \mathbb{R}_0^+$ en f tweemaal afleidbaar. Als f convex is op I, dan is*

$$\frac{k_1 \cdot f(a_1) + \cdots + k_n \cdot f(a_n)}{k_1 + \cdots + k_n} \geq f\left(\frac{k_1 a_1 + \cdots + k_n a_n}{k_1 + \cdots + k_n}\right).$$

Als f concaaf is op I, dan is

$$\frac{k_1 \cdot f(a_1) + \cdots + k_n \cdot f(a_n)}{k_1 + \cdots + k_n} \leq f\left(\frac{k_1 a_1 + \cdots + k_n a_n}{k_1 + \cdots + k_n}\right).$$

Gelijkheid treedt op als en slechts als ofwel alle a_i gelijk zijn, ofwel de functie een rechte is.

gevolgen

Stelling 3.10. *(gewogen AM-GM) Voor alle $a_i, k_i > 0$ geldt dat*

$$\frac{k_1 a_1 + k_2 a_2 + \cdots + k_n a_n}{k_1 + k_2 + \cdots + k_n} \geq \sqrt[k_1+\ldots+k_n]{a_1^{k_1} a_2^{k_2} \cdots a_n^{k_n}}.$$

Stelling 3.11. *(gewogen QM-AM) Voor alle $a_i, k_i > 0$ geldt dat*

$$\sqrt{\frac{k_1 a_1^2 + k_2 a_2^2 + \cdots + k_n a_n^2}{k_1 + \cdots + k_n}} \geq \frac{k_1 a_1 + k_2 a_2 + \cdots + k_n a_n}{k_1 + k_2 + \cdots + k_n}.$$

Stelling 3.12. *(gewogen GM-HM) Voor alle $a_i, k_i > 0$ geldt dat*

$$\sqrt[k_1+\ldots+k_n]{a_1^{k_1} a_2^{k_2} \cdots a_n^{k_n}} \geq \frac{k_1 + \cdots + k_n}{\frac{k_1}{a_1} + \frac{k_2}{a_2} + \cdots + \frac{k_n}{a_n}}$$

Stelling 3.13. *(Gewogen Power-Mean Ongelijkheid) Als $i > j$ dan is:*

$$f_i(k_m, a_m) \geq f_j(k_m, a_m),$$

gelijkheid als en slechts als alle a_m gelijk zijn. Hierbij staat

$$f_j = \sqrt[j]{\frac{k_1 a_1^j + \cdots + k_n a_n^j}{k_1 + k_2 + \cdots + k_n}}$$

met f_0 het gewogen GM, en $f_{\pm\infty}$ gewoon minimum en maximum resp.

1. Zij $r_1, r_2, \ldots, r_n$ rele getallen groter of gelijk aan 1. Bewijs dat

$$\frac{1}{r_1+1}+\frac{1}{r_2+1}+\cdots+\frac{1}{r_n+1}\geq\frac{n}{\sqrt[n]{r_1r_2\ldots r_n}+1}.$$

klik

Stelling van Karamata

Deze stelling is de algemenere stelling van Jensen en tonen we hier op zijn algemeenst en zodoende zeer sterke ongelijkheid, de werkelijke stelling van Karamata zegt normaal enkel dat
als $x_1, x_2, \cdots, x_n$ $y_1, y_2, \cdots, y_n$ majoriseert, dat dan geldt dat
$f(x_1) + f(x_2) + \cdots + f(x_n) \geq f(y_1) + f(y_2) + \cdots + f(y_n)$ als f convex is over het interval $[x_1, x_n]$.
De volgende veralgemening die we geven noemen we de uitgebreidde stelling van gewogen Karamata.

Stelling 3.14. *(stelling van Karamata)*
gewogen majorisatie:
Zij $x = \begin{pmatrix} x_1 & x_2 & ... & x_n \\ a_1 & a_2 & ... & a_n \end{pmatrix}$ *en* $y = \begin{pmatrix} y_1 & y_2 & ... & y_m \\ b_1 & b_2 & ... & b_m \end{pmatrix}$ *2 gewogen "reeksen", dan majoriseert* x *de "reeks"*y *als geldt dat*
$a_1 + a_2 + ... + a_n = b_1 + b_2 + ... + b_m$, $a_1x_1 + a_2x_2 + ... + a_nx_n = b_1y_1 + b_2y_2 + ... + b_my_m$, $x_1 \geq x_2 \geq ... \geq x_n$ *and* $y_1 \geq y_2 \geq ... \geq y_m$, *alsook*
voor alle indices u *and* v *en alle* $\alpha, \beta \in \mathbb{R}$ *waarvoor geldt dat* $0 \leq \alpha \leq 1$ *en* $0 \leq \beta \leq 1$ *gekozen zodat*
$a_1 + a_2 + ... + a_{u-1} + \alpha a_u = b_1 + b_2 + ... + b_{v-1} + \beta b_v$, *geldt dat*
$a_1x_1 + a_2x_2 + ... + a_{u-1}x_{u-1} + \alpha a_u x_u \geq b_1y_1 + b_2y_2 + ... + b_{v-1}y_{v-1} + \beta b_v y_v$.
We schrijven dit als $(x) \succ (y)$
Nu zegt de uitgebreidde stelling van gewogen Karamata :
Neem de gewogen "reeks" $\begin{pmatrix} x_1 & x_2 & ... & x_n \\ a_1 & a_2 & ... & a_n \end{pmatrix}$ *die de gewogen "reeks"* $\begin{pmatrix} y_1 & y_2 & ... & y_m \\ b_1 & b_2 & ... & b_m \end{pmatrix}$ *majoriseert.*
Zij I *een interval die de getallen* x_1, x_2, ..., x_n, y_1, y_2, ..., y_m *bevat en* f *een functie die tweemaal afleidbaar is. Als* f *convex is op* I, *dan is*
$a_1 f(x_1) + a_2 f(x_2) + ... + a_n f(x_n) \geq b_1 f(y_1) + b_2 f(y_2) + ... + b_m f(y_m)$.
Bij een concave functie geldt dit in de andere richting.

hint

Probeer een schets/grafiek te maken om andere eigenschappen van de functie te bekomen die handig zijn om de vraag op te lossen.

1. Zij I een interval en $f : I \to \mathbb{R}$ een convexe functie. D.w.z. $\forall a, b \in I$ en $\forall \lambda \in [0, 1]$ geldt

$$f(\lambda a + (1 - \lambda)b)) \leq \lambda f(a) + (1 - \lambda)f(b).$$

   Bewijs dan dat $\forall a, b, c \in I$ met $a < b < c$ geldt dat

$$f(b) + f(a + c - b) \leq f(a) + f(c).$$

   klik

2. Er geldt dat $a + b + c = 1$ waarbij $a, b, c \in \mathbb{R}^+$.
   Vind het maximum van $\frac{1}{a^2-4a+9} + \frac{1}{b^2-4b+9} + \frac{1}{c^2-4c+9}$.

   klik

3. Als $n \geq 2$ een natuurlijk getal is en $0 \leq a_1 \leq a_2 \leq \ldots \leq a_{2n+1}$ zijn reele getallen, bewijs dan dat

$$\sqrt[n]{a_1} - \sqrt[n]{a_2} + \sqrt[n]{a_3} - \ldots - \sqrt[n]{a_{2n}} + \sqrt[n]{a_{2n+1}} \leq \sqrt[n]{a_1 - a_2 + a_3 - \ldots - a_{2n} + a_{2n+1}}.$$

   klik

Stelling 3.15. *(Orde-ongelijkheid voor sommen) Zij* $a_1 \geq ... \geq a_n \geq 0$, $b_1 \geq ... \geq b_n \geq 0$, $c_1 \geq ... \geq c_n \geq 0$, $\cdots$,$x_1 \geq ... \geq x_n \geq 0$ *dan geldt voor alle permutaties* σ, τ*:*

$$\sum \prod \begin{pmatrix} a_1 & a_2 & \cdots & a_n \\ b_1 & b_2 & \cdots & b_n \\ c_1 & c_2 & \cdots & c_n \\ \cdots & & & \\ x_1 & x_2 & \cdots & x_n \end{pmatrix} \geq \sum \prod \begin{pmatrix} a_1 & a_2 & \cdots & a_n \\ b_{\sigma(1)} & b_{\sigma(2)} & \cdots & b_{\sigma(n)} \\ c_{\tau(1)} & c_{\tau(2)} & \cdots & c_{\tau(n)} \\ \cdots & & & \\ x_{\tau(1)} & x_{\tau(2)} & \cdots & x_{\tau(n)} \end{pmatrix}.$$

Stelling 3.16. *(Chebychev) Zij* $a_1 \geq ... \geq a_n \geq 0$, $b_1 \geq ... \geq b_n \geq 0$, $c_1 \geq ... \geq c_n \geq 0$, $\cdots$,$x_1 \geq ... \geq x_n \geq 0$ k *gelijkgesorteerde rijen, dan geldt :*

$$(\sum_{i=1}^{n} a_i)(\sum_{i=1}^{n} b_i) \cdots (\sum_{i=1}^{i=n} x_i) \leq n^{k-1} \sum_{i=1}^{n} a_i b_i \cdots x_i.$$

Stelling 3.17. *(Orde-ongelijkheid voor producten) Zij* $a_1 \geq ... \geq a_n \geq 0$, $b_1 \geq ... \geq b_n \geq 0$, $c_1 \geq ... \geq c_n \geq 0$, $\cdots$,$x_1 \geq ... \geq x_n \geq 0$ *dan geldt voor alle permutaties* σ, τ*:*

$$\prod \sum \begin{pmatrix} a_1 & a_2 & \cdots & a_n \\ b_1 & b_2 & \cdots & b_n \\ c_1 & c_2 & \cdots & c_n \\ \cdots & & & \\ x_1 & x_2 & \cdots & x_n \end{pmatrix} \leq \prod \sum \begin{pmatrix} a_1 & a_2 & \cdots & a_n \\ b_{\sigma(1)} & b_{\sigma(2)} & \cdots & b_{\sigma(n)} \\ c_{\tau(1)} & c_{\tau(2)} & \cdots & c_{\tau(n)} \\ \cdots & & & \\ x_{\tau(1)} & x_{\tau(2)} & \cdots & x_{\tau(n)} \end{pmatrix}.$$

terug een rij helpers om in te oefenen

1. Zij $x_1 \leq ... \leq x_n$ en $y_1 \leq ... \leq y_n$. Bewijs dat voor elke permutatie σ geldt dat

$$\sum_{i=1}^{n}(x_i - y_i)^2 \leq \sum_{i=1}^{n}(x_i - y_{\sigma(i)})^2.$$

   klik

2. Zij $a_k > 0 \in \mathbb{N}$ met alle a_k verschillend. Bewijs dat

$$\sum_{k=1}^{n}\frac{a_k}{k^2} \geq \sum_{k=1}^{n}\frac{1}{k}.$$

   klik

3. Zij $x_1, \ldots, x_n$ rele getallen zodat $|x_1 + \cdots + x_n| = 1$ en $|x_i| \leq \dfrac{n+1}{2}$ voor alle i. Bewijs dat er een permutatie σ bestaat zodat

$$\left|\sum_{i=1}^{n} i x_{\sigma(i)}\right| \leq \frac{n+1}{2}.$$

   klik

Stelling 3.18. *(Schur)* $\forall a,b,c,r>0$ $a^r(a-b)(a-c)+b^r(b-a)(b-c)+c^r(c-a)(c-b)\geq 0$ *met gelijkheid a.e.s.a* $a=b=c$
of $f(a)(a-b)(a-c)+f(b)(b-a)(b-c)+f(c)(c-a)(c-b)\geq 0$ *met* f *een strikt stijgende functie.*

Stelling 3.19. *(Muirhead) Als* $(a_1,a_2,\cdots,a_n)\succ(b_1,b_2,\cdots,b_n)$*, en* $x_i>0$*, dan geldt:* $\forall x_i\in\mathbb{R}^+:\sum_{sym}x_1^{a_1}\cdots x_n^{a_n}\geq\sum_{sym}x_1^{b_1}\cdots x_n^{b_n}$

minder belangrijke stellingen:

Stelling 3.20. *(Bernoulli)* $\forall x_i\geq -1$ *met alle* x_i *hetzelfde teken geldt dat* $\prod_{i=1}^n(1+x_i)\geq 1+\sum_{i=1}^n x_i$

Stelling 3.21. *(Maclaurin) Zij* $S_k=\frac{a_{i1}a_{i2}\cdots a_{ik}}{\binom{n}{k}}$ *(gemiddelde van de termen van de ontbinding van* $(1+a_1)(1+a_2)\cdots(1+a_n)$ *met graad* k*) Dan geldt dat* $f_i\geq f_j$ *als* $i\leq j$ *met* $f_i=\sqrt[i]{s_i}$

Stelling 3.22. *(Newton) Bij de eig. van Maclaurin geldt ook dat* $S_k^2\geq S_{k-1}S_{k+1}$

1. Zij x, y, z positieve reele getallen zodat $xyz = 1$. Bewijs dat

$$\frac{x^3}{(1+y)(1+z)} + \frac{y^3}{(1+z)(1+x)} + \frac{z^3}{(1+x)(1+y)} \geq \frac{3}{4}.$$

klik

2. Zij $a_0, a_1, a_2, \ldots$ een willekeurige oneindige rij van positieve getallen. Toon aan dat de ongelijkheid $1 + a_n > a_{n-1}\sqrt[n]{2}$ opgaat voor oneindig veel natuurlijke getallen n. klik

3. $a, b, c > 0$ zodat $min(a+b, b+c, a+c) > \sqrt{2}$ en $a^2 + b^2 + c^2 = 3$. Bewijs dat

$$\frac{a}{(b+c-a)^2} + \frac{c}{(b-c+a)^2} + \frac{b}{(c+a-b)^2} \geq \frac{3}{(abc)^2}.$$

klik

Stelling 3.23. *(extremeaaltechniek)*
Als een functie f convex of lineair is in alle variabelen $x_1, \cdots, x_n$,
geldt dat het maximum optreedt wanneer alle variabelen gelijk zijn aan het minimum of maximum.

Voorbeeld 3.24. *(IMOLL Peter Vandendriessche)*
Zij $n \geq 2$ een natuurlijk getal en zij $x_1, \ldots, x_n, y_1, \ldots, y_n \geq 0$. Toon aan dat

$$x_1^2 + \cdots + x_n^2 + \left(\frac{y_1 + \cdots + y_n}{n}\right)^2 \geq \sqrt[n]{(x_1^2 + y_1^2)(x_2^2 + y_2^2)\cdots(x_n^2 + y_n^2)}.$$

Bewijs. Merk eerst op dat de ongelijkheid homogeen is.
Veronderstel dat we een tegenvoorbeeld hebben $(x_1, x_2, \cdots, x_n, y_1 \cdots, y_n)$.
Schrijf $S_i = x_i^2 + y_i^2 > 0$.
De ongelijkheid wordt

$$F(y_1, y_2, \cdots, y_n) = S_1 - y_1^2 + \cdots + S_n - y_n^2 + (\frac{y_1 + y_2 + \cdots + Y_n}{n})^2 \geq \sqrt[n]{S_1 \cdot S_2 \cdots S_n}.$$

Stel alle variabelen y_j vast voor $j \in \{1, 2, \cdots i-1, i+1, i+2 \cdots n\}$ en laat y_i varieren.

Het rechterlid is constant, maar als we $f(y_i) = S_1 - y_1^2 + \cdots + S_n - y_n^2 + (\frac{y_1+y_2+\cdots+Y_n}{n})^2$ schrijven, geldt $f"(y_i) = -2 + \frac{2}{n^2} < 0$ zodat het linkerlid concaaf is in y_i.
Dit betekent dat het linkerlid minimaal is als $y_i \in \{0, \sqrt{S_i}\}$ wanneer we het extremaalprincipe toepassen.

We bekijken nu $F(0,0,0,0\cdots,0), F(\sqrt{S_1}, 0, \cdots, 0), F(\sqrt{S_1}, \sqrt{S_2}, 0, \cdots, 0), \cdots, F(\sqrt{S_1}, \sqrt{S_2}, \cdots, \sqrt{S_n}))$.
Vervolgens passen we $AM - GM$ toe op (x_i) en op (y_j), de rijen van de getallen die positief zijn ($(x_j), (y_i)$ zijn 0).
Fixeer nu de rij (y_j) en we bekijken (x_i) met $\sum x_i^2 = N$, met $AM - GM$ zien we dat het linkerlid constant is en het rechterlid maximaal, wanneer alle x_i gelijk zijn.
Analoog zetten we (x_i) vast en schrijven $\sum y_i = N$ zodat het linkerlid constant is en het rechterlid maximaal als alle y_i gelijk zijn.
We vinden dus het extremum als er j keer geldt dat x vast is, en $n - j$ keer y.
Er valt nu te bewijzen dat

$$jx^2 + (\frac{n-j}{n}y)^2 \geq \sqrt[n]{(x^2)^j (y^2)^{n-j}}$$

Voor $j = n$ is het triviaal, in het ander geval kunnen we $AM - GM$ teopassen:

$$j \times x^2 + (n-j) \times \frac{(n-j)y^2}{n^2} \geq n \sqrt[n]{\frac{(n-j)^{n-j}}{(n^2)^{n-j}}(x^2)^j (y^2)^{n-j}}$$

Het volstaat nu nog te bewijzen dat $n^n \frac{(n-j)^{n-j}}{(n^2)^{n-j}} \geq 1$
Schrijf $n - j = n\alpha$, verkrijgen we dat er te bewijzen valt dan $n^{1-\alpha}\alpha^\alpha \geq 1$
Als $n \geq e$ geldt dit altijd (enkel $\alpha = 1$ trviale gelijkheid met de afgeleide) en als $n = 2$ kon $t = 0, 1, 2$ handmatig gecheckt worden.

□

oefenen

1. Als $k \geq v, w, x, y, z \geq h > 0$, toon dan aan dat

$$(v+w+x+y+z)\left(\frac{1}{v}+\frac{1}{w}+\frac{1}{x}+\frac{1}{y}+\frac{1}{z}\right) \leq 25+6\left(\sqrt{\frac{h}{k}}-\sqrt{\frac{k}{h}}\right)^2.$$

   Wanneer treedt gelijkheid op?

   klik

2. (*)

   Voor iedere $i \in \{1, 2, \cdots, n\}$ geldt dat $x_i > 0, x_i y_i > z_i^2$ waarbij alle getallen reel zijn. Bewijs dat: $\frac{n^3}{(\sum_{i=1}^n x_i)(\sum_{i=1}^n y_i)-(\sum_{i=1}^n z_i)^2} \leq \sum_{i=1}^n \frac{1}{x_i y_i - z_i^2}$ ** Op de IMO was dit voor het specifieke geval $n = 2$.

   klik

de laatste loodjes bij onze ongelijkheden

Stelling 3.25. *(Lagrange multipliers)*
Zij gegeven dat $f(a_1, ..., a_n) = r$ *dan kan men extremums van de functie* $g(a, b \cdots, x)$ *bekomen door het oplossen van* $g(a_1, ..., a_n) + \lambda[f(a_1, ..., a_n) - r]$ *af te leiden naar* 1 *variablele, wiens afgeleide* 0 *moet zijn om een extremum te bekomen, waardoor bij symmetrische functies de juiste waarde van* λ *het gevraagde simple aantoont.*

Stelling 3.26. *(EMV-stelling) Zij* f *een continu functie en* $[f]$ *is de som van de afgeleiden in iedere variabele. De ongelijkheid* $f(x_1, x_2, ..., x_n) \geq 0$ *met* $x_i \geq 0$ *geldt als :*
$(i).f(x_1, x_2, ..., x_n) \geq 0$ *if* $x_1x_2...x_n = 0$. $(ii).[f] \geq 0$ $\forall x_1, x_2, ..., x_n \geq 0$.
(geldt niet noodzakelijk in omgekeerd)

Voorbeeld 3.27.

$$a^4 + b^4 + c^4 + d^4 + 2abcd \geq a^2(b^2 + c^2 + d^2) + b^2(c^2 + d^2)) + c^2d^2$$

Equivalent met $F = \sum_{cyc} a^4 + 2abcd - \sum_{cyc} a^2b^2 \geq 0$

Bewijs. i $d = 0$, dan $AM - GM$ geeft dat het klopt.
ii $[F] = 4\sum_{cyc} a^3 + 2\sum_{cyc} abc - 2\sum_{cyc} ab(a + b) \geq 0$ (sommatie van Schur-ongelijkheden)

□

1. $a, b, c, d \in \mathbb{R} : a + b + c + d = 19$ en $a^2 + b^2 + c^2 + d^2 = 91$.Vind het maximum dat mogelijk is voor de uitdrukking $\frac{1}{a} + \frac{1}{b} + \frac{1}{c} + \frac{1}{d}$?

   klik

Diverse Ongelijkheden oefeningen

1. Zij $a_1, a_2, \ldots, a_n$ positieve rele getallen zodat $a_1 + a_2 + \cdots + a_n \leq 1$. Bewijs dat

$$\frac{a_1 a_2 \ldots a_n (1 - (a_1 + a_2 + \cdots + a_n))}{(a_1 + a_2 + \cdots + a_n)(1 - a_1)(1 - a_2)\ldots(1 - a_n)} \leq \frac{1}{n^{n+1}}.$$

   klik

2. $a^2 + b^2 + c^2 = 2abc + 1$ geldt voor de positieve getallen a, b, c.

   Vind het maximum dat $(a - 2bc)(b - 2ca)(c - 2ab)$ kan aannemen.

   klik

3. Gegeven zijn reele getallen $a_1, a_2, \cdots, a_n$.
   Definieer $d_i = max\{a_j | 1 \leq j \leq i\} - min\{a_j | i \leq j \leq n\}$ voor elke i tussen 1 en n en laat $d = max\{d_i | 1 \leq i \leq n\}$.

   (a) Bewijs dat voor alle getallen $x_1 \leq x_2 \leq \cdots \leq x_n \in \mathbb{R}$ geldt dat

$$max\{|x_i - a_i| | 1 \leq i \leq n\} \geq \frac{d}{2}$$

   [1]

   (b) Bewijs dat er zo'n rij $(x_n)_n$ was zodat er gelijkheid geldde in [1].

   klik

3.2 polynoom vergelijkingen

Bij zowel polynoom- als functievergelijkingen zijn er 2 grote dingen die men moet doen:
*aantonen dat de andere mogelijkheden niet voldoen
* bewijzen dat de gevonden functies altijd voldoen aan je vergelijking.

Bij veeltermvergelijkingen geeft men het voordeel dat men de hoogstegraadsterm kan bekijken en hieruit conclusies trekken en als dit begrensd is, de volledige polynoom te kunnen invullen. (wat men niet kan bij een functievergelijking)

oefeningen

1. Vind alle veeltermen $P(x)$ met rele cofficinten die voldoen aan

$$P(a-b)+P(b-c)+P(c-a)=2P(a+b+c)$$

voor alle drietallen a,b,c van rele getallen met $ab+bc+ca=0$.

klik

2. Bewijs dat iedere monische veelterm van graad n met rele cofficinten het rekenkundig gemiddelde is van twee monische veeltermen van graad n met n rele wortels.

klik

3. Vind alle rele veeltermen $p(x,y)$ zodanig dat $p(x,y)p(u,v)=p(xu+yv,xv+yu)$ voor alle $x,y,u,v\in\mathbb{R}$.

klik

4. Vind alle polynomen $a+b+c|P(a)+P(b)+P(c)$ met gehele cofficinten als $a+b+c\neq 0$ en alle 3 gehele getallen zijn.

klik

5. Bepaal alle homogene veelteren $F(x,y)\in\mathbb{R}[\mathbb{X}]$ zodat $f(1,0)=0$ en zodat $\forall a,b,c\in\mathbb{R}$ geldt dat $f(a+b,c)+f(b+c,a)+f(a+c,b)=0$. klik

In het algemeen kan men HIER oefeningen over veeltermen bekijken.

3.3 functievergelijkingen

Het is vooral creatief zijn en volgende middelen kunnen helpen:

* waarden zoeken die de functievergelijking reduceren tot iets dat gemakkelijk aantoont dat bepaalde oplossingen niet kunnen voldoen/ met kleine waarden op ideen komen,
transformaties uitvoeren die de voorwaarden behouden maar de vergelijking behouden en eventueel een goed gevalsonderscheid

* kijken naar sur-,bi- of injectiviteit:

injectieve functies beelden alle originelen op verschillende functiewaarden af, dus $a \neq b \Rightarrow f(a) \neq f(b)$

surjectieve functies: er is geen enkele waarde uit het codomein die geen functiewaarde is

bijectiviteit: injectiviteit en surjectiviteit ineen

* $f(x) = g(x) + h(x)$ stellen, dit kan helpen als er slechts 1 functie g is die voldoet, bij het invullen van $g(x) + h(x)$ kan men men proberen aan te tonen dat $h(x)$ de nulfunctie is.

* dek- of fixpunten zoeken, dit zijn waarden die gelijk zijn aan hun functiewaarden: $f(x) = x$ Dit is vooral handig bij functies waar het aantal fixpunten klein is, aangezien anders niet veel te concluderen valt.

* $f^{n+1}(x)$ op meerdere manieren opvatten, waardoor bepaalde dingen duidelijk komen: $f(f^n(x)) = f^n(f(x))$

* eventueel in een bepaalde basis kijken naar de getallen en voorwaarden in functie van die representatie te bekomen.

* de waarden zoeken waarvoor geldt dat $f(x) = 0$, indien dit slechts 1 waarde is, helpt het vaak deze waarde elders in te vullen om de vergelijking te verkorten.

* bij polynoomvergelijkingen de hoogstegraadsterm vinden om de algemene formule kort te kunnen gebruiken.

* de Cauchy-vgl'en:

$f(x+y) = f(x) + f(y)$ geeft enkel de opl. $f(x) = cx$
$f(xy) = f(x) + f(y)$ "$f(x) = cln|x|$
$f(xy) = f(x)f(y)$ "$f(x) = x^c$ of $\equiv 0$.
$f(x+y) = f(x)f(y)$ "$f(x) = c^x$
met $x, y \in \mathbb{R}$
als er geweten is dat de functie 1 van volgende eigenschappen geeft:

monotoom/continu/stijgend of dalend/begrensd

* kijk eventueel ook naar de moduloresten in het domein/codomein

1. Zoek alle functies f$\mathbb{R} \to \mathbb{R}$ zodat er geldt dat

   $f([x]y) = f(x)[f(y)]$.

   *** Hierbij wordt met $[x]$ de entierfunctie bedoelt die een getal afrond naar beneden op zijn geheel deel.

   klik

2. Vind alle functies $f\mathbb{R} \to \mathbb{R}$ die voldoen aan

   $$f(xy)(f(x) - f(y)) = (x - y)f(x)f(y)$$

   voor alle rele x, y.

   klik

3. Vind alle functies $f\mathbb{N} \to \mathbb{N}$, zodat $\forall a, b \in \mathbb{N}$ geldt dat er een niet-ontaarde driehoek bestaat met lengten $a, f(b)$ en $f(b + f(a) - 1)$.

   klik

4. $f\mathbb{Q}^+ \to \mathbb{Q}^+$ voldoet aan $f(f(x)^2 y) = x^3 f(xy) \forall x, y \in \mathbb{Q}^+$ (verzameling van strikt positieve rationale getallen). klik

5. Vind alle functies $fR \to R$ waarvoor geldt dat $f(yf(x + y) + f(x)) = 4x + 2yf(x + y)$ voor alle$x, y \in R$.

   klik

6. Zij f een functie $\mathbb{R} \to \mathbb{R}$, bewijs dat er $x, y \in \mathbb{R}$ bestaan waarvoor geldt dat $f(x - f(y)) > yf(x) + x$

   klik

7. Bepaal alle functies $f\mathbb{R} \to \mathbb{R}$ zodat voor alle $x, y \in \mathbb{R}$ geldt dat

   $$f(x - f(y)) = f(f(y)) + xf(y) + f(x) - 1.$$

   klik

8. Bepaal alle functies $f : \mathbb{Z} \to \mathbb{Z}$ die voldoen aan de gelijkheid

   $$f(a)^2 + f(b)^2 + f(c)^2 = 2[f(a)f(b) + f(a)f(c) + f(c)f(b)]$$

   $\forall a, b, c \in \mathbb{Z}$ zodat $a + b + c = 0$. klik

9. Nind alle functies $f\mathbb{R} \to \mathbb{R}$ die voldoet aan $\forall x, y \in \mathbb{R}$ $f(xf(x + y)) = f(yf(x)) + x^2$

   klik

10. $f\mathbb{R} \to \mathbb{R}$ $f(x + y) \leq yf(x) + f(f(x)) \forall x, y \in \mathbb{R}$ TB: $f(x) = 0$ $\forall x \leq 0$.

    klik

Voor meer voorbeelden om te proberen, kun je hier f-vgl.'en vinden.

Het gebeurt vrij regelmatig dat een functievergelijking moet opgelost worden met eigenschappen, lemma's en werkwijzes uit de getaltheorie.
Functievergelijkingen in meerdere functies zijn speciale, waarbij terug elegant moet gedacht worden.
Het focussen op de vorm van 1 v.d. functies is soms handig, maar ook ongelijkheden en inductie kunnen helpen.
Het is hier terug belangrijker voorbeelden te zien, omdat er niet echt veel specifieke theorie over is.

1. $\forall a, b \in \mathbb{N}$ geldt $(a-b)|f(a)-f(b)$ waarbij f geen constante functie is. Bewijs dat er oneindig veel priemgetallen zijn, die een deler zijn van een functiewaarde.

   klik

2. $f, g\mathbb{N} \to \mathbb{N} f(g(n)) = f(n)+1, g(f(n)) = g(n)+1, n \in \mathbb{N}$ TB: $f(n) = g(n) \forall n \in \mathbb{N}$

   klik

3. Vind alle surjectieve functies die voldoen aan $p|f(m+n)$ a.e.s.a. $p|f(m)+f(n)$ waarbij $m, n \in \mathbb{N}$ zijn.

   klik

4. Vind alle paren (f, g) van functies $\mathbb{R} \to \mathbb{R}$ zodat

   $$g(f(x+y)) = f(x) + (2x+y)g(y)$$

   $\forall x, y \in \mathbb{R}$. klik

5. Vind alle paren (f, g) van functies $\mathbb{N} \to \mathbb{N}$ zodat

   $$f^{g(n)+1}(n) + g^{f(n)}(n) = f(n+1) - g(n+1) + 1$$

   $\forall n \in \mathbb{N}$. Hierbij is $f^k(n) = \underbrace{f(f(\cdots f(n))\cdots)}_{k*f}$. klik

6. Zij $\mathbb{N}_0$ de verzameling van de natuurlijke getallen zonder 0. Zoek alle functies $g\mathbb{N}_0 \to \mathbb{N}_0$ zodat $(g(m)+n)(g(n)+m)$ een volkomen kwadraat is$\forall m, n \in \mathbb{N}_0$

   klik

7. $f\mathbb{Z} \to \mathbb{N}$ voldoet aan $f(m-n)|f(m)-f(n) \forall m, n \in \mathbb{Z}$. Bewijs dat als $f(m) \leq f(n)$ dat $f(m)|f(n)$.

   klik

3.4 rijen

Een onderwerp met geen specifieke theorie.

Vaak komt er iets voor uit de algemene combinatoriek aan te pas zoals inducties, contradictie en dergelijke.

Een recursie op stellen en dergelijke komt niet puur voor.

Er is dus niks beter dan er goede voorbeelden van te zien:

1. Vind het kleinste natuurlijk getal met de volgende eigenschap: er bestaat geen rekenkundige rij van 1999 rele getallen die precies n gehele getallen bevat.

   klik

2. Definieren we een rij van rijen als volgt: $R_1 = 1$ en als $R_{n-1} = (a_1, \ldots, a_s)$, dan is $R_n = (1, 2, \ldots, a_1, 1, 2, \ldots, a_2, 1, 2, \ldots, \ldots, 1, 2, \ldots, a_s, n)$. Bijvoorbeeld, $R_2 = (1, 2)$ en $R_3 = (1, 1, 2, 3)$. Bewijs dat als $n \geq$, dan is de k-de term van links in de rij R_n gelijk aan 1 als en slechts als de k-de term van rechts in de rij R_n verschillend is van 1.

   klik

3. Zij $a_1, a_2, \cdots$ een rij van positieve rele getallen. Veronderstel dat er een natuurlijk getal s is zodat $a_n = max\{a_k + a_{(n-k)} | 1 \leq k \leq n-1\}$ voor alle $n > s$. Bewijs dat er natuurlijke getallen l en N bestaan met $l \leq s$ en zodat $a_n = a_l + a_{(n-l)}$ voor alle $n \in \mathbb{N}$.

   klik

4. Zij $s_1, s_2, s_3, \ldots$ een strikt stijgende rij van natuurlijke getallen zodat de subrijen $s_{s_1}, s_{s_2}, s_{s_3}, \ldots$ en $s_{s_1+1}, s_{s_2+1}, s_{s_3+1}, \ldots$ beiden rekenkundige rijen zijn. Bewijs dat de rij$s_1, s_2, s_3, \ldots$ zelf een rekenkundige rij is.

   klik

5. Zij n een natuurlijk getal en zij $a_1, a_2, \cdots, a_n$ verschillende natuurlijke getallen zijn. Er zijn $n-1$ getallen tussen 1 en $\sum_{i=1}^{i=n} a_i - 1$ gekozen in de verzameling M waar mensen hem willen vangen. De sprinkhaan start in het punt 0 en maakt n sprongen met de lengten a_1 tot a_n, bewijs dat hij die volgorde kan kiezen zodat hij nergens wordt gevangen in een punt van M.

   klik

Verder kan men zoeken voor RIJvoorbeelden voor extra problemen indien gewenst.

4 getaltheorie

4.1 basis

De basis bij getaltheorie bestaat uit o.a. de eenduidige priemontbinding, kgv, ggd, aantal delers en dergelijke kennen en toepassen.

De stelling van Bezout-Bachet zegt in het algemeen dat voor veeltermen $P_1, P_2, \cdots$ in $\mathbb{Z}[x]$ geldt dat hun ggd te schrijven is in de vorm $\sum P_i Q_i$ waarbij $Q_i \in \mathbb{Q}[x]$.

De stelling van Euler zegt dat $n^{\phi m} \equiv 1 \pmod{m}$ als $ggd(n, m) = 1$.
Als $m = \prod p_i^{k_i}$ is $\phi(m) = \prod (p_i - 1) p_i^{k_i - 1}$.
Herbij is $a \equiv b \pmod{m}$ aesa $m | a - b$

chinese reststelling (CRS) Als m_1 tot m_k gehele getallen die paarsgewijs relatief priem zijn en a_1 tot a_k zijn gehele getallen.
Dan bestaat er 1 oplossing x modulo $\prod m_i$ zodat $x \equiv a_i \pmod{m_i}$ $\forall i \in \{1, 2 \cdots, k\}$.

stelling van Wilson Voor ieder priemgetal geldt $(p-1)! \equiv -1 \pmod{p}$.

hint Werken met lineaire veelvouden helpt vaak.

Voorbeeld 4.1. *Bewijs dat de breuk $\frac{21n+4}{14n+3}$ voor geen enkel natuurlijk getal n vereenvoudigbaar is.*

Bewijs. Merk op dat $3(14n + 3) - 2(21n + 4) = 1$ en dus is het ggd $= 1$.

□

Voorbeeld 4.2. *Er zijn oneindig veel getallel $n \in \mathbb{N}$ zodat $n^2 + 1 | n!$ (en oneindig veel die hier niet aan voldoen ook, zie verder)*

Bewijs. Schrijf $n = 2x^2$, dan hebben we $n^2 + 1 = 4x^4 + 1 = (2x^2 + 2x + 1)(2x^2 - 2x + 1)$.
We willen dat $2x^2 + 2x + 1$ niet priem is, door $x \equiv 1 \pmod{5}$ met $x > 1$ te nemen, geldt dat $\frac{(2x^2+2x+1)}{5}, 5, (2x^2 - 2x + 1)$ kleiner zijn dan $n = 2x^2$ en dus allen verschillende delers zijn van $n!$.

□

Indien er nog vragen waren over de basis, is volgende bestand altijd handig om te helpen in dit getaltheorie-hoofdstuk (voor dit deel vooral de eerste 2 pagina's : getallenleer

1. Bepaal het grootste natuurlijk getal n met de eigenschap dat n deelbaar is door alle positieve natuurlijke getallen kleiner dan $\sqrt[3]{n}$. link

2. Zij f een tweedegraadspolynoom met gehele cofficinten, waarbij $5|f(k)$ voor alle $k \in \mathbb{Z}$. Toon aan dat alle cofficinten van f deelbaar zijn door 5. link

3. Zij $x, y, z \in \mathbb{Z}$ met $29|x^4 + y^4 + z^4$. Toon aan dat $29^4|x^4 + y^4 + z^4$. link

4. Vind de kleinste $x \in \mathbb{N}$ zodat $\frac{7x^{25}-10}{83}$ geheel is. link

5. Zij $\tau(n)$ de functie die $n \in \mathbb{N}$ afbeeldt op het aantal verschillende positieve delers van n. Bewijs dat er oneindig veel natuurlijke getallen a bestaan zodat de vergelijking $\tau(an) = n$ geen natuurlijk oplossing n heeft. link

6. We noemen een getal alternatief als al zijn cijfers afwisselend oneven en even zijn. Vind alle natuurlijke getallen n zodat n een alternatief veelvoud heeft. link

7. Zijn er ∞ veel oplossingen voor $n! + 1|(2012n)!$? link

8. $a^n + n|b^n + n, \forall n \in \mathbb{N}$, bewijs dat $a = b$. link

4.2 priemgetalstellingen

stelling van Euclides
Er zijn oneindig veel priemgetallen.
postulaat van Bertrand
Voor iedere $n \in \mathbb{N}_0$ is er een priemgetal tussen n en $2n$
stelling van Dirichlet
Als $ggd(a, b) = 1$ bestaan er oneindig veel priemgetallen van de vorm $an + b$ met $n \in \mathbb{N}$.
stelling van Green-Tao
Er bestaan rekenkundige rijen van iedere lengte die bestaan uit enkel priemgetallen.
In formulevorm bestaan er $a, b, n \in \mathbb{N}$ zodat $a, a + b, a + 2b, \cdots, a + (n - 1)b$ allen priem zijn.

Stelling 4.3. *(LTE: Lifting The Exponent Lemma)*
Er zijn enkele gevallen die we hier opsommen, die samen het totale LTE geven. Hierbij wordt met het symbool $v_p(x)$ het (exacte) aantal factoren p bedoeld in het getal x. bvb. $v_3(63) = 2$.
Zij p een priemgetal number en $x, y \in \mathbb{Z}$ die geen veelvoud zijn van p. Dan geldt dat
a) $\forall n \in \mathbb{N}$ *als*

- $p \neq 2$ *en* $p \mid x - y$, *dan*
$$v_p(x^n - y^n) = v_p(x - y) + v_p(n).$$
- $p = 2$ *en* $2 \mid x - y$, *dan*
$$v_2(x^n - y^n) = v_2(x - y) + v_2(x + y) + v_2(n) - 1.$$

b) $\forall n \in \mathbb{N}$ *die oneven zijn en zodat* $p \mid x + y$, *geldt er dat*
$$v_p(x^n + y^n) = v_p(x + y) + v_p(n).$$

Stelling 4.4. *(stelling van Zsigmondy)*

De stelling van Zsigmondy voor verschillen, zegt dat als $a > b > 0$ natuurlijke getallen zijn die relatief priem zijn en zei $f(n) = a^n - b^n$ met $n \in \mathbb{N} > 0$. Dan heeft $f(n)$ een priemfactor die niet voorkomt in $f(k)$ voor iedere $k \in \{1, 2, \cdots, n - 1\}$ uitgezonderd in enkele speciale gevallen:

$a + b = 2^z$ en $n = 2$, want $a^2 - b^2 = (a - b)(a + b)$ en de factor 2 zit in $a - b$.

$a = 2, b = 1$ en $n = 6$ want 63 bevat enkel priemfactoren $2^3 - 1 = 7, 2^2 - 1 = 3$

de stelling van Zsigmondy voor sommen:
op analoge wijze geeft $a^n + b^n$ namelijk een priemfactor die voor geen enkele kleinere $k \in \mathbb{N}$ in $a^k + b^k$ zit, met uitzondering van $a = 2, b = 1, n = 3$.

Voorbeeld 4.5. *Bewijs dat er oneindig veel priemgetallen zijn van de vorm $4k + 3$.*

Bewijs. Dit is een voorbeeld van de stelling van Direchlet, dat combinatorisch te bewijzen is:
Uit het ongerijmde zijn het er een eindig aantal n, schrijf dat P_1, P_2 tot P_n al de priemgetallen zijn van die vorm.
Bekijk nu $4(\prod P_i) - 1 \equiv 3 \pmod 4$ dat relatief priem is met de n andere priemgetallen, contradictie aangezien het niet het product kan zijn met priemgetallen die $\equiv 1 \pmod 4$ zijn.

□

Voor meer info over LTE; volledigere LTE-bestand

1. Vind de hoogste waarde van k zodat 1991^k een deler is van $1990^{1991^{1992}} + 1992^{1991^{1990}}$.

   link

2. Paul Erdos bewees in 1932 het zogenaamde postulaat van Bertrand, dit is de volgende bewering:

   "voor ieder natuurlijk getal $n > 1$ ligt er minstens n priemgetal tussen n en $2n$".

   Je mag deze stelling zonder bewijs aannemen. Gebruik nu het postulaat van Bertrand om volgende stelling te bewijzen: "Voor elke $k \in \mathbb{N}$ kan de verzameling $\{1, 2, \ldots, 2k-1, 2k\}$ opgedeeld worden in k paren, waarvan de som telkens een priemgetal is."

   Voor $k = 4$ krijgen we bijvoorbeeld $\{1, 6\}$, $\{2, 3\}$, $\{4, 7\}$ en $\{5, 8\}$.

   link

3. Zij b, m, n natuurlijke getallen zodat $b > 1$ en $m \neq n$. Bewijs dat als $b^m - 1$ en $b^n - 1$ dezelfde priemdelers hebben, dan $b + 1$ dan een macht van 2 is.

   link

4. Zij n een positief geheel getal en zij k een oneven natuurlijk getal. Laat bovendien a, b en c gehele getallen zijn (niet noodzakelijk positief) waarvoor geldt: $a^n + kb = b^n + kc = c^n + ka$: Bewijs dat $a = b = c$. link

5. Bepaal alle viertallen (a, b, p, n) van positieve gehele getallen $(a > 0, b > 0, p > 0, n > 0)$, zodanig dat p een priemgetal is en $a^3 + b^3 = p^n$.

   link

6. Bestaat er een natuurlijk getal n zodat n precies 2000 priemdelers heeft en $n|2^n + 1$?

   link

7. Bepaal alle drietallen (a, m, n) van natuurlijke getallen zodat $a^m + 1|(a + 1)^n$.

   link

8. Vind alle viertallen (p, a, b, c) met p priem en $a, b, c > 0$ gehele getallen zodat geldt dat $a^p + b^p = p^c$.

   link

9. Zij $p_1, p_2, \ldots, p_n$ verschillende priemgetallen groter dan 3.
   Toon aan dat $2^{p_1 p_2 \cdots p_n} + 1$ minimum 4^n delers heeft.

   link

10. $P(x) \in \mathbb{Z}[X]$ is een veelterm van graad ≥ 2. Bewijs dat er een natuurlijk getal m bestaat zodat $P(m!)$ een samengesteld getal is. link

4.3 kwadratische stellingen

Fermat's kerstmisstelling: Voor ieder priemgetal $\equiv 1 \pmod 4$ bestaat er 1 en slechts 1 oplossing $a > b \in \mathbb{N}$ zodat $p = a^2 + b^2$.

Ieder getal waarvor geldt dat de priemfactoren $\equiv 1 \pmod 4$ een even aantal keer voorkomt, kan geschreven worden als de som van 2 volkomen kwadraten $a^2 + b^2$.

Er zijn exact $\prod_{p \in P}(v_p(n) + 1)$ verschillende oplossing voor a, b waarbij enkel $p \in P$ als $p \equiv 1 \pmod 4$.

Ieder getal die niet van de vorm $\equiv 7 \pmod 8$ is, kan worden geschreven als de som van 3 volkomen kwadraten.

De vierkwadratenstelling toont dat alle getallen kunnen worden geschreven als de som van 4 volkomen kwadraten.

Ieder getal groter dan 100 tot slot kan worden geschreven als de som van exact 5 volkomen kwadraten (waarmee we bedoelen dat er geen 0 wordt gebruikt zoals bij de anderegevallen kan zijn)

som-van-n-machtenfeit

Zij p een priemgetal, dan geldt voor ieder geheel getal a dat de congruentie $x_1^n + x_2^n + ... + x_s^n \equiv a (mod\ p)$ een oplossing geeft als $s \geq n$.

lemma in kwadratische vergelijking mod p Als a, b onderling priem zijn met p

* als er geldt dat $p | a^2 - D \cdot b^2$, dan geldt dat D een kwadraatrest is modulo p.

* $ax^2 + bx + c \equiv 0 \pmod p$ heeft $1 + \left(\frac{b^2-4ac}{p}\right)$ oplossingen modulo p

Stelling 4.6. *(Pellvergelijkingen)*
De vergelijking $x^2 - dy^2 = 1$ heeft oneindig veel oplossingen als $d \in \mathbb{N}$ en geen volkomen kwadraat is. Deze oplossingen zijn van de vorm (x_n, y_n) met
$x_n = \frac{(x_1+y_1\sqrt{d})^n+(x_1-y_1\sqrt{d})^n}{2}$ *en* $y_n = \frac{(x_1+y_1\sqrt{d})^n(x_1-y_1\sqrt{d})^n}{2\sqrt{d}}$

uitgebreide pellvgl:
als $x^2 - ky^2 = -1$ een oplossing heeft (dit kan als $k \equiv 1 \pmod 4$)
Dan noemen we de primitieve oplossing (x_0, y_0) zoals bij de normale pellvgl.
We hebben dat $x_m + y_m\sqrt{k} = (x_0 + y_0\sqrt{k})^{2m+1}$ voor $m \geq 1$.
Bij $x^2 - ky^2 = n$ noteren we (a_0, b_0) als primitieve oplossing van deze vergelijking,
en (x_m, y_m) de m^{de} oplossing van $x^2 - ky^2 = 1$,
dan moeten we $a_m + b_m\sqrt{k} = (a_0 \pm b_0\sqrt{k})(x_m \pm y_m\sqrt{k})$ uitwerken voor de andere oplossingen van de eerste pellvergelijking: a_m, b_m.
Extra info te vinden via Pellvergelijking-bijlage

Stelling 4.7. *(priemdelers van $(a^p - 1)$-lemma)*
Dit lemma zegt dat als $q | \frac{a^p-1}{a-1}$ waarbij p, q beide priemgetallen zijn, dat dan geldt dat $p = q$ of $q \equiv 1 \pmod p$.

extraatje: de vergelijking $P(x) = x^{p-1} - 1$ heeft $p - 1$ wortels in $\mathbb{Z}_{p^2}$.

Stelling 4.8. *(Mihaleiscu)*
De vergelijking $y^m = x^n + 1$ heeft enkel de oplossing $3^2 = 2^3 + 1$ bij de voorwaarden $m, n, x, y \in \mathbb{N} > 1$

1. Toon aan dat de verzameling van de natuurlijke getallen die niet kunnen voorgesteld worden als de som van verschillende volkomen kwadraten eindig is. link

2. Bewijs dat er oneindig veel natuurlijke getallen n bestaan zodat $p = nr$, met p en r respectievelijk de halve omtrek en de straal van de ingeschreven cirkel een driehoek met gehele zijdelengtes.

   link

3. Zij p een priemgetal. Bewijs dat er een priemgetal q bestaat zodat voor ieder natuurlijk getal n geldt dat $n^p - p$ niet deelbaar is door p. link

4.4 Vieta-jumping

De methode van *Vieta-jumping* (vaak *root-flipping* genoemd) is toepasbaar op een zeer herkenbare klasse van problemen, die vaak van een bijzonder hoge moeilijkheidsgraad zijn. De problemen waarop deze methode werkt zijn meestal gekarakteriseerd door het concept van deelbaarheid van natuurlijke getallen en het veelvuldig voorkomen van kwadraten.

Het basisidee is geworteld in Fermat's methode van de oneindige afdaling. Er wordt een fictieve oplossing gekozen die een bepaalde grootheid minimaliseert, maar waarvoor de te bewijzen eigenschap niet geldt. De gegeven relatie wordt dan als een kwadratische vergelijking in een van de variabelen herschreven. Met behulp van de formules van Vieta wordt vervolgens een oplossing geconstrueerd die de grootheid nog kleiner maakt. De methode algemeen omschrijven is weinig nuttig, ze laat het duidelijkst haar kracht zien wanneer we ze leren kennen in de context van concrete problemen. De problemen die met deze methode oplosbaar zijn, zijn naast zeer herkenbaar ook zeer dun gezaaid. We beginnen met het klassieke voorbeeld, waar overigens een verhaal aan verbonden is.

Voorbeeld 4.9. *Zijn x en y natuurlijke getallen zodat xy een deler is van x^2+y^2+1. Bewijs dat*

$$\frac{x^2+y^2+1}{xy} = 3.$$

Oplossing. Stel $\frac{x^2+y^2+1}{xy} = k$. Fixeer nu k en beschouw de verzameling

$$\mathcal{A} = \left\{ (x,y) \in \mathbb{N}^2 \;\mid\; \frac{x^2+y^2+1}{xy} = k \right\}.$$

Neem nu de het koppel (X,Y) in $\mathcal{A}$ dat de minimale waarde van $x+y$ oplevert. Als er meerdere koppels zijn neem er dan willekeurig een van.

Veronderstel dat $X > Y$ en beschouw nu de volgende vergelijking in t:

$$\frac{t^2+Y^2+1}{tY} = k \iff t^2 - kYt + Y^2 + 1 = 0$$

We weten dat $t = X$ een oplossing is van deze vergelijking, en uit de formules van Vieta halen we dat

$$x_2 = kY - X = \frac{Y^2+1}{X}$$

ook een oplossing is. Uit $x_2 = kY - X$ halen we dat x_2 een geheel getal is. Verder leiden we af dat $x_2 = \frac{Y^2+1}{X} < X$ en dat bovendien x_2 positief is. Uit al deze informatie besluiten we dat $(x_2, Y) \in \mathcal{A}$, en dat bovendien $x_2 + Y < X + Y$, hetgeen een contradictie is.

Uit het voorgaande leiden we af dat de oplossing die $x+y$ minimaliseert noodzakelijk $x = y$ moet hebben. We krijgen dat $k = \frac{2X^2+1}{X^2}$ en dus dat $X = 1$. Hieruit volgt het gestelde. □

1. (klassieker)

   Gegeven zijn positieve gehele getallen a en b waarvoor geldt dat $ab+1$ een deler is van a^2+b^2. Bewijs dat $\frac{a^2+b^2}{ab+1}$ het kwadraat van een geheel getal is. link

2. Voor welke natuurlijke waarden van n heeft de vergelijking

   $$a+b+c+d = n\sqrt{abcd}$$

   oplossingen met a, b, c, d natuurlijke getallen? link

3. $a, b \in \mathbb{N}$. Bewijs dat $4ab-1 | (4a^2-1)^2$ aesa $a = b$. link

4. Bepaal alle koppels natuurlijke getallen (a, b) zodat

   $$\frac{a^2}{2ab^2 - b^3 + 1}$$

   een natuurlijk getal is.

   link

5. Bestaat er een natuurlijk getal m zodat de vergelijking

   $$\frac{1}{a} + \frac{1}{b} + \frac{1}{c} + \frac{1}{abc} = \frac{m}{a+b+c}$$

   oneindig veel oplossingen heeft in natuurlijke getallen a, b, c? link

6. $k \in \mathbb{N}$ TB: $(4k^2-1)^2$ hebben een deler van de vorm $8kn-1$ aesa k even is. link

4.5 orde en kleinste elementen

De orde van een bewerking is het kleinste aantal keer dat je het moet gebruiken om tot het eenheidelement te komen.

Bij machten zeggen we dat de orde o het kleinste natuurlijke getal is zodat $a^o \equiv \pmod{b}$ $(ggd(a, 0) = 1)$

Kijken naar ordes en de kleinste priemdelers, kan vele vragen helpen op te lossen.

Primitieve wortels g mod m zijn getallen die orde ϕm hebben.

Een getal m geeft enkel een primitieve wortels als m v.d. vorm $2p^k, p^k, 2$ of 4 met p priem.

Als a een primitieve wortel is mod. p, dan is a of $a + p$ de primitieve wortel voor p^k met $k \geq 2$.

1. (a) Zij $n > 1$ een natuurlijk getal en $a \in \mathbb{Z}$ zodat $n|a^n - 1$. Bewijs dat de kleinste priemdeler van n een deler van $a - 1$ is. (b) Vind alle natuurlijke n $n^2|3^n + 1$

   link

2. Bepaal alle gehele getallen $n > 1$, zodanig dat $\frac{2^n+1}{n^2}$ een geheel getal is.

   link

3. Vind alle koppels natuurlijke getallen (x, p) met p een priemgetal, $x \leq p$ en $x^{p-1}|(p-1)^x + 1$.
   link

Bij problemen bij het kiezen van het juiste element dat men moet beschouwen als orde, deler etc, help getal kiezen

4.6 kwadraatresten

Het legendresymbool werkt als volgt:

$$\left(\frac{n}{p}\right) = \begin{cases} 0 \text{ als } p|n \\ 1 \text{ als } n \text{ een kwadraatrest is van } p, \text{ nl. er bestaat een } x \in \mathbb{N} \text{ bestaat zodat } x^2 \equiv n \pmod{p} \\ \text{-1 als } n \text{ geen kwadraatrest is modulo } p \end{cases}$$

Euler's criterium zegt dat $\left(\frac{n}{p}\right) = n^{\frac{p-1}{2}}$ waardoor $\left(\frac{n}{p}\right)\left(\frac{m}{p}\right) = \left(\frac{mn}{p}\right)$ direct een gevolg is.

Wegens $p|(j-i)(j+i)$ zien we dat slechts de helft van de getallen tussen 1 en $p-1$ een kwadraatrest kunnen zijn.

De kwadratische receprociteitswet zegt dat $\left(\frac{n}{m}\right)\left(\frac{m}{n}\right) = (-1)^{(m-1)(n-1)}$ als m, n priem zijn.

Met de Chinese reststelling is het gemakkelijk te bewijzen dat voor $n = \prod_{i=1}^{i=m} p_i^{\alpha_i}$ a een kwadraatrest is, als en slechts als het een kwadraatrest is van iedere P_i.

Voor meer info, kun je naar de kwadratische resten-bijlage gaan.

1. Zij b een natuurlijk getal groter dan 5. Voor ieder natuurlijk getal n, beschouw het getal

$$x_n = \underbrace{11\ldots1}_{n-1}\underbrace{22\ldots2}_{n}5,$$

geschreven in basis b. Bewijs dat de volgende eigenschap geldt, als en slechts als $b = 10$: er bestaat een natuurlijk getal M zodat voor elk natuurlijk getal $n > M$ het getal x_n een volkomen kwadraat is. link

2. Bewijs dat er oneindig veel getallen n bestaan zodat $n^2 + 1$ een priemfactor groter dan $2n + \sqrt{2n}$ heeft.

   link

3. $a, b > 1$ met $a \neq b$ zijn natuurlijke getallen zodat ab geen volkomen kwadraat is, bewijs dat er een natuurlijk getal n bestaat zodat $(a^n - 1)(b^n - 1)$ geen volkomen kwadraat is.

   link

4.7 willekeurige getaltheorie

Dit omdat geen enkele olympiade zegt welke stellingen je zou kunnen gebruiken: mooie getaltheorie-vragen

5 meetkunde

5.1 basis

De grootste brok theorie en herhaling; Kort de basis overlopen en alle naamgeving gebeurt tot op pagaina 8 voor de volledigheid.
(zij die de baisis kennen, kunnen de rode woorden nog eens herhalen of slaan dit over , doch er nog interessante eigenschappen tussen zitten)

Een driehoek $\triangle ABC$ is gelijkbenig met $|AB| = |AC|$ als en slechts als de twee hoeken ("basishoeken") B en C gelijk zijn.
In gelijkvormige driehoeken $\triangle ABC$ en$\triangle XYZ$ zijn de overeenkomstige hoeken gelijk en de overeenkomstige zijden hebben een constante verhouding ("de zijden zijn evenredig").
I.e.$A = X, B = Y, C = Z$, $\frac{AB}{XY} = \frac{BC}{YZ} = \frac{CA}{ZX}$. Notatie $ABC \sim XYZ$
Twee driehoeken $\triangle ABC$ en $\triangle XYZ$ zijn congruent (gelijkvormig en de overeenkomstige zijden zijn even lang) als een van de volgende voldaan is ("congruentiekenmerken")
Alle overeenkomstige zijden even lang zijn ("ZZZ")
Twee overeenkomstige zijden even lang zijn, en de ingesloten hoeken gelijk zijn ("ZHZ")
Twee overeenkomstige hoeken gelijk zijn, en n overeenkomstig paar zijden even lang is ("HZHën "ZHH")
Voor gelijkvomigheid is het voldoende dat de verhouding van de zijden gelijk is (ZHH wordt dan HH)
Stelling van Pythagoras: in een rechthoekige driehoek met rechte hoek A geldt $|AB|^2 + |AC|^2 = |BC|^2$.
Als twee driehoeken een gemeenschappelijke top hebben, en een basis met dezelfde drager (de drager van een lijnstuk AB is de rechte AB), dan verhouden hun oppervlakten zich als de lengten van hun basissen.
Dus: voor driehoeken ABC en ADE, met B, C, D, E colineair (op dezelfde rechte) geldt $[ABC]/[ADE] = BC/DE$.
Dit volgt onmiddellijk uit öppervlakte driehoek = basis * hoogte / 2". Hierbij staat $[ABC]$ voor de oppervlakte van de driehoek ABC.

Nog wat naamkennis herhalen:

* De zwaartelijnen van een driehoek (uit een hoekpunt naar het midden van de overstaande zijde) snijden elkaar in 1 punt, het zwaartepunt Z van de driehoek, in de theorie staat er G

* De hoogtelijnen van een driehoek (uit een hoekpunt loodrecht op de overstaande zijde) snijden elkaar in n punt, het hoogtepunt H van de driehoek.

* De middelloodlijnen van een driehoek (de middelloodlijnen van de zijden) snijden elkaar in n punt, het omcentrum O van de driehoek.

* De bissectrices van een driehoek (de bissectrices van de hoeken) snijden elkaar in n punt, het incentrum I van de driehoek.

* (de negenpuntscirkel) het centrum van de negenpuntscirkel wordt met E of N dit is het middelpunt van de cirkel (de negenpuntscirkel) door:
- de middens van de zijden van de driehoek
- de middens van $[AH], [BH], [CH]$
- de voetpunten van de hoogtelijnen

eig. Vliegers:

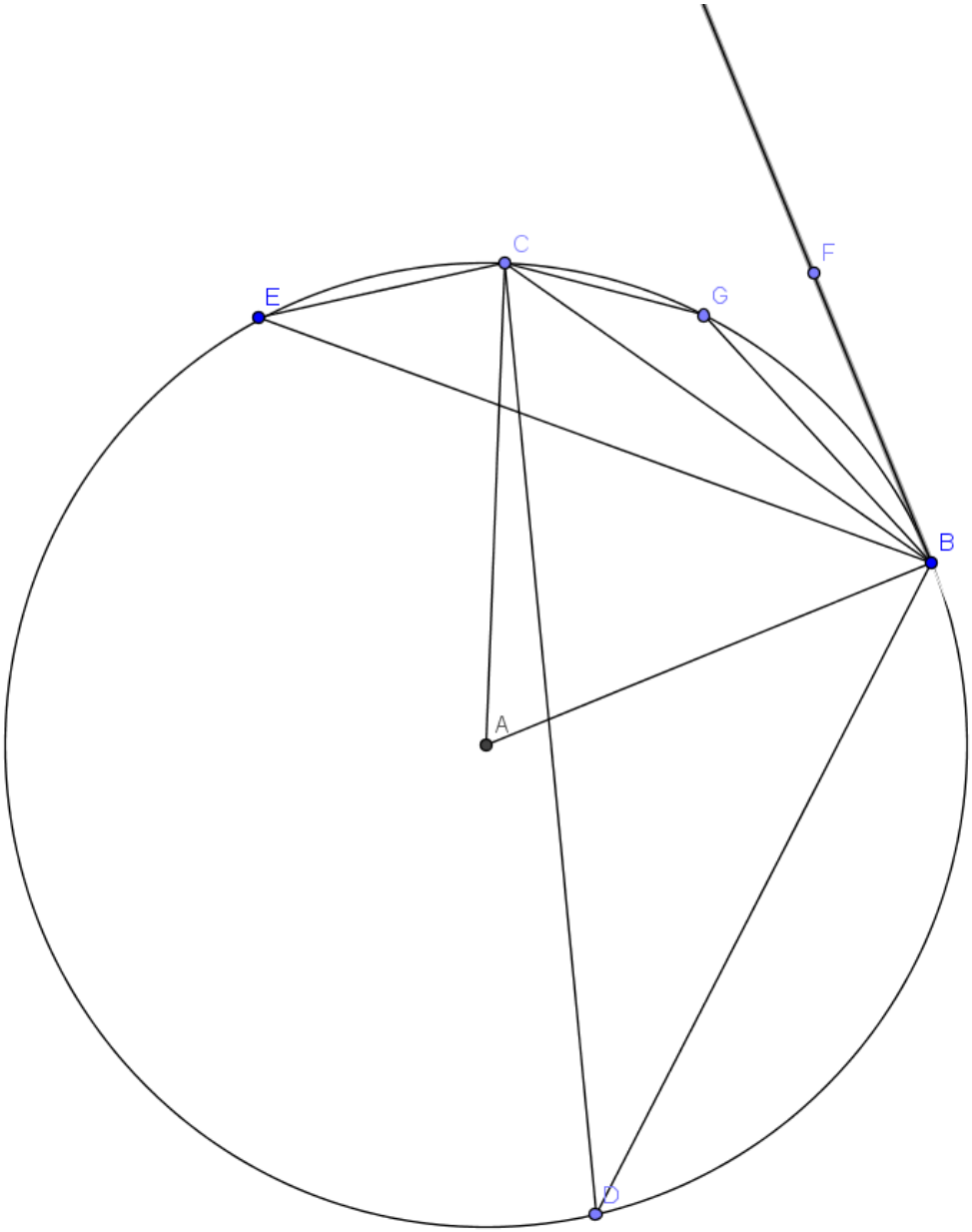
*2 paar aangrenzende zijden zijn even lang en *De diagonalen van een vlieger staan loodrecht op elkaar.

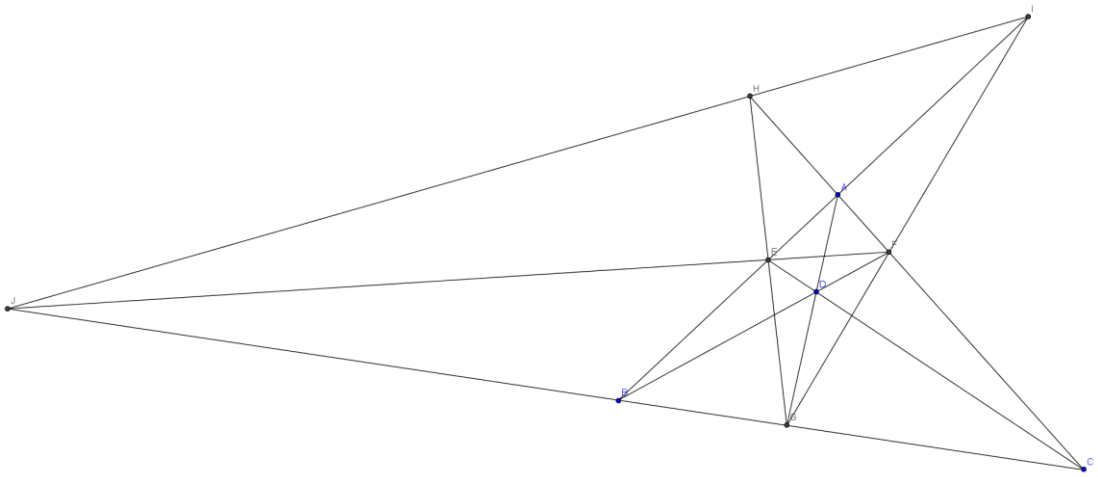
basiseigenschappen over koordenvierhoeken:

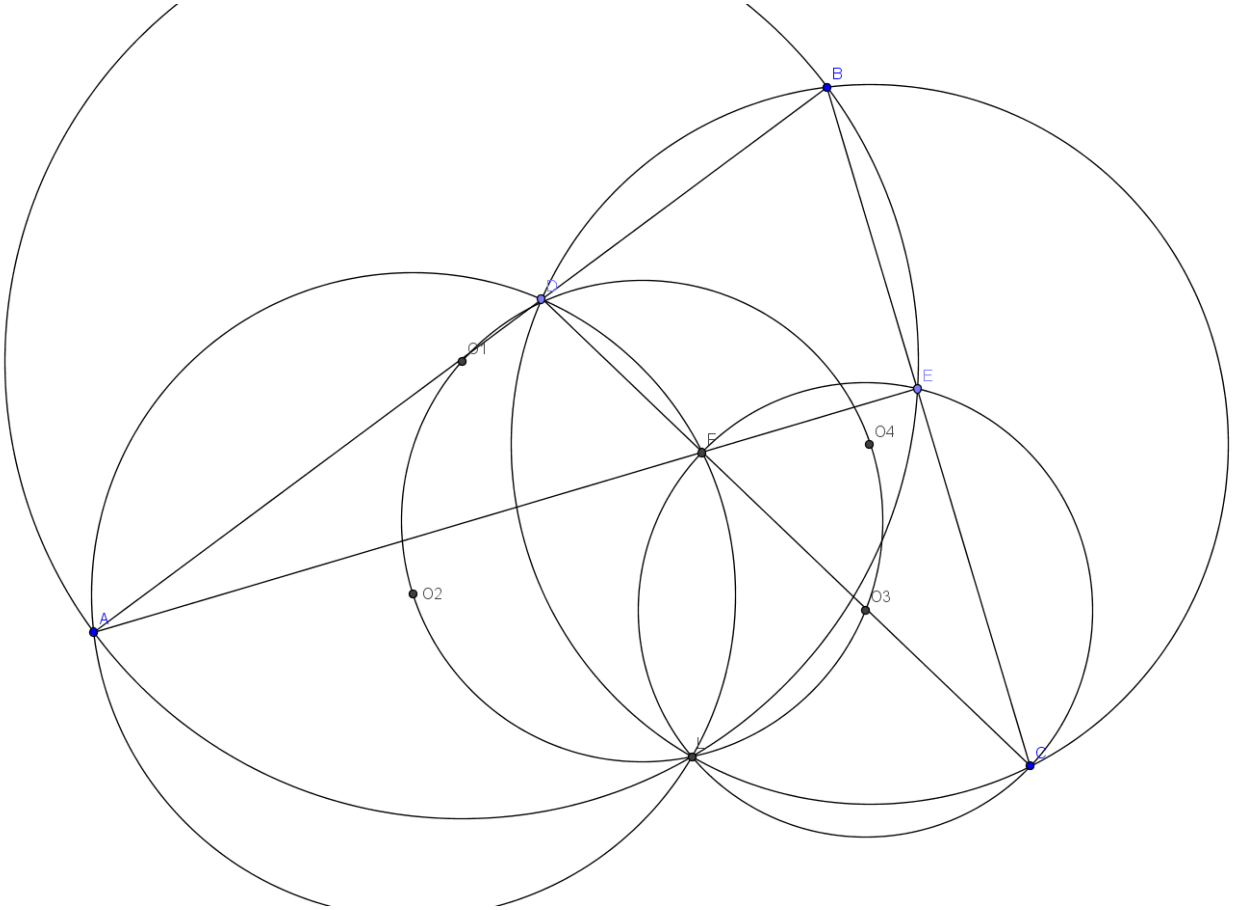
*de overstaande hoeken zijn supplementair

*omtrekshoeken op een gelijke boog zijn gelijk, met een waarde die de helft is van de middelpuntshoek op de boog.

Op volgende tekening zien we dus duidelijk dat $\angle CBF = \angle CDB = \angle CEB = 0.5\angle CAB$ met die eigenschappen.


C
F
E
G
B
A
D


I
H
K
F
G
J
B
D
C


B
D
O1
E
F
O4
O2
O3
A
K
C

vervollediging naamgeving

- de trilineaire pool en poollijn van een driehoek: zie bovenste tekening op vorige bladzijde. In een driehoek ABC is er een punt P (de trilineaire pool), AP, BP, CP snijden hun overstaande zijden in $A", B", C"$ dan snijden $A'B', AB$ en analoog op een rechte, deze noemen we de trilineaire poollijn (bestaat door de stelling van Desargues)

  Stelling 5.1. *(stellingen van Miquel bij vierhoeken)*

  M is het Miquel punt tov AB, CD, AC, BD wanneer geldt dat X het snijpunt is van de omgeschreven cirkels , van de snijpunten van ieder triplet lijnen. Stel $P = AC \cap BD, R = AB \cap CD$. Er geldt met de stelling dat M op de omgeschreven cirkels van ABP, CPD, RBD en RAC ligt. De hoogtelijnen van de 4 driehoeken liggen op een rechte (gevolg van de steinerlijn). De 4 omcentra zijn cyclisch met het Miquelpunt

  De stelling van Miquel zegt ook dat als in een driehoek $\triangle ABC$ met $D, E, F \in [AB], [AC], [CB]$, als 2 van de 3 omgeschreven cirkels $\odot CEF, \odot ADE, \odot BDF$ elkaar snijden in een punt M inwendig in de driehoek, dan ligt dat punt ook op de derde cirkel.

- ceviaandriehoek:

  $\triangle A"B"C"$ is de ceviaandriehoek van $\triangle ABC$ tov het punt P als

  $A", B", C"$ gelijk zijn aan $AP \cup BC, BP \cup AC$ en $CP \cup AB$ resp.

- anticeviaandriehoek, als $\triangle A'B'C'$ is de anticeviaandriehoek van $\triangle ABC$ tov het punt P als

  A op de rechte $B'C'$ ligt en zo ook $C', A,' B$ collineair zijn en $C \in A'B'$

  $P = AA' \cup BB' \cup CC'$

  $\triangle ABC$ de ceviaandriehoek is van $A'B'C'$ tov P

  Er geldt dat $(AA"PA') = -1$ en de cyclische viertallen zijn harmonisch (zie sectie deelverhoudingen).

  Enkele anticeviaandriehoeken van belangrijke punten:

  I : excentral triangle

  Z: anticomplementaire driehoek

  symmediaanpunt: raaklijnendriehoek (driehoek gevormd door de raaklijnen)

- circumceviaandriehoek / circumceviantriangle

  In $\triangle ABC$ is P een inwendig punt (geen hoekpunt), dan is $A'B'C'$ de circumceviaandriehoek als $A' = PA \cup \tau$ en analoog met τ de omgescreven cirkel. Als $O = P$ is A' de antipode / diametraal overgesteld punt van A.

- contactdriehoek (intouchtriangle) voetpuntsdriehoek van I

- aanraakdriehoek: driehoek gevormd door de raakpunten van $\triangle ABC$ met zijn aangeschreven cirkels

- isogonaal verwant: 2 punten zijn isogonaal verwant als de rechten door deze punten en een hoekpunt van de driehoek, symmetrisch zijn tov de bissectrice uit dat hoekpunt en dit voor ieder hoekpunt (hun overeenkomstige cevianen/ hoektransversalen hebben een omgekeerde verhouding tot de zijden)

  De voetpuntsdriehoeken van 2 isogonaal verwante punten liggen op 1 cirkel, waarvan het middelpunt het midden is van die 2 punten.

- isotomisch verwantschap:2 punten zijn isotomischl verwant als de loodlijnen op een zijde van de driehoek, symmetrisch zijn tov de middelloodlijnen uit dat hoekpunt en dit voor iedere zijde (vb. het punt van Gergonne en het punt van Nagel zijn isotomisch verwant)

- symmedianen: De symmedianen zijn de spiegelbeelden van de zwaartelijnen in de bissectrices =de rechten isogonaal verwant met de zwaartelijnen
- Het punt van Lemoine in een driehoek is het snijpunt van de symmedianen. Het punt van Lemoine is het punt dat de som van de kwadraten van de afstanden tot de zijden van de driehoek minimaliseert en is isogonaal verwant met het zwaartepunt.
- het punt van Nagel: X, Y, Z zijn de raakpunten van I_a, I_b, I_c met de driehoek $\triangle ABC$, dan is het N_a het punt van concurrenctie van AX, BY en CZ.
- rechte van Nagel: N_a, I, S (punt van Spieker dat het incentrum is van de middendriehoek) en Z liggen op deze rechte in de verhouding $|NS| : |SZ| : |ZI| = 3 : 1 : 2$.
- (rechte van Euler): H, N, Z, O op 1 rechte liggen in die volgorde en de verhoudingen zijn: $HN : NZ : ZO = 3 : 1 : 2$
- het inwendig punt van Gergonne: het snijpunt van de lijnen door de hoekpunten en de raakpunten van de ingeschreven cirkel aan de overstaande zijden. uitwendig punt van Gergonne: snijpunten van cevianen door de raakpunten van 1 aangeschreven cirkel
- de Longchapspunt L: het hoogtepunt v.d. anticomplementaire driehoek en is het punt zodat O het midden is van HL.
- Bevanpunt: omcentrum V van de aangeschreven driehoek/excentral triangle $I_aI_bI_c$ ($\triangle$ gevormd door aancentra)

  voetpuntsdriehoek van V is de aanraakdriehoek

  V = het midden van $[NaL]$ en ligt er samen met V' (isogonaal verwante punt van V')

  O is het midden van $[IV]$

  het spiekerpunt is het midden van $[HV]$
- Het punt van Fermat:

  In het geval dat de grootste hoek van de driehoek kleiner is dan 120, is de totale afstand van het punt naar de drie hoekpunten minimaal.

  De binnenste hoeken, gevormd door dit punt: $\angle AFB, \angle BFC, \angle CFA$ zijn alle gelijk aan 120°

  De omschreven cirkels van de drie gelijkzijdige driehoeken van de constructie snijden in dit punt

  De driehoek, gevormd door de centra van de drie gelijkzijdige driehoeken in de constructie is ook een gelijkzijdige driehoek (Stelling van Napoleon) en het centrum van de omgeschreven cirkel van deze driehoek is het punt van Fermat van de originele driehoek

  Als $\triangle ABZ, \triangle ACY, \triangle BCX$ de uitwendige gelijkzijdige driehoeken zijn, geldt dat $F = CZ \cup BY \cup AX$. F is het punt waarvoor $AP + BP + CP$ minimaal is, bewijs door Z' te nemen door P 60 te draaien rond A richting Z, waarna $PZ' = AP$ en $|ZZ'| = |PB|$.
- Brocardpunten: zij $O1$ het eerste punt van Brocard van driehoek $\triangle ABC$, dan geldt dat $\angle O1AB = \angle O1BC = \angle O1CA = \gamma = \angle ABO_2 = \angle BCO_2 = \angle CAO_2$ met O_2 het tweede Brocardpunt, dat het isogonaal geconjugeerd punt van O_1 is.
- **De rechte van Simson**

  Drie punten zijn collineair als en slechts als de driehoek gevormd door deze drie punten een oppervlakte heeft die nul is. Ga nu zelf met behulp van eigenschap 3 uit de vorige paragraaf de volgende stelling na:

  De projecties van een punt P op de zijden van ABC zijn dan en slechts dan collineair als P op de omgeschreven cirkel van ABC ligt.

De rechte die de drie projecties van het punt P bevat noemt men de *rechte van Simson* van punt P t.o.v. ABC.

Deze stelling zal je soms van pas komen wanneer je een probleem te lijf gaat. Tracht ook als oefening eens een rechtstreeks bewijs te vinden, dus zonder de uitdrukking voor de oppervlakte van een voetpuntsdriehoek te gebruiken.

- **de Steinerlijn** is de lijn l gevormd door een punt P op de omgeschreven cirkel te spiegelen over AB, BC, AC en gaat door H.
  Het is de homothetie met center in P van de Simsonlijn met een factor 2.
  Het punt P wordt het antiSteinerpunt van l tov $\triangle ABC$ genoemd

De basis om een meetkundevraag op een problem-solving-wedstrijd bestaat uit het volgende;

- isometrieen zoals homotheties, verschuivingen, verdraaiingen
- gelijkvormigheid en congruentie
- de eigenschappen van koordenvierhoeken en omtrekshoeken: omtrekshoeken op eenzelfde boog zijn gelijk en dus ook de hoeken op dezelfde zijde binnen een koordenvierhoek, wiens overstaande hoeken een som van 180 graden heeft.
- vectoren.
- angle-chasing: hoekjagen, vaak start men met hoekjagen om hieruit de conclusies i.v.m. cirkelbogen te kunnen trekken
- constructie van interessante punten die helpen het probleem op te lossen.
  Wanneer men iets moet bewijzen dat niet direct meetkundig te interpreteren is, kan een constructie van een nieuw punt vaak helpen, bvb. als de som van 2 lengtes gelijk moet zijn aan 1 lengte.
  Bij collineairiteit, kan het helpen, wanneer men een nieuw punt P beschouwt, als P, A, B en P, A, C collineair zijn, zijn A, B, C ook collineair.
- Als het niet duidelijk is, hoe iets te bewijzen valt, moet men het goede idee/ oplossing soms zien van het meetkundig probleem en dan bewijzen dat die oplossing idd voldoet.

Wie het toch nodig vindt, kan de uitgebreide bijlage basis meetkunde eerst eens doornemen.
Stellingen kennen, is zeker handig, maar toch zal er eerst wat moeten geoefend worden om de inzichten te hebben.

1. In $\triangle ABC$ is $\widehat{A} = 66°$ en $|AB| < |AC|$. De buitenbissectrice in A snijdt BC in D en $|BD| = |AB| + |AC|$. Bepaal de hoeken van $\triangle ABC$. link

2. $\triangle ABC$ is een scherphoekige driehoek en ω is een cirkel met middelpunt $L \in [BC]$ die raakt aan $[AB], [AC]$ in B', C'.
   Het omcentrum O van $\triangle ABC$ ligt op de kleine boog $B'C'$ van ω.
   Bewijs dat de omcirkel $\odot O$ en ω snijden in 2 verschillende punten. link

3. Zij ABC een scherphoekige driehoek met omgeschreven cirkel τ en hoogtepunt H. Zij K een punt op τ aan de andere kant van BC dan A. Zij L het spiegelbeeld van K in de lijn AB en zij M het spiegelbeeld van K in de lijn BC. Zij E het tweede snijpunt van τ met de omgeschreven cirkel van driehoek BLM. Bewijs dat de lijnen KH, EM en BC door n punt gaan. (Het hoogtepunt van een driehoek is het punt dat op alle drie de hoogtelijnen ligt.)

   link

4. Gegeven zijn vijf punten A, B, C, Den E zodanig dat $ABCD$ een parallellogram is en $BCED$ een koordenvierhoek.
   Zij l een lijn (een rechte) door A, die het inwendige van het lijnstuk DC snijdt in F en die de lijn BC snijdt in G. Veronderstel dat $|EF| = |EG| = |EC|$.

   Bewijs dat l de bissectrice is van hoek DAB. link

5. Zij $ABCD$ een convexe vierhoek met AB niet parallel aan CD, en X een inwendig punt zodat $\angle ADX = \angle BCX < 90°$ en $\angle DAX = \angle CBX < 90°$.
   Als Y het snijpunt is van de middelloodlijnen van AB en CD, bewijs dan dat $\angle AYB = 2\angle ADX$. link

6. Zij AH_1, BH_2, CH_3 de hoogtelijnen van een scherphoekige driehoek ABC.
   De ingeschreven cirkel raakt de zijden BC, AC, AB in T_1, T_2, T_3 respectievelijk.
   Beschouw de spiegelbeelden van de rechten H_1H_2, H_2H_3, H_3H_1 ten opzichte van de rechten T_1T_2, T_2T_3, T_3T_1.
   Bewijs dat deze beelden een driehoek vormen waarvan de hoekpunten op de ingeschreven cirkel van ABC liggen.link

7. Zij ABC een driehoek met $\angle BAC = 90°$. Zij AP de bissectrice van $\angle BAC$ en BQ de bissectrice van $\angle ABC$, met P op BC en Q op AC. Als $AB + BP = AQ + QB$, wat zijn dan de hoeken van de driehoek? link

8. $ABCD$ is een convexe vierhoek met Γ_1 en Γ_2 de incirkels van $\triangle ABC$ en $\triangle ADC$ respectievelijk.

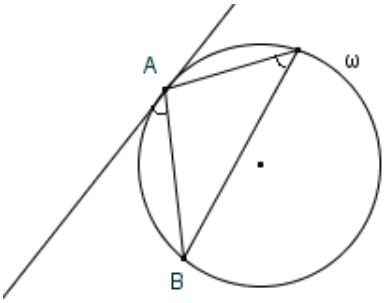
   Stel dat er een cirkel Γ is die raakt aan de (verlengde) zijden BC, AB, AD, CD.

   Bewijs dat de gemeenschappelijke uitwendige raaklijnen van Γ_1 en Γ_2 elkaar snijden op Γ

   link

5.2 lemma's

In deze paragraaf bekijken we enkele lemma's van dichterbij die ongewoon vaak hun intrede deden in IMO-problemen de voorbije jaren. Waar Lemma 1 misschien ook als fundamentele stelling bestempeld zou kunnen worden, zijn Lemma 2 en vooral Lemma 3 ongemeen belangrijk voor elke IMO-deelnemer.



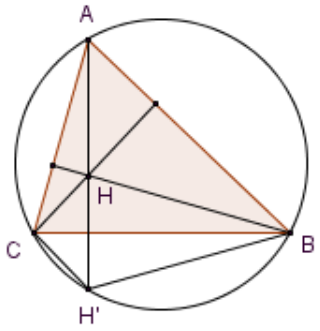
Figuur 1: De raakomtrekshoek

Lemma 1 (Raakomtrekshoek) *Beschouw een cirkel ω die de punten A en B bevat. De raaklijn aan ω in A sluit een hoek in met AB die in grootte gelijk is aan een van beide omtrekshoeken op AB in ω.*

(Bewijs als oefening)

Lemma 2 *De reflecties van het hoogtepunt H van ABC ten opzichte van de zijden liggen op de omgeschreven cirkel van ABC.*

Het bewijs van dit lemma is eenvoudig en kan als oefening dienen voor de lezer.

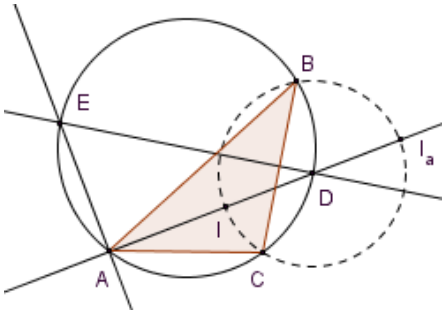


Figuur 2: Lemma 2

Lemma 3 *In driehoek ABC noemen we I het middelpunt van de ingeschreven cirkel, en I_a het middelpunt van de aangeschreven cirkel tegenover A.*

- *De binnen(resp. buiten)bissectrice van A snijdt de middelloodlijn van BC in het punt D (resp. het punt E) op de omgeschreven cirkel.*
- *De cirkel met diameter II_a bevat B en C en heeft D als middelpunt.*
- *De cirkel door B, C, I_c, I_b geeft E als middelpunt*

Dit lemma is allicht het belangrijkste uit deze hele paragraaf, en een van de vaakst terugkerende lemmata in oplossingen van IMO-problemen.
Het bewijs is een goede oefening angle-chasing, maar he is ook een gevolg van het feit dat ABC de negenpuntscirkel is van $I_aI_bI_c$.



Figuur 3: Lemma 3

Lemma 4 *In driehoek ABC zijn O en H isogonaal verwant*
Bewijs. met angle-chasing: $\angle BAH = \angle CAO = 90 - \alpha$ en analoog □

uitgebreide bijlage over vele lemma's

1. Twee cirkels G_1 en G_2 snijden in M en N. Zij AB de rechte die raakt aan beide cirkels in A en B respectievelijk, zodat M dichter bij AB ligt dan N. Zij CD de rechte parallel met AB die door M gaat en C op G_1 en D op G_2. De rechten AC en BD snijden in E; de rechten AN en CD snijden in P; de rechten BN en CD snijden in Q. Toon aan dat $EP = EQ$.

   link

Lemma 5 *In een driehoek $\triangle ABC$ met omgeschreven cirkel $\odot O$ snijden de raaklijnen in B en C in D, dan geldt dat AD een symmediaan is in de driehoek*

Lemma 6 *In een driehoek $\triangle ABC$ met ingeschreven cirkel $\odot I$ die $[BC]$ raakt in D, is DE de diameter van $\odot I$.*
De lijn AE snijdt $[BC]$ in F. Dan geldt dat $|BD| = |CF|$.
Bewijs. Merk op dat een homothetie uit A die $\odot I$ naar $\odot I_a$ afbeeldt, ook E op F afbeeldt. □

lemma 7 *$[AB]$ en $[CD]$ zijn 2 lijnstukken. Er bestaat dan een spiral similarity (samenstelling van rotatie met een draaiing) die het ene lijnstuk op het andere afbeeldt.*
Zij $X = AC \cup BD$, dan is $O = \odot ABX \cup \odot CDX$ het centrum van die spiral silmilarity.

lemma 8A *In een driehoek $\triangle ABC$ met incenter I construeren we de mixtilineaire incirkel die raakt aan $[AB], [AC]$ in X, Y en aan de omcirkel $\odot O$ in P.*
Dan is I het midden van $[XY]$.

lemma 8B *In een driehoek $\triangle ABC$ met incenterl I en $D \in [BC]$ construeren we de cirkel die raakt aan $[AD], [DC]$ in X, Y en aan de omcirkel $\odot O$ in P.*
Dan is $I \in [XY]$.

lemma 9a *De incirkel $\odot I$ raakt de zijden $[AB], [AC], [BC]$ in F, E, D resp.*
Als M het midden is van $[BC]$ geldt dat EF, DI, AM concurrent zijn

lemma 9b *De incirkel $\odot I$ raakt de zijden $[AB], [AC], [BC]$ in F, E, D resp.*
Als M het midden is van $[BC]$ en $T = CI \cup EF$, geldt dat B, D, I, T, F cyclisch zijn

lemma 10 *2 cirkels $\odot I, \odot O$ raken elkaar inwendig in P, waarbij $\odot I$ de kleinste cirkel is.*
$A, B \in \odot O$ verschillend van P, AA', BB' zijn raaklijnen aan $\odot I$ met $A', B' \in \odot I$. Dan geldt dat $\frac{|AA'|}{|AP|} = \frac{|BB'|}{|BP|}$

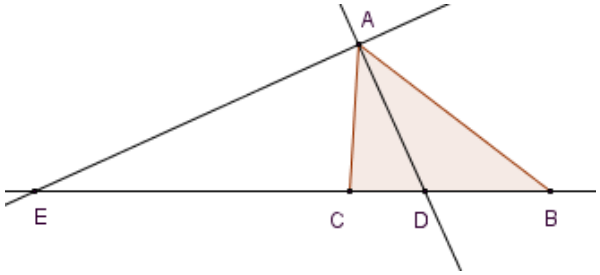
Via voorbeelen genoeg met enkel lemma 10 zie je hoe zinvol ieder lemma apart kan zijn.

De oefensectie:

1. A_1, B_1, C_1 zijn punten die gekozen zijn op de zijdne $BC, AC; AB$ van een driehoek ABC. De omgeschreven cirkels van AB_1C1, A_1BC_1, A_1B_1C snijden de omgeschreven cirkel van $\triangle ABC$ in A_2, B_2, C_2 resp. Punten A_3, B_3, C_3 zijn de spiegelbeelden van A_1, B_1, C_1 tov de middens van BC, AC, AB resp. Bewijs dat $\triangle A_2B_2C_2 \sim \triangle A_3B_3C_3$.

basisstellingen

De bissectricestelling



Figuur 4: De bissectricestelling

Een eigenschap van de bissectrices van een driehoek ABC die in talloze problemen als een zeer fundamentele stelling opduikt is de volgende:

De binnen- en buitenbissectrice van hoek α snijdt BC in D en E respectievelijk.
Er geldt dat

$$\frac{AB}{AC} = \frac{BD}{CD} = \frac{BE}{CE}$$

Reim's stelling

Zij $ABCD$ een koordenvierhoek en X, Y op de rechten AC, BD zodat $XY//BC$, dan is $ADXY$ cyclisch.
Merk op dat dit ook geldt voor X, Y op AB, CD met $XY//BC$ dat $ADXY$ ook cyclisch is en dit ook geldt in omgekeerde zin (2 koordenvierhoeken, rechte door snijpunten snijdt in 4 punten, waarvan er 2 bij 2 evenwijdig zijn).

Voetpuntsdriehoeken Wanneer je in een probleem een punt binnen een driehoek ontmoet, kan het vaak nuttig zijn om de zogenaamde *voetpuntsdriehoek* te beschouwen. De voetpuntsdriehoek van een punt P in een driehoek ABC is per definitie de driehoek gevormd door de loodrechte projecties van P op de zijden van ABC. Zo zullen bijvoorbeeld de middens van de zijden van ABC de hoekpunten van de voetpuntsdriehoek van O zijn. De voetpuntsdriehoek heeft enkele eigenschappen die af en toe tot een zeer korte oplossing van een probleem kunnen leiden.

Eigenschappen *In ΔABC tekent men de voetpuntsdriehoek DEF van P. Er geldt dat*

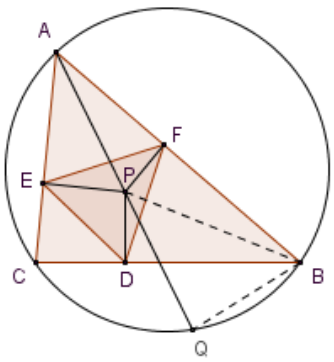
- $\angle EDF = \angle CPB - \angle CAB$
- $|DE| = |CP| \cdot \sin C$
- $\text{Opp(DEF)} = \frac{1}{4}\left|1 - \frac{|OP|^2}{R^2}\right| \cdot \text{Opp(ABC)}.$

Essentieel bestaat het bewijs van de eerste twee eigenschappen uit het optellen van hoeken en het gebruiken van de sinusregel. De derde eigenschap is minder eenvoudig en verdient speciale aandacht. We geven eerst een bewijs:

Noem Q het tweede snijpunt van AP en de omgeschreven cirkel van ABC. We merken op dat $BDPF$ en $AEPF$ koordenvierhoeken zijn, en dus vinden we dat $\angle EFD = \angle EFP + \angle PFD = \angle EAP + \angle PBD = \angle CBQ + \angle PBC = \angle PBQ$. Voor de oppervlakte van DEF vinden we ten slotte:

$$\begin{aligned} 2 \cdot \text{Opp}(DEF) &= |EF| \cdot |DF| \,.\, \sin EFD \\ &= |AP| \cdot |BP| \cdot \sin A \cdot \sin B \cdot \sin PBQ \\ &= |AP| \cdot |PQ| \cdot \sin A \cdot \sin B \cdot \sin PQB \\ &= |\left(|OP|^2 - R^2\right)| \cdot \sin A \cdot \sin B \cdot \sin C \end{aligned}$$

Hieruit volgt het gestelde.



Figuur 5: De voetpuntsdriehoek

1. Zij $ABCD$ een cyclische vierhoek en P, Q, R de voetpunten van de loodrechten uit D op de rechten BC, CA, AB respectievelijk. Toon aan dat $PQ = QR$ als en slechts als de bissectrices van $\angle ABC$ en $\angle ADC$ concurrent zijn met AC.

   link

2. Van een driehoek ABC is I het middelpunt van de ingeschreven cirkel. De binnenbissectrices van de hoeken A, B en C snijden de overstaande zijden respectievelijk in A_0, B_0 en C_0. Bewijs dat $\frac{1}{4} < \frac{|AI||BI||CI|}{|AA_0||BB_0||CC_0|} \leq \frac{8}{27}$ link

5.3 goniometrie

Goniometrische formules onder de knie hebben, biedt soms een voordeel om aan te tonen dat een uitdrukking klopt, wanneer men een meetkundige vraag uitwerkte met de goniometrische formules.

Het kan ook helpen bij een ongelijkheid met speciale voorwaarden, die aan een goniometrische substitutie voldoen.

Een voorbeeld hiervan, is $xyz = x + y + z$, waarna $x = tanA, y = tanB, z = tanC$ soms kan helpen.

De substitutie van Ravi, zegt dat driehoeksgetallen a, b, c geschreven kunnen worden als $x+y, x+z, y+z$ waarbij $x, y, z > 0$.
Dit komt door de raakpunten van de incirkel te beschouwen.

Er bestaan ook enkele formules voor de oppervlakte van een veelhoek in het gecoordinaliseerde vlak:

Stelling 5.2. *(stelling van Pick)*
Een roosterveelhoek (iedere hoek ligt op een roosterpunt), bevat I punten inwendig en O punten op de omtrek, da n is de oppervlakte van deze veelhoek gelijk aan $I + \frac{O}{2} - 1$.

Stelling 5.3. *(shoelace formule)*
$P_1P_2 \cdots P_n$ is een veelhoek waarbij de coordinaten van punt $P_i = (x_i, y_i)$ zijn.
Dan geldt dat de oppervlakte $= \frac{1}{2}|\sum x_i y_{i+1} - \sum x_{i+1} y_i|$.

Geometry : metric properties

1. General trigonometry

   - $\sin(X \pm Y) = \sin X \cos Y \pm \sin Y \cos X$
   - $\cos(X \pm Y) = \cos X \cos Y \mp \sin X \sin Y$
   - $\tan(X \pm Y) = \frac{\tan X \pm \tan Y}{1 \mp \tan X \tan Y}$
   - $\sin 2X = 2 \sin X \cos X$
   - $\cos 2X = \cos^2 X - \sin^2 X = 2\cos^2 X - 1 = 1 - 2\sin^2 X$
   - $\sin X = \frac{2t}{1+t^2}$, $\cos X = \frac{1-t^2}{1+t^2}$, $\tan X = \frac{2t}{1-t^2}$ $\qquad (t = \tan \frac{X}{2})$
   - $\sin P \pm \sin Q = 2 \sin \frac{P \pm Q}{2} \cos \frac{P \mp Q}{2}$
   - $\cos P + \cos Q = 2 \cos \frac{P+Q}{2} \cos \frac{P-Q}{2}$, $\cos P - \cos Q = -2 \sin \frac{P+Q}{2} \sin \frac{P-Q}{2}$

2. If $A + B + C = 180°$, then

   - $\sin A + \sin B + \sin C = 4 \cos \frac{A}{2} \cos \frac{B}{2} \cos \frac{C}{2}$
   - $\cos A + \cos B + \cos C = 1 + 4 \sin \frac{A}{2} \sin \frac{B}{2} \sin \frac{C}{2}$
   - $\tan(kA) + \tan(kB) + \tan(kC) = \tan(kA) \tan(kB) \tan(kC)$ $\qquad (k \in \mathbb{Z})$
   - $\sin 2A + \sin 2B + \sin 2C = 4 \sin A \sin B \sin C$
   - $\cos 2A + \cos 2B + \cos 2C = -(1 + 4 \cos A \cos B \cos C)$
   - $\cot \frac{A}{2} + \cot \frac{B}{2} + \cot \frac{C}{2} = \cot \frac{A}{2} \cot \frac{B}{2} \cot \frac{C}{2}$
   - $\sin^2 A + \sin^2 B + \sin^2 C = 2.(1 + \cos A \cos B \cos C)$
   - $\cos^2 A + \cos^2 B + \cos^2 C = 1 - 2 \cos A \cos B \cos C$

3. Area of a triangle

$$
\begin{aligned}
S &= \frac{a.h_A}{2} = \frac{b.h_B}{2} = \frac{c.h_C}{2} \\
&= \frac{ab}{2} \sin C = \frac{bc}{2} \sin A = \frac{ca}{2} \sin B \\
&= pr \\
&= (p-a)r_A = (p-b)r_B = (p-c)r_C \\
&= \sqrt{p(p-a)(p-b)(p-c)} \\
&= \sqrt{r.r_A.r_B.r_C} \\
&= \frac{abc}{4R} \\
&= 2R^2 \sin A \sin B \sin C \\
&= 4rR \cos \frac{A}{2} \cos \frac{B}{2} \cos \frac{C}{2} \\
&= p^2 \tan \frac{A}{2} \tan \frac{B}{2} \tan \frac{C}{2}
\end{aligned}
$$

1. Vind alle $k \in \mathbb{N}$ zodat als geldt dat $k(ab+bc+ca) > 5(a^2+b^2+c^2)$ voor $a, c, b > 0$, dat dan geldt dat a, b, c de zijden van een driehoek kunnen zijn.

   link

2. Vind alle drietallen $x, y, z \in \mathbb{R}$ zodat geldt dat

$$5\left(x+\frac{1}{x}\right) = 12\left(y+\frac{1}{y}\right) = 13\left(z+\frac{1}{z}\right), xy+yz+zx = 1$$

   link

3. De raaklijnen in B en A aan de omgeschreven cirkel van een scherphoekige driehoek ABC snijden de raaklijn in C in T en U respectievelijk.
   AT snijdt BC in P en Q is het midden van AP.
   BU snijdt CA in R en S is het midden van BR. Bewijs dat $\angle ABQ = \angle BAS$.
   Bepaal, in termen van verhoudingen van de zijdelengtes, de driehoek waarvoor deze hoek maximaal is. link

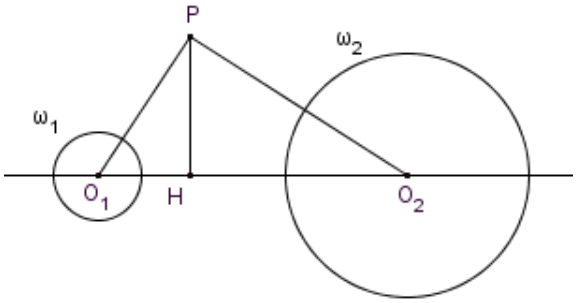
5.4 Macht van een punt

Wanneer we een punt P en een cirkel ω met middelpunt O en straal R beschouwen, en we tekenen een willekeurige rechte door P die ω snijdt in A en B, dan merken we op dat de grootte van $|PA|\,.\,|PB|$ onafhankelijk is van de gekozen rechte. (Bewijs dit) We definiëren P_ω^P :de *macht van* P *t.o.v.* ω als

- $-\,|PA|\,.\,|PB|$ als P binnen de cirkel ligt.
- 0 als P op de cirkel ligt.
- $|PA|\,.\,|PB|$ als P buiten de cirkel ligt.

Toon nu aan dat de macht van P t.o.v. ω ook gegeven wordt door $|OP|^2 - R^2$. Nu wordt de betekenis van het minteken in de definitie van de macht van P t.o.v. ω duidelijk, ze was nodig om de zonet gegeven uitdrukking steeds te doen kloppen.

Machtlijn



Figuur 6: De machtlijn

Wanneer er twee cirkels in het spel zijn zouden we ons kunnen afvragen wat de *meetkundige plaats* is van alle punten P die t.o.v. beide cirkels dezelfde macht hebben. Beschouw daartoe twee cirkels ω_1, ω_2 met straal r_1, r_2 en middelpunt O_1, O_2. Zij P een punt dat gelijke macht ten opzichte van beide cirkels heeft, en noem H de projectie van P op O_1O_2. Volgende gelijkheden zijn nu equivalent:

$$\begin{aligned} |O_1P|^2 - r_1^2 &= |O_2P|^2 - r_2^2 \\ |O_1H|^2 + |HP|^2 - r_1^2 &= |O_2H|^2 + |HP|^2 - r_2^2 \\ |O_1H|^2 - r_1^2 &= (|O_2O_1| - |HO_1|)^2 - r_2^2 \\ 2\cdot|HO_1|\cdot|O_2O_1| &= |O_2O_1|^2 + r_1^2 - r_2^2 \end{aligned}$$

Merk op dat de onderste vergelijking enkel afhankelijk is van de positie van H. Het punt P zal met andere woorden dan en slechts dan een gelijke macht hebben t.o.v. beide cirkels als H dat ook heeft. Wanneer $|O_1O_2| \neq 0$ is de laatste vergelijking een eerstegraadsvergelijking die een unieke H oplevert (Merk op dat we met geöriënteerde lengtes werken). Bijgevolg is de meetkundige plaats die we zochten een rechte loodrecht op O_1O_2. Deze rechte wordt ook wel de *machtlijn* van beide cirkels genoemd.

Merk op dat we eenvoudig de machtlijn van twee snijdende cirkels kunnen terugvinden als de rechte door beide snijpunten (of de gemeenschappelijk raaklijn indien de cirkels raken in een punt). Ga na waarom dat zo is.

Machtpunt

Wanneer we drie cirkels beschouwen, dan kunnen we voor elk paar cirkels de machtlijn gaan beschouwen. Bewijs nu zelf de volgende stelling:
Gegeven zijn drie cirkels ω_1, ω_2 en ω_3. De drie machtlijnen die we krijgen door telkens twee verschillende cirkels uit de gegeven drie cirkels te beschouwen zijn concurrent.
Het punt van concurrentie van deze 3 machtlijnen wordt vaak het *machtpunt* van de drie cirkels genoemd.

Voorbeeld 5.4. *(stelling van Cayce)*
Zij O_1, O_2 de centra van cirkelsω_1, ω_2 met machtlijn l, dan geldt dat $|P^P_{\omega_1} - P^P_{\omega_2}| = 2|O_1O_2| \cdot d(P, l)$.

Bewijs. We zeggen dat P aan dezelfde kant als ω_2 van l ligt uit symmetrieredenen.
Merk ten eerste op dat we terug de projectie van P op O_1O_2 kunnen nemen:
P' en dat $|P^{P'}_{\omega_1} - P^{P'}_{\omega_2}| = |P^P_{\omega_1} - P^P_{\omega_2}|$.
Net zoals op de figuur van vorige bladzijde is $H = l \cup O_1O_2$.
$P^{P'}_{\omega_1} = |O_1P'|^2 - r_1^2 = (|O_1H| + |HP'|)^2 - r_1^2 = |HP'|^2 + 2|O_1H||HP'| + P^H_{\omega_1}$
$P^{P'}_{\omega_2} = |O_2P'|^2 - r_1^2 = (|O_2H| - |HP'|)^2 - r_2^2 = |HP'|^2 - 2|O_2H||HP'| + P^H_{\omega_2}$
De leden van elkaar aftrekken en opmerken dat $P^H_{\omega_1} = P^H_{\omega_2}$ geeft nu het gevraagde.

□

1. Zijn A, B, C en D vier verschillende punten op een rechte, in die volgorde. De cirkels met diameters AC en BD snijden in X en Y. O is een willekeurig punt op XY maar niet op AD. CO snijdt de cirkel met diameter AC opnieuw in N. Bewijs dat AM, DN en XY concurrent zijn. link

2. In het vlak zijn twee cirkels gegeven die snijden in de punten X en Y. Bewijs dat er vier punten bestaan met de volgende eigenschap: voor iedere cirkel die de twee gegeven cirkels raakt in A en B, en die de rechte XY in C en D snijdt, passeert ieder van de rechten AC, AD, BC, BD door n van deze punten.

   link

5.5 meetkundige relaties

Onder meetkundige relaties verstaan we zowel enkele relaties binnen de meetkunde ivm lengten, alsook meetkundige plaatsen.

Stelling 5.5. *(De cirkel van Apollonius)*
Zij $[AB]$ *een lijnstuk en* k *een positief reel getal ongelijk aan* 1.
De meetkundige plaats van alle punten P *waarvoor geldt* $\frac{|PA|}{|PB|} = k$ *is een cirkel met middelpunt op de rechte* AB.

opmerking: indien $k = 1$ is de meetkundige plaats de middelloodlijn van het lijnstuk, mbv de bissectricestelling kennen we 3 punten en is de cirkel dus te construeren
extra: wanneer men de 3 Apoloniuscirkels tekent in een driehoek,
zijn hun middelpunten collineair en snijden ze elkaar in 2 gemeenschappelijke punten.
Andere meetkundige plaatsen, kunnen met oppervlakten of met hoeken werken,
bvb. een lijnstuk $[AB]$ is gegeven , alle punten C zodat $[\triangle ABC] = C^{te}$ liggen op 2 evenwijdige rechten.
alle punten D zodat $\angle ADB = C^{te}$ liggen op 2 cirkelbogen.

Stelling 5.6. *(Stewart's relatie)*
In een driehoek ABC *met een punt* $P \in [BC]$ *geldt dat* $|AC|^2|BP| + |AB|^2|PC| = |BC|(|AP|^2 + |CP||PB|)$

Stelling 5.7. *(Brahmagupta)*

De oppervlakte in een vierhoek $ABCD : S = \sqrt{(s-a)(s-b)(s-c)(s-d) - abcd\cos^2\frac{\angle ABC + \angle BCD}{2}}$
(vooral gekend als die cosinus 0 *is bij een koordenvierhoek, als generalisatie van de formule van Heroon)*

Stelling 5.8. *(stellingen van Carnot)*

- *Zij* O *het omcentrum van* $\triangle ABC$ *dan geldt dat de som van de afstanden van* O *tot de middens van de zijden van de driehoek gelijk is aan* $R + r$ *waarbij een afstand afgetrokken werd als ze buiten de driehoek ligt.*
- X, Y, Z *zijn punten op de rechten* AB, AC, BC *van driehoek* $\triangle ABC$.
  De loodlijnen uit deze 3 *punten zijn concurrent a.e.s.a.*
  $|BX|^2 + |AY|^2 + |ZC|^2 = |BZ|^2 + |CY|^2 + |AX|^2$
- *Een triviale stelling van Carnot:* A, B, X, Y *zijn* 4 *punten, dan geldt dat* $AB \perp XY \Leftrightarrow AX^2 - BX^2 = AY^2 - BY^2$.

Voorbeeld 5.9. *(stelling van Newton) Het middelpunt* I *van de ingeschreven cirkel van een raaklijnenvierhoek* $[ABCD]$ *is collineair met de middens* M, N *van de diagonalen*$[AC], [BD]$.

Bewijs. Beschouw de meetkundige plaats van alle punten P waarvoor geldt dat $[ABP]+[CDP] = [ADP] + [BCP]$.
Ten eerste merkten we natuurlijk al op dat I, M, N hieraan voldoen :
$2[ADN] = [ADB], 2[BNC] = [BDC]$ en analoog voor M.
$[AID] + [BIC] = \frac{r}{2}(|AD| + |BC|) = 0.5[ABCD]$.
Tot slot is het voldoende te bewijzen dat onze meetkundige plaats een rechte is,
neem hiervoor $E = AD \cup BC$ en laat $F, G \in AD, BC$ zodat $|AD| = |EF|, |BC| = |EG|$. De som van de oppervlakten is nu gelijk aan $[EFG] + [FGP]$ en dus moet P op een evenwijdige reche met EF liggen.

□

4. Other relations

- $\frac{a}{\sin A} = \frac{b}{\sin B} = \frac{c}{\sin C} = 2R$
- $a^2 = b^2 + c^2 - 2bc\cos A,\ b^2 = c^2 + a^2 - 2ca\cos B,\ c^2 = a^2 + b^2 - 2ab\cos C$
- $\frac{r}{R} = 4\sin\frac{A}{2}\sin\frac{B}{2}\sin\frac{C}{2} = \cos A + \cos B + \cos C - 1$
- $\frac{r_A}{R} = 4\sin\frac{A}{2}\cos\frac{B}{2}\cos\frac{C}{2},\ \frac{r_B}{R} = 4\cos\frac{A}{2}\sin\frac{B}{2}\cos\frac{C}{2},\ \frac{r_C}{R} = 4\cos\frac{A}{2}\cos\frac{B}{2}\sin\frac{C}{2}$
- $\frac{r}{p-a} = \frac{r_A}{p} = \tan\frac{A}{2},\ \frac{r}{p-b} = \frac{r_B}{p} = \tan\frac{B}{2},\ \frac{r}{p-c} = \frac{r_C}{p} = \tan\frac{C}{2}$
- $\frac{r}{p} = \tan\frac{A}{2}\tan\frac{B}{2}\tan\frac{C}{2}$
- $r = a\frac{\sin\frac{B}{2}\sin\frac{C}{2}}{\cos\frac{A}{2}} = b\frac{\sin\frac{C}{2}\sin\frac{A}{2}}{\cos\frac{B}{2}} = c\frac{\sin\frac{A}{2}\sin\frac{B}{2}}{\cos\frac{C}{2}}$
- $r_A + r_B + r_C - r = 4R$
- $\frac{1}{r_A} + \frac{1}{r_B} + \frac{1}{r_C} - \frac{1}{r} = 0$

5. Distances between vertices and tangency points

	Incircle		
	T_A	T_B	T_C
A	•	$p-a$	$p-a$
B	$p-b$	•	$p-b$
C	$p-c$	$p-c$	•

	Excircles								
	T_{AA}	T_{AB}	T_{AC}	T_{BA}	T_{BB}	T_{BC}	T_{CA}	T_{CB}	T_{CC}
A	•	p	p	•	$p-c$	$p-c$	•	$p-b$	$p-b$
B	$p-c$	•	$p-c$	p	•	p	$p-a$	•	$p-a$
C	$p-b$	$p-b$	•	$p-a$	$p-a$	•	p	p	•

6. Other distances and powers with respect to the circumcircle

X	$\lvert OX\rvert^2$	$\pi_\Gamma(X) = \lvert OX\rvert^2 - R^2$
H	$9R^2 - (a^2+b^2+c^2)$	$-8R^2\cos A\cos B\cos C$
G	$R^2 - \frac{1}{9}(a^2+b^2+c^2)$	$-\frac{1}{9}(a^2+b^2+c^2)$
E	$\frac{9}{4}R^2 - \frac{1}{4}(a^2+b^2+c^2)$	$\frac{5}{4}R^2 - \frac{1}{4}(a^2+b^2+c^2)$
I	$R^2 - 2Rr$	$-2Rr$
I_A	$R^2 + 2Rr_A$	$2Rr_A$
I_B	$R^2 + 2Rr_B$	$2Rr_B$
I_C	$R^2 + 2Rr_C$	$2Rr_C$

7. Feuerbach's Theorem : Γ_E is internally tangent to γ and externally tangent to γ_A, γ_B and γ_C. Therefore, we have

$$|EI| = \frac{R}{2} - r,\ |EI_A| = \frac{R}{2} + r_A,\ |EI_B| = \frac{R}{2} + r_B,\ |EI_C| = \frac{R}{2} + r_C.$$

1. $ABCD$ is een vierkant in het vlak. Vind alle punten $P \neq A, B, C, D$ die voldoen aan $\widehat{APB} + \widehat{CPD} = 180°$.

   link

2. Vind de meetkundige plaats van alle punten P in $\triangle ABC$ zodat de driehoek met zijdelengten $|PA|, |PB|, |PC|$ een oppervlakte geeft van $\frac{1}{3}[ABC]$.

   link

3. In een driehoek $\triangle ABC$ geldt dat AD, BE, CF de drie hoogtelijnen zijn. Bewijs dat de omtrek van $\triangle DEF \leq$ is dan de helft van de omtrek van $\triangle ABC$.

   link

4. A en B zijn twee opeenvolgende hoekpunten van een regelmatige veelhoek met n zijden, $n > 5$, en middelpunt O.
   Een driehoek XYZ, die congru- ent is met de driehoek OAB, wordt om te beginnen zo geplaatst dat de hoekpunten X, Y en Z samenvallen met respectievelijk O, A en B.
   De driehoek XYZ verplaatst zich vervolgens in het vlak van de veelhoek zo, dat de punten Y en Z op de zijden van de veelhoek blijven liggen en X binnen de veelhoek blijft.
   Welke figuur wordt door het punt X beschreven wanneer Y de hele omtrek van de veelhoek doorloopt? link

5.6 de echte stellingen

2 stellingen die zeer frequent toepasbaar zijn, zijn de volgende:

Stelling 5.10. *(Menelaos)*
Zij D, E, F punten die resp. op AB, AC, BC liggen, dan geldt dat D, E, F collineair zijn a.e.s.a. $\frac{\vec{AD}\cdot\vec{BF}\cdot\vec{CE}}{\vec{DB}\cdot\vec{FC}\cdot\vec{EA}} = -1$

Stelling 5.11. *(Ceva)*
Zij D, E, F punten die resp. op AB, AC, BC liggen, dan geldt dat D, E, F collineair zijn a.e.s.a. $\frac{\vec{AD}\cdot\vec{BF}\cdot\vec{CE}}{\vec{DB}\cdot\vec{FC}\cdot\vec{EA}} = 1$

Beide zijn te bewijzen door te werken met oppervlakten.
gevolgen van stelling van Ceva

Stelling 5.12. *(de goniometrische vorm van Ceva)*
Zij D, E, F punten die resp. op AB, AC, BC liggen, dan geldt dat AD, BE, CF concurrent zijn a.e.s.a.
$\frac{sin(BAF)\cdot sin(ACD)\cdot sin(CBE)}{sin(DCB)\cdot sin(FAC)\cdot sin(EBA)} = 1$

Stelling 5.13. *(stelling van Jacobi)*
Gegegen een driehoek $\triangle ABC$ en 3 punten X, Y, Z die alle 3 inwendig zijn of alle 3 uitwendig liggen). Als $\angle ZAB = \angle YAC, \angle ZBA = \angle XBC$ en $\angle XCB = \angle YCA$ dan zijn de lijnen AX, BY, CZ concurrent.

1. Zij A_1 het midden van het vierkant ingeschreven in de scherphoekige driehoek ABC met twee hoekpunten van het vierkant op de zijde BC en dus de twee andere op AB en AC. De punten B_1 en C_1 worden analoog gedefinieerd voor de ingeschreven vierkanten met twee hoekpunten op AC en AB respectievelijk. Bewijs dat de rechten AA_1, BB_1, CC_1 concurrent zijn.

   link

2. Zij O het midden van de omgeschreven cirkel en H het hoogtepunt van een scherphoekige driehoek ABC. Toon aan dat er punten D, E, F bestaan op de zijden BC, CA, AB respectievelijk, zodat $OD + DH = OE + EH = OF + FH$, en de rechten AD, BE en CF concurrent zijn.

   link

3. Zij ABC een driehoek en P een uitwendig punt in het vlak van de driehoek. Veronderstel dat de rechten AP, BP, CP de rechten BC, CA, AB snijden in D, E, F respectievelijk. Veronderstel verder dat de oppervlakte van de driehoeken PBD, PCE, PAF allemaal gelijk zijn. Bewijs dat ieder van deze oppervlaktes gelijk is aan de oppervlakte van de driehoek ABC zelf.link

5.7 stellingen die gevolg zijn van de stelling van Menelaos

Stelling 5.14. *(Gauss) Een rechte snijdt de zijden van een driehoek in punten A', B', C' , dan geldt dat de middens van $[AA'], [BB'], [CC']$ collineair zijn.*

Stelling 5.15. *(stelling van Monge)*
De stelling van Monge-d'Alembert zegt dat als we 3 cirkels hebben, de gemeenschappelijke uitwendige raaklijnen snijden in 3 punten die collineair zijn.
Hierbij bedoelen we dat de uitwendige raaklijnen van Γ_i en γ_{i+1} snijden in P_i en dat P_1, P_2, P_3 op 1 rechte liggen.
alternatieve vorm:
Zij O_1, O_2, O_3 de 3 centra van resp. $\Gamma_1, \Gamma_2, \Gamma_3$, zij P_1 het snijpunt van de uitwendige raaklijnen van Γ_1, Γ_2 en P_2, P_3 de snijpunten van de inwendige raaklijnen van resp. Γ_3, Γ_2 en Γ_1, Γ_3. Dan geldt dat P_1, P_2, P_3 opnieuw collineair zijn.

Stelling 5.16. *(transversaalstelling)*
A, B, C zijn 3 punten op een lijn en P is een punt die niet op die lijn ligt: A', B', C' zijn 3 punten die liggen op AP, BP, CP resp., dan geldt er dat de punten A', B', C' collineair zijn als en slechts als $\frac{\vec{BC}\vec{AP}}{\vec{A'P}} + \frac{\vec{CA}\vec{BP}}{\vec{B'P}} \cdot \frac{\vec{AB}\cdot\vec{CP}}{\vec{C'P}} = 0$

Stelling 5.17. *(Menelaos voor vierhoeken)*
Als X, Y, Z, W punten zijn op AB, BC, CD, AD van een vierhoek $ABCD$ en deze 4 punten liggen op een rechte, dan geldt dat
$\frac{\vec{AX}}{\vec{XB}} \cdot \frac{\vec{BY}}{\vec{YC}} \cdot \frac{\vec{CZ}}{\vec{ZD}} \cdot \frac{\vec{DW}}{\vec{WA}} = 1$
(deze stelling geldt niet in beide richtingen)

Stelling 5.18. *(Cevian Nests theorem)*
$\triangle ABC$ is een willekeurige driehoek met A', B', C' 3 punten op BC, AC, AB resp. en A", B", C" 3 punten op de rechten $B'C', A'C', A'B'$ resp., dan gelden de volgende 3 uitspraken als er 2 waar zijn:

- *AA', BB', CC' zijn concurrent*
- *AA", BB", CC" zijn concurrent*
- *A'A", B'B", C'C" zijn concurrent*

Stelling 5.19. *(stelling van Desargues)*
Twee driehoeken, ABC en XYZ, noemen we puntperspectief als AX, BY en CZ door 1 punt gaan en we noemen ze lijnperspectief als de snijpunten van AB en XY , BC en YZ, en CA en ZX op 1 lijn liggen. De stelling van Desargues zegt dat twee driehoeken lijnperspectief zijn dan en slechts als ze puntperspectief zijn.

1. Een punt P ligt willekeurig op $[AB]$ van de convexe vierhoek $ABCD$.
   ω is de incirkel van $\triangle CPD$ met I als incenter.
   ω is rakend aan de incirkels van $\triangle APD, BPC$ in K, L resp.
   $E = AC \cap BD, F = AK \cap BL$.
   TB:E, I, F zijn collineair. link

5.8 overige synthetische stellingen

Stelling 5.20. *(De vlinderstelling) Laat M het midden zijn van een koorde PQ van een cirkel en AB en CD twee andere koorden door M. Noem X het snijpunt zijn van AD en PQ en Y van BC en PQ. Dan is M het midden van XY.*

Stelling 5.21. *(Poncelet) Als er een veelhoek tegelijk ingeschreven is in kegelsnede Γ_1 als kegelsnede Γ_2 omschrijft, bestaan er oneindig veel zo'n veelhoeken.*

Stelling 5.22. *(Taylorcirkel)*
Laat D, E, F de voetpunten zijn van A, B, C en zij D_1, D_2 de voetpunten van de loodlijnen uit D op AC, AB en analoog, dan gaat de Taylorcirkel door $D_1, D_2, E_1, E_2, F_1, F_2$.

Stelling 5.23. *(Morley's driehoek)*
De eerste snijpunten van de trisectrices vormen in iedere driehoek een gelijkzijdige driehoek.

Stelling 5.24. *(stelling van Steiner)*
In een $\triangle ABC$ geldt dat als $D, E \in [BC]$ en AD, AE isogonaal geconjugeerd zijn, geldt dat $\frac{|BD||BE|}{|DC||EC|} = \frac{|AB|^2}{|CA|^2}$
Een ander resultaat van Steiner is dat de ingeschreven n-hoek met de grootste oppervlakte (in een vaste cirkel), de regelmatige $n-$hoek is.

Stelling 5.25. *(stelling van Pappos)*
Deze stelling luidt: Liggen A_1, B_1 en C_1 op een rechte d_1 en liggen A_2, B_2 en C_2 op een rechte d_2 , dan zijn de punten A : snijpunt van B_1C_2 en B_2C_1, B : snijpunt van A_1C_2 en A_2C_1 en C : snijpunt van A_1B_2 en A_2B_1 collineair.

Stelling 5.26. *(stelling van Pascal)*
Neem zes willekeurige punten op een cirkel of andere kegelsnede, zeg $A, B, C, D, EenF$. Het snijpunt van delijnen AB en DE noemen we P, het snijpunt van BC en EF noemen we Q en het snijpunt van CD en FA noemen we R. Dan liggen P, Q en R op 1 lijn.

Stelling 5.27. *(gegeneraliseerde stelling van Pascal door Mobius) stel dat een veelhoek met $4n+2$ zijden ingeschreven wordt in een kegelsnede, en paren van tegenoverstaande zijden worden verlengd totdat zij elkaar ontmoeten in $2n+1$ punten, dan zal, als $2n$ van die punten op 1 lijn liggen, het laatste punt ook op die lijn liggen.*

Stelling 5.28. *(stelling van Brianchon)*
Neem een zeshoek $ABCDEF$ van zes raaklijnen aan een kegelsnede. Dan zijn de lijnenAD, BE en CF concurrent.

Terug wat oefeningen, de enige goede leerschool.

1. Gegeven een scherphoekige driehoek ABC met $|AC| > |BC|$ en F als voetpunt van C op $[AB]$. Laat P een punt zijn op AB, $\neq A$ zodat $|AF| = |PF|$. Zij H, O, M het hoogtepunt,omcentrum en midden van $[AC]$. Zij X het snijpunt van BC en HP en Y 't snijpunt van OM en FX, laat OF snijden met AC in Z. Bewijs dat F, M, Y, Z een koordenvierhoek vormen. link

5.9 meetkundige ongelijkheden

Bepaalde meetkundige ongelijkheden, kunnen opgelost worden via gewone meetkunde (constructies en lemma's opstellen).
Er zijn enkele ongelijkheden die wel handig zijn te kennen:

Stelling 5.29. *(meetkundige ongelijkheden)*
voor 4 punten in de ruimte:
parallellogramongelijkheid: $|AB|^2 + |BC|^2 + |CD|^2 + |AD|^2 \geq |AC|^2 + |BD|^2$ *met enkel gelijkheid als* $ABCD$ *een parallellogram is.*
Ptolemaeus: $|AB||CD| + |BC||AD| \geq |AC||BD|$ *met gelijkheid als* $ABCD$ *een koordenvierhoek is.*
Binnen een driehoek:
Euler: $R \geq 2r$ *en Leibniz:* $9R^2 \geq a^2 + b^2 + c^2$ *met gelijkheid a.e.s.a.* $\triangle ABC$ *gelijkzijdig is.*
Erdos-Mordell: Zij $P \in \triangle ABC$ *met* DEF *de voetpuntsdriehoek van* P *in die driehoek, dan geldt dat* $|PA| + |PB| + |PC| \geq 2(|PD| + |PE| + |PF|)$

Stelling 5.30. *(stelling van Casey)*
Zij $\Gamma_1, \Gamma_2, \Gamma_3, \Gamma_4$ 4 *cirkels en* L_{ij} *zijn de lengtes van de in- of uitwendige raaklijnen tussen* Γ_i *en* Γ_j*. Dan geldt dat*

$$L_{12}L_{34} \pm L_{13}L_{24} \pm L_{14}L_{23} = 0$$

uitgebreide bijlage over de stelling van Casey

1. Een soldaat moet zich ervan overtuigen dat er binnen een gebied, dat de vorm heeft van een gelijkzijdige driehoek, geen mijnen liggen.
   De actieradius van zijn detector is gelijk aan de helft van de hoogte van de driehoek.
   Hij begint in een hoekpunt.
   Welke weg moet hij kiezen opdat de afstand die hij aflegt tot het moment dat het gebied geheel is gecontroleerd, zo klein mogelijk is? link

2. O is een punt op een rechte g. $\vec{OP_1} \cdots \vec{OP_n}$ zijn eenheidsvectoren, gelegen in een vlak door g, allemaal aan dezelfde kant van g. Bewijs: als n oneven is, dan geldt $|\vec{OP_1} + \cdot + \vec{OP_n}| > 1$
   link

3. Zij $A = A_1A_2 \ldots A_n$ een convexe n-hoek, $n \geq 4$. Bewijs dat A cyclisch is als en slechts als men aan ieder hoekpunt A_i een koppel (b_i, c_i) van rele getallen kan toekennen zodat $A_iA_j = b_jc_i - b_ic_j$, $1 \leq i < j < n$.
   link

4. $\triangle ABC$ is een scherphoekige driehoek met omgeschreven cirkel ω waar t een raaklijn aan is. t_a, t_b, t_c zijn de lijnen bekomen door t te spiegelen in BC, AC, AB resp.
   TB: Bewijs dat de omgeschreven cirkel van de driehoek gevormd door de snijpunten van t_a, t_b, t_c raakt aan ω. link

Stelling 5.31. *(inversie)*
Bij inversie wordt een punt O als centrum gekozen en ieder punt X wordt getransformeerd naar een punt Y zodat O, X, Y op de zelfde halfrechte liggen en $|OX||OY| = c$ waarbij c een reel getal is.
Indien f de inverterende functie is binnen deze meetkunde, geldt $f(X) = Y, f(Y) = X$ in dit voorbeeld, wat algemeen logisch $f(f(X)) = X$ heeft voor ieder voorwerp.
We zullen vanaf nu voor ieder punt A $f(A) = A'$ noteren om de eigenschappen op te sommen:

1. *een lijn door O wordt op zichzelf afgebeeld*
2. *een cirkel door O wordt geprojecteerd op een lijn die 0 niet bevat*
3. *een cirkel die niet door O gaat, wordt geprojecteerd op een andere cirkel die niet door O gaat.*
4. *hoeken worden behouden, maar er geldt wel dat $\angle OAB = \angle OB'A'$*
5. *lengtes van lijnstukken veranderen in volgende verhouding: $|A'B'| = \frac{|c||AB|}{|OA||OB|}$*

Met deze eigenschappen kunnen problemen vanuit een heel andere hoek worden opgelost en op een zeer ingenieuze manier opgelost worden.
voor meet info
Er is nog een andere transformatie om bepaalde gevallen simpeler te maken (opgelet met zo'n transformaties te combineren!!!)

Stelling 5.32. *(affiene meetkunde)*
Een affiene transformatie bestaat uit een afbeelding $(x, y) \to (ax + by + c, dx + ey + f)$.
Binnen de affiene meetkunde kunnen we met zo'n afbeelding 3 niet-collineaire punten vervangen door 3 andere niet-collineaire punten op een manier zoals je ze zelf kiest.
De affiene transformaties behouden

- *evenwijdigheid van lijnen*
- *collineairiteit van punten*
- *concurrentie van lijnen*
- *verhouding oppervlakten*

De hoeken als ook verhouding van lijnen zijn niet strikt noodzakelijk behouden.
Deze transformatie kan dus enkel helpen wanneer 1 van de andere 4 punten te bewijzen valt of te gebruiken is.

1. In een driehoek ABC is I het middelpunt van de ingeschreven cirkel. De bissectrices AI, BI en CI snijden respectievelijk de zijden BC, CA en AB in de punten D, E en$F$. De middelloodlijn van het lijnstuk AD snijdt de lijnen BI en CI respectievelijk in M en N. Bewijs datA, I, M en N op 1 cirkel liggen.

   link

6 deelverhoudingen en polen

Een deelverhouding wordt als volgt gedefinieerd; P is een punt op AB, dan is de unieke deelverhouding $(ABP) = \frac{\vec{PA}}{\vec{PB}}$. Een dubbelverhouding is het product van 2 deelverhoudingen:

A, B, C, D zijn 4 punten op een rechte, dan is de dubbelverhouding $(ABCD) = \frac{(ABC)}{(ABD)} = \frac{\vec{CA} * \vec{DB}}{\vec{CB}\vec{DA}}$. Als $(ABCD) = k$, dan geldt dat $(ABDC) = \frac{1}{k}, (ACBD) = 1 - k$. Een vierstraal = 4 rechten door

1 punt. Wordt een vierstraal gesneden door een rechte, is de dubbelverhouding constant. In een

cirkel geldt $\forall P$ op de cirkel en A, B, C, D vaste punten op die cirkel: (PA, PB, PC, PD) constant is voor alle P. (Hiermee bedoelen we dat een rechte die de vierstraal snijdt in 4 punten, deze verdeelt in een constante dubbelverhouding)

Gegeven een cirkel c en punt P ,
de middellijn van c door P snijdt de cirkel in A en B.
Kiest men het punt P' op deze lijn zodat $(ABPP') = -1$, dan is de loodlijn p op de middellijn door P' de poollijn v.h. punt P tov cirkel c.
P is de pool v.d. rechte p tov c. De volgende eigenschappen gelden:

- Als Q op de poollijn p van P ligt, ligt P op q.
  Dit is een gevolg van de stelling van Salmon: $\frac{d(P,q)}{d(Q,p)} = \frac{|OP|}{|OQ|}$
- Een willekeurige lijn door P en c snijdt die cirkel in A, B, de poollijn wordt gesneden in X, dan geldt dat $(ABXP) = -1$.
- A, B, C, D liggen op een rechte in die volgorde , X ligt niet op die rechte.
  Als 2 van de volgende 3 eigenschappen gelden, klopt de derde ook.
  $*(ACBD) = -1$
  * Er geldt dat $\angle BXD = 90°$
  $*BX$ de bissectrice is van $\angle AXC$.
  Bvb.volgt hieruit dat $(AXII_a) = -1$ met X het snijpunt van BC met de bissectrice van $\angle BAC$.
  gevolg:De loodlijnen uit I, I_a zijn resp. X, Y en Z is het spiegelbeeld van X tov I. Dan geldt dat A, Y, Z op een rechte liggen.
- ABCD is een koordenvierhoek, E, K, J zijn de snijpunten van AC, BD,AB, CD en AD, BC, dan geldt dat E de pool is van de poollijn JK en analoog J van poollijn EK en is de poollijn van $K = EJ$. gevolg: O is de omcentrum van $ABCD$ en is het hoogtepunt van $\triangle EJK$ andere stelling hiermee: laat KE AB en CD snijden in P, Q dan geldt dat $(KEPQ) = -1$
- De pool van een lijn door2 polen ligt op het snijpunt van de 2 poollijnen (duaal/ geldt in de 2 richtingen)
- De poollijnen van 3 collineaire punten zijn concurrent (opnieuw een duale stelling)
- Als AX, BY, CZ drie cevianen zijn in een driehoek $\triangle ABC$, waarbij X, Y, Z op BC, AC, AB resp. liggen en $T = YZ \cup BC$, dan is $(BXCT) = -1$ aesa AX, BY, CZ concurrent zijn.
- DEF is de orthic triangle van $\triangle ABC$ (AD tot CF zijn hoogtelijnen) en $D' = EF \cup BC$, $E' = AC \cup FD$ en $F' = AB \cup DE$ dan geldt dat D, E, F op een rechte liggen.

Stelling 6.1. *(harmonische vierhoeken)*
Binnen een koordenvierhoek $PQRS$ zijn volgende eigenschappen equivalent:

1. *$PQRS$ is harmonisch*
2. *$|PQ||RS| = |PS||RQ|$*
3. *PR is de P-symmedian of $\triangle QPS$*
4. *QS is de Q-symmedian van $\triangle PQR$*
5. *de raaklijnen in P, R aan de omgeschreven cirkel snijden op QS*
6. *de raaklijnen in Q, S aan de omgeschreven cirkel snijden op PR*
7. *TA, TB, TC, TD of $T(abcd)$ is een harmonische vierstraal waarbij T een ander punt is op de omgeschreven cirkel*

Een uitgebreide bijlage over projectieve meetkunde

1. Let $\triangle ABC$ een willekeurige driehoek zijn. De raaklijnen aan de negenpuntscirkel in het voetpunt van de hoogtelijn uit A op $[BC]$ en het middenpunt van de zijde BC snijden in A'. Analoog worden B', C' geconstrueerd. Bewijs dat AA', BB', CC' concurrent zijn. link

2. De incirkel van $\triangle ABC$ raakt de zijdenBC, CA, AB inD, E, Frespectivelijk. X is een punt in$\triangle ABC$ zodat de incirkel van $\triangle XBC[BC]$ in$D$ raakt enCX, XB inY, Z resp. Bewijs dat $ZEYF$ cyclisch is. link

3. $\triangle ABC$ geeft $\angle A = 90°$ en$D \in [AC]$. E is de reflectie van$A$ inBD en F is het snijpunt van CE met de loodlijn uit D op BC. Bewijs datAF, DE, BC concurrent zijn.

   link

4. E, F zijn de snijpunten van de overstaande zijden van vierkant $ABCD$.
   De 2 diagonalen snijden in P en O is het voetpunt van de loodlijn uit P op EF.
   Bewijs dat $\angle BOC = \angle AOD$. link

Tot slot zijn er nog heel wat diverse meetkundevragen te vinden via deze link

We hebben nog enkele noodoplossingsbijlangen voor zij zonder meetkunding inzicht: calculusmeetkunde ter volledigheid.

7 varia : mooie vragen door elkaar

Tot slot nog enkele vragen van alle onderwerpen om zeker te zijn dat alles toegepast kan worden op een echte wedstrijd.
(top-Belgisch)

1.

2.

3.

4.

5.

6.

Verder raden we aan de gewenste olympiade (en gelijkaardige) via het archief op te zoeken, voor de beste voorbereiding.
Veel succes op die olympiades om het land goed voorbereid te vertegenwoordigen!

8 bijlages en bronnen

Hier staan PDF's die uitgebreid handelen over een specifiek deel.
Er staan nog enkele nieuwe dingen tussen ter volledigheid zoals complexe meetkunde, die soms een oplossing brute-force/minder elegant oplossen, wat we afraden.
Naar de interessante stukken wordt op het juiste moment verwezen doorheen de echte bundel.

Getallenleer
IMO-stage Beersel 2011

Arne Smeets - arne.smeets@wis.kuleuven.be

4 april 2011

1 Deelbaarheid en priemgetallen

We beginnen met een herhaling van de belangrijkste definities en eigenschappen. We laten de bewijzen daarbij achterwege, maar het is uiteraard een goed idee om te proberen om die bewijzen zelf te geven...

Definities

(1) Als $m, n \in \mathbb{Z}$, dan zeggen we dat m een deler is van n (notatie: $m \mid n$) als er een $a \in \mathbb{Z}$ bestaat zodat $n = am$.

(2) Een natuurlijk getal p is priem als p precies twee *positieve* delers heeft.

(3) Gegeven $a \in \mathbb{Z}$ en $b \in \mathbb{N}_0$, dan bestaan er twee (unieke) gehele getallen q en r zodanig dat $a = qb + r$ met $0 \leq r < b$. We noemen q en r respectievelijk het quotiënt en de rest bij deling van a door b.

(4) Zijn $m, n \in \mathbb{Z}$. De grootste gemene deler van m en n is het (unieke) natuurlijk getal d dat voldoet aan $d \mid m$, $d \mid n$ en de volgende voorwaarde: als e een geheel getal is met $e \mid m$ en $e \mid n$, dan is $e \mid d$. Het kleinste gemene veelvoud van m en n is het (unieke) natuurlijk getal a dat voldoet aan $m \mid a$, $n \mid a$ en de volgende voorwaarde: als b een geheel getal is met $m \mid b$ en $n \mid b$, dan is $a \mid b$. We noteren $d = \text{ggd}(m, n)$ en $a = \text{kgv}(m, n)$. Op analoge wijze kunnen we de grootste gemene deler en het kleinste gemene veelvoud van meer dan twee gehele getallen definiëren.

(5) Twee gehele getallen zijn onderling ondeelbaar als hun grootste gemene deler gelijk is aan 1.

Eigenschappen

(1) Als m en n gehele getallen zijn met $m \mid n$, dan is $|m| \leq |n|$.

(2) Elk natuurlijk getal kan op unieke wijze worden geschreven als een product van priemgetallen: gegeven een natuurlijk getal n, dan bestaan er priemgetallen $p_1, p_2, \cdots, p_r$ en natuurlijke getallen $a_1, a_2, \cdots, a_r \geq 1$ zodat $n = p_1^{a_1} p_2^{a_2} \cdots p_r^{a_r}$, en deze schrijfwijze is uniek op permutatie van de factoren na.

(3) (Euclides) Er bestaat oneindig veel priemgetallen. Sterker nog (postulaat van Bertrand): voor elke $n \in \mathbb{N}_0$ bestaat er een priemgetal p met $n \leq p \leq 2n$. Een ander nuttig resultaat (stelling van Dirichlet): als a en b onderling ondeelbaar zijn, dan bestaan er oneindig veel priemgetallen van de vorm $an + b$.

(4) Als m en n natuurlijke getallen zijn, dan is $\text{ggd}(m, n) \cdot \text{kgv}(m, n) = mn$. De grootste gemene deler van twee natuurlijke getallen kan worden berekend met het algoritme van Euclides. Als $m = p_1^{a_1} p_2^{a_2} \cdots p_r^{a_r}$ en $n = p_1^{b_1} p_2^{b_2} \cdots p_r^{a_r}$ met $p_1, p_2, \cdots, p_r$ verschillende priemgetallen en $a_1, a_2, \cdots, a_r \geq 0$, dan is

$$\text{ggd}(m, n) = p_1^{\min(a_1, b_1)} p_2^{\min(a_2, b_2)} \cdots p_r^{\min(a_r, b_r)}, \quad \text{kgv}(m, n) = p_1^{\max(a_1, b_1)} p_2^{\max(a_2, b_2)} \cdots p_r^{\max(a_r, b_r)}.$$

(5) Als $a, b, c \in \mathbb{Z}$, $a \mid bc$ en $\text{ggd}(a, b) = 1$, dan $a \mid c$. Dus als $m, n \in \mathbb{Z}$, als p priem is en $p \mid mn$, dan is $p \mid m$ of $p \mid n$.

(6) Als $a, b, c \in \mathbb{Z}$ zodat $a \mid c$, $b \mid c$ en $\text{ggd}(a, b) = 1$, dan is $ab \mid c$.

(7) (Bézout) Zijn $a_1, a_2, \cdots, a_n \in \mathbb{Z}$. Dan bestaan er $m_1, m_2, \cdots, m_n \in \mathbb{Z}$ met $a_1m_1 + \cdots + a_nm_n = \text{ggd}(a_1, \cdots, a_n)$.

(8) Als $d, m, n \in \mathbb{Z}$ met $d \mid m$ en $d \mid n$, dan geldt ook dat $d \mid am + bn$ voor alle $a, b \in \mathbb{Z}$. Met andere woorden: een deler van twee gehele getallen is ook een deler van elke lineaire combinatie van deze twee getallen. Deze eigenschap kan makkelijk worden veralgemeend naar een gemeenschappelijke deler van drie of meer gehele getallen.

(9) Zij n een geheel getal met $n = p_1^{a_1}p_2^{a_2}\cdots p_r^{a_r}$ als unieke ontbinding in priemfactoren. Zij $\tau(n)$ het aantal positieve delers van n en zij $\sigma(n)$ de som van de positieve delers van n. Dan hebben we de volgende gelijkheden:

$$\tau(n) = (a_1 + 1)(a_2 + 1)\cdots(a_r + 1), \;\; \sigma(n) = \left(\frac{p_1^{a_1+1} - 1}{p_1 - 1}\right)\left(\frac{p_2^{a_2+1} - 1}{p_2 - 1}\right)\cdots\left(\frac{p_r^{a_r+1} - 1}{p_r - 1}\right).$$

(10) Zij n een natuurlijk getal en p een priemgetal. De exponent van p in de priemfactorenontbinding van $n!$ is gelijk aan

$$\left\lfloor \frac{n}{p} \right\rfloor + \left\lfloor \frac{n}{p^2} \right\rfloor + \left\lfloor \frac{n}{p^3} \right\rfloor + \cdots$$

met $\lfloor x \rfloor$ het grootste geheel getal kleiner dan of gelijk aan x.

Voorbeelden

Voorbeeld 1. Bepaal alle natuurlijke getallen n zodat $7^n \mid 9^n - 1$.

Oplossing. Merk op dat $9^n - 1 = (3^n - 1)(3^n + 1)$. Nu is $\text{ggd}(3^n - 1, 3^n + 1) = 2$, en daaruit volgt dat $7^n \mid 3^n - 1$ of $7^n \mid 3^n + 1$ (waarom?). Maar $7^n > 3^n + 1 > 3^n - 1$ voor $n \geq 1$, dus de enige oplossing is $n = 0$.

Voorbeeld 2. Zij $a, m, n \in \mathbb{N}$ met $a \geq 2$. Dan is $\text{ggd}(a^m - 1, a^n - 1) = a^{\text{ggd}(m,n)} - 1$.

Bewijs. Zij $d = \text{ggd}(m, n)$. Het is duidelijk dat $a^d - 1 \mid a^m - 1$ en $a^d - 1 \mid a^n - 1$, dus $a^d - 1 \mid \text{ggd}(a^m - 1, a^n - 1)$. Kies nu gehele getallen p en q met $pm + qn = d$. Veronderstel - zonder verlies van de algemeenheid - dat $p > 0$ en $q < 0$. Zij s een gemeenschappelijke deler van $a^m - 1$ en $a^n - 1$. We moeten nagaan dat $s \mid a^d - 1$. Maar $s \mid a^{pm} - 1$ en $s \mid a^{-qn} - 1$, dus $s \mid (a^{pm} - 1) - (a^{-qn} - 1)$, of nog, $s \mid a^{-qn}(a^{pm+qn} - 1)$. Maar $\text{ggd}(s, a) = 1$ (waarom?), dus $s \mid a^d - 1$. □

Voorbeeld 3. (IMO 1994) Bepaal alle koppels (m, n) van natuurlijke getallen m en n zodat $mn - 1 \mid m^3 + 1$.

Oplossing. Merk op dat $mn - 1 \mid n(m^3 + 1) - m^2(mn - 1)$, dus $mn - 1 \mid m^2 + n$. Stel $m^2 + n = a(mn - 1)$. Dan geldt $m^2 - amn + n + a = 0$. Bekijk de vierkantsvergelijking $x^2 - anx + a + n = 0$. Dan is m zeker een oplossing van de vergelijking - zij p de andere oplossing (het is mogelijk dat $p = m$). Dan geldt er $m + p = an$ en $mp = a + n$. Als $a, m, n, p \geq 2$, dan geldt $mp \geq m + p = an \geq a + n = mp$. Bijgevolg moet er gelijkheid optreden, dus $m = n = a = p = 2$.

Veronderstel dus dat één van de getallen a, m, n, p gelijk is aan 1. Stel eerst $a = 1$. Dan is $mp = n + 1$ en $m + p = n$, dus $mp = m + p + 1$, of nog, $(m - 1)(p - 1) = 2$. Bijgevolg is $(m, p) = (2, 3)$ of $(m, p) = (3, 2)$, en $n = 5$. Wegens symmetrie geeft de veronderstelling $n = 1$ dezelfde oplossingen voor (m, p). Veronderstel nu dat $p = 1$. Dan is analoog $(a, n) = (2, 3)$ of $(a, n) = (3, 2)$, en $m = 5$. Wegens symmetrie geeft de veronderstelling $m = 1$ dezelfde oplossingen voor (a, n).

Samenvattend: de mogelijke oplossingen zijn $(m, n) \in \{(2, 2), (2, 5), (3, 5), (5, 2), (5, 3), (1, 2), (1, 3), (2, 1), (3, 1)\}$, en een eenvoudige controle leert ons dat elk van deze koppels inderdaad een oplossing is. □

Voorbeeld 4. (IMO 2002) Zij $n \geq 2$ een natuurlijk getal met positieve delers $1 = d_1 < d_2 < \cdots < d_k = n$. Toon aan dat $d_1d_2 + d_2d_3 + \cdots + d_{k-1}d_k < n^2$. Wanneer geldt $d_1d_2 + d_2d_3 + \cdots + d_{k-1}d_k \mid n^2$?

Oplossing. Merk op dat $d_id_{k+i-1} = n$. We moeten dus bewijzen dat

$$\frac{1}{d_1d_2} + \frac{1}{d_2d_3} + \cdots + \frac{1}{d_{k-1}d_k} < 1.$$

Maar $d_i \geq i$, dus

$$\frac{1}{d_1 d_2}+\frac{1}{d_2 d_3}+\cdots+\frac{1}{d_{k-1}d_k} \leq \frac{1}{1\cdot 2}+\frac{1}{2\cdot 3}+\cdots+\frac{1}{(k-1)\cdot k} = \left(1-\frac{1}{2}\right)+\left(\frac{1}{2}-\frac{1}{3}\right)+\cdots+\left(\frac{1}{k-1}-\frac{1}{k}\right) = 1-\frac{1}{k} < 1.$$

Nu is $d_2 = p$ priem (waarom?). Dan is $d_{k-1} = n/p$, dus $d_1 d_2 + d_2 d_3 + \cdots + d_{k-1}d_k \geq d_{k-1}d_k = n^2/p$. Maar de grootste echte deler van n^2 is n^2/p. Dus $d_1 d_2 + d_2 d_3 + \cdots d_{k-1}d_k$ deelt n^2 a.s.a. $k = 2$, m.a.w. als $n = p$ een priemgetal is.

Opgaven

(1) Zijn x en y gehele getallen. Bewijs dat $17 \mid 2x + 3y$ als en slechts als $17 \mid 9x + 5y$.

(2) Bewijs dat het product van n opeenvolgende natuurlijke getallen steeds deelbaar is door $n!$.

(3) Zij $n \geq 5$ een natuurlijk getal. Bewijs: n is niet priem $\iff n \mid (n-1)!$.

(4) Zijn $a, b, m, n \in \mathbb{N}$ zodat $a^m - 1$ en $b^n + 1$ priem zijn. Geef zoveel mogelijk informatie over a, b, m en n.

(5) Zij $n \in \mathbb{N}$ zodat $24 \mid n + 1$. Bewijs dat de som van de positieve delers van n ook deelbaar is door 24.

(6) (IMO 1972) Bewijs dat de volgende uitdrukking een natuurlijk getal is voor alle $m, n \in \mathbb{N}_0$:

$$\frac{(2m)!(2n)!}{m!n!(m+n)!}.$$

(7) Bewijs dat de volgende uitdrukking een natuurlijk getal is voor alle $m, n \in \mathbb{N}_0$:

$$\frac{\text{ggd}(m,n)}{n}\binom{n}{m}.$$

(8) (IMO 1992) Bepaal alle natuurlijke getallen $1 < a < b < c$ zodat $(a-1)(b-1)(c-1) \mid abc - 1$.

(9) (IMO 2009) Zij n een natuurlijk getal. Zijn $a_1, a_2, \cdots, a_k$ (met $k \geq 2$) verschillende elementen van de verzameling $\{1, 2, \cdots, n\}$ zodanig dat $n \mid a_i(a_{i+1} - 1)$ voor $i = 1, 2, \cdots, k-1$. Bewijs dat n geen deler is van $a_k(a_1 - 1)$.

(10) (IMO 1998) Bepaal alle natuurlijke getallen a en b zodat $ab^2 + b + 7 \mid a^2 b + a + b$.

2 Modulo-rekenen

Gegeven gehele getallen a, b en m met $m \geq 2$, dan zeggen we dat a en b congruent zijn modulo m - of nog, $a \equiv b \pmod{m}$ - indien $m \mid a - b$, of nog, indien a en b dezelfde rest geven bij deling door m. Op deze manier krijgen we een equivalentierelatie (een transitieve, symmetrische en reflexieve relatie) op de verzameling $\mathbb{Z}$ van gehele getallen die ons toelaat om met "restklassen modulo m" te rekenen in plaats van met alle gehele getallen. Het grote voordeel van deze operatie is natuurlijk dat er slechts eindig veel restklassen modulo m zijn... De restklassen modulo m vormen een algebraïsche structuur die we de *ring* $\mathbb{Z}/m\mathbb{Z}$ noemen. Optelling en vermenigvuldiging zijn in deze ring gedefinieerd op de evidente manier: als $a \equiv b \pmod{m}$ en $c \equiv d \pmod{m}$, dan geldt $a + c \equiv b + d \pmod{m}$ en $ac \equiv bd \pmod{m}$. (Ga dat na!)

Voor een gegeven $a \in \mathbb{Z}$ en $m \geq 2$ zeggen we dat x een inverse is voor a modulo m als $ax \equiv 1 \pmod{m}$.

Bewering. Er bestaat een inverse voor a modulo m als en slechts als $\text{ggd}(a, m) = 1$.

Bewijs. Stel dat x een inverse is voor a. Dan is $ax \equiv 1 \pmod{m}$, m.a.w er bestaat een $p \in \mathbb{Z}$ met $ax = 1 + pm$. Dus $ax - pm = 1$. Maar $\text{ggd}(a, m)$ deelt $ax - pm$, dus $\text{ggd}(a, m) = 1$. Omgekeerd, als $\text{ggd}(a, m) = 1$, dan bestaat er volgens de stelling van Bézout een geheel getal p zodat $ax - pm = 1$, dus $ax \equiv 1 \pmod{m}$. □

Het aantal restklassen modulo m die een inverse hebben kan dus worden geïdentificeerd met de verzameling van natuurlijke getallen a met $0 < a < m$ en ggd$(a, m) = 1$. We noteren $\varphi(m)$ voor het aantal natuurlijke getallen a met die eigenschappen (φ is de *Euler-functie*). Indien $m = p_1^{a_1} p_2^{a_2} \cdots p_r^{a_r}$ de unieke ontbinding van m in priemfactoren is, dan geldt er

$$\varphi(m) = (p_1 - 1)p_1^{a_1 - 1}(p_2 - 1)p_2^{a_2 - 1} \cdots (p_r - 1)p_r^{a_r - 1}.$$

(Oefening: probeer deze gelijkheid zelf af te leiden!)

Eigenschappen

(1) Zij $m \geq 2$ een natuurlijk getal en zij $a \in \mathbb{Z}$ met ggd$(a, m) = 1$. Er bestaat een kleinste natuurlijk getal $q \neq 0$ met de eigenschap dat $a^q \equiv 1 \pmod{m}$. Dit getal q heet de *orde* van a modulo m en is een deler van $\varphi(m)$. Als r een willekeurig natuurlijk getal is zodanig dat $a^r \equiv 1 \pmod{m}$, dan geldt $q \mid r$.

(2) (Fermat) Als p priem is en $a \in \mathbb{N}$ is niet deelbaar door p, dan is $a^{p-1} \equiv 1 \pmod{p}$. (Dit is een speciaal geval van (1)!)

(3) (Wilson) Als p priem is, dan is $(p - 1)! \equiv -1 \pmod{p}$.

(4) (Primitieve wortels) Zij $m \geq 2$. Een natuurlijk getal a waarvan de orde modulo m gelijk is $\varphi(m)$ noemen we een *primitieve wortel* modulo m. Een primitieve wortel bestaat als en slechts als $m = 2$, $m = 4$, $m = p^k$ of $m = 2p^k$ met p priem. Als a een primitieve wortel is modulo p en $a^{p-1} \not\equiv 1 \pmod{p^2}$, dan is a ook een primitieve wortel modulo p^2; indien $a^{p-1} \equiv 1 \pmod{p^2}$, dan is $a + p$ een primitieve wortel modulo p^2. Als a een primitieve wortel is modulo p^k met $k \geq 2$, dan is a ook een primitieve wortel modulo p^ℓ voor alle $\ell \geq k$.

(5) (Chinese reststelling) Zijn $m_1, m_2, \cdots, m_k$ gehele getallen die paarsgewijs onderling ondeelbaar zijn, m.a.w. zodat ggd$(m_i, m_j) = 1$ als $i \neq j$. Zijn $a_1, a_2, \cdots, a_k \in \mathbb{Z}$ willekeurig. Dan bestaat er een natuurlijk getal x zodat $x \equiv a_1 \pmod{m_1}$, $x \equiv a_2 \pmod{m_2}$, $\cdots$, $x \equiv a_k \pmod{m_k}$. Dat getal x is bovendien uniek modulo $m_1 m_2 \cdots m_k$.

(6) (Kwadraten) Volkomen kwadraten zijn congruent met 0 of 1 modulo 4, congruent met 0 of 1 modulo 3, congruent met 0, 1 of 4 modulo 8, ... In het algemeen geldt: als p een oneven priemgetal is, dan bestaan er precies $\frac{1}{2}(p+1)$ restklassen modulo p (inclusief 0) die een volkomen kwadraat zijn modulo p. Als p een willekeurig priemgetal is, dan is -1 een kwadraat modulo p als en slechts als $p = 2$ of $p \equiv 1 \pmod{4}$, en 2 is een kwadraat modulo p als en slechts als $p = 2$ of $p \equiv \pm 1 \pmod{8}$. Dus elke priemdeler van een natuurlijk getal van de vorm $n^2 + 1$ is gelijk aan 2 of congruent met 1 modulo 4, en elke priemdeler van een getal van de vorm $n^2 - 2$ is gelijk aan 2 of congruent met ± 1 modulo 8.

Voorbeelden

Voorbeeld 1. Bestaat er een rij van 2011 opeenvolgende natuurlijke getallen zodanig dat elk van deze getallen deelbaar is door de 2011-de macht van een natuurlijk getal?

Oplossing. Het antwoord is ja. Zijn $2 = p_1 < p_2 < \cdots < p_{2011}$ de eerste 2011 priemgetallen. Volgens de Chinese reststelling bestaat er een natuurlijk getal n zodanig dat $n \equiv -1 \pmod{p_1^{2011}}$, $n \equiv -2 \pmod{p_2^{2011}}$, $\cdots$, $n \equiv -2011 \pmod{p_{2011}^{2011}}$. Dan is $n + 1, n + 2, \cdots, n + 2011$ de gevraagde rij. □

Voorbeeld 2. Bepaal alle oplossingen (in gehele getallen) van de vergelijking $x^2 = y^5 - 4$.

Oplossing. Een klein beetje rekenwerk leert ons dat een kwadraat steeds congruent is met 0, 1, 3, 4, 5 of 9 modulo 11, en dat een vijfdemacht steeds congruent is met -1, 0 of 1 modulo 11.[1] Dus $y^5 - 4$ is steeds congruent met 6, 7 of 8 modulo 11. Daaruit volgt dat er geen oplossingen zijn. □

Voorbeeld 3. Bepaal alle natuurlijke getallen x, y en z zodat $3^x + 4^y = 5^z$.

[1] Trucje voor de vijfdemachten: als y niet deelbaar is door 11, dan is $y^{10} \equiv 1 \pmod{11}$, dus $11 \mid y^{10} - 1 = (y^5 - 1)(y^5 + 1)$, dus $y^5 \equiv \pm 1 \pmod{11}$.

Oplossing. Modulo 4 wordt de vergelijking $(-1)^x \equiv 1 \pmod 4$. Bijgevolg is x even. Modulo 3 wordt de vergelijking $1 \equiv (-1)^z \pmod 3$. Bijgevolg is ook z even. Dus $4^y = 2^{2y} = (5^{z/2} - 3^{x/2})(5^{z/2} + 3^{x/2})$. Bijgevolg geldt $5^{z/2} - 3^{x/2} = 2^k$ en $5^{z/2} + 3^{x/2} = 2^\ell$ met $k < \ell$ en $k + \ell = 2y$. Dus $2 \cdot 5^{z/2} = 2^k + 2^\ell = 2^k(1 + 2^\ell)$. Daaruit volgt dat $k = 1$. Dus $5^{z/2} = 1 + 2^{\ell-1}$ en $3^{x/2} = -1 + 2^{\ell-1}$. Nu geldt dat machten van 3 steeds congruent zijn met 1 of 3 modulo 8 - het rechterlid van de laatste gelijkheid is echter congruent met -1 modulo 8 tenzij $\ell \leq 3$. Nu geeft $\ell = 3$ dat $x = y = z = 2$, en $\ell = 2$ geeft geen oplossing. Bijgevolg is de enige oplossing $(x, y, z) = (2, 2, 2)$.

Voorbeeld 4. (IMO 1999) Bepaal alle paren (n, p) van natuurlijke getallen n en p waarvoor geldt: p is een priemgetal, $n < 2p$ en $(p-1)^n + 1$ is deelbaar door n^{p-1}.

Oplossing. Zij q de kleinste priemdeler van n. Dan is $q \mid n^{p-1} \mid (p-1)^n + 1$, dus $(p-1)^n \equiv -1 \pmod q$. Bijgevolg is $p - 1$ niet deelbaar door q. Zij α de orde van $p - 1$ modulo q. Dan is $\alpha \mid q - 1$ omdat $(p-1)^{q-1} \equiv 1 \pmod q$ (Fermat). Verder is ook $(p-1)^{2n} \equiv 1 \pmod q$, dus $\alpha \mid 2n$. Dus $\alpha \mid \text{ggd}(q-1, 2n)$. Maar elke deler van $q - 1$ is kleiner dan q, en dus geen deler van n - want q is de *kleinste* priemdeler van n. Dus $\alpha = 1$ (als n even is) of $\alpha = 2$ (als n oneven is). Als $\alpha = 1$, dan is $q = 2$ (want n is even), maar dan moet natuurlijk ook $p = 2$, en dus $n = 2$. Als $\alpha = 2$, dan is $(p-1)^2 \equiv 1 \pmod q$, dus $q \mid p(p-2)$. Maar $p \not\equiv 2 \pmod q$ (anders zou $\alpha = 1$), dus $q \mid p$ en $q = p$. Omdat $n < 2p$ volgt daaruit dat $n = p$. Het is duidelijk dat $n = p = 3$ ook een oplossing is. Veronderstel nu dat $p \geq 5$. Dan geldt dat p^{p-1} - en dus ook p^3 - een deler is van $(p-1)^p + 1$. Maar met het binomium van Newton zien we dat deze uitdrukking gelijk is aan p^2 modulo p^3, contradictie! □

Voorbeeld 5. Zijn m en n gehele getallen. Bewijs dat $4mn - m - n$ geen volkomen kwadraat is.

Oplossing. Stel dat $4mn - m - n = a^2$. Dan is $4a^2 + 1 = (4m - 1)(4n - 1)$. Daaruit volgt dat $4a^2 + 1$ minstens één priemdeler heeft die congruent is met 3 modulo 4 - immers, niet alle priemdelers van $4m - 1$ kunnen congruent zijn met 1 modulo 4. Maar uit eigenschap (6) hierboven volgt dat dat niet kan. □

Voorbeeld 6. Zij $m \neq 0$ een veelvoud van 8. Hoeveel oplossingen (modulo m) heeft de kwadratische vergelijking $x^2 \equiv 1 \pmod m$ dan? Druk je antwoord uit in functie van het aantal priemdelers van m.

Oplossing. Zij $m = p_1^{a_1} p_2^{a_2} \cdots p_r^{a_r}$ de ontbinding van m, met dus $p_1 = 2$ en $a_1 \geq 3$. De congruentie $x^2 \equiv 1 \pmod m$ is volgens de Chinese reststelling equivalent met het stelsel van congruenties $x^2 \equiv 1 \pmod{p_i^{a_i}}$ voor $1 \leq i \leq r$. Dus $p_i^{a_i} \mid x^2 - 1 = (x-1)(x+1)$ voor alle i. Als $i \geq 2$ (m.a.w. als $p_i \geq 3$) dan volgt daaruit dat $x \equiv \pm 1 \pmod{p_i^{a_i}}$, want $p_i^{a_i}$ kan slechts één van de factoren $x - 1$ en $x + 1$ delen. Verder moet dus $2^{a_1} \mid (x-1)(x+1)$. Nu zal $\text{ggd}(x-1, x+1) = 2$, en dus moet x congruent zijn met 1, -1, $1 + 2^{a_1 - 1}$ of $-1 + 2^{a_1 - 1}$ modulo 2^{a_1}. Volgens de Chinese reststelling kunnen we nu de aantallen oplossingen modulo elke factor $p_i^{a_i}$ gewoon vermenigvuldigen om het aantal oplossingen modulo m te bekomen - denk daar even over na! - en we zien dus dat het aantal oplossingen gelijk is aan $2^{r-1} \cdot 4 = 2^{r+1}$. □

Opgaven

(1) Bepaal alle natuurlijke getallen n zodat $2^n \mid 3^n - 1$.

(2) Zij $p \geq 5$ een priemgetal. Bewijs dat $7^p - 6^p - 1$ deelbaar is door 43.

(3) Zij m een geheel getal zodat er een primitieve wortel a modulo m bestaat. Bewijs: $a^{\varphi(m)/2} \equiv -1 \pmod m$.

(4) Toon aan dat $2^{2 \cdot 3^{n-1}} \equiv 1 + 3^n \pmod{3^{n+1}}$ voor alle n en dat 2 een primitieve wortel is modulo 3^n, voor $n \geq 1$.

(5) Bepaal de grootste gemene deler van alle getallen van de vorm $n^{13} - n$, voor $n \in \mathbb{Z}$.

(6) Zij $n \geq 2$ een natuurlijk getal. Bewijs dat $2^n - 1$ niet deelbaar is door n.

(7) (IMO 2006) Beschouw de rij $(a_n)_{n \geq 1}$ gegeven door $a_n = 2^n + 3^n + 6^n - 1$. Bepaal alle natuurlijke getallen die onderling ondeelbaar zijn met elke term van deze rij.

(8) Bepaal de drie laatste cijfers van het getal $2003^{2002^{2001}}$.

(9) (IMO 1976) Wanneer 4444^{4444} in decimale schrijfwijze wordt geschreven, dan is de som van de cijfers gelijk aan A. Zij B de som de cijfers van A. Wat is de som van de cijfers van B?

(10) Zij n een natuurlijk getal en zij $p = 2^n + 1$. Veronderstel dat $p \mid 3^{(p-1)/2} + 1$. Bewijs dat p dan een priemgetal is.

(11) (LIMO 2007) Zij n een natuurlijk getal, p een priemgetal en d een deler van $(n+1)^p - n^p$. Bewijs dat $d \equiv 1 \pmod{p}$.

(12) Zij n een natuurlijk getal en zij p een priemgetal met $p \leq n$. Bewijs dat

$$\binom{n}{p} \equiv \left\lfloor \frac{n}{p} \right\rfloor \pmod{p}.$$

(13) (IMO 1996) Zijn a en b natuurlijke getallen (verschillend van 0) zodat $15a + 16b$ en $16a - 15b$ volkomen kwadraten zijn. Bepaal de kleinst mogelijke waarde van het kleinste van deze twee kwadraten.

(14) Zij $n \geq 3$ een oneven getal. Beschouw de verzameling S van gehele getallen x zodat $1 \leq x \leq n$ en zodat x en $x + 1$ allebei onderling ondeelbaar zijn met n. Bewijs dat het product van de elementen van S congruent is met 1 modulo n.

(15) (IMO 1990) Bepaal alle natuurlijke getallen n zodat $n^2 \mid 2^n + 1$. (*Hint:* 2 *een primitieve wortel is modulo* 3^ℓ.)

3 Uitsmijter: meer over kwadraatresten

Voor diegenen die de bovenstaande theorie al hebben gezien, een "uitsmijter": kwadratische reciprociteit in een notendop... Zij p een oneven priemgetal en zij n een geheel getal. Als $p \mid n$, dan stellen we $(\frac{n}{p}) = 0$. Als p geen deler is van n, dan noteren we $(\frac{n}{p}) = 1$ indien n een kwadraat is modulo p en $(\frac{n}{p}) = -1$ als n geen kwadraat is modulo p. We noemen $(\frac{n}{p})$ het *Légendre-symbool*. Bewijs eerst zelf de volgende eigenschappen:

- Er geldt $(\frac{n}{p}) \equiv n^{(p-1)/2} \pmod{p}$.
- Voor alle $m, n \in \mathbb{Z}$ is $(\frac{mn}{p}) = (\frac{m}{p})(\frac{n}{p})$.
- Er geldt $(\frac{1}{p}) + (\frac{2}{p}) + (\frac{3}{p}) + \cdots + (\frac{p-1}{p}) = 0$ - er zijn dus evenveel kwadraten modulo p als niet-kwadraten modulo p.
- Er geldt $(\frac{-1}{p}) = (-1)^{(p-1)/2}$ en $(\frac{2}{p}) = (-1)^{(p^2-1)/8}$.

De volgende stelling moet je niet proberen te bewijzen:

Stelling. (Kwadratische reciprociteit) Voor oneven priemgetallen p en q geldt dat

$$\left(\frac{p}{q}\right)\left(\frac{q}{p}\right) = (-1)^{\frac{1}{2}(p-1)\cdot\frac{1}{2}(q-1)}.$$

Voorbeeld 1. Zij $n \geq 3$ oneven en zij p een priemdeler van $2^n - 1$. Bewijs dat $p \equiv \pm 1 \pmod{8}$.

Oplossing. Stel $n = 2m + 1$. Dan is $2 \cdot (2^m)^2 \equiv 1 \pmod{p}$. Daaruit volgt dat $(\frac{2}{p}) \equiv 1 \pmod{p}$ - immers, stel α is de inverse van 2^m modulo p, dan is $\alpha^2 \equiv 2 \pmod{p}$. Uit de bovenstaande eigenschappen volgt dan dat $p \equiv \pm 1 \pmod{8}$. □

Voorbeeld 2. Voor welke priemgetallen p heeft de congruentie $x^2 \equiv -3 \pmod{p}$ een oplossing?

Oplossing. Voor $p = 2$ en $p = 3$ is er uiteraard een oplossing. Stel $p \geq 5$. We moeten alle p vinden zodat $(\frac{-3}{p}) = 1$, m.a.w. zodat $(\frac{-1}{p})(\frac{3}{p}) = 1$. Maar $(\frac{3}{p}) = (\frac{p}{3}) \cdot (-1)^{(p-1)/2}$ wegens de kwadratische reciprociteitswet, en $(\frac{-1}{p}) = (-1)^{(p-1)/2}$, dus we besluiten dat $(\frac{-3}{p}) = (\frac{p}{3})(-1)^{p-1} = (\frac{p}{3})$. Bijgevolg voldoen alle p met $p \equiv 1 \pmod{3}$. □

Voorbeeld 3. Voor welke natuurlijke getallen n bestaat er een natuurlijk getal m zodat $2^n - 1 \mid m^2 + 9$?

Oplossing. We bewijzen dat m bestaat als en slechts als $n = 2^k$ (met $k \geq 0$). Stel eerst dat n geen macht van 2 is. Kies dus een oneven priemdeler p van n: dan geldt $2^p - 1 \mid m^2 + 9$. Kies een priemfactor q van $2^p - 1$ met $q \equiv 3 \pmod 4$ en $q \neq 3$ (ga na dat q bestaat!). Dan $1 = (\frac{-9}{q}) = (\frac{9}{q})(\frac{-1}{q}) = (\frac{-1}{q}) = -1$ (omdat $q \equiv 3 \pmod 4$), contradictie. Bijgevolg bestaat m niet. Stel nu dat $n = 2^k$. Stel $T_r = 2^{2^r} + 1$, dan is $2^n - 1 = T_0 T_1 T_2 \cdots T_{k-1}$. Merk nu op dat ggd$(T_i, T_j) = 1$ als $i \neq j$ (ga dat na als oefening!). Kies een natuurlijk getal m zodat $m \equiv 2 \pmod{T_1}$, $m \equiv 2^2 \pmod{T_2}$, $\ldots$, $m \equiv 2^{2^{k-2}} \pmod{T_{k-1}}$ (Chinese reststelling). Dan is $m^2 + 1$ deelbaar door $T_1 T_2 \cdots T_{k-1}$, en dus is $(3m)^2 + 9$ deelbaar door $2^n - 1$. $\square$

Opgaven

(1) Bewijs dat 16 een volkomen achtste macht is modulo p, voor elk priemgetal p.

(2) Zijn a, b, c paarsgewijs onderling ondeelbare natuurlijke getallen met $c^2 = a^2 - ab + b^2$. Zij p een priemdeler van c. Bewijs dat $p \equiv 1 \pmod 6$.

(3) Zij F_n het n-de Fibonacci-getal. Bewijs dat voor elk priemgetal $p \geq 7$ geldt dat $F_p \equiv (\frac{p}{5}) \pmod p$.

4 Meer oefenmateriaal!

Hierna volgt nog een lijst van 35 leuke problemen die kan dienen als extra oefenmateriaal. Voor sommige opgaven zal de theorie die hierboven werd aangehaald erg nuttig zijn, maar er zitten ook opgaven bij die niet rechtstreeks aansluiten op de theorie. Ik heb de opgaven gerangschikt op moeilijkheidsgraad, maar die rangschikking is natuurlijk subjectief...

(1) Drie Amerikaanse wiskundigen gaven een tegenvoorbeeld voor een bekend vermoeden van Euler (in de jaren 1980) door aan te tonen dat er een natuurlijk getal n bestaat zodat $n^5 = 133^5 + 110^5 + 84^5 + 27^5$. Wat is de waarde van n?

(2) Het getal 219821459173083304870133 69 is de dertiende macht van een natuurlijk getal. Welk getal?

(3) Stel $34! = 95232799cd960414084761860964 35ab000000$. Bepaal de cijfers a, b, c en d.

(4) Laat zien dat de vergelijking $x^2 + y^5 = z^3$ oneindig veel gehele oplossingen heeft met $x, y, z \neq 0$.

(5) Zijn a en b natuurlijke getallen zodat $2^n a + b$ een volkomen kwadraat is voor alle natuurlijke getallen n. Bewijs: $a = 0$.

(6) Toon aan dat oneindig veel natuurlijke getallen niet kunnen worden geschreven als $x^2 + y^3 + z^7$, met $x, y, z \in \mathbb{N}$.

(7) Zij n een natuurlijk getal zodat $N = 2 + 2\sqrt{28n^2 + 1}$ een natuurlijk getal is. Bewijs dat N een volkomen kwadraat is.

(8) Bepaal alle $a, b \in \mathbb{N}$ zodat $(a + 19b)^{18} + (a + b)^{18} + (19a + b)^{18}$ een volkomen kwadraat is.

(9) Bewijs dat voor alle natuurlijke getallen n geldt: $7 \mid n^3 + 3^n \iff 7 \mid n^3 3^n + 1$.

(10) Definieer voor elk natuurlijk getal n het getal $p(n)$ als de grootste oneven deler van n. Bewijs:

$$\frac{1}{2^k} \sum_{n=1}^{2^k} \frac{p(n)}{n} > \frac{2}{3}.$$

(11) (IMO 1986) Zij d een natuurlijk getal met $d \notin \{0, 2, 5, 13\}$. Bewijs dat er in de verzameling $\{2, 5, 13, d\}$ steeds twee getallen a en b zitten zodanig dat het getal $ab - 1$ geen volkomen kwadraat is.

(12) Zijn n en q natuurlijke getallen met $n \geq 5$ en $2 \leq q \leq n$. Bewijs dat $q - 1$ een deler is van $\lfloor \frac{(n-1)!}{q} \rfloor$.

(13) Bepaal alle natuurlijke oplossingen van de vergelijking $a! \cdot b! = a! + b! + c!$.

(14) Toon aan dat elk geheel getal de som is van vijf volkomen derdemachten.

(15) We noemen $n \in \mathbb{N}$ *machtig* als n de volgende eigenschap heeft: als $n \mid a^n - 1$ voor een zekere $a \in \mathbb{N}$, dan geldt $n^2 \mid a^n - 1$. Bewijs dat priemgetallen machtig zijn, en dat oneindig veel niet-priemgetallen machtig zijn.

(16) Bepaal alle natuurlijke getallen a en b zodat $(\sqrt[3]{a} + \sqrt[3]{b} - 1)^2 = 49 + 20\sqrt[3]{6}$.

(17) Bepaal alle rekenkundige rijtjes van drie natuurlijke getallen met de eigenschap dat het product van de drie termen van het rijtje geen priemfactor heeft die strikt groter is dan 3.

(18) Zij α de grootste wortel van de vergelijking $x^3 - 3x^2 + 1 = 0$. Bewijs dat $\lfloor \alpha^{1788} \rfloor$ en $\lfloor \alpha^{1988} \rfloor$ deelbaar zijn door 17.

(19) Bewijs dat voor elk natuurlijk getal n geldt dat $\lfloor \sqrt[3]{n} + \sqrt[3]{n+1} \rfloor = \lfloor \sqrt[3]{8n+3} \rfloor$.

(20) Zij N een natuurlijk getal. Bewijs dat er een rij van N *opeenvolgende* natuurlijke getallen bestaat zodanig dat de j-de term van deze rij de som is van j *verschillende* volkomen kwadraten.

(21) Definieer $a_0 = 0$, $a_1 = 1$ en $a_{n+2} = 2a_{n+1} + a_n$ voor $n \geq 0$. Bewijs: $2^k \mid a_n \iff 2^k \mid n$.

(22) Definieer een rij van natuurlijke getallen door $u_0 = 1$ en $u_{n+1} = au_n + b$, waarbij a en b willekeurige natuurlijke getallen zijn. Bewijs dat deze rij (voor elke keuze van a en b) oneindig veel termen heeft die niet priem zijn.

(23) Bewijs dat er oneindig veel natuurlijke getallen n bestaan met $n^2 + 1 \mid n!$.

(24) Definieer $y_0 = 1$ en $y_{n+1} = \frac{1}{2}(3y_n + \sqrt{5y_n^2 - 4})$. Bewijs dat $y_n \in \mathbb{N}$ voor alle n.

(25) (IMO 2006) Bepaal alle natuurlijke getallen x en y zodat $y^2 = 1 + 2^x + 2^{2x+1}$.

(26) (IMO 1997) Bepaal alle gehele getallen a en b zodat $a^{b^2} = b^a$.

(27) (IMO 2003) Bepaal alle gehele getallen a en b zodat $a^2/(2ab^2 - b^3 + 1)$ een natuurlijk getal is.

(28) Een deelnemer aan het IMO-stageweekend die niet goed heeft opgelet herinnert zich de kleine stelling van Fermat als volgt: als p een priemgetal is en a een natuurlijk getal, dan is $a^{p+1} \equiv a \pmod{p}$. Dat slaat natuurlijk nergens op, en iedereen die opgelet heeft weet dat de juiste congruentie de volgende is: $a^p \equiv a \pmod{p}$. Maar toch de volgende vraag: welke natuurlijke getallen p hebben de eigenschap dat $a^{p+1} \equiv a \pmod{p}$ voor alle natuurlijke getallen a?

(29) Zijn $p_1, p_2, \cdots, p_k \geq 5$ verschillende priemgetallen. Bewijs dat $\tau(2^{p_1 p_2 \cdots p_k} + 1) \geq 4^k$.

(30) Zijn x en y natuurlijke getallen zodat xy een deler is van $x^2 + y^2 + 1$. Bewijs dat $x^2 + y^2 + 1 = 3xy$.

(31) (IMO 2007) Zijn a en b natuurlijke getallen zodat $4ab - 1 \mid (4a^2 - 1)^2$. Bewijs dat $a = b$.

(32) Bestaat er een natuurlijk getal m zodanig dat de vergelijking

$$\frac{1}{a} + \frac{1}{b} + \frac{1}{c} + \frac{1}{abc} = \frac{m}{a+b+c}$$

oneindig veel natuurlijke oplossingen heeft?

(33) (IMO 1988) Zijn a en b natuurlijke getallen zodat $ab + 1$ een deler is van $a^2 + b^2$. Bewijs dat

$$\frac{a^2 + b^2}{ab + 1}$$

een volkomen kwadraat is.

(34) (IMO 2000) Bestaat er een natuurlijk getal n met precies 2000 priemdelers zodat n een deler is van $2^n + 1$?

(35) (IMO 1998) Voor welke natuurlijke getallen k bestaat er een natuurlijk getal k zodat $\tau(n^2) = k\tau(n)$?

terug naar echt bestand

Winter Camp 2009
Number Theory Tips and Tricks

David Arthur
darthur@gmail.com

1 Introduction

This handout is about some of the key techniques for solving number theory problems, especially Diophantine equations (equations with integer variables). Some of this stuff is pretty advanced, so if you have trouble following something, it's okay. Don't be afraid to ask questions!

I'm going to assume you already know some of the basics of number theory, especially modular arithmetic. I am also not going to spend much time covering theorems. If you want theorems or more background, I suggest checking out Naoki Sato's handout:

http://www.artofproblemsolving.com/Resources/Papers/SatoNT.pdf

And once again, don't be afraid to ask questions!

2 Reduce mod n

Most IMO-level students will be familiar with the idea that an equation (or system of equations) can sometimes be solved by first reducing mod n, and then showing it has no solutions mod n. Here is one such problem:

Example: (*IMO 1986, #1*) Let d be any positive integer not equal to 2, 5, or 13. Show that one can find distinct a, b in the set $\{2, 5, 13, d\}$ such that $ab - 1$ is not a perfect square.

Solution: The quadratic residues mod 16 are $\{0, 1, 4, 9\}$. Therefore, $2d - 1$ can only be a perfect square if $d \in \{1, 5, 9, 13\} \pmod{16}$, and $5d - 1$ can only be a perfect square if $d \in \{1, 2, 10, 13\} \pmod{16}$, and $13d - 1$ can only be a perfect square if $d \in \{2, 5, 9, 10\} \pmod{16}$. There is no d that simultaneously satisfies all three conditions, and the result follows. □

If you have done a lot of number theory before, this may seem like a fairly standard problem to you, and you might wonder that it made it onto the IMO. And it is a standard problem... sort of. We need to show a set of equations: $2d-1 = x^2, 5d-1 = y^2, 13d-1 = z^2$ has no integer solutions. So we reduce the equations mod 16, and by checking every possible value for d, we confirm that in fact, there are no possible solutions. There is one tricky part though: why should we choose 16 in particular? Nothing smaller works.

The art for these problems is choosing the right n. Here are some tips:

- Make n a prime power. The Chinese Remainder Theorem guarantees that looking at any polynomial mod xy is no better than looking at it mod x and then looking at it mod y

(assuming x and y are relatively prime). If there are variables in the exponents, you might want to break this rule, but even then, prime powers are still usually the right choice.

- If there are perfect squares in the equations, try n a power of 2. The fewer quadratic residues there are mod n, the better off you will be. If n is a power of 2, the number of quadratic residues mod n is $\lceil \frac{n}{6} \rceil + 1$. If n is a power of $p \neq 2$, the number of quadratic residues mod n is $\left\lceil \frac{pn}{2(p+1)} \right\rceil \geq \lceil \frac{n}{3} \rceil$. So, as you can see, powers of 2 are basically twice as good as the other choices! In practice, 4, 8, and sometimes 16 are good numbers to try.

- If there are m^{th} powers in the equation, the key is to choose $n = p^k$ so that $g = \gcd(m, (p-1)p^{k-1})$ is as large as possible. This is because the number of m^{th} powers mod n is approximately $\frac{n}{g}$. Usually you want to choose p, k so that $m|(p-1)p^{k-1}$.

- Make sure there is something to gain from doing modular arithmetic in the first place! The technique is very useful for showing an equation has no solutions. But if it has even one solution, it will also have a solution mod n for all n. Even if you can show all solutions are 1 mod 1,000,000,000, that still leaves an infinite number of possibilities to check!

- If the sum of the digits of an integer $S(n)$ is involved, always consider mod 9, because $S(n) \equiv n \pmod 9$.

3 Check the size of things

Almost as important as modular arithmetic in number theory is the fact that distinct integers differ by at least 1. An obvious fact, but a useful one nonetheless!

Example: (*see APMO 1999, #4*[1]) Find all positive integers (a, b) such that $a^2 + 4b$ and $b^2 + 4a$ are both perfect squares.

Solution: Suppose (a, b) is a solution. Assume without loss of generality that $a \leq b$. Then $b^2 < b^2 + 4a \leq b^2 + 4b < (b+2)^2$. It follows that $b^2 + 4a = (b+1)^2$, which implies $a = \frac{2b+1}{4}$. However, this is impossible because $\frac{2b+1}{4}$ is not an integer. Therefore, there are no solutions. □

In general, you should just always keep in mind approximately how large the quantities you are working with are. Take a step back, and ask if these bounds are pretty restrictive. If they are, you should probably investigate them pretty carefully. Here are a couple things to keep in mind:

- If $a|b$, then $b = 0$ or $|a| \leq |b|$.

- For all x, we have $x - 1 < \lfloor x \rfloor \leq x$, and $x \leq \lceil x \rceil < x + 1$.

- If x is a known integer and y is an unknown integer with $y \approx x$, then there aren't very many possibilities for y!

- Remember the division algorithm! Given integers n, m, there are unique integers a, b with $0 \leq b < m$, so that $n = am + b$. Sometimes, you can make this substitution, deal with the am part trivially, and then use inequality techniques to deal with the b part.

[1]The real APMO problem asks you to look for negative solutions as well. The same approach works, but there are more cases that you have to consider.

4 Factor

For many an Olympiad problem, the key step is a clever factoring or rewriting of the equation. Here are some useful things you can say after writing an expression as $x \cdot y$ for integers x, y:

- If $x, y > 1$, then xy is composite.
- If $xy = 0$, then $x = 0$ or $y = 0$.
- If xy is a power of a prime p, then x and y are powers of p as well.
- If x, y are relatively prime, and xy is a perfect k^{th} power, then x and y are perfect k^{th} powers as well.
- If $xy = a^2 + b^2$ with $\gcd(a, b) = 1$, then $x, y \not\equiv 3 \pmod 4$. (See example below.)

Example: Prove that there are no integer solutions (x, y) to $y^2 = x^3 + 23$.

Solution: The solution is based on the fact that if a prime p is congruent to 3 (mod 4), then -1 is not a quadratic residue modulo p. Remember that you were asked to prove this very useful fact during the pre-camp problem set!

In the given equation, note that if y is odd, then $x^3 \equiv 1 - 23 \equiv 2 \pmod 4$, which is impossible. If y is even, then $x^3 \equiv -23 \equiv 1 \pmod 4$. This leaves only the possibility that y is even and $x \equiv 1 \pmod 4$.

Now, write the equation as $4\left(\left(\frac{y}{2}\right)^2 + 1\right) = (x + 3)(x^2 - 3x + 9)$, and note that $x^2 - 3x + 9 \equiv 3 \pmod 4$. Therefore, there exists a prime $p \equiv 3 \pmod 4$ that divides $x^2 - 3x + 9$. This prime must also divide $\left(\frac{y}{2}\right)^2 + 1$. However, this would imply that -1 is a quadratic residue modulo p, which is impossible. □

Factoring is used in many places. Here are some things to watch out for:

- Completing the square: $x^2 + ax = \left(x + \frac{a}{2}\right)^2 - \frac{a^2}{4}$. You can then look at quadratic residues, or you can use this as part of a larger factoring. This can also set up a Pell's equation[2] (e.g. $x^2 + x = 2y^2$).
- Difference and sums of nth powers:
  $x^n - y^n = (x - y)(x^{n-1} + x^{n-2}y + x^{n-3}y^2 + \ldots + xy^{n-2} + y^{n-1})$,
  $x^n + y^n = (x + y)(x^{n-1} - x^{n-2}y + x^{n-3}y^2 - \ldots - xy^{n-2} + y^{n-1})$ if n is odd.
- Sophie Germain's identity: $x^4 + 4y^4 = (x^2 - 2xy + 2y^2)(x^2 + 2xy + 2y^2)$.
- Polynomials evaluated at two different points: There are many problems that revolve around the identity $x - y | P(x) - P(y)$ for P a polynomial with integer coefficients.
- Diophantine equations with variables in the exponents (e.g. $3^x - 2^y = 1$). For these problems, you almost always want to use modular arithmetic to show a couple exponents have to be multiples of some integer n, and then factor the equation as a difference (or sum) of n^{th} powers.

[2] A Pell's equation is an equation of the form $x^2 - Dy^2 = 1$, where D is a constant non-square. Such an equation always has an infinite number of integer solutions. If (x_1, y_1) is the smallest solution with $x > 1$, then the full set of solutions is given by $(\pm x_n, \pm y_n)$ where x_n and y_n are defined by $x_n + y_n\sqrt{D} = (x_1 + y_1\sqrt{D})^n$.

5 Use infinite descent

Suppose you want to show an equation has no positive integer solutions, or it has no positive integer solutions of a certain type. With infinite descent, you do a proof by contradiction. Let x be the smallest "bad" solution. Then show there is another bad solution x' with $0 < x' < x$, giving a contradiction.

Warning: This only works for positive integers. You cannot assume an equation has a smallest rational solution, a smallest real solution, or a smallest negative solution! If an equation involves positive and negative integers, you can still do infinite descent if you take x minimizing $|x|$, but do *not* forget the absolute value!

If you want to do infinite descent, the real question is: how do you find x'?

Example: 2009 stones are given, each with positive integer weight. If any one stone is removed, the remaining stones can be split into two heaps with the same total weight, each containing exactly 1004 stones. Prove that all the stones weigh the same.

Solution: Let $\{w_i\}$ denote the weight of the stones, and let $W = \sum_{i=1}^{2009} w_i$. Assume the claim is false, and consider a counterexample minimizing W. If w_i is removed, then the remaining stones can be split into two equal-weight piles, so $W - w_i$ must be even. Since W is a constant, it follows that each w_i has the same parity.

If each w_i is even, then 2009 stones of weight $\{\frac{w_i}{2}\}$ is also a counterexample, but with smaller W. If each w_i is odd, then 2009 stones of weight $\{\frac{w_i+1}{2}\}$ is also a counterexample, but with smaller W (since, by assumption, we did not have every $w_i = 1$). Either way, we have a contradiction, and the result is proven. □

Here are some tips on how to set up infinite descent:

- If you can show every variable in an equation must be even (or a multiple of n), then you can often divide everything by n to get a smaller solution. *Example:* Solve $x^2 + y^2 = 3z^2$.
- Suppose you have an equation that is quadratic in one or more variables: e.g. $a^2+b^2 = abc+c$. Then if you know one root, you can find the other. For example, if $a = x$ is a root of the above equation, then $a = bc - x^2$ is another root. If the new value is smaller than the old one (but still positive), you can do infinite descent!

  This is an extremely important technique for the IMO, and it is called *root flipping* or *Vieta jumping*. Even outside of infinite descent solutions, you can still often learn something by looking at the other root of a quadratic.
- Sometimes it is easy to find larger solutions. You can then try to reverse the construction to get smaller solutions. *Example:* Call a positive integer "good" if it can be written in the form $a^2 + 3b^2$. Show that the product of two good numbers is good. Then reverse this construction to show that if $7n$ is good, then n is good.

6 Look at the order of elements mod n

You will often find yourself dealing with terms of the form x^y for various values y. To work with such expressions, it is helpful to remember the following:

Theorem 6.1. *Fix x and n with $\gcd(x, n) = 1$. There exists an integer m, called the order of x modulo n, such that $x^y \equiv 1$ (mod n) if and only if $m|y$.*

In these terms, the very popular Fermat's little theorem states that if $x \not\equiv 0 \pmod{p}$, then the order of x mod p divides $p-1$.

Example: (*IMO Shortlist 2006, N5*) Find all integer solutions of the equation $\frac{x^7-1}{x-1} = y^5 - 1$.

Solution: This question is similar to $y^2 = x^3 + 23$, which was covered earlier.

Suppose $p \not\equiv 1 \pmod{7}$ is a prime divisor of $\frac{x^7-1}{x-1} = x^6+x^5+\ldots+1$, and let m denote the order of x modulo p. We know $x^7 \equiv 1 \pmod{p}$ so $m|7$. Also, by Fermat's little theorem, $m|p-1$. Since $p \not\equiv 1 \pmod{7}$, we know 7 and $p-1$ are relatively prime, so $m = 1$. Therefore, $x^1 \equiv 1 \pmod{p}$. Plugging this in, we have $0 \equiv x^6 + x^5 + \ldots + 1 \equiv 1 + 1 + \ldots + 1 \equiv 7 \pmod{p}$, and hence $p = 7$. It follows that if $z|\frac{x^7-1}{x-1}$, then z is congruent to 0 or 1 mod 7.

If $\frac{x^7-1}{x-1} = (y-1)(y^4+y^3+y^2+y+1)$, we therefore have $y \equiv 1, 2 \pmod{7} \implies y^4+y^3+y^2+y+1 \equiv 5, 3 \pmod{7}$. Either way, we have a contradiction. Therefore, the equation has no integer solutions. □

There are a couple generalizations of Fermat's little theorem that you should also know. First of all, what if p is not prime? To cover this case, we define $\phi(n)$ to be the number of positive integers less than n relatively prime to n. You can check that $\phi(p_1^{e_1} \cdot p_2^{e_2} \cdot \ldots \cdot p_k^{e_k}) = (p_1 - 1)p_1^{e_1-1} \cdot (p_2 - 1)p_2^{e_2-1} \cdot \ldots \cdot (p_k - 1)p_k^{e_k-1}$. Then:

Theorem 6.2. *(Euler's theorem) Fix x and n with $\gcd(x, n) = 1$. Then the order of x modulo n divides $\phi(n)$. Equivalently, $x^{\phi(n)} \equiv 1 \pmod{n}$.*

If n is a prime, then $\phi(n) = n - 1$, so this reduces exactly to Fermat's little theorem.

The next theorem is very powerful, and you should absolutely keep it in mind as the right intuition for how things work modulo prime powers. However, it is considered an advanced theorem, so quote it directly on contests at your own risk![3]

Theorem 6.3. *(Primitive roots) Suppose $n = p^k$ or $2p^k$ for some odd prime p and some positive integer k. Then, there exists x so that the order of x modulo n is exactly $\phi(n)$.*

Why is this so useful? It means that the set of integers relatively prime to n is precisely the set $\{1, x, x^2, \ldots, x^{\phi(n)-1}\}$ modulo n and the set of k^{th} powers relatively prime to n is precisely the set $\{1, x^k, x^{2k}, \ldots\}$. As an exercise, you might try using primitive roots to prove the following facts:

- If p is an odd prime, then -1 is a quadratic residue mod p^k if and only if $p \equiv 1 \pmod{4}$.
- Fix a prime p. Then p divides $1^k + 2^k + \ldots + p^k$ if and only if $p - 1$ does not divide k.

[3]On the IMO, if you use a known theorem and you have a correct solution, you will likely get your 7 points no matter what. The bad news is if you use a theorem or technique they don't like (e.g. coordinate geometry or Lagrange multipliers) and then make a mistake, you may not get much partial credit. Other Olympiads, including the CMO, are sometimes less generous even on correct solutions.

7 Problems

I have divided the problems into 3 types. The A problems are short (but not necessarily easy) problems that illustrate the concepts in these notes. The B problems are real Olympiad problems, most of which are quite challenging. The C problems are comparable to the hardest IMO number theory problems. If you can get them, you can get anything!

There are hints at the back, but only look at them after seriously trying the problems first.

A1. Let a, b, c, d be positive integers with $ab = cd$. Prove that $a + b + c + d$ is composite.

A2. Show that $4^n + n^4$ is composite for all integers $n \geq 2$.

A3. Prove that the system of equations:

$$x^2 + 6y^2 = z^2$$
$$6x^2 + y^2 = t^2$$

has no non-trivial integer solutions.

A4. Show that 19^{19} cannot be written as $m^4 + n^3$ for any integers m and n.

A5. Recall that e is given by the infinite sum $e = 1 + \frac{1}{1!} + \frac{1}{2!} + \frac{1}{3!} + \cdots$. Show that e is irrational.

B1. For which positive integers n do there exist positive integers a, b satisfying $a+b+n \cdot \gcd(a, b) = \text{lcm}(a, b)$?

B2. (*Korean Math Olympiad 1998, #1*) Find all pairwise relatively prime positive integers l, m, n such that

$$(l + m + n) \cdot \left(\frac{1}{l} + \frac{1}{m} + \frac{1}{n} \right)$$

is an integer.

B3. (*USAMO 2005, #2*) Prove that the system

$$\begin{aligned} x^6 + x^3 + x^3y + y &= 147^{157} \\ x^3 + x^3y + y^2 + y + z^9 &= 157^{147} \end{aligned}$$

has no solutions in integers x, y, and z.

B4. (*Bulgarian Math Olympiad 1981, #4*) Prove that, if $1 + 2^n + 4^n$ is prime, then $n = 3^k$ for some integer k.

B5. (*IMO 1989, #5*) Prove that for each positive integer n, there exist n consecutive positive integers none of which is an integral power of a prime number.

B6. (*Russian Math Olympiad 1999, Grade 11, #5*) Four natural numbers have the property that the square of the sum of any two of the numbers is divisible by the product of the other two. Show that at least three of the four numbers are equal.

B7. (*IMO Shortlist 1996, N4*) Find all positive integers m and n such that $\left\lfloor \frac{m^2}{n} \right\rfloor + \left\lfloor \frac{n^2}{m} \right\rfloor = \left\lfloor \frac{m}{n} + \frac{n}{m} \right\rfloor + mn$.

B8. Prove that if $n \geq 3$, then $\lceil (3+\sqrt{5})^n \rceil$ is divisible by 8.

B9. (*IMO 2003, #2*) Determine all pairs of positive integers (a, b) such that

$$\frac{a^2}{2ab^2 - b^3 + 1}$$

is a positive integer.

B10. (*Romania 1997*) Let $P(x), Q(x)$ be monic irreducible polynomials over the rational numbers. Suppose P and Q have respective roots α and β such that $\alpha + \beta$ is rational. Prove that the polynomial $P(x)^2 - Q(x)^2$ has a rational root.

B11. Prove that if n is a positive integer with the property that both $3n+1$ and $4n+1$ are perfect squares, then n is divisible by 7.

B12. (*IMO 1998, #3*) For any positive integer n, let $d(n)$ denote the number of positive divisors of n (including 1 and n itself).

Determine all positive integers k such that

$$\frac{d(n^2)}{d(n)} = k$$

for some n.

C1. (*IMO Shortlist 1998, N5*) Find all positive integers n for which there is an integer m with $2^n - 1 | m^2 + 9$.

C2. (a) (*IMO 2007, #5*) Let a and b be positive integers. Show that if $4ab-1$ divides $(4a^2-1)^2$, then $a = b$.

(b) (*IMO 1988, #6*) Let a and b be positive integers such that $ab+1$ divides a^2+b^2. Show that

$$\frac{a^2+b^2}{ab+1}$$

is the square of an integer.

C3. (*IMO Shortlist 2005, N6*) Let a and b be positive integers such that $a^n + n$ divides $b^n + n$ for every positive integer n. Show that $a = b$.

C4. (*IMO 1987, #6*) Let n be an integer greater than or equal to 2. Prove that if $k^2 + k + n$ is prime for all integers k such that $0 \leq k \leq \sqrt{n/3}$, then $k^2 + k + n$ is prime for all integers k such that $0 \leq k \leq n - 2$.

C5. (*IMO 1990, #6*) Prove that there exists a convex 1990-gon with the following two properties: (a) all angles are equal; (b) the lengths of the 1990 sides are the numbers $1^2, 2^2, 3^2, \ldots, 1990^2$ in some order.

C6. (*IMO Shortlist 2001, N6*) Is it possible to find 100 positive integers not exceeding 25,000, such that all pairwise sums of them are different?

C7. (*Chinese Math Olympiad 2006, #3*) Positive integers k, m, n satisfy $mn = k^2 + k + 3$. Prove there exist odd integers x, y so that either $x^2 + 11y^2 = 4m$ or $x^2 + 11y^2 = 4n$.

8 Selected Hints

A1. Substitute $d = \frac{ab}{c}$.

A2. Use Sophie Germain's identity.

A3. Add the equations, and do infinite descent.

A4. You want to reduce mod x. Remember the tips for choosing x.

A5. Assume that $e = \frac{m}{n}$ and multiply through by $n!$.

B1. Let $p = \gcd(a, b), q = \frac{a}{p}, r = \frac{b}{p}$.

B2. Show $m|n + l$. If $m \geq n, l$, that is pretty restrictive.

B3. Add the equations, and factor.

B4. Work out some examples. You should be able to see what must divide what.

B5. Use the Chinese Remainder Theorem.

B6. If p divides one of the numbers, what can you say about p?

B7. This is basically an inequality problem. If $m \geq n$, first show $m^2 > n$.

B8. Show $\lceil (3 + \sqrt{5})^n \rceil = (3 + \sqrt{5})^n + (3 - \sqrt{5})^n$.

B9. By root-flipping on $a^2 = 2ab^2x - b^3x + x$, it suffices to focus on the case $a \leq b$.

B10. Prove $Q(x) = \pm P(\alpha + \beta - x)$.

B11. Surprisingly, reducing mod 7^k doesn't work. But you can find *all* solutions using a Pell's equation. See the earlier footnote about them.

B12. Recall that if $n = \prod p_i^{e_i}$, then $d(n) = \prod(e_i + 1)$. Now try to do small values of k. Can you generalize your construction?

C1. You will need to show $2^{2^a} - 1$ has no prime divisors other than 3 that are 3 (mod 4).

C2. These problems both rely on infinite descent with root-flipping. For (a), first make the numerator of $\frac{(4a^2-1)^2}{4ab-1}$ more manageable. For (b), the tricky part is showing the new solution is positive.

C3. You don't need to consider every n. Choose one (depending on a, b) that is easy to work with.

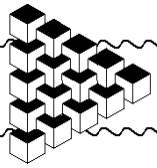
C4. If x is a solution to $k^2 + k + n \equiv 0 \pmod{p}$, then so is $p - 1 - x$.

C5. Let ω_n denote a complex nth root of unity. Find an ordering $\ell_{i,j,k}$ of $\{1^2, 2^2, \ldots, 1990^2\}$ for which $\sum_{i=0}^{1} \sum_{j=0}^{4} \sum_{k=0}^{198} \omega_2^i \omega_5^j \omega_{199}^k \ell_{i,j,k} = 0$. (A simple ordering works – no need to be fancy.)

C6. The answer is yes, even for 101 numbers (hint, hint). Try setting it up so you can deduce $i + j$ and ij from $x_i + x_j$.

C7. Use infinite descent (but not root flipping for once) to prove that if $k^2 + 11 = 4mn$, then there exist a, b, c, d, not all even, satisfying $4m = a^2 + 11b^2$, $4n = c^2 + 11d^2$, $2k = ac + 11bd$.

terug naar echt bestand



©2007 The Author(s) and The IMO Compendium Group

Quadratic Congruences

Dušan Djukić

Contents

1 Quadratic Congruences to Prime Moduli

Definition 1. *Let m, n and a be integers, $m > 1$, $n \geq 1$ and $(a,m) = 1$. We say that a is a residue of n-th degree modulo m if congruence $x^n \equiv a \pmod{m}$ has an integer solution; else a is a nonresidue of n-th degree.*

Specifically, for $n = 2,3,4$ the residues are called quadratic, cubic, biquadratic, respectively. This text is mainly concerned with quadratic residues.

Theorem 1. *Given a prime p and an integer a, the equation $x^2 \equiv a$ has zero, one, or two solutions modulo p.*

Proof. Suppose that the considered congruence has a solution x_1. Then so clearly is $x_2 = -x_1$. There are no other solutions modulo p, because $x^2 \equiv a \equiv x_1^2 \pmod{p}$ implies $x \equiv \pm x_1$. □

As a consequence of the above simple statement we obtain:

Theorem 2. *For every odd positive integer p, among the numbers $1, 2, \ldots, p-1$ there are exactly $\frac{p-1}{2}$ quadratic residues (and as many quadratic nonresidues).* □

Definition 2. *Given a prime number p and an integer a, Legendre's symbol $\left(\frac{a}{p}\right)$ is defined as*

$$\left(\frac{a}{p}\right) = \begin{cases} 1, & \text{if } p \nmid a \text{ and } a \text{ is a quadratic residue (mod } p\text{);} \\ -1, & \text{if } p \nmid a \text{ and } a \text{ is a quadratic nonresidue (mod } p\text{);} \\ 0, & \text{if } p \mid a. \end{cases}$$

Example 1. *Obviously, $\left(\frac{x^2}{p}\right) = 1$ for each prime p and integer x, $p \nmid x$.*

Example 2. *Since 2 is a quadratic residue modulo 7 ($3^2 \equiv 2$), and 3 is not, we have $\left(\frac{2}{7}\right) = 1$ and $\left(\frac{3}{7}\right) = -1$.*

From now on, unless noted otherwise, p is always an odd prime and a an integer. We also denote $p' = \frac{p-1}{2}$.

Clearly, a is a quadratic residue modulo p if and only if so is $a + kp$ for some integer k. Thus we may regard Legendre's symbol as a map from the residue classes modulo p to the set $\{-1, 0, 1\}$.

Fermat's theorem asserts that $a^{p-1} \equiv 1 \pmod{p}$, which implies $a^{p'} \equiv \pm 1 \pmod{p}$. More precisely, the following statement holds:

Theorem 3 (Euler's Criterion). $a^{p'} \equiv \left(\frac{a}{p}\right)$ *(mod p).*

Proof. The statement is trivial for $p \mid a$. From now on we assume that $p \nmid a$.

Let g be a primitive root modulo p. Then the numbers g^i, $i = 0, 1, \ldots, p-2$ form a reduced system of residues modulo p. We observe that $(g^i)^{p'} = g^{ip'} \equiv 1$ if and only if $p-1 \mid ip'$, or equivalently, $2 \mid i$.

On the other hand, g^i is a quadratic residue modulo p if and only if there exists $j \in \{0, 1, \ldots, p-2\}$ such that $(g^j)^2 \equiv g^i \pmod{p}$, which is equivalent to $2j \equiv i \pmod{p-1}$. The last congruence is solvable if and only if $2 \mid i$, that is, exactly when $(g^i)^{p'} \equiv 1 \pmod{p}$. □

The following important properties of Legendre's symbol follow directly from Euler's criterion.

Theorem 4. *Legendre's symbol is multiplicative, i.e.* $\left(\frac{ab}{p}\right) = \left(\frac{a}{p}\right)\left(\frac{b}{p}\right)$ *for all integers* a, b *and prime number* $p > 2$. □

Problem 1. *There exists a natural number* $a < \sqrt{p} + 1$ *that is a quadratic nonresidue modulo p.*

Solution. Consider the smallest positive quadratic nonresidue a modulo p and let $b = \left[\frac{p}{a}\right] + 1$. Since $0 < ab - p < a$, $ab - p$ must be a quadratic residue. Therefore

$$1 = \left(\frac{ab-p}{p}\right) = \left(\frac{a}{p}\right)\left(\frac{a}{p}\right) = -\left(\frac{b}{p}\right).$$

Thus b is a quadratic nonresidue and hence $a \le b < \frac{p}{a} + 1$, which implies the statement.

Theorem 5. *For every prime number* $p > 2$, $\left(\frac{-1}{p}\right) = (-1)^{\frac{p-1}{2}}$.

In other words, the congruence $x^2 \equiv -1$ modulo a prime p is solvable if and only if $p = 2$ or $p \equiv 1 \pmod{4}$. △

Problem 2. *If* p *is a prime of the form* $4k+1$, *prove that* $x = (p')!$ *is a solution of the congruence* $x^2 + 1 \equiv 0$ *(mod p).*

Solution. Multiplying the congruences $i \equiv -(p-i) \pmod{p}$ for $i = 1, 2, \ldots, p'$ yields $(p')! \equiv (-1)^{p'}(p'+1)\cdots(p-2)(p-1)$. Note that p' is even by the condition of the problem. We now have

$$x^2 = (p')!^2 \equiv (-1)^{p'} p' \cdot (p'+1)\cdots(p-2)(p-1) = (-1)^{p'}(p-1)! \equiv (-1)^{p'+1} = -1 \pmod{p}$$

by Wilson's theorem. △

One can conclude from Problem 1 that every prime factor of number $x^2 + y^2$ (where $x, y \in \mathbb{N}$ are coprime) is either of the form $4k+1$, $k \in \mathbb{N}$, or equal to 2. This conclusion can in fact be generalized.

Theorem 6. *Let* x, y *be coprime integers and* a, b, c *be arbitrary integers. If* p *is an odd prime divisor of number* $ax^2 + bxy + cy^2$ *which doesn't divide abc, then*

$$D = b^2 - 4ac$$

is a quadratic residue modulo p.

In particular, if $p \mid x^2 - Dy^2$ *and* $(x, y) = 1$, *then D is a quadratic residue (mod p).*

Proof. Denote $N = ax^2 + bxy + cy^2$. Since $4aN = (2ax + by)^2 - Dy^2$, we have

$$(2ax + by)^2 \equiv Dy^2 \pmod{p}.$$

Furthermore, y is not divisible by p; otherwise so would be $2ax+by$ and therefore x itself, contradicting the assumption.

There is an integer y_1 such that $yy_1 \equiv 1 \pmod p$. Multiplying the above congruence by y_1^2 gives us $(2axy_1+byy_1)^2 \equiv D(yy_1)^2 \equiv D \pmod p$, implying the statement. □

For an integer a, $p \nmid a$ and $k = 1,2,\ldots,p'$ there is a unique $r_k \in \{-p',\ldots,-2,-1,1,2,\ldots,p'\}$ such that $ka \equiv r_k \pmod p$. Moreover, no two of the r_k's can be equal in absolute value; hence $|r_1|,|r_2|,\ldots,|r_{p'}|$ is in fact a permutation of $\{1,2,\ldots,p'\}$. Then

$$a^{p'} = \frac{a\cdot 2a\cdot\cdots\cdot p'a}{1\cdot 2\cdot\cdots\cdot p'} \equiv \frac{r_1 r_2\ldots r_{p'}}{1\cdot 2\cdot\cdots\cdot p'}.$$

Now, setting $r_k = \varepsilon_k|r_k|$ for $k = 1,\ldots,p'$, where $\varepsilon_k = \pm 1$, and applying Euler's criterion we obtain:

Theorem 7. $\left(\dfrac{a}{p}\right) = \varepsilon_1\varepsilon_2\cdots\varepsilon_{p'}$. □

Observe that $r_k = -1$ if and only if the remainder of ka upon division by p is greater than p', i.e. if and only if $\left[\frac{2ka}{p}\right] = 2\left[\frac{ka}{p}\right]+1$. Therefore, $r_k = (-1)^{\left[\frac{2ka}{p}\right]}$. Now Theorem 7 implies the following statement.

Theorem 8 (Gauss' Lemma). $\left(\dfrac{a}{p}\right) = (-1)^S$, *where* $S = \displaystyle\sum_{k=1}^{p'}\left[\frac{2ka}{p}\right]$. □

Gauss' lemma enables us to easily compute the value of Legendre's symbol $\left(\frac{a}{p}\right)$ for small a or small p. If, for instance, $a = 2$, we have $\left(\frac{2}{p}\right) = (-1)^S$, where $S = \sum_{k=1}^{p'}\left[\frac{4k}{p}\right]$. Exactly $\left[\frac{1}{2}p'\right]$ summands in this sum are equal to 0, while the remaining $p' - \left[\frac{1}{2}p'\right]$ are equal to 1. Therefore $S = p' - \left[\frac{1}{2}p'\right] = \left[\frac{p+1}{4}\right]$, which is even for $p \equiv \pm 1$ and odd for $p \equiv \pm 3 \pmod 8$. We have proven the following

Theorem 9. $\left(\dfrac{2}{p}\right) = (-1)^{\left[\frac{p+1}{4}\right]}$.

In other words, 2 is a quadratic residue modulo a prime $p > 2$ if and only if $p \equiv \pm 1$ (mod 8).

The following statements can be similarly shown.

Theorem 10. *(a) -2 is a quadratic residue modulo p if and only if $p \equiv 1$ or $p \equiv 3$ (mod 8);*

(b) -3 is a quadratic residue modulo p if and only if $p \equiv 1$ (mod 6);

(c) 3 je quadratic residue modulo p if and only if $p \equiv \pm 1$ (mod 12);

(d) 5 is a quadratic residue modulo p if and only if $p \equiv \pm 1$ (mod 10). □

Problem 3. *Show that there exist infinitely many prime numbers of the form (a) $4k+1$; (b) $10k+9$.*

Solution. (a) Suppose the contrary, that $p_1, p_2, \ldots, p_n$ are all such numbers. Then by Theorem 5, all prime divisors of $N = (2p_1p_2\cdots p_n)^2+1$ are of the form $4k+1$. However, N is not divisible by any of $p_1, p_2, \ldots, p_n$, which is impossible.

Part (b) is similar to (a), with number $N = 5(2p_1p_2\cdots p_n)^2 - 1$ being considered instead. △

Problem 4. *Prove that for $n \in \mathbb{N}$ every prime divisor p of number n^4-n^2+1 is of the form $12k+1$.*

Solution. We observe that

$$n^4-n^2+1 = (n^2-1)^2+n^2 \quad \text{i} \quad n^4-n^2+1 = (n^2+1)^2-3n^2.$$

In view of theorems 5, 6, and 10, the first equality gives us $p \equiv 1 \pmod 4$, whereas the other one gives us $p \equiv \pm 1 \pmod{12}$. These two congruences together yield $p \equiv 1 \pmod{12}$. △

Problem 5. *Evaluate*

$$\left[\frac{1}{2003}\right]+\left[\frac{2}{2003}\right]+\left[\frac{2^2}{2003}\right]+\cdots+\left[\frac{2^{2001}}{2003}\right].$$

Solution. Note that 2003 is prime. It follows from Euler's criterion and Theorem 10 that $2^{1001}\equiv\left(\frac{2}{2003}\right)=-1 \pmod{2003}$. Therefore $2003 \mid 2^i(2^{1001}+1)=2^{1001+i}+2^i$; since 2^i and 2^{1001+i} are not multiples of 2003, we conclude that

$$\left[\frac{2^i}{2003}\right]+\left[\frac{2^{1001+i}}{2003}\right]=\frac{2^i+2^{1001+i}}{2003}-1.$$

Summing up these equalities for $i=0,1,\ldots,1000$ we obtain that the desired sum equals

$$\frac{1+2+2^2+\cdots+2^{2001}}{2003}-1001=\frac{2^{2002}-1}{2003}-1001. \ \triangle$$

The theory we have presented so far doesn't really facilitate the job if we need to find out whether, say, 814 is a quadratic residue modulo 2003. That will be done by the following theorem, which makes such a verification possible with the amount of work comparable to that of the Euclidean algorithm.

Theorem 11 (Gauss' Reciprocity Law). *For any different odd primes p and q,*

$$\left(\frac{p}{q}\right)\left(\frac{q}{p}\right)=(-1)^{p'q'},$$

where $p'=\frac{p-1}{2}$ and $q'=\frac{q-1}{2}$.

Proof. Define $S(p,q)=\sum_{k=1}^{q'}\left[\frac{kp}{q}\right]$. We start by proving the following auxiliary statement.

Lemma 1. $S(p,q)+S(q,p)=p'q'$.

Proof of the Lemma. Given $k\in\mathbb{N}$, we note that $\left[\frac{kp}{q}\right]$ is the number of integer points (k,l) in the coordinate plane with $0<l<kp/q$, i.e. such that $0<ql<kp$. It follows that the sum $S(p,q)$ equals the number of integer points (k,l) with $0<k<p'$ and $0<ql<kp$. Thus $S(p,q)$ is exactly the number of points with positive integer coordinates in the interior or on the boundary of the rectangle $ABCD$ that lie *below* the line AE, where $A(0,0)$, $B(p',0)$, $C(p',q')$, $D(0,q')$, $E(p,q)$.

Analogously, $S(q,p)$ is exactly the number of points with positive integer coordinates in the interior or on the boundary of the rectangle $ABCD$ that lie *above* the line AE. Since there are $p'q'$ integer points in total in this rectangle, none of which is on the line AE, it follows that $S(p,q)+S(q,p)=p'q'$. $\triangledown$

We now return to the proof of the theorem. We have

$$S(p+q,q)-S(p,q)=1+2+\cdots+p'=\frac{p^2-1}{8}.$$

Since Theorem 9 is equivalent to $\left(\frac{2}{p}\right)=(-1)^{\frac{p^2-1}{8}}$, Gauss' lemma gives us

$$\left(\frac{2}{q}\right)\left(\frac{p}{q}\right)=\left(\frac{2p}{q}\right)=\left(\frac{2(p+q)}{q}\right)=\left(\frac{\frac{p+q}{2}}{q}\right)=(-1)^{S(p+q,q)}=\left(\frac{2}{q}\right)(-1)^{S(p,q)},$$

hence $\left(\frac{p}{q}\right)=(-1)^{S(p,q)}$. Analogously, $\left(\frac{q}{p}\right)=(-1)^{S(q,p)}$. Multiplying the last two inequalities and using the lemma yields the desired equality. $\square$

Let us now do the example mentioned before the Reciprocity Law.

Example 3. $\left(\frac{814}{2003}\right) = \left(\frac{2}{2003}\right)\left(\frac{11}{2003}\right)\left(\frac{37}{2003}\right) = -\left(\frac{11}{2003}\right)\left(\frac{37}{2003}\right)$.

Furthermore, the Reciprocity Law gives us

$$\left(\frac{11}{2003}\right) = -\left(\frac{2003}{11}\right) = \left(\frac{1}{11}\right) = 1 \quad \textit{and} \quad \left(\frac{37}{2003}\right) = \left(\frac{2003}{37}\right) = \left(\frac{5}{37}\right) = \left(\frac{37}{5}\right) = -1.$$

Thus $\left(\frac{814}{2003}\right) = 1$, *i.e.* 814 *is a quadratic residue modulo* 2003.

Problem 6. *Prove that an integer a is a quadratic residue modulo every prime number if and only if a is a perfect square.*

Solution. Suppose that a is not a square. We may assume w.l.o.g. (why?) that a is square-free.

Suppose that $a > 0$. Then $a = p_1 p_2 \cdots p_k$ for some primes $p_1, \ldots, p_k$. For every prime number p it holds that

$$\left(\frac{a}{p}\right) = \prod_{i=1}^{k}\left(\frac{p_i}{p}\right) \quad \text{and} \quad \left(\frac{p_i}{p}\right) = (-1)^{p_i' p'}\left(\frac{p}{p_i}\right). \tag{1}$$

If $a = 2$, it is enough to choose $p = 5$. Otherwise a has an odd prime divisor, say p_k. We choose a prime number p such that $p \equiv 1 \pmod 8$, $p \equiv 1 \pmod{p_i}$ for $i = 1, 2, \ldots, k-1$, and $p \equiv a \pmod{p_k}$, where a is an arbitrary quadratic nonresidue modulo p_k. Such prime number p exists according to the Dirichlet theorem on primes in an arithmetic progression. Then it follows from (1) that $p_1, \ldots, p_{k-1}$ are quadratic residues modulo p, but p_k is not. Therefore a is a quadraic nonresidue modulo p.

The proof in the case $a < 0$ is similar and is left to the reader. △

2 Quadratic Congruences to Composite Moduli

Not all moduli are prime, so we do not want to be restricted to prime moduli. The above theory can be generalized to composite moduli, yet losing as little as possible. The following function generalizes Legendre's symbol to a certain extent.

Definition 3. *Let a be an integer and b an odd number, and let $b = p_1^{\alpha_1} p_2^{\alpha_2} \cdots p_r^{\alpha_r}$ be the factorization of b onto primes. Jakobi'a symbol $\left(\frac{a}{b}\right)$ is defined as*

$$\left(\frac{a}{b}\right) = \left(\frac{a}{p_1}\right)^{\alpha_1}\left(\frac{a}{p_2}\right)^{\alpha_2}\cdots\left(\frac{a}{p_r}\right)^{\alpha_r}.$$

Since there is no danger of confusion, Jacobi's and Legendre's symbol share the notation.

It is easy to see that $\left(\frac{a}{b}\right) = -1$ implies that a is a quadratic nonresidue modulo b. Indeed, if $\left(\frac{a}{b}\right) = -1$, then by the definition $\left(\frac{a}{p_i}\right) = -1$ for at least one $p_i \mid b$; hence a is a quadratic nonresidue modulo p_i.

However, the converse is *false*, as seen from the following example.

Example 4. *Although*

$$\left(\frac{2}{15}\right) = \left(\frac{2}{3}\right)\left(\frac{2}{5}\right) = (-1)\cdot(-1) = 1,$$

2 is not a quadratic residue modulo 15, as it is not so modulo 3 and 5.

In fact, the following weaker statement holds.

Theorem 12. *Let a be an integer and b a positive integer, and let $b = p_1^{\alpha_1} p_2^{\alpha_2} \cdots p_r^{\alpha_r}$ be the faktorization of b onto primes. Then a is a quadratic residue modulo b if and only if a is a quadratic residue modulo $p_i^{\alpha_i}$ for each $i = 1, 2, \ldots, r$.*

Proof. If a quadratic residue modulo b, it is clearly so modulo each $p_i^{\alpha_i}$, $i = 1,2,\ldots,r$.

Assume that a is a quadratic residue modulo each $p_i^{\alpha_i}$ and that x_i is an integer such that $x_i^2 \equiv a \pmod{p_i^{\alpha_i}}$. According to Chinese Remainder Theorem there is an x such that $x \equiv x_i \pmod{p_i^{\alpha_i}}$ for $i = 1,2,\ldots,r$. Then $x^2 \equiv x_i^2 \equiv a \pmod{p_i^{\alpha_i}}$ for each i, and therefore $x^2 \equiv a \pmod{b}$. □

Theorem 13. *The number of quadratic residues modulo p^n ($n > 0$) is equal to*

$$\left[\frac{2^{n-1}-1}{3}\right]+2 \text{ for } p=2, \quad \text{and} \quad \left[\frac{p^{n+1}-1}{2(p+1)}\right]+1 \text{ for } p>2.$$

Proof. Let k_n denote the number of quadratic residues modulo p^n.

Let p be odd and $n \geq 2$. Number a is a quadratic residue modulo p^n if and only if either $p \nmid a$ and a is a quadratic residue modulo p, or $p^2 \mid a$ and a/p^2 is a quadratic residue modulo p^{n-2}. It follows that $k_n = k_{n-2} + p'p^{n-1}$.

Let $p = 2$ and $n \geq 3$. Number a is a quadratic residue modulo 2^n if and only if either $a \equiv 1 \pmod 8$ or $4 \mid a$ and $a/4$ is a quadratic residue modulo 2^{n-2}. We obtain $k_n = k_{n-2} + 2^{n-3}$.

Now the statement is shown by simple induction on n. □

Many properties of Legendre's symbols apply for Jacobi's symbols also. Thus the following statements hold can be easily proved by using the definition of Jacobi's symbol and the analogous statements for Legendre's symbols.

Theorem 14. *For all integers a,b and odd numbers c,d the following equalities hold:*

$$\left(\frac{a+bc}{c}\right)=\left(\frac{a}{c}\right), \quad \left(\frac{ab}{c}\right)=\left(\frac{a}{c}\right)\left(\frac{b}{c}\right), \quad \left(\frac{a}{cd}\right)=\left(\frac{a}{c}\right)\left(\frac{a}{d}\right). \ \square$$

Theorem 15. *For every odd integer a,*

$$\left(\frac{-1}{a}\right)=(-1)^{\frac{a-1}{2}}, \quad \left(\frac{2}{a}\right)=(-1)^{\left[\frac{a+1}{4}\right]}. \ \square$$

Theorem 16 (The Reciprocity Rule). *For any two coprime odd numbers a,b it holds that*

$$\left(\frac{a}{b}\right)\left(\frac{b}{a}\right)=(-1)^{\frac{a-1}{2}\cdot\frac{b-1}{2}}. \ \square$$

Problem 7. *Prove that the equation $x^2 = y^3 - 5$ has no integer solutions (x,y).*

Solution. For even y we have $x^2 = y^3 - 5 \equiv 3 \pmod 8$, which is impossible.

Now let y be odd. If $y \equiv 3 \pmod 4$, then $x^2 = y^3 - 5 \equiv 3^3 - 5 \equiv 2 \pmod 4$, impossible again. Hence y must be of the form $4z+1$, $z \in \mathbb{Z}$. Now the given equation transforms into

$$x^2 + 4 = 64z^3 + 48z^2 + 12z = 4z(16z^2 + 12z + 3).$$

It follows that $x^2 \equiv 4 \pmod{16z^2+12z+3}$.

However, the value of Jacobi's symbol

$$\left(\frac{-4}{16z^2+12z+3}\right)=\left(\frac{-1}{16z^2+12z+3}\right)$$

equals -1 because $16z^2 + 12z + 3 \equiv 3 \pmod 4$. Contradiction. △

Problem 8. *Prove that $4kxy - 1$ does not divide the number $x^m + y^n$ for any positive integers x,y,k,m,n.*

Solution. Note that $(x^m, y^n, 4kxy - 1) = 1$. Let us write $m' = [m/2]$ and $n' = [n/2]$. We need to investigate the following cases.

1° $m = 2m'$ and $n = 2n'$. Then $4kxy-1 \mid (x^{m'})^2 + (y^{n'})^2$ by Theorem 6 implies $\left(\frac{-1}{4kxy-1}\right) = 1$, which is false.

2° $m = 2m'$ and $n = 2n'+1$ (the case $m = 2m'+1$, $n = 2n'$ is analogous). Then $4kxy-1 \mid (x^{m'})^2 + y(y^{n'})^2$ and hence $\left(\frac{-y}{4kxy-1}\right) = 1$. We claim this to be impossible.

Suppose that y is odd. The Reciprocity Rule gives us

$$\left(\frac{-y}{4kxy-1}\right) = \left(\frac{-1}{4kxy-1}\right)\left(\frac{y}{4kxy-1}\right) = (-1)\cdot(-1)^{\frac{y-1}{2}}\left(\frac{-1}{y}\right) = -1.$$

Now assume that $y = 2^t y_1$, where $t \geq 1$ is an integer and $y_1 \in \mathbb{N}$. According to Theorem 15, we have $\left(\frac{2}{4kxy-1}\right) = 1$, whereas, like in the case of odd y, $\left(\frac{-y_1}{4kxy-1}\right) = \left(\frac{-y_1}{4\cdot 2^t kxy_1-1}\right) = -1$. It follows that

$$\left(\frac{-y}{4kxy-1}\right) = \left(\frac{2}{4kxy-1}\right)^t\left(\frac{-y_1}{4kxy-1}\right) = -1.$$

3° $m = 2m'+1$ and $n = 2n'+1$. Then $4kxy-1 \mid x(x^{m'})^2 + y(y^{n'})^2$, and hence $\left(\frac{-xy}{4kxy-1}\right) = 1$. On the other hand,

$$\left(\frac{-xy}{4kxy-1}\right) = \left(\frac{-4xy}{4kxy-1}\right) = \left(\frac{-1}{4kxy-1}\right) = -1,$$

a contradiction.

This finishes the proof. △

3 Some Sums of Legendre's symbols

Finding the number of solutions of a certain conguence is often reduced to counting the values of $x \in \{0,1,\ldots,p-1\}$ for which a given polynomial $f(x)$ with integer coefficients is a quadratic residue modulo an odd prime p. The answer is obviously directly connected to the value of the sum

$$\sum_{x=0}^{p-1}\left(\frac{f(x)}{p}\right).$$

In this part we are interested in sums of this type.

For a linear polynomial f, the considered sum is easily evaluated:

Theorem 17. *For arbitrary integers a, b and a prime $p \nmid a$,*

$$\sum_{x=0}^{p-1}\left(\frac{ax+b}{p}\right) = 0.$$

Proof. Since $p \nmid a$, the numbers $ax+b$, $x = 0,1,\ldots,p-1$ form a complete system of residues modulo p. Exactly $\frac{p-1}{2}$ of them are quadratic residues, exactly $\frac{p-1}{2}$ are quadratic nonresidues, and one is divisible by p. It follows that

$$\sum_{x=0}^{p-1}\left(\frac{ax+b}{p}\right) = \frac{p-1}{2}\cdot 1 + \frac{p-1}{2}\cdot(-1) + 0 = 0. \ \square$$

To evaluate the desired sum for quadratic polynomials f, we shall use the following proposition.

Theorem 18. *Let $f(x)^{p'} = a_0 + a_1x + \cdots + a_{kp'}x^{kp'}$, where k is the degree of polynomial f. We have*

$$\sum_{x=0}^{p-1}\left(\frac{f(x)}{p}\right) \equiv -(a_{p-1} + a_{2(p-1)} + \cdots + a_{k'(p-1)}) \ (mod\ p), \quad \text{where } k' = \left[\frac{k}{2}\right].$$

Proof. Define $S_n = \sum_{x=0}^{p-1} x^n$ ($n \in \mathbb{N}$) and $S_0 = p$. It can be shown that $S_n \equiv -1 \pmod p$ for $n > 0$ and $p-1 \mid n$, and $S_n \equiv 0 \pmod p$ otherwise. Now Euler's Criterion gives us

$$\sum_{x=0}^{p-1}\left(\frac{f(x)}{p}\right) \equiv \sum_{x=0}^{p-1} f(x)^{p'} = \sum_{i=0}^{kp'} a_i S_i \equiv -(a_{p-1} + a_{2(p-1)} + \cdots + a_{k'p-1}) \pmod p. \ \square$$

Theorem 19. *For any integers a, b, c and a prime $p \nmid a$, the sum*

$$\sum_{x=0}^{p-1}\left(\frac{ax^2+bx+c}{p}\right)$$

equals $-\left(\frac{a}{p}\right)$ if $p \nmid b^2-4ac$, and $(p-1)\left(\frac{a}{p}\right)$ if $p \mid b^2-4ac$.

Proof. We have

$$\left(\frac{4a}{p}\right)\sum_{x=0}^{p-1}\left(\frac{ax^2+bx+c}{p}\right) = \sum_{x=0}^{p-1}\left(\frac{(2ax+b)^2-D}{p}\right),$$

where $D = b^2-4ac$. Since numbers $ax+b$, $x = 0, 1, \ldots, p-1$ comprise a complete system of residues modulo p, we obtain

$$\left(\frac{a}{p}\right)\sum_{x=0}^{p-1}\left(\frac{ax^2+bx+c}{p}\right) = \sum_{x=0}^{p-1}\left(\frac{x^2-D}{p}\right) = S.$$

Theorem 18 gives us $S \equiv -1 \pmod p$, which together with $|S| \le p$ yields $S = -1$ or $S = p-1$.

Suppose that $S = p-1$. Then $p-1$ of the numbers $\left(\frac{x^2-D}{p}\right)$ are equal to 1, and exactly one, say for $x = x_0$, is equal to 0, i.e. $p \mid x_0^2 - D$. Since this implies $p \mid (-x_0)^2 - D = x_0^2 - p$ also, we must have $x_0 = 0$ and consequently $p \mid D$. Conversely, if $p \mid D$, we have $S = p-1$; otherwise $S = -1$, which finishes the proof. $\square$

Problem 9. *The number of solutions (x, y) of congruence*

$$x^2 - y^2 = D \ (mod\ p),$$

where $D \not\equiv 0$ (mod p) is given, equals $p-1$.

Solution. This is an immediate consequence of the fact that, for fixed x, the number of solutions y of the congruence $y^2 \equiv x^2 - D \pmod p$ equals $\left(\frac{x^2-D}{p}\right)+1$. $\triangle$

Evaluating the sums of Legendre's symbols for polynomials $f(x)$ of degree greater than 2 is significantly more difficult. In what follows we investigate the case of cubic polynomials f of a certain type.

For an integer a, define

$$K(a) = \sum_{x=0}^{p-1}\left(\frac{x(x^2+a)}{p}\right).$$

Assume that $p \nmid a$. We easily deduce that for each $t \in \mathbb{Z}$,

$$K(at^2) = \left(\frac{t}{p}\right)\sum_{x=0}^{p-1}\left(\frac{\frac{x}{t}((\frac{x}{t})^2+a)}{p}\right) = \left(\frac{t}{p}\right)K(a).$$

Therefore $|K(a)|$ depends only on whether a is a quadratic residue modulo p or not.

Now we give one non-standard proof of the fact that every prime $p \equiv 1 \pmod 4$ is a sum of two squares.

Theorem 20 (Jacobstal's identity). *Let a and b be a quadratic residue and nonresidue modulo a prime number p of the form $4k+1$. Then $|K(a)|$ and $|K(b)|$ are even positive integers that satisfy*

$$\left(\frac{1}{2}|K(a)|\right)^2 + \left(\frac{1}{2}|K(b)|\right)^2 = p.$$

Proof. The previous consideration gives us $p'(K(a)^2 + K(b)^2) = \sum_{n=1}^{p-1} K(n)^2 = \sum_{n=0}^{p-1} K(n)^2$, since $K(0) = 0$. Let us determine $\sum_{n=0}^{p-1} K(n)^2$. For each n we have

$$K(n)^2 = \sum_{x=0}^{p-1}\sum_{y=0}^{p-1}\left(\frac{xy(x^2+n)(y^2+n)}{p}\right),$$

which implies

$$\sum_{n=0}^{p-1} K(n)^2 = \sum_{x=0}^{p-1}\sum_{y=0}^{p-1}\left(\frac{xy}{p}\right)\sum_{n=0}^{p-1}\left(\frac{(n+x^2)(n+y^2)}{p}\right).$$

Note that by the theorem 19, $\sum_{n=0}^{p-1}\left(\frac{(n+x^2)(n+y^2)}{p}\right)$ equals $p-1$ if $x = \pm y$, and -1 otherwise. Upon substituting these values the above equality becomes

$$\sum_{n=0}^{p-1} K(n)^2 = p(2p-2) - \sum_{x=0}^{p-1}\sum_{y=0}^{p-1}\left(\frac{xy}{p}\right) = 4pp'.$$

We conclude that $K(a)^2 + K(b)^2 = 4p$. Furthermore, since $K(a)^2 + K(b)^2$ is divisible by 4, both $K(a)$ and $K(b)$ must be even, and the statement follows. □

4 Problems

10. Let p be a prime number. Prove that there exists $x \in \mathbb{Z}$ for which $p \mid x^2 - x + 3$ if and only if there exists $y \in \mathbb{Z}$ for which $p \mid y^2 - y + 25$.

11. Let $p = 4k-1$ be a prime number, $k \in \mathbb{N}$. Show that if a is an integer such that the congruence $x^2 \equiv a \pmod p$ has a solution, then its solutions are given by $x = \pm a^k$.

12. Show that all odd divisors of number $5x^2 + 1$ have an even tens digit.

13. Show that for every prime number p there exist integers a, b such that $a^2 + b^2 + 1$ is a multiple of p.

14. Prove that $\frac{x^2+1}{y^2-5}$ is not an integer for any integers $x, y > 2$.

15. Let $p > 3$ be a prime and let $a, b \in \mathbb{N}$ be such that

$$1 + \frac{1}{2} + \cdots + \frac{1}{p-1} = \frac{a}{b}.$$

Prove that $p^2 \mid a$.

16. Consider $P(x) = x^3 + 14x^2 - 2x + 1$. Show that there exists a natural number n such that for each $x \in \mathbb{Z}$,

$$101 \mid \underbrace{P(P(\ldots P}_{n}(x)\ldots)) - x.$$

17. Determine all $n \in \mathbb{N}$ such that the set $A = \{n, n+1, \ldots, n+1997\}$ can be partitioned into at least two subsets with equal products of elements.

18. (a) Prove that for no $x, y \in \mathbb{N}$ is $4xy - x - y$ a square;

    (b) Prove that for no $x, y, z \in \mathbb{N}$ is $4xyz - x - y$ a square.

19. If $n \in \mathbb{N}$, show that all prime divisors of $n^8 - n^4 + 1$ are of the form $24k+1$, $k \in \mathbb{N}$.

20. Suppose that m, n are positive integers such that $\varphi(5^m - 1) = 5^n - 1$. Prove that $(m, n) > 1$.

21. Prove that there are no positive integers a, b, c for which

$$\frac{a^2 + b^2 + c^2}{3(ab + bc + ca)}$$

    is an integer.

22. Prove that, for all $a \in \mathbb{Z}$, the number of solutions (x, y, z) of the congruence

$$x^2 + y^2 + z^2 \equiv 2axyz \pmod{p}$$

    equals $\left(p + (-1)^{p'}\right)^2$.

5 Solutions

10. The statement is trivial for $p \leq 3$, so we can assume that $p \geq 5$.

    Since $p \mid x^2 - x + 3$ is equivalent to $p \mid 4(x^2 - x + 3) = (2x-1)^2 + 11$, integer x exists if and only if -11 is a quadratic residue modulo p. Likewise, since $4(y^2 - y + 25) = (2y-1)^2 + 99$, y exists if and only if -99 is a quadratic residue modulo p. Now the statement of the problem follows from

$$\left(\frac{-11}{p}\right) = \left(\frac{-11 \cdot 3^2}{p}\right) = \left(\frac{-99}{p}\right).$$

11. According to Euler's criterion, the existence of a solution of $x^2 \equiv a \pmod{p}$ implies $a^{2k-1} \equiv 1 \pmod{p}$. Hence for $x = a^k$ we have $x^2 \equiv a^{2k} \equiv a \pmod{p}$.

12. If $p \mid 5x^2 + 1$, then $\left(\frac{-5}{p}\right) = 1$. The Reciprocity rule gives us

$$\left(\frac{-5}{p}\right) = \left(\frac{-1}{p}\right)\left(\frac{5}{p}\right) = (-1)^{\frac{p-1}{2}}\left(\frac{p}{5}\right).$$

    It is easy to verify that the last expression has the value 1 if and only if p is congruent to $1, 3, 7$ or 9 modulo 20.

13. Clearly, $p \mid a^2 + b^2 + 1$ if and only if $a^2 \equiv -b^2 - 1 \pmod{p}$.

    Both sets $\{a^2 \mid a \in \mathbb{Z}\}$ and $\{-b^2 - 1 \mid b \in \mathbb{Z}\}$ modulo p are of cardinality exactly $\frac{p+1}{2}$, so they have an element in common, i.e. there are $a, b \in \mathbb{Z}$ with a^2 and $-b^2 - 1$ being equal modulo p.

14. If y is even, $y^2 - 5$ is of the form $4k+3$, $k \in \mathbb{Z}$ and thus cannot divide $x^2 + 1$ for $x \in \mathbb{Z}$. If y is odd, then $y^2 - 5$ is divisible by 4, while $x^2 + 1$ is never a multiple of 4.

15. It suffices to show that $\frac{2(p-1)!a}{b} = \sum_{i=1}^{p-1} \frac{2(p-1)!}{i}$ is divisible by p^2. To start with,

$$\frac{2(p-1)!a}{b} = \sum_{i=1}^{p-1} \left(\frac{(p-1)!}{i} + \frac{(p-1)!}{p-i} \right) = \sum_{i=1}^{p-1} \frac{p(p-1)!}{i(p-i)}.$$

Therefore, $p \mid a$. Moreover, if for $i \in \{1,2,\ldots,p-1\}$ i' denotes the inverse of i modulo p, we have

$$\frac{2(p-1)!a}{pb} = \sum_{i=1}^{p-1} \frac{(p-1)!}{i(p-i)} \equiv \sum_{i=1}^{p-1} i'^2 (p-1)! \equiv 0 \pmod{p}.$$

It follows that $p^2 \mid 2(p-1)!a$.

16. All congruences in the solution will be modulo 101.

It is clear that $P(x) \equiv P(y)$ for integers x, y with $x \equiv y$.

We claim that the converse holds: $P(x) \not\equiv P(y)$ if $x \not\equiv y$. We have

$$\frac{4[P(x)-P(y)]}{x-y} = 4(x^2+xy+y^2+14x+14y-2) \equiv (2x+y+14)^2 + 3(y-29)^2.$$

Since -3 is not a quadratic residue modulo 101, the left hand side is not divisible by 101 unless if $2x+y+14 \equiv y-29 \equiv 0$, i.e. $x \equiv y \equiv 29$. This justifies our claim.

We now return to the problem. The above statement implies that $P(0), P(1), \ldots, P(100)$ is a permutation of $0, 1, \ldots, 100$ modulo 101. We conclude that for each $x \in \{0, 1, \ldots, 100\}$ there is an n_x such that $P(P(\ldots P(x)\ldots)) \equiv x$ (with P applied n_x times).

Any common multiple of the numbers $n_0, n_1, \ldots, n_{100}$ is clearly a desired n.

17. Suppose that A can be partitioned into k subsets $A_1, \ldots, A_k$, each with the same product of elements m. Since at least one and at most two elements of A are divisible by the prime 1997, we have $1997 \mid m$ and hence $k = 2$. Furthermore, since the number of elements divisible by the prime 1999 is at most one, we have $1999 \nmid m$; hence no elements of A is divisible by 1999, i.e. the elements of A are congruent to $1, 2, 3, \ldots, 1998$ modulo 1999. Then $m^2 \equiv 1 \cdot 2 \cdot 3 \cdots 1998 \equiv -1 \pmod{1999}$, which is impossible because -1 is a quadratic nonresidue modulo $1999 = 4 \cdot 499 + 3$.

18. Part (a) is a special case of (b).

(b) Suppose $x, y, z, t \in \mathbb{N}$ are such that $4xyz - x - y = t^2$. Multiplying this equation by $4z$ we obtain

$$(4xz-1)(4yz-1) = 4zt^2 + 1.$$

Therefore, $-4z$ is a quadratic residue modulo $4xz-1$. However, it was proved in problem 8 that the value of Legendre's symbol $\left(\frac{-z}{4xz-1}\right)$ is -1 for all x, z, yielding a contradiction.

19. Consider an arbitrary prime divisor p of $n^8 - n^4 + 1$. It follows from problem 4 that p is congruent to 1 or 13 (mod 24). Furthermore, since

$$n^8 - n^4 + 1 = (n^4 + n^2 + 1) - 2(n^3+n)^2,$$

2 is a quadratic residue modulo p, excluding the possibility $p \equiv \pm 13 \pmod{24}$.

20. Suppose that $(m, n) = 1$. Let

$$5^m - 1 = 2^\alpha p_1^{\alpha_1} \cdots p_k^{\alpha_k} \tag{1}$$

be the factorization of $5^m - 1$ onto primes, where $p_i > 2$ za $i = 1, \ldots, k$. By the condition of the problem,

$$5^n - 1 = \varphi(5^m - 1) = 2^{\alpha-1} p_1^{\alpha_1 - 1} \cdots p_k^{\alpha_k - 1} (p_1 - 1) \cdots (p_k - 1). \tag{2}$$

Obviously, $2^\alpha \mid 5^n-1$. On the other hand, it follows from $(5^m-1,5^n-1)=5^1-1=4$ that $\alpha_i=1$ for each $i=1,\ldots,k$ and $\alpha=2$. Since $2^3 \mid 5^x-1$ for every even x, m must be odd: $m=2m'+1$ for some $m'\in\mathbb{N}_0$.

Since $p_i \mid 5\cdot(5^{m'})^2-1$ for $i=1,\ldots,k$, 5 is a quadratic residue modulo p_i, and consequently $p_i\equiv\pm1 \pmod 5$. However, (2) implies that none of p_i-1 is divisible by 5. We thus obtain that $p_i\equiv -1 \pmod 5$ for all i.

Reduction of equality (1) modulo 5 yields $(-1)^k=1$. Thus k is even. On the other hand, equality (2) modulo 5 yields $(-2)^{k+1}\equiv 1 \pmod 5$, and therefore $k\equiv 3 \pmod 4$, contradicting the previous conclusion.

Remark. Most probably, m and n do not even exist.

21. Suppose that a,b,c,n are positive integers such that $a^2+b^2+c^2=3n(ab+bc+ca)$. This equality can be rewritten as

$$(a+b+c)^2=(3n+2)(ab+bc+ca).$$

Choose a prime number $p\equiv 2 \pmod 3$ which divides $3n+2$ with an odd exponent, i.e. such that $p^{2i-1} \mid 3n+2$ and $p^{2i}\nmid 3n+2$ for some $i\in\mathbb{N}$ (such p must exist). Then $p^i \mid a+b+c$ and therefore $p \mid ab+bc+ca$. Substituting $c\equiv -a-b \pmod p$ in the previous relation we obtain

$$p \mid a^2+ab+b^2 \quad\Rightarrow\quad p \mid (2a+b)^2+3b^2.$$

It follows that $\left(\frac{-3}{p}\right)=1$, which is false because $p\equiv 2 \pmod 3$.

22. The given congruence is equivalent to

$$(z-axy)^2\equiv (a^2x^2-1)y^2-x^2 \pmod p. \tag{1}$$

For any fixed $x,y\in\{0,\ldots,p-1\}$, the number of solutions z of (1) equals

$$1+\left(\frac{(a^2x^2-1)y^2-x^2}{p}\right).$$

Therefore the total number of solutions of (1) equals

$$N=p^2+\sum_{x=0}^{p-1}\sum_{y=0}^{p-1}\left(\frac{(a^2x^2-1)y^2-x^2}{p}\right).$$

According to theorem 19, $\sum_{y=0}^{p-1}\left(\frac{(a^2x^2-1)y^2-x^2}{p}\right)$ is equal to $-\left(\frac{a^2x^2-1}{p}\right)$ if $ax\not\equiv\pm1 \pmod p$, and to $p\left(\frac{-1}{p}\right)$ if $ax\equiv\pm1 \pmod p$. Therefore

$$N=p^2+2p\left(\frac{-1}{p}\right)-\sum_{x=0}^{p-1}\left(\frac{a^2x^2-1}{p}\right)=\left(p+\left(\frac{-1}{p}\right)\right)^2.$$

terug naar echt bestand

Lifting The Exponent Lemma (LTE)

Version 6 - Amir Hossein Parvardi

April 7, 2011

Lifting The Exponent Lemma is a powerful method for solving exponential Diophantine equations. It is pretty well-known in the Olympiad folklore (see, e.g., [3]) though its origins are hard to trace. Mathematically, it is a close relative of the classical Hensel's lemma (see [2]) in number theory (in both the statement and the idea of the proof). In this article we analyze this method and present some of its applications.

We can use the Lifting The Exponent Lemma (this is a long name, let's call it **LTE**!) in lots of problems involving exponential equations, especially when we have some prime numbers (and actually in some cases it "explodes" the problems). This lemma shows how to find the greatest power of a prime p – which is often ≥ 3 – that divides $a^n \pm b^n$ for some positive integers a and b. The proofs of theorems and lemmas in this article have nothing difficult and all of them use elementary mathematics. Understanding the theorem's usage and its meaning is more important to you than remembering its detailed proof.

I have to thank Fedja, darij grinberg(Darij Grinberg), makar and ZetaX(Daniel) for their notifications about the article. And I specially appreciate JBL(Joel) and Fedja helps about TeX issues.

1 Definitions and Notation

For two integers a and b we say a is divisible by b and write $b \mid a$ if and only if there exists some integer q such that $a = qb$.

We define $v_p(x)$ to be the greatest power in which a prime p divides x; in particular, if $v_p(x) = \alpha$ then $p^\alpha \mid x$ but $p^{\alpha+1} \nmid x$. We also write $p^\alpha \| x$, if and only if $v_p(x) = \alpha$. So we have $v_p(xy) = v_p(x) + v_p(y)$ and $v_p(x+y) \geq \min\{v_p(x), v_p(y)\}$.

Example. The greatest power of 3 that divides 63 is 3^2. because $3^2 = 9 \mid 63$ but $3^3 = 27 \nmid 63$. in particular, $3^2 \| 63$ or $v_3(63) = 2$.

Example. Clearly we see that if p and q are two different prime numbers, then $v_p(p^\alpha q^\beta) = \alpha$, or $p^\alpha \| p^\alpha q^\beta$.

Note. We have $v_p(0) = \infty$ for all primes p.

2 Two Important and Useful Lemmas

Lemma 1. *Let x and y be (not necessary positive) integers and let n be a positive integer. Given an arbitrary prime p (in particular, we can have $p = 2$) such that $\gcd(n, p) = 1$, $p \mid x - y$ and neither x, nor y is divisible by p (i.e., $p \nmid x$ and $p \nmid y$). We have*

$$v_p(x^n - y^n) = v_p(x - y).$$

Proof. We use the fact that

$$x^n - y^n = (x - y)(x^{n-1} + x^{n-2}y + x^{n-3}y^2 + \cdots + y^{n-1}).$$

Now if we show that $p \nmid x^{n-1} + x^{n-2}y + x^{n-3}y^2 + \cdots + y^{n-1}$, then we are done. In order to show this, we use the assumption $p \mid x - y$. So we have $x - y \equiv 0 \pmod{p}$, or $x \equiv y \pmod{p}$. Thus

$$\begin{aligned} x^{n-1}+x^{n-2}y + x^{n-3}y^2 + \cdots + y^{n-1} & \\ &\equiv x^{n-1} + x^{n-2} \cdot x + x^{n-3} \cdot x^2 + \cdots + x \cdot x^{n-2} + x^{n-1} \\ &\equiv nx^{n-1} \\ &\not\equiv 0 \pmod{p}. \end{aligned}$$

This completes the proof. □

Lemma 2. *Let x and y be (not necessary positive) integers and let n be an odd positive integer. Given an arbitrary prime p (in particular, we can have $p = 2$) such that $\gcd(n, p) = 1$, $p \mid x + y$ and neither x, nor y is divisible by p, we have*

$$v_p(x^n + y^n) = v_p(x + y).$$

Proof. Since x and y can be negative, using **Lemma 1** we obtain

$$v_p(x^n - (-y)^n) = v_p(x - (-y)) \implies v_p(x^n + y^n) = v_p(x + y).$$

Note that since n is an odd positive integer we can replace $(-y)^n$ with $-y^n$. □

3 Lifting The Exponent Lemma (LTE)

Theorem 1 (First Form of LTE)**.** *Let x and y be (not necessary positive) integers, let n be a positive integer, and let p be an odd prime such that $p \mid x - y$ and none of x and y is divisible by p (i.e., $p \nmid x$ and $p \nmid y$). We have*

$$v_p(x^n - y^n) = v_p(x - y) + v_p(n).$$

Proof. We may use induction on $v_p(n)$. First, let us prove the following statement:

$$v_p(x^p - y^p) = v_p(x - y) + 1. \tag{1}$$

In order to prove this, we will show that

$$p \mid x^{p-1} + x^{p-2}y + \cdots + xy^{p-2} + y^{p-1} \tag{2}$$

and

$$p^2 \nmid x^{p-1} + x^{p-2}y + \cdots + xy^{p-2} + y^{p-1}. \tag{3}$$

For **(2)**, we note that

$$x^{p-1} + x^{p-2}y + \cdots + xy^{p-2} + y^{p-1} \equiv px^{p-1} \equiv 0 \pmod{p}.$$

Now, let $y = x + kp$, where k is an integer. For an integer $1 \leq t < p$ we have

$$\begin{aligned} y^t x^{p-1-t} &\equiv (x+kp)^t x^{p-1-t} \\ &\equiv x^{p-1-t}\left(x^t + t(kp)(x^{t-1}) + \frac{t(t-1)}{2}(kp)^2(x^{t-2}) + \cdots\right) \\ &\equiv x^{p-1-t}\left(x^t + t(kp)(x^{t-1})\right) \\ &\equiv x^{p-1} + tkpx^{p-2} \pmod{p^2}. \end{aligned}$$

This means

$$y^t x^{p-1-t} \equiv x^{p-1} + tkpx^{p-2} \pmod{p^2}, \quad t = 1, 2, 3, 4, \ldots, p-1.$$

Using this fact, we have

$$\begin{aligned} &x^{p-1} + x^{p-2}y + \cdots + xy^{p-2} + y^{p-1} \\ &\quad\equiv x^{p-1} + (x^{p-1} + kpx^{p-2}) + (x^{p-1} + 2kpx^{p-2}) + \cdots + (x^{p-1} + (p-1)kpx^{p-2}) \\ &\quad\equiv px^{p-1} + (1 + 2 + \cdots + p - 1)kpx^{p-2} \\ &\quad\equiv px^{p-1} + \left(\frac{p(p-1)}{2}\right)kpx^{p-2} \\ &\quad\equiv px^{p-1} + \left(\frac{p-1}{2}\right)kp^2x^{p-1} \\ &\quad\equiv px^{p-1} \not\equiv 0 \pmod{p^2}. \end{aligned}$$

So we proved **(3)** and the proof of **(1)** is complete. Now let us return to our problem. We want to show that

$$v_p(x^n - y^n) = v_p(x-y) + v_p(n).$$

Suppose that $n = p^\alpha b$ where $\gcd(p, b) = 1$. Then

$$\begin{aligned} v_p(x^n - y^n) &= v_p((x^{p^\alpha})^b - (y^{p^\alpha})^b) \\ &= v_p(x^{p^\alpha} - y^{p^\alpha}) = v_p((x^{p^{\alpha-1}})^p - (y^{p^{\alpha-1}})^p) \\ &= v_p(x^{p^{\alpha-1}} - y^{p^{\alpha-1}}) + 1 = v_p((x^{p^{\alpha-2}})^p - (y^{p^{\alpha-2}})^p) + 1 \\ &= v_p(x^{p^{\alpha-2}} - y^{p^{\alpha-2}}) + 2 \\ &\;\;\vdots \\ &= v_p((x^{p^1})^1 - (y^{p^1})^1) + \alpha - 1 = v_p(x-y) + \alpha \\ &= v_p(x-y) + v_p(n). \end{aligned}$$

Note that we used the fact that if $p \mid x-y$, then we have $p \mid x^k - y^k$, because we have $x - y \mid x^k - y^k$ for all positive integers k. The proof is complete. □

Theorem 2 (Second Form of LTE)**.** *Let x, y be two integers, n be an odd positive integer, and p be an odd prime such that $p \mid x + y$ and none of x and y is divisible by p. We have*

$$v_p(x^n + y^n) = v_p(x + y) + v_p(n).$$

Proof. This is obvious using **Theorem 1**. See the trick we used in proof of **Lemma 2**. □

4 What about $p = 2$?

Question. Why did we assume that p is an odd prime, i.e., $p \neq 2$? Why can't we assume that $p = 2$ in our proofs?

Hint. Note that $\frac{p-1}{2}$ is an integer only for $p > 2$.

Theorem 3 (LTE for the case $p = 2$)**.** *Let x and y be two odd integers such that $4 \mid x - y$. Then*

$$v_2(x^n - y^n) = v_2(x - y) + v_2(n).$$

Proof. We showed that for any prime p such that $\gcd(p, n) = 1, p \mid x - y$ and none of x and y is divisible by p, we have

$$v_p(x^n - y^n) = v_p(x - y)$$

So it suffices to show that

$$v_2(x^{2^n} - y^{2^n}) = v_2(x - y) + n.$$

Factorization gives

$$x^{2^n} - y^{2^n} = (x^{2^{n-1}} + y^{2^{n-1}})(x^{2^{n-2}} + y^{2^{n-2}}) \cdots (x^2 + y^2)(x + y)(x - y)$$

Now since $x \equiv y \equiv \pm 1 \pmod 4$ then we have $x^{2^k} \equiv y^{2^k} \equiv 1 \pmod 4$ for all positive integers k and so $x^{2^k} + y^{2^k} \equiv 2 \pmod 4, k = 1, 2, 3, \ldots$. Also, since x and y are odd and $4 \mid x - y$, we have $x + y \equiv 2 \pmod 4$. This means the power of 2 in all of the factors in the above product (except $x - y$) is one. We are done. □

Theorem 4. *Let x and y be two odd integers and let n be an even positive integer. Then*

$$v_2(x^n - y^n) = v_2(x - y) + v_2(x + y) + v_2(n) - 1.$$

Proof. We know that the square of an odd integer is of the form $4k+1$. So for odd x and y we have $4 \mid x^2 - y^2$. Now let m be an odd integer and k be a positive integer such that $n = m \cdot 2^k$. Then

$$\begin{aligned} v_2(x^n - y^n) &= v_2(x^{m\cdot 2^k} - y^{m\cdot 2^k}) \\ &= v_2((x^2)^{2^{k-1}} - (y^2)^{2^{k-1}}) \\ &\vdots \\ &= v_2(x^2 - y^2) + k - 1 \\ &= v_2(x-y) + v_2(x+y) + v_2(n) - 1. \end{aligned}$$

□

5 Summary

Let p be a prime number and let x and y be two (not necessary positive) integers that are not divisible by p. Then:

a) For a positive integer n

- if $p \neq 2$ and $p \mid x - y$, then

$$v_p(x^n - y^n) = v_p(x-y) + v_p(n).$$

- if $p = 2$ and $4 \mid x - y$, then

$$v_2(x^n - y^n) = v_2(x-y) + v_2(n).$$

- if $p = 2$, n is even, and $2 \mid x - y$, then

$$v_2(x^n - y^n) = v_2(x-y) + v_2(x+y) + v_2(n) - 1.$$

b) For an odd positive integer n, if $p \mid x + y$, then

$$v_p(x^n + y^n) = v_p(x+y) + v_p(n).$$

c) For a positive integer n with $\gcd(p, n) = 1$, if $p \mid x - y$, we have

$$v_p(x^n - y^n) = v_p(x-y).$$

If n is odd, $\gcd(p, n) = 1$, and $p \mid x + y$, then we have

$$v_p(x^n + y^n) = v_p(x+y).$$

Note. The most common mistake in using LTE is when you don't check the $p \mid x \pm y$ condition, so always remember to check it. Otherwise your solution will be completely wrong.

6 Problems with Solutions

Problem 1 (Russia 1996)**.** *Find all positive integers n for which there exist positive integers x, y and k such that $\gcd(x, y) = 1, k > 1$ and $3^n = x^k + y^k$.*

Solution. k should be an odd integer (otherwise, if k is even, then x^k and y^k are perfect squares, and it is well known that for integers a, b we have $3 \mid a^2 + b^2$ if and only if $3 \mid a$ and $3 \mid b$, which is in contradiction with $\gcd(x, y) = 1$.). Suppose that there exists a prime p such that $p \mid x + y$. This prime should be odd. So $v_p(3^n) = v_p(x^k + y^k)$, and using **Theorem 2** we have $v_p(3^n) = v_p(x^k + y^k) = v_p(k) + v_p(x + y)$. But $p \mid x + y$ means that $v_p(x + y) \geq 1 > 0$ and so $v_p(3^n) = v_p(k) + v_p(x + y) > 0$ and so $p \mid 3^n$. Thus $p = 3$. This means $x + y = 3^m$ for some positive integer m. Note that $n = v_3(k) + m$. There are two cases:

- $m > 1$. We can prove by induction that $3^a \geq a + 2$ for all integers $a \geq 1$, and so we have $v_3(k) \leq k - 2$ (why?). Let $M = \max(x, y)$. Since $x + y = 3^m \geq 9$, we have $M \geq 5$. Then

$$x^k + y^k \geq M^k = \underbrace{M}_{\geq \frac{x+y}{2} = \frac{1}{2} \cdot 3^m} \cdot \underbrace{M^{k-1}}_{\geq 5^{k-1}} > \frac{1}{2} 3^m \cdot 5^{k-1}$$
$$> 3^m \cdot 5^{k-2} \geq 3^{m+k-2} \geq 3^{m+v_3(k)} = 3^n$$

  which is a contradiction.

- $m = 1$. Then $x + y = 3$, so $x = 1, y = 2$ (or $x = 2, y = 1$). Thus $3^{1+v_3(k)} = 1 + 2^k$. But note that $3^{v_3(k)} \mid k$ so $3^{v_3(k)} \leq k$. Thus

$$1 + 2^k = 3^{v_3(k)+1} = 3 \cdot \underbrace{3^{v_3(k)}}_{\leq k} \leq 3k \implies 2^k + 1 \leq 3k.$$

  And one can check that the only odd value of $k > 1$ that satisfies the above inequality is $k = 3$. So $(x, y, n, k) = (1, 2, 2, 3), (2, 1, 2, 3)$ in this case.

Thus, the final answer is $n = 2$.

Problem 2 (Balkan 1993)**.** *Let p be a prime number and $m > 1$ be a positive integer. Show that if for some positive integers $x > 1, y > 1$ we have*

$$\frac{x^p + y^p}{2} = \left(\frac{x + y}{2} \right)^m,$$

then $m = p$.

Solution. One can prove by induction on p that $\frac{x^p + y^p}{2} \geq \left(\frac{x+y}{2}\right)^p$ for all positive integers p. Now since $\frac{x^p + y^p}{2} = \left(\frac{x+y}{2}\right)^m$, we should have $m \geq p$. Let $d = \gcd(x, y)$, so there exist positive integers x_1, y_1 with $\gcd(x_1, y_1) = 1$ such that $x = dx_1$, $y = dy_1$ and $2^{m-1}(x_1^p + y_1^p) = d^{m-p}(x_1 + y_1)^m$. There are two cases:

Assume that p is odd. Take any prime divisor q of x_1+y_1 and let $v = v_q(x_1+y_1)$. If q is odd, we see that $v_q(x_1^p+y_1^p) = v+v_q(p)$ and $v_q(d^{m-p}(x_1+y_1)^m) \geq mv$ (because q may also be a factor of d). Thus $m \leq 2$ and $p \leq 2$, giving an immediate contradiction. If $q = 2$, then $m-1+v \geq mv$, so $v \leq 1$ and $x_1+y_1 = 2$, i.e., $x = y$, which immediately implies $m = p$.

Assume that $p = 2$. We notice that for $x+y \geq 4$ we have $\frac{x^2+y^2}{2} < 2\left(\frac{x+y}{2}\right)^2 \leq \left(\frac{x+y}{2}\right)^3$, so $m = 2$. It remains to check that the remaining cases $(x,y) = (1,2),(2,1)$ are impossible.

Problem 3. *Find all positive integers a, b that are greater than 1 and satisfy*

$$b^a | a^b - 1.$$

Solution. Let p be the least prime divisor of b. Let m be the least positive integer for which $p|a^m-1$. Then $m|b$ and $m \mid p-1$, so any prime divisor of m divides b and is less than p. Thus, not to run into a contradiction, we must have $m = 1$. Now, if p is odd, we have $av_p(b) \leq v_p(a-1)+v_p(b)$, so $a-1 \leq (a-1)v_p(b) \leq v_p(a-1)$, which is impossible. Thus $p = 2$, b is even, a is odd and $av_2(b) \leq v_2(a-1)+v_2(a+1)+v_2(b)-1$ whence $a \leq (a-1)v_2(b)+1 \leq v_2(a-1)+v_2(a+1)$, which is possible only if $a = 3$, $v_2(b) = 1$. Put $b = 2B$ with odd B and rewrite the condition as $2^3B^3 \mid 3^{2B}-1$. Let q be the least prime divisor of B (now, surely, odd). Let n be the least positive integer such that $q \mid 3^n-1$. Then $n \mid 2B$ and $n \mid q-1$ whence n must be 1 or 2 (or B has a smaller prime divisor), so $q \mid 3-1 = 2$ or $q \mid 3^2-1 = 8$, which is impossible. Thus $B = 1$ and $b = 2$.

Problem 4. *Find all positive integer solutions of the equation $x^{2009}+y^{2009} = 7^z$*

Solution. Factor 2009. We have $2009 = 7^2 \cdot 41$. Since $x+y \mid x^{2009}+y^{2009}$ and $x+y > 1$, we must have $7 \mid x+y$. Removing the highest possible power of 7 from x, y, we get $v_7(x^{2009}+y^{2009}) = v_7(x+y)+v_7(2009) = v_7(x+y)+2$, so $x^{2009}+y^{2009} = 49 \cdot k \cdot (x+y)$ where $7 \nmid k$. But we have $x^{2009}+y^{2009} = 7^z$, which means the only prime factor of $x^{2009}+y^{2009}$ is 7, so $k = 1$). Thus $x^{2009}+y^{2009} = 49(x+y)$. But in this equation the left hand side is much larger than the right hand one if $\max(x,y) > 1$, and, clearly, $(x,y) = (1,1)$ is not a solution. Thus the given equation does not have any solutions in the set of positive integers.

7 Challenge Problems

1. Let k be a positive integer. Find all positive integers n such that $3^k \mid 2^n - 1$.

2 (UNESCO Competition 1995)**.** Let a, n be two positive integers and let p be an odd prime number such that

$$a^p \equiv 1 \pmod{p^n}.$$

Prove that

$$a \equiv 1 \pmod{p^{n-1}}.$$

3 (Iran Second Round 2008)**.** Show that the only positive integer value of a for which $4(a^n + 1)$ is a perfect cube for all positive integers n, is 1.

4. Let $k > 1$ be an integer. Show that there exists infinitely many positive integers n such that

$$n | 1^n + 2^n + 3^n + \cdots + k^n.$$

5 (Ireland 1996)**.** Let p be a prime number, and a and n positive integers. Prove that if

$$2^p + 3^p = a^n$$

then $n = 1$.

6 (Russia 1996)**.** Let x, y, p, n, k be positive integers such that n is odd and p is an odd prime. Prove that if $x^n + y^n = p^k$, then n is a power of p.

7. Find the sum of all the divisors d of $N = 19^{88} - 1$ which are of the form $d = 2^a 3^b$ with $a, b \in \mathbb{N}$.

8. Let p be a prime number. Solve the equation $a^p - 1 = p^k$ in the set of positive integers.

9. Find all solutions of the equation

$$(n-1)! + 1 = n^m$$

in positive integers.

10 (Bulgaria 1997)**.** For some positive integer n, the number $3^n - 2^n$ is a perfect power of a prime. Prove that n is a prime.

11. Let m, n, b be three positive integers with $m \neq n$ and $b > 1$. Show that if prime divisors of the numbers $b^n - 1$ and $b^m - 1$ be the same, then $b + 1$ is a perfect power of 2.

12 (IMO ShortList 1991)**.** Find the highest degree k of 1991 for which 1991^k divides the number

$$1990^{1991^{1992}} + 1992^{1991^{1990}}.$$

13. Prove that the number $a^{a-1} - 1$ is never square-free for all integers $a > 2$.

14 (Czech Slovakia 1996)**.** Find all positive integers x, y such that $p^x - y^p = 1$, where p is a prime.

15. Let x and y be two positive rational numbers such that for infinitely many positive integers n, the number $x^n - y^n$ is a positive integer. Show that x and y are both positive integers.

16 (IMO 2000)**.** Does there exist a positive integer n such that n has exactly 2000 prime divisors and n divides $2^n + 1$?

17 (China Western Mathematical Olympiad 2010)**.** Suppose that m and k are non-negative integers, and $p = 2^{2^m} + 1$ is a prime number. Prove that

- $2^{2^{m+1}p^k} \equiv 1 \pmod{p^{k+1}}$;
- $2^{m+1}p^k$ is the smallest positive integer n satisfying the congruence equation $2^n \equiv 1 \pmod{p^{k+1}}$.

18. Let $p \geq 5$ be a prime. Find the maximum value of positive integer k such that

$$p^k | (p-2)^{2(p-1)} - (p-4)^{p-1}.$$

19. Let a, b be distinct real numbers such that the numbers

$$a - b,\ a^2 - b^2,\ a^3 - b^3, \ldots$$

are all integers. Prove that a, b are both integers.

20 (MOSP 2001)**.** Find all quadruples of positive integers (x, r, p, n) such that p is a prime number, $n, r > 1$ and $x^r - 1 = p^n$.

21 (China TST 2009)**.** Let $a > b > 1$ be positive integers and b be an odd number, let n be a positive integer. If $b^n \mid a^n - 1$, then show that $a^b > \frac{3^n}{n}$.

22 (Romanian Junior Balkan TST 2008)**.** Let p be a prime number, $p \neq 3$, and integers a, b such that $p \mid a + b$ and $p^2 \mid a^3 + b^3$. Prove that $p^2 \mid a + b$ or $p^3 \mid a^3 + b^3$.

23. Let m and n be positive integers. Prove that for each odd positive integer b there are infinitely many primes p such that $p^n \equiv 1 \pmod{b^m}$ implies $b^{m-1} \mid n$.

24 (IMO 1990)**.** Determine all integers $n > 1$ such that

$$\frac{2^n + 1}{n^2}$$

is an integer.

25. Find all positive integers n such that

$$\frac{2^{n-1} + 1}{n}.$$

is an integer.

26. Find all primes p, q such that $\dfrac{(5^p - 2^p)(5^q - 2^q)}{pq}$ is an integer.

27. For some natural number n let a be the greatest natural number for which $5^n - 3^n$ is divisible by 2^a. Also let b be the greatest natural number such that $2^b \leq n$. Prove that $a \leq b + 3$.

28. Determine all sets of non-negative integers x, y and z which satisfy the equation

$$2^x + 3^y = z^2.$$

29 (IMO ShortList 2007)**.** Find all surjective functions $f : \mathbb{N} \to \mathbb{N}$ such that for every $m, n \in \mathbb{N}$ and every prime p, the number $f(m+n)$ is divisible by p if and only if $f(m) + f(n)$ is divisible by p.

30 (Romania TST 1994)**.** Let n be an odd positive integer. Prove that $((n-1)^n + 1)^2$ divides $n(n-1)^{(n-1)^n+1} + n$.

31. Find all positive integers n such that $3^n - 1$ is divisible by 2^n.

32 (Romania TST 2009)**.** Let $a, n \geq 2$ be two integers, which have the following property: there exists an integer $k \geq 2$, such that n divides $(a-1)^k$. Prove that n also divides $a^{n-1} + a^{n-2} + \cdots + a + 1$.

33. Find all the positive integers a such that $\frac{5^a+1}{3^a}$ is a positive integer.

8 Hints and Answers to Selected Problems

1. Answer: $n = 2 \cdot 3^{k-1}s$ for some $s \in \mathbb{N}$.

2. Show that $v_p(a-1) = v_p(a^p - 1) - 1 \geq n - 1$.

3. If $a > 1$, a^2+1 is not a power of 2 (because it is > 2 and either 1 or 2 modulo 4). Choose some odd prime $p|a^2 + 1$. Now, take some $n = 2m$ with odd m and notice that $v_p(4(a^n + 1)) = v_p(a^2 + 1) + v_p(m)$ but $v_p(m)$ can be anything we want modulo 3.

5. $2^p + 3^p$ is not a square, and use the fact that $v_5(2^p + 3^p) = 1 + v_5(p) \leq 2$.

8. Consider two cases : $p = 2$ and p is an odd prime. The latter does not give any solutions.

9. $(n, m) = (2, 1)$ is a solution. In other cases, show that n is an odd prime and m is even. The other solution is $(n, m) = (5, 2)$.

12. Answer: $\max(k) = 1991$.

13. Take any odd prime p such that $p \mid a - 1$. It's clear that $p^2 \mid a^{a-1} - 1$.

14. Answer: $(p, x, y) = (2, 1, 1), (3, 2, 1)$.

18. Let $p - 1 = 2^s m$ and show that $v_p(2^{s-1}m) = 0$. The maximum of k is 1.

19. Try to prove Problem **15** first.

20. Show that $p = 2$ and r is an even positive integer.

22. If $p \mid a, p \mid b$, then $p^3 \mid a^3 + b^3$. Otherwise LTE applies and $v_p(a + b) = v_p(a^3 + b^3) \geq 2$.

24. The answer is $n = 1$ or $n = 3$.

26. Answer: $(p, q) = (3, 3), (3, 13)$.

27. If n is odd, then $a = 1$. If n is even, then $a = v_2(5^n - 3^n) = v_2(5 - 3) + v_2(5 + 3) + v_2(n) - 1 = 3 + v_2(n)$. But, clearly, $b \geq v_2(n)$.

30. $n \mid (n-1)^n + 1$, so for every $p \mid (n-1)^n + 1$, we have

$$\begin{aligned} v_p((n-1)^{(n-1)^n+1} + 1) &= v_p((n-1)^n + 1) + v_p\left(\frac{(n-1)^{n+1} + 1}{n}\right) \\ &= 2v_p((n-1)^n + 1) - v_p(n) \end{aligned}$$

which completes the proof.

31. $n \leq v_2(3^n - 1) \leq 3 + v_2(n)$, so $n \leq 4$.

33. a must be odd (otherwise the numerator is $2 \mod 3$). Then $a \leq v_3(5^a+1) = 1 + v_3(a)$ giving $a = 1$ as the only solution.

References

[1] Sepehr Ghazi Nezami, **Leme Do Khat** (in English: Lifting The Exponent Lemma) published on October 2009.

[2] Kurt Hensel, **Hensel's lemma**, WikiPedia.

[3] Santiago Cuellar, Jose Alejandro Samper, *A nice and tricky lemma (lifting the exponent)*, Mathematical Reflections **3** - 2007.

[4] Amir Hossein Parvardi, Fedja et al., AoPS **topic** #393335, *Lifting The Exponent Lemma (Containing PDF file).*

[5] Orlando Doehring et al., AoPS **topic** #214717, *Number* $\mod (f(m+n), p) = 0$ *iff* $\mod (f(m) + f(n), p) = 0$.

[6] Fang-jh et al., AoPS **topic** #268964, *China TST, Quiz 6, Problem 1.*

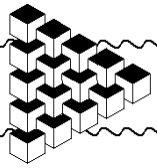
[7] Valentin Vornicu et al., AoPS **topic** #57607, *exactly 2000 prime divisors (IMO 2000 P5).*

[8] Orlando Doehring et al., AoPS **topic** #220915, *Highest degree for 3-layer power tower.*

[9] Sorush Oraki, Johan Gunardi, AoPS **topic** #368210, *Prove that* $a = 1$ *if* $4(a^n + 1)$ *is a cube for all* n.

terug naar echt bestand

The IMO Compendium Group
Olympiad Training Materials
www.imomath.com



©2007 The Author(s) and The IMO Compendium Group

Arithmetic in Extensions of $\mathbb{Q}$

Dušan Djukić

Contents

1 General Properties

What makes work with rational numbers and integers comfortable are the essential properties they have, especially the unique factorization property (the Main Theorem of Arithmetic). However, the might of the arithmetic in $\mathbb{Q}$ is bounded. Thus, some polynomials, although they have zeros, cannot be factorized into polynomials with rational coefficients. Nevertheless, such polynomials can always be factorized in a wider field. For instance, the polynomial x^2+1 is irreducible over $\mathbb{Z}$ or $\mathbb{Q}$, but over the ring of the so called *Gaussian integers* $\mathbb{Z}[i]=\{a+bi \mid a,b\in\mathbb{Z}\}$ it can be factorized as $(x+i)(x-i)$. Sometimes the wider field retains many properties of the rational numbers. In particular, it will turn out that the Gaussian integers are a unique factorization domain, just like the (rational) integers $\mathbb{Z}$. We shall first discuss some basics of higher algebra.

Definition 1. *A number $\alpha\in\mathbb{C}$ is* algebraic *if there is a polynomial $p(x)=a_nx^n+a_{n-1}x^{n-1}+\cdots+a_0$ with integer coefficients such that $p(\alpha)=0$. If $a_n=1$, then α is an* algebraic integer.

Further, $p(x)$ is the minimal polynomial *of α if it is irreducible over $\mathbb{Z}[x]$ (i.e. it cannot be written as a product of nonconstant polynomials with integer coefficients).*

Example 1. *The number i is an algebraic integer, as it is a root of the polynomial x^2+1 which is also its minimal polynomial. Number $\sqrt{2}+\sqrt{3}$ is also an algebraic integer with the minimal polynomial x^4-10x^2+1 (verify!).*

Example 2. *The minimal polynomial of a rational number $q=a/b$ ($a\in\mathbb{Z}$, $b\in\mathbb{N}$, $(a,b)=1$) is $bx-a$. By the definition, q is an algebraic integer if and only if $b=1$, i.e. if and only if q is an integer.*

Definition 2. *Let α be an algebraic integer and $p(x)=x^n+a_{n-1}x^{n-1}+\cdots+a_0$ ($a_i\in\mathbb{Z}$) be its minimal polynomial. The* extension *of a ring A by the element α is the set $A[\alpha]$ of all complex numbers of the form*

$$c_0+c_1\alpha+\cdots+c_{n-1}\alpha^{n-1}\ \ (c_i\in A), \qquad (*)$$

with all the operations inherited from A. The degree *of the extension is the degree n of the polynomial $p(x)$.*

The theme of this text are extensions of the ring $\mathbb{Z}$ of degree 2, so called *quadratic extensions*. Thus, for example, the polynomials x^2+1 and x^2+x+1 determine the extensions $\mathbb{Z}[i]$ and $\mathbb{Z}[\omega]$, where $\omega = \dfrac{-1+\sqrt{3}}{2}$ (this notation will be used later).

All elements of a quadratic extension of $\mathbb{Z}$ are algebraic integers with the minimal polynomial of second degree. Two elements having the same minimal polynomials are said to be *conjugates*. Each nonrational element z of the quadratic extension has exactly one conjugate, called the conjugate of z and denoted $\overline{z}$. For a rational integer z we define $\overline{z} = z$.

Definition 3. *The* norm *of an element z of a quadratic extension of $\mathbb{Z}$ is $N(z) = z\overline{z}$.*

The norm is always an integer. Roughly speaking, it is a kind of equivalent of the absolute value in the set of integers $\mathbb{Z}$.

Example 3. *If $z \in \mathbb{Z}[\sqrt{d}]$, $z = a+b\sqrt{d}$ $(a,b \in \mathbb{Z})$, then $\overline{z} = a-b\sqrt{d}$ and $N(z) = a^2-db^2$. In particular, in $\mathbb{Z}[i]$ the norm of element $a+bi$ $(a,b \in \mathbb{N})$ is $N(a+bi) = a^2+b^2$.*

If $z = a+b\omega \in \mathbb{Z}[\omega]$ $(a,b \in \mathbb{Z})$, then $\overline{z} = a-b-b\omega$ and $N(z) = a^2-ab+b^2$.

In every complex quadratic field the conjugation corresponds to the complex conjugation.

The following two propositions follow directly from definition.

Theorem 1. *The conjugation is multiplicative, i.e. for arbitrary elements z_1, z_2 of a quadratic extension of $\mathbb{Z}$ it holds that $\overline{z_1 z_2} = \overline{z_1}\,\overline{z_2}$.* □

Theorem 2. *The norm is multiplicative, i.e. for arbitrary elements z_1, z_2 of a quadratic extension of $\mathbb{Z}$ it holds that $N(z_1z_2) = N(z_1)N(z_2)$.* □

An element $\varepsilon \in \mathbb{Z}[\alpha]$ is called a *unit* if there exists $\varepsilon' \in \mathbb{Z}[\alpha]$ such that $\varepsilon\varepsilon' = 1$. In that case $N(\varepsilon)N(\varepsilon') = N(1) = 1$, so $N(\varepsilon) = \pm 1$. In fact, ε is a unit if and only if its norm is ± 1: indeed, if $N(\varepsilon) = \pm 1$ then $\varepsilon\overline{\varepsilon} = \pm 1$ by definition.

Example 4. *The only units in $\mathbb{Z}$ are ± 1.*

Let us find the units in $\mathbb{Z}[i]$. If $a+bi$ $(a,b \in \mathbb{Z})$ is a unit, then $N(a+bi) = a^2+b^2 = \pm 1$, which implies $a+bi \in \{\pm 1, \pm i\}$.

All units in $\mathbb{Z}[\omega]$ are $\pm 1, \pm\omega, \pm(1+\omega)$. Indeed, if $a+b\omega$ is a unit then $a^2-ab+b^2 = 1$, i.e. $(2a-b)^2+3b^2 = 4$ and he result follows. Note that ω^2 equals $-(1+\omega)$.

Problem 1. *Let p be a prime number and $N = \prod_{k=1}^{p-1}(k^2+1)$. Determine the remainder of N upon division by p.*

Solution. Denote $P(x) = (1+x)(2+x)\ldots(p-1+x)$. We know that $P(x) = x^{p-1}-1+pQ(x)$ for some polynomial $Q(x)$ with integer coefficients.

Since $k^2+1 = (k+i)(k-i)$ for each k, we immediately obtain that

$$\begin{aligned} N &= P(i)P(-i) = \left(i^{p-1}-1+pQ(i)\right)\left((-i)^{p-1}-1+pQ(-i)\right) \\ &\equiv \begin{cases} 4, & \text{if } p \equiv 3 \pmod 4; \\ 0, & \text{otherwise.} \end{cases} \end{aligned}$$ △

The divisibility and congruences in an extension K of the ring $\mathbb{Z}$ is defined in the usual way: $x \in K$ is divisible by $y \in K$ (denoted $y \mid x$) if there exists $z \in K$ such that $x = yz$, and $x \equiv y \pmod z$ if $z \mid x-y$.

Since every element of a quadratic ring is divisible by every unit, the definition of the notion of a prime must be adjusted to the new circumstances.

Definition 4. *An element y of a quadratic ring K is* adjoint *to element x (denoted $y \sim x$) if there exists a unit ε such that $y = \varepsilon x$.*

Definition 5. *A nonzero element $x \in K$ which is not a unit is* prime *if it has no other divisors but the units and elements adjoint to itself.*

We have the following simple proposition.

Theorem 3. *Let $x \in K$. If $N(x)$ is a prime, then x is prime.*

Proof. Suppose that $x = yz$, $y,z \in K$. Then $N(x) = N(y)N(z)$, so at least one of $N(y), N(z)$ equals ± 1, i.e. either y or z is a unit, while the other is (by definition) adjoint to x. □

The converse does not hold, as 3 is a prime in $\mathbb{Z}[i]$, but $N(3) = 9$ is composite.

Of course, the elements conjugate or adjoint to a prime are also primes. Therefore the smallest positive rational integer divisible by a prime z equals $z\overline{z} = N(z)$.

Consider an arbitrary nonzero and nonunit element $x \in K$. If x is not prime then there are nonunit elements $y,z \in K$ such that $yz = x$. Hereby $N(y)N(z) = N(x)$ and $N(y), N(z) > 1$. Hence $N(y), N(z) < N(x)$. Continuing this procedure we end up with a factorization $x = x_1 x_2 \cdots x_k$ in which all elements are prime. This shows that:

Theorem 4. *Every nonzero and nonunit $x \in K$ can be factorized into primes.* □

Problem 2. *Given a nonzero and nonunit element $z \in K$, find the number of equivalence classes in K modulo z.*

Solution. Let $K = \mathbb{Z}[\alpha]$, where $\alpha^2 = p\alpha + q$, $p,q \in \mathbb{Z}$, and let $z = a + b\alpha$ ($a,b \in \mathbb{Z}$). If $b = 0$ then $a_1 + b_1\alpha \equiv a_2 + b_2\alpha \pmod{z}$ if and only if $a_1 \equiv a_2$ and $b_1 \equiv b_2 \pmod{z}$. Thus there are $N(z) = z^2$ equivalence classes.

Now suppose that $b \neq 0$ and that $(a,b) = d$. Then $\alpha z = (a + pb)\alpha + qb$. Since $(a + pb, b) = d$, the coefficient at α in xz ($x \in K$) can be any integer divisible by d and no other integer. Moreover, the smallest natural number divisible by z is $|(a + b\alpha)(\overline{a + b\alpha})|/d = |N(z)|/d$. We conclude that for every $x \in K$ there is a unique $X = A + B\alpha \in K$ with $A,B \in \mathbb{Z}$, $0 \leq A < |N(z)|/d$, $0 \leq B < d$ such that $x \equiv X \pmod{z}$. Therefore the required number of equivalence classes equals $|N(z)|$. △

Naturally, we would like to know when the factorization into primes is unique, i.e. when the Fundamental Theorem of Arithmetic holds. But let us first note that, by the above definition, the primes of $\mathbb{Z}$ are $\pm 2, \pm 3, \pm 5$, etc, so the factorization into primes is not exactly unique, as e.g. $2 \cdot 3 = (-2)(-3)$. Actually, in this case the uniqueness of factorization is true in the following wording.

Definition 6. *FTA, or "The Fundamental Theorem of Arithmetic" means: Each nonzero and nonunit element of $\mathbb{Z}$ or of its quadratic extension K can be written as a product of primes. This factorization is unique up to the order of the factors and adjoining between corresponding factors.*

The division with remainder in a quadratic extension K of $\mathbb{Z}$ can be formulated as follows:

Definition 7. *DWR means: For each $a,b \in K$ with $b \neq 0$ there exist $p,q \in K$ such that $a = pb + q$ and $N(q) < N(b)$.*

Obviously, such a division, if it exists, is not necessarily unique - it is not so even in $\mathbb{Z}$ itself. Moreover, it does not exist in some quadratic extensions, as we shall see later. The significance of the existence of a division with remainder, however, lies in the fact that it implies the uniqueness of factorization:

Theorem 5. *If the division with remainder in a quadratic ring K is always possible, then FTA holds in K.*

Proof. If the division with remainder is possible in K, then the Euclidean algorithm ends in a finite number of steps. A simple implication of the Euclidean algorithm is that if p is a prime, $a,b \in K$ and $p \mid ab$, then $p \mid a$ or $p \mid b$. The uniqueness of factorization into primes (FTA) now easily follows. □

There are quadratic rings in which FTA holds despite the nonexistence of a division with remainder. However, FTA is an exception rather than a rule.

Example 5. *FTA is false in $\mathbb{Z}[\sqrt{-5}]$, as 9 has two factorizations into primes: $9 = 3\cdot 3 = (2+\sqrt{-5})(2-\sqrt{-5})$, which are not equivalent since $2 \pm \sqrt{-5} \not\sim 3$.*

Example 6. *The factorizations of the element $4-\omega$ in $\mathbb{Z}[\omega]$ as $(1-\omega)(3+\omega) = (-2-3\omega)(1+2\omega)$ are considered the same, since $1+2\omega = \omega(1-\omega) \sim 1-\omega$ and $-2-3\omega = -(1+\omega)(3+\omega) \sim 3+\omega$. We shall show later that FTA is true in $\mathbb{Z}[\omega]$.*

2 Arithmetic in the Gaussian Integers $\mathbb{Z}[i]$

We have already seen that the norm of element $a+bi \in \mathbb{Z}[i]$ ($a,b \in \mathbb{Z}$) is $N(a+bi) = a^2+b^2$ and the units are ± 1 and $\pm i$. Therefore, all divisors of a prime element $\pi \in \mathbb{Z}[i]$ are $\pm 1, \pm i, \pm\pi, \pm i\pi$.

Theorem 6. *The Fundamental Theorem of Arithmetic (FTA) holds in the set of Gaussian integers $\mathbb{Z}[i]$.*

Proof. Based on theorem 5, it is enough to show that for all $a,b \in \mathbb{Z}[i]$ with $b \neq 0$ there exists an element $p \in \mathbb{Z}[i]$ such that $N(a-pb) < N(b)$.

Let $\sigma,\tau \in \mathbb{R}$ be such that $a/b = \sigma+\tau i$, and let $s,t \in \mathbb{Z}$ be such that $|\sigma - s| \leq 1/2$ and $|\tau - t| \leq 1/2$. Setting $p = s+ti$ yields $a-pb = (\sigma+\tau i)b - pb = [(\sigma-s)+(\tau-t)i]b$, which implies

$$N(a-pb) = N[(\sigma-s)+(\tau-t)i]N(b) = [(\sigma-s)^2+(\tau-t)^2]N(b) \leq N(b)/2 < N(b).$$

This proves the theorem. □

The following proposition describes all prime elements in the set of Gaussian integers.

Theorem 7. *An element $x \in \mathbb{Z}[i]$ is prime if and only if $N(x)$ is a prime or $|x|$ is a prime integer of the form $4k-1$, $k \in \mathbb{N}$.*

Proof. Consider an arbitrary prime $x = a+bi \in \mathbb{Z}[i]$ ($a,b \in \mathbb{Z}$). Element $\overline{x}$ is prime also (indeed, if $\overline{x} = yz$, then $x = \overline{y}\,\overline{z}$), so $N(x)$ factorizes into primes as $x\overline{x}$.

Suppose that $N(x)$ is composite, $N(x) = mn$ for some two natural numbers $m,n > 1$. It follows from $x\overline{x} = mn$ and the FTA that $x \sim m$ or $x \sim n$, so we may suppose w.l.o.g. that x is a prime integer. If $x = 2$ or $x \equiv 1 \pmod 4$, then there exist integers $a,b \in \mathbb{Z}$ such that $N(a+bi) = (a+bi)(a-bi) = a^2+b^2 = x$; hence x is composite in $\mathbb{Z}[i]$. On the other hand, if x is a prime integer with $x \equiv 3 \pmod 4$, then x is also prime in $\mathbb{Z}[i]$. Indeed, if $x = uv$ for some nonunit elements $u,v \in \mathbb{Z}[i]$, then $x^2 = N(x) = N(u)N(v)$ implies $N(u) = N(v) = x$ which is impossible. This proves our claim. □

Problem 3. *Solve the equation $x^5 - 1 = y^2$ in integers.*

Solution. Rewrite the equation in the form $x^5 = (y+i)(y-i)$. Note that x is not even, as otherwise $y^2 \equiv -1 \pmod 4$. Thus y is even and consequently the elements $y+i$ and $y-i$ are coprime in $\mathbb{Z}[i]$. Since $(y+i)(y-i)$ is a fifth power, it follows that $y+i$ and $y-i$ are both fifth powers. Let $a,b \in \mathbb{Z}$ be such that $y+i = (a+bi)^5 = a(a^4-10a^2b^2+5b^4) + b(5a^4-10a^2b^2+b^4)i$. It holds that $b(5a^4-10a^2b^2+b^4) = 1$, and therefore $b = \pm 1$. It is easily seen that in both cases we have $a = 0$; hence $y = 0$, $x = \pm 1$ are the only solutions. △

Problem 4. *Suppose that x,y and z are natural numbers satisfying $xy = z^2+1$. Prove that there exist integers a,b,c,d such that $x = a^2+b^2$, $y = c^2+d^2$ and $z = ac+bd$.*

Solution. We use the following important fact: If m,n,p,q are elements of a unique factorization domain K (in this case, $K=\mathbb{Z}[i]$) satisfying $mn=pq$, then there exist $u_1,u_2,v_1,v_2\in K$ such that $m=u_1v_1$, $n=u_2v_2$, $p=u_1v_2$, $q=u_2v_1$. The proof of this fact is the same as in the case of $m,n,p,q\in\mathbb{Z}$ and follows directly from the factorizations of m,n,p,q into primes.

Since $xy=z^2+1=(z+i)(z-i)$, the above fact gives us

$$x=u_1v_1\quad(1),\quad y=u_2v_2\quad(2),\quad z+i=u_1v_2\quad(3),\quad z-i=u_2v_1\quad(4)$$

for some $u_1,u_2,v_1,v_2\in\mathbb{Z}[i]$. The numbers x,y are real, and therefore $v_1=q_1\overline{u_1}$, $v_2=q_2\overline{u_2}$ for some rational numbers q_1,q_2. From (3) and (4) we easily conclude that $q_1=q_2=1$. Now putting $u_1=a+bi$, $u_2=c+di$ yields $x=u_1\overline{u_1}=a^2+b^2$, $y=c^2+d^2$ and $z=ac+bd$. $\triangle$

3 Arithmetic in the ring $\mathbb{Z}[\omega]$

Here ω denotes a primitive cubic root of unity. Then the norm of an element $a+b\omega\in\mathbb{Z}[\omega]$ $(a,b\in\mathbb{Z})$ is $N(a+b\omega)=a^2-ab+b^2$ and the units are ±1, $\pm\omega$ and $\pm(1+\omega)=\mp\omega^2$.

Theorem 8. *FTA holds in the ring* $\mathbb{Z}[\omega]$.

Proof. By the theorem 5, it suffices to show that a division with remainder is possible, i.e. for all $a,b\in\mathbb{Z}[\omega]$, $b\neq0$ there exist $p\in\mathbb{Z}[\omega]$ such that $N(a-pb)<N(b)$.

Like in the proof for the Gaussian integers, let $\sigma,\tau\in\mathbb{R}$ be such that $a/b=\sigma+\tau i$, and let $s,t\in\mathbb{Z}$ be such that $|\sigma-s|\le1/2$ and $|\tau-t|\le1/2$. Setting $p=s+ti$ gives us $N(a-pb)\le3N(b)/4<N(b)$, implying the theorem. □

Problem 5. *If* $p\equiv1$ *(mod 6) is a prime number, prove that there exist* $a,b\in\mathbb{Z}$ *such that* $p=a^2-ab+b^2$.

Solution. It suffices to show that p is composite in $\mathbb{Z}[\omega]$. Indeed, if there is a prime element $z=a+b\omega\in\mathbb{Z}[\omega]$ $(a,b\in\mathbb{Z})$ that divides p, then also $\overline{z}\mid\overline{p}=p$. Note that z and $\overline{z}$ are coprime; otherwise $z\mid\overline{z}$, so there exists a unit element ε with $\overline{z}=\varepsilon z$, and hence $z\sim(1-\omega)\mid3$, which is false. Therefore $a^2-ab+b^2=z\overline{z}\mid p$, which implies $a^2-ab+b^2=p$.

Thus we need to prove that p is composite in $\mathbb{Z}[\omega]$. It follows from the condition on p that -3 is a quadratic residue modulo p, so there exist $m,n\in\mathbb{Z}$ which are not divisible by p such that $p\mid(2m-n)^2+3n^2=4(m^2-mn+n^2)$, i.e. $p\mid(m-n\omega)\overline{m-n\omega}$. However, p does not divide any of the elements $(m-n\omega),\overline{m-n\omega}$, so it must be composite. $\triangle$

Theorem 9. *Element* $x\in\mathbb{Z}[\omega]$ *is prime if and only if* $N(x)$ *is prime or* $|x|$ *is a prime integer of the form* $3k-1$, $k\in\mathbb{N}$.

Proof. Number $x=3$ is composite, as $N(1-\omega)=(1-\omega)(2+\omega)=3$. Moreover, by problem 4, every prime integer $p\equiv1$ (mod 6) is composite in $\mathbb{Z}[\omega]$.

The rest of the proof is similar to the proof of Theorem 7 and is left as an exercise. □

Maybe the most famous application of the elementary arithmetic of the ring $\mathbb{Z}[\omega]$ is the Last Fermat Theorem for the exponent $n=3$. This is not unexpected, having in mind that x^3+y^3 factorizes over $\mathbb{Z}[\omega]$ into linear factors:

$$x^3+y^3=(x+y)(x+\omega y)(x+\omega^2y)=(x+y)(\omega x+\omega^2y)(\omega^2x+\omega y).\qquad(1)$$

The proof we present was first given by Gauss.

Theorem 10. *The equation*

$$x^3+y^3=z^3\qquad(*)$$

has no nontrivial solutions in $\mathbb{Z}[\omega]$, *and consequently has none* $\mathbb{Z}$ *either.*

Proof. Suppose that x,y,z are nonzero elements of $\mathbb{Z}[\omega]$ that satisfy $(*)$. We can assume w.l.o.g. that x,y,z are pairwise coprime.

Consider the number $\rho = 1-\omega$. It is prime, as its norm equals $(1-\omega)(1-\omega^2) = 3$. We observe that $\overline{\rho} = 1-\omega^2 = (1-\omega)(1+\omega) \sim \rho$; hence $\alpha \in \mathbb{Z}[\omega]$ is divisible by ρ if and only if so is $\overline{alpha}$. Each element in $\mathbb{Z}[\omega]$ is congruent to $-1, 0$ or $1 \pmod{\rho}$: indeed, $a+b\omega \equiv a+b = 3q+r \equiv r \pmod{\rho}$ for some $q \in \mathbb{Z}$ and $r \in \{-1,0,1\}$.

The importance of number ρ lies in the following property:

$$\alpha \equiv \pm 1 \pmod{\rho}\ (\alpha \in \mathbb{Z}[\omega]) \text{ implies } \alpha^3 \equiv \pm 1 \pmod{\rho^4}. \tag{2}$$

Indeed, if $\alpha = \pm 1 + \beta\rho$, we have $a^3 \mp 1 = (a \mp 1)(a \mp \omega)(a \mp \omega^2) = \rho^3\beta(\beta \pm 1)(\beta \pm (1+\omega))$, where the elements b, $b \pm 1$, $b \pm (1+\omega)$ leave three distinct remainders modulo ρ, implying that one of them is also divisible by ρ, thus justifying our claim.

Among the numbers x,y,z, (exactly) one must be divisible by ρ: otherwise, by (2), x^3, y^3, z^3 would be congruent to $\pm 1 \pmod{\rho^4}$, which would imply one of the false congruences $0 \equiv \pm 1$, $\pm 1 \equiv \pm 2 \pmod{\rho^4}$. We assume w.l.o.g. that $\rho \mid z$. Moreover, (2) also gives us that $\rho^2 \mid z$.

Let $k \geq 2$ be the smallest natural number for which there exists a solution to $(*)$ with $(x,y,z) = 1$ and $\rho^k \mid z$, $\rho^{k+1} \nmid z$. Consider this solution (x,y,z).

The factors $x+y$, $\omega x + \omega^2 y$, $\omega^2 x + \omega y$ from (1) are congruent modulo ρ and have the sum 0. It follows from $\rho \mid z$ that each of them is divisible by ρ and that ρ is their greatest common divisor. Let

$$x+y = A\rho, \quad \omega x + \omega^2 y = B\rho, \quad \omega^2 x + \omega y = C\rho,$$

where $A,B,C \in \mathbb{Z}[\omega]$ are pairwise coprime and $A+B+C = 0$. The product ABC is a perfect cube (equal to $(z/\rho)^3$), and hence each of A,B,C is adjoint to a cube:

$$A = \alpha\zeta^3, \quad B = \beta\eta^3, \quad C = \gamma\xi^3$$

for some pairwise coprime $\zeta, \eta, \xi \in \mathbb{Z}[\omega]$ and units α, β, γ. Therefore,

$$\alpha\zeta^3 + \beta\eta^3 + \gamma\xi^3 = 0. \tag{3}$$

Since $\alpha\beta\gamma$ is a unit and a perfect cube, we have $\alpha\beta\gamma = \pm 1$. Furthermore, $ABC = (z/\rho)^3$ is divisible by ρ (since $\rho^2 \mid z$), so (exactly) one of the numbers ζ, η, ξ, say ξ, is divisible by ρ. In fact, ξ^3 divides ABC which is divisible by ρ^{3k-3} and not by ρ^{3k-2}, so ρ^{k-1} is the greatest power of ρ that divides ξ. The numbers ζ and η are not divisible by ρ; consequently, ζ^3 and η^3 are congruent to ± 1 modulo ρ^4. Thus the equality $A+B+C = 0$ gives us $\alpha \pm \beta \equiv 0 \pmod{\rho^4}$, therefore $\beta = \pm\alpha$; now $\alpha\beta\gamma = \pm 1$ implies $\gamma = \pm\alpha$.

Canceling α in equation (3) yields $\zeta^3 \pm \eta^3 \pm \xi^3 = 0$, which gives us another nontrivial solution to $(*)$ with $(\zeta, \eta, \xi) = 1$. However, in this solution we have $\rho^{k-1} \mid \xi$ and $\rho^k \nmid \xi$, which contradicts the choice of k. □

4 Arithmetic in other quadratic rings

Every quadratic ring belongs to one of the two classes:

1° Extensions of the form $K = \mathbb{Z}[\sqrt{d}]$, where $d \neq 1$ is a squarefree integer. The conjugation and norm are given by the formulas $\overline{x+y\sqrt{d}} = x - y\sqrt{d}$ and $N(x+y\sqrt{d}) = x^2 - dy^2$, where $x,y \in \mathbb{Z}$.

2° Extensions of the form $K = \mathbb{Z}[\alpha]$ for $\alpha = \dfrac{-1+\sqrt{d}}{2}$, where $d = 4k+1$ ($k \in \mathbb{Z}$) is a squarefree integer with $d \neq 1$ (then α is an algebraic integer: $\alpha^2 + \alpha - k = 0$). The conjugation and norm are given by $\overline{x+y\alpha} = x - y - y\alpha$ and $N(x+y\alpha) = x^2 - xy - ky^2$, where $x,y \in \mathbb{Z}$.

Some of these rings are Euclidean, such as $\mathbb{Z}[\sqrt{d}]$ for $d = -2,-1,2,3,6,7$ and $\mathbb{Z}\left[\frac{-1+\sqrt{d}}{2}\right]$ for $d = -7,-3,5$.

Determining all quadratic unique factorization rings (including the non-Euclidean ones) is extremely serious. Among the rings of the type 1° and 2° with $d < 0$, apart from the ones mentioned already, the FTA holds in only five other rings: namely, the rings of the type 2° for $d = -11,-19,-43,-67,-163$. Gauss' conjecture that the FTA holds in infinitely many quadratic rings with a positive d has not been proved nor disproved until today.

Problem 6. *Find all integer solutions of the equation* $x^2+2 = y^3$.

Solution. Let us write the equation as $(x+\sqrt{-2})(x-\sqrt{-2}) = y^3$. For x even we have $y^3 \equiv 2 \pmod 4$, which is impossible; therefore x is odd. Then $x+\sqrt{-2}$ and $x-\sqrt{-2}$ are coprime elements of $\mathbb{Z}[\sqrt{-2}]$ whose product is a perfect cube. Using the FTA in $\mathbb{Z}[\sqrt{-2}]$ we conclude that $x+\sqrt{-2}$ and $x-\sqrt{-2}$ are both perfect cubes. Hence there exist $a,b \in \mathbb{Z}$ such that $(a+b\sqrt{-2})^3 = x+\sqrt{-2}$. Comparing the coefficients at $\sqrt{-2}$ yields $b(3a^2-2b^2) = 1$; therefore $b = 1$ and $a = \pm 1$. Now we easily obtain that $x = \pm 5$ and $y = 3$ is the only integral solution of the equation. $\triangle$

Problem 7. *Consider the sequence* $a_0, a_1, a_2, \ldots$ *given by* $a_0 = 2$ *and* $a_{k+1} = 2a_k^2 - 1$ *for* $k \geq 0$. *Prove that if an odd prime number* p *divides* a_n, *then* $p \equiv \pm 1$ *(mod* 2^{n+2}*).*

Solution. Consider the sequence x_k of positive numbers given by $a_k = \cosh x_k$ (cosh is the *hyperbolic cosine*, defined by $\cosh t = \frac{e^t+e^{-t}}{2}$). It is easily verified that $\cosh(2x_k) = 2a_k^2 - 1 = \cosh x_{k+1}$, so $x_{k+1} = 2x_k$, i.e. $x_k = \lambda \cdot 2^k$ for some $\lambda > 0$. The condition $a_0 = 2$ gives us $\lambda = \log(2+\sqrt{3})$. Therefore

$$a_n = \frac{(2+\sqrt{3})^{2^n} + (2-\sqrt{3})^{2^n}}{2}.$$

Let $p > 2$ be a prime number such $p \mid a_n$. We distinguish two cases.

1° $m^2 \equiv 3 \pmod p$ for some $m \in \mathbb{Z}$. Then

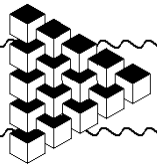
$$(2+m)^{2^n} + (2-m)^{2^n} \equiv 0 \pmod p. \tag{1}$$

Since $(2+m)(2-m) = 4-m^2 \equiv 1 \pmod p$, multiplying both sides of (1) by $(2+m)^{2^n}$ yields $(2+m)^{2^{n+1}} \equiv -1 \pmod p$. It follows that the order of number $(2+m)$ modulo p equals 2^{n+2}, from which we conclude $2^{n+2} \mid p-1$.

2° The congruence $m^2 \equiv 3 \pmod p$ has no solutions. We work in the quadratic extension $\mathbb{Z}_p(\sqrt{3})$ of the field $\mathbb{Z}_p$ in which number $\sqrt{3}$ actually plays the role of the number m from case (1). As in case (1) we have $(2+\sqrt{3})^{2^{n+1}} = -1$, which means that the order of $2+\sqrt{3}$ in the multiplicative group $\mathbb{Z}_p(\sqrt{3})^*$ equals 2^{n+2}. This is not enough to finish the proof as in case (1), as the group $\mathbb{Z}_p(\sqrt{3})^*$ has p^2-1 elements; instead, we only get that $2^{n+2} \mid p^2-1$. However, we shall be done if we find $u \in \mathbb{Z}_p(\sqrt{3})$ for which $u^2 = 2+\sqrt{3}$: indeed, then the order of u is 2^{n+3}, so $2^{n+3} \mid p^2-1$ and therefore $2^{n+2} \mid p-1$, since $4 \nmid p+1$.

Note that $(1+\sqrt{3})^2 = 2(2+\sqrt{3})$. Now it is enough to show that $1/2$ is a perfect square in the field $\mathbb{Z}_p(\sqrt{3})$. This immediately follows from the relation $a_n = 0 = 2a_{n-1}^2 - 1$, as $1/2 = a_{n-1}^2$. This finishes the proof. $\triangle$

terug naar echt bestand



©2007 The Author(s) and The IMO Compendium Group

Generating Functions

Milan Novaković

Contents

1 Introduction

Generating functions are powerful tools for solving a number of problems mostly in combinatorics, but can be useful in other branches of mathematics as well. The goal of this text is to present certain applications of the method, and mostly those using the high school knowledge.

In the beginning we have a formal treatement of generating functions, i.e. power series. In other parts of the article the style of writing is more problem-soving oriented. First we will focus on solving the reccurent equations of first, second, and higher order, after that develope the powerful method of „the snake oil“, and for the end we leave some other applications and various problems where generating functions can be used.

The set of natural numbers will be denoted by $\mathbb{N}$, while $\mathbb{N}_0$ will stand for the set of non-negative integers. For the sums going from 0 to $+\infty$ the bounds will frequently be omitted – if a sum is without the bounds, they are assumed to be 0 and $+\infty$.

2 Theoretical Introduction

In dealing with generating functions we frequently want to use different transformations and manipulations that are illegal if the generating functions are viewed as analytic functions. Therefore they will be introduced as algebraic objects in order to obtain wider range of available methods. The theory we will develope is called the *formal theory of power series*.

Definition 1. A formal power series *is the expression of the form*

$$a_0 + a_1 x + a_2 x^2 + \cdots = \sum_{i=0}^{\infty} a_i x^i.$$

A sequence of integers $\{a_i\}_0^\infty$ *is called* the sequence of coefficients.

Remark. We will use the other expressions also: series, generating function...

For example the series

$$A(x) = 1 + x + 2^2 x^2 + 3^3 x^3 + \cdots + n^n x^n + \cdots$$

converges only for $x=0$ while, in the formal theory this is well defined formal power series with the corresponginds sequence of coefficients equal to $\{a_i\}_0^\infty, a_i = i^i$.
Remark. Sequences and their elements will be most often denoted by lower-case latin letters (a, b, a_3 $\cdots$), while the power series generated by them (unless stated otherwise) will be denoted by the corresponding capital letters (A, B, $\cdots$).

Definition 2. *Two series* $A=\sum_{i=0}^{\infty} a_i x^i$ *and* $B=\sum_{i=0}^{\infty} b_i x^i$ *are called* equal *if their corresponding sequences of coefficients are equal, i.e.* $a_i = b_i$ *for every* $i \in \mathbb{N}_0$.

Remark. The coefficient near x^n in the power series F will be denoted by $[x^n]F$.

We can define the *sum* and the *difference* of power series in the following way

$$\sum_n a_n x^n \pm \sum_n b_n x^n = \sum_n (a_n \pm b_n) x^n$$

while the *product* is defined by

$$\sum_n a_n x^n \sum_n b_n x^n = \sum_n c_n x^n, \qquad c_n = \sum_i a_i b_{n-i}$$

Instead of $F \cdot F$ we write F^2, and more generally $F^{n+1} = F \cdot F^n$. We see that the neutral for addition is 0, and 1 is the neutral for multiplication. Now we can define the following term:

Definition 3. *The formal power series* G *is* reciprocal *to the formal power series* F *if* $FG = 1$.

The generating function reciprocal to F will be usually denoted by $1/F$. Since the multiplication is commutative we have that $FG=1$ is equivalent to $GF=1$ hence F and G are *mutually reciprocal*. We also have $(1-x)(1+x+x^2+\cdots) = 1+\sum_{i=1}^{\infty}(1\cdot 1 - 1\cdot 1)x^i = 1$ hence $(1-x)$ and $(1+x+x^2+\cdots)$ are mutually reciprocal.

Theorem 1. *Formal power series* $F=\sum_n a_n x^n$ *has a reciprocal if and only if* $a_0 \neq 0$. *In that case the reciprocal is unique.*

Proof. Assume that F has a reciprocal given by $1/F = \sum_n b_n x^n$. Then $F \cdot (1/F) = 1$ implying $1 = a_0 b_0$ hence $a_0 \neq 0$. For $n \geqslant 1$ we have $0 = \sum_k a_k b_{n-k}$ from where we conclude.

$$b_n = -\frac{1}{a_0}\sum_k a_k b_{n-k}.$$

The coefficients are uniquely determined by the prefious formula.

On the other hand if $a_0 \neq 0$ we can uniquely determine all coefficients b_i using the previously established relations which gives the series $1/F$. □

Now we can conclude that the set of power series with the above defined operation forms a ring whose invertible elements are precisely those power series with the non-zero first coefficient.

If $F=\sum_n f_n x^n$ is a power series, $F(G(x))$ will denote the series $F(G(x)) = \sum_n f_n G(x)^n$. This notation will be used also in the case when F is a polynomial (i.e. when there are only finitely many non-zero coefficients) or if the free term of G equals 0. In the case that the free term of G equal to 0, and F is not a polynomial, we can't determine the particular element of the series $F(G(x))$ in finitely many steps.

Definition 4. *A formal power series* G *is said to be an inverse of* F *if* $F(G(x)) = G(F(x)) = x$.

We have a symmetry here as well, if G is inverse of F than F is inverse of G as well.

Theorem 2. *Let F and G be mutually inverse power series. Then $F = f_1x + f_2x^2 + \cdots$, $G = g_1x + g_2x^2 + \cdots$, and $f_1g_1 \neq 0$.*

Proof. In order for $F(G(x))$ and $G(F(x))$ to be defined we must have 0 free terms. Assume that $F = f_rx^r + \cdots$ and $G = g_sx^s + \cdots$. Then $F(G(x)) = x = f_rg_s^rx^{rs} + \cdots$, thus $rs = 1$ and $r = s = 1$. □

Definition 5. *The derivative of a power series $F = \sum_n f_nx^n$ is $F' = \sum_n nf_nx^{n-1}$. The derivative of order $n > 1$ is defined recursively by $F^{(n+1)} = (F^{(n)})'$.*

Theorem 3. *The following properties of the derivative hold:*

- $(F+G)^{(n)} = F^{(n)} + G^{(n)}$
- $(FG)^{(n)} = \sum_{i=0}^{n} \binom{n}{i} F^{(i)}G^{(n-i)}$

The proof is very standard as is left to the reader. □

We will frequently associate the power series with its generating sequence, and to make writing more clear we will define the the relation $\overset{osr}{\leftrightarrow}$ in the following way:

Definition 6. *$A \overset{osr}{\leftrightarrow} \{a_n\}_0^\infty$ means that A is a usual power series which is generated by $\{a_n\}_0^\infty$, i.e. $A = \sum_n a_nx^n$.*

Assume that $A \overset{osr}{\leftrightarrow} \{a_n\}_0^\infty$. Then

$$\sum_n a_{n+1}x^n = \frac{1}{x}\sum_{n>0} a_nx^n = \frac{A(x) - a_0}{x}$$

or equivalently $\{a_{n+1}\}_0^\infty \overset{osr}{\leftrightarrow} \dfrac{A - a_0}{x}$. Similarly

$$\{a_{n+2}\}_0^\infty \overset{osr}{\leftrightarrow} \frac{(A-a_0)/x - a_1}{x} = \frac{A - a_0 - a_1x}{x^2}.$$

Theorem 4. *If $\{a_n\}_0^\infty \overset{osr}{\leftrightarrow} A$ the for $h > 0$:*

$$\{a_{n+h}\}_0^\infty \overset{osr}{\leftrightarrow} \frac{A - a_0 - a_1x - \cdots - a_{h-1}x^{h-1}}{x^h}.$$

Proof. We will use the induction on h. For $h = 1$ the statement is true and that is shown before. If the statement holds for some h then

$$\begin{aligned}\{a_{n+h+1}\}_0^\infty &\overset{osr}{\leftrightarrow} \{a_{(n+h)+1}\}_0^\infty \overset{osr}{\leftrightarrow} \frac{\dfrac{A - a_0 - a_1x - \cdots - a_{h-1}x^{h-1}}{x^h} - a_h}{x} \\ &\overset{osr}{\leftrightarrow} \frac{A - a_0 - a_1x - \cdots - a_hx^h}{x^{h+1}},\end{aligned}$$

which finishes the proof. □

We already know that $\{(n+1)a_{n+1}\}_0^\infty \overset{osr}{\leftrightarrow} A'$. Our goal is to obtain the sequence $\{na_n\}_0^\infty$. That is exactly the sequence xA'. We will define the operator xD in the following way:

Definition 7. *$xDA = xA'$ i.e. $xDA = x\frac{dA}{dx}$.*

The following two theorems are obvious consequences of the properties of the derivative:

Theorem 5. *Let $\{a_n\}_0^\infty \overset{osr}{\leftrightarrow} A$. Then $\{n^ka_n\}_0^\infty \overset{osr}{\leftrightarrow} (xD)^kA$.*

Theorem 6. *Let $\{a_n\}_0^\infty \overset{osr}{\leftrightarrow} A$ and P be a polynomial. Then*

$$P(xD)A \overset{osr}{\leftrightarrow} \{P(n)a_n\}_0^\infty$$

Let us consider the generating function $\dfrac{A}{1-x}$. It can be written as $A\frac{1}{1-x}$. As we have shown before the reciprocal to the series $1-x$ is $1+x+x^2+\cdots$, hence $\frac{A}{1-x} = (a_0+a_1x+a_2x^2+\cdots)(1+x+x^2+\cdots) = a_0+(a_0+a_1)x+(a_0+a_1+a_2)x^2+\cdots$.

Theorem 7. *If $\{a_n\}_0^\infty \overset{osr}{\leftrightarrow} A$ then*

$$\frac{A}{1-x} \quad \overset{osr}{\leftrightarrow} \quad \left\{\sum_{j=0}^{n} a_j\right\}_{n\geqslant 0}.$$

Now we will introduce the new form of generating functions.

Definition 8. *We say that A is* exponential generating function *(or series, power series) of the sequence $\{a_n\}_0^\infty$ if A is the ususal generating function of the sequence $\{\frac{a_n}{n!}\}_0^\infty$, or equivalently*

$$A = \sum_n \frac{a_n}{n!}x^n.$$

If B is exponential generating function of the series $\{b_n\}_0^\infty$ we can also write $\{b_n\}_0^\infty \overset{esr}{\leftrightarrow} B$.

If $B \overset{esr}{\leftrightarrow} \{b_n\}_0^\infty$, we are interested in B'. It is easy to see that

$$B' = \sum_{n=1}^{\infty} \frac{nb_nx^{n-1}}{n!} = \sum_{n=1}^{\infty} \frac{b_nx^{n-1}}{(n-1)!} = \sum_{n=0}^{\infty} \frac{b_{n+1}x^n}{n!}$$

hence $B' \overset{esr}{\leftrightarrow} \{b_{n+1}\}_0^\infty$.

Theorem 8. *If $\{b_n\}_0^\infty \overset{esr}{\leftrightarrow} B$ then for $h \geqslant 0$:*

$$\{b_{n+h}\}_0^\infty \quad \overset{osr}{\leftrightarrow} \quad B^{(h)}.$$

We also have an equivalent theorem for exponential generating functions.

Theorem 9. *Let $\{b_n\}_0^\infty \overset{esr}{\leftrightarrow} B$ and let P be a polynomial. Then*

$$P(xD)B \overset{esr}{\leftrightarrow} \{P(n)b_n\}_0^\infty$$

The exponential generating functions are useful in combinatorial identities because of the following property.

Theorem 10. *Let $\{a_n\}_0^\infty \overset{esr}{\leftrightarrow} A$ and $\{b_n\}_0^\infty \overset{esr}{\leftrightarrow} B$. Then the generating function AB generates the sequence*

$$\left\{\sum_k \binom{n}{k} a_kb_{n-k}\right\}_{n=0}^{\infty}.$$

Proof. We have that

$$AB = \left\{\sum_{i=0}^{\infty} \frac{a_ix^i}{i!}\right\}\left\{\sum_{j=0}^{\infty} \frac{b_jx^j}{j!}\right\} = \sum_{i,j\geqslant 0} \frac{a_ib_j}{i!\,j!}x^{i+j} = \sum_n x^n \left\{\sum_{i+j=n} \frac{a_ib_j}{i!\,j!}\right\},$$

or

$$AB = \sum_n \frac{x^n}{n!}\left\{\sum_{i+j=n} \frac{n!a_ib_j}{i!\,j!}\right\} = \sum_n \frac{x^n}{n!}\sum_k \binom{n}{k} a_kb_{n-k},$$

and the proof is complete. □

We have listed above the fundamental properties of generating functions. New properties and terms will be defined later.

Although the formal power series are defined as solely algebraic objects, we aren't giving up their analytical properties. We will use the well-known Taylor's expansions of functions into power series. For example, we will treat the function e^x as a formal power series obtained by expanding the function into power series, i.e. we will identify e^x with $\sum_{n=0}^{\infty} \frac{x^n}{n!}$. We will use the converse direction also. Here we will list the Taylor expansions of most common functions.

$$\frac{1}{1-x} = \sum_{n\geqslant 0} x^n$$

$$\ln \frac{1}{1-x} = \sum_{n\geqslant 1} \frac{x^n}{n}$$

$$e^x = \sum_{n\geqslant 0} \frac{x^n}{n!}$$

$$\sin x = \sum_{n\geqslant 0} (-1)^n \frac{x^{2n+1}}{(2n+1)!}$$

$$\cos x = \sum_{n\geqslant 0} (-1)^n \frac{x^{2n}}{(2n)!}$$

$$(1+x)^\alpha = \sum_k \binom{\alpha}{k} x^k$$

$$\frac{1}{(1-x)^{k+1}} = \sum_n \binom{n+k}{n} x^n$$

$$\frac{x}{e^x-1} = \sum_{n\geqslant 0} \frac{B_n x^n}{n!}$$

$$\arctan x = \sum_{n\geqslant 0} (-1)^n \frac{x^{2n+1}}{2n+1}$$

$$\frac{1}{2x}(1-\sqrt{1-4x}) = \sum_n \frac{1}{n+1}\binom{2n}{n} x^n$$

$$\frac{1}{\sqrt{1-4x}} = \sum_n \binom{2n}{n} x^n$$

$$x\cot x = \sum_{n\geqslant 0} \frac{(-4)^n B_{2n}}{(2k)!} x^{2n}$$

$$\tan x = \sum_{n\geqslant 1} (-1)^{n-1} \frac{2^{2n}(2^{2n}-1)B_{2n}}{(2n)!} x^{2n-1}$$

$$\frac{x}{\sin x} = \sum_{n\geqslant 0} (-1)^{n-1} \frac{(4^n-2)B_{2n}}{(2n)!} x^{2n}$$

$$\frac{1}{\sqrt{1-4x}} \left(\frac{1-\sqrt{1-4x}}{2x}\right)^k = \sum_n \binom{2n+k}{n} x^n$$

$$\left(\frac{1-\sqrt{1-4x}}{2x}\right)^k = \sum_{n\geqslant 0} \frac{k(2n+k-1)!}{n!(n+k)!}x^n$$

$$\arcsin x = \sum_{n\geqslant 0} \frac{(2n-1)!!x^{2n+1}}{(2n)!!(2n+1)}$$

$$e^x \sin x = \sum_{n\geqslant 1} \frac{2^{\frac{n}{2}} \sin\frac{n\pi}{4}}{n!}x^n$$

$$\ln^2 \frac{1}{1-x} = \sum_{n\geqslant 2} \frac{H_{n-1}}{n}x^n$$

$$\sqrt{\frac{1-\sqrt{1-x}}{x}} = \sum_{n=0}^{\infty} \frac{(4n)!}{16^n\sqrt{2}(2n)!(2n+1)!}x^n$$

$$\left(\frac{\arcsin x}{x}\right)^2 = \sum_{n=0}^{\infty} \frac{4^n n!^2}{(k+1)(2k+1)!}x^{2n}$$

Remark: Here $H_n = \sum_{i=1}^{n} \frac{1}{i}$, and B_n is the n-th Bernoulli number.

3 Recurrent Equations

We will first solve one basic recurrent equation.

Problem 1. *Let a_n be a sequence given by $a_0 = 0$ and $a_{n+1} = 2a_n + 1$ for $n \geqslant 0$. Find the general term of the sequence a_n.*

Solution. We can calculate the first several terms 0, 1, 3, 7, 15, and we are tempted to guess the solution as $a_n = 2^n - 1$. The previous formula can be easily established using mathematical induction but we will solve the problem using generating functions. Let $A(x)$ be the generating function of the sequence a_n, i.e. let $A(x) = \sum_n a_n x^n$. If we multiply both sides of the recurrent relation by x^n and add for all n we get

$$\sum_n a_{n+1}x^n = \frac{A(x)-a_0}{x} = \frac{A(x)}{x} = 2A(x) + \frac{1}{1-x} = \sum_n (2a_n+1)x^n.$$

From there we easily conclude

$$A(x) = \frac{x}{(1-x)(2-x)}.$$

Now the problem is obtaining the general formula for the elements of the sequence. Here we will use the famous trick of decomposing A into two fractions each of which will have the corresponding generating function. More precisely

$$\frac{x}{(1-x)(2-x)} = x\left(\frac{2}{1-2x} - \frac{1}{1-x}\right) = (2x + 2^2x^2 + \cdots) - (x + x^2 + \cdots).$$

Now it is obvious that $A(x) = \sum_{n=0}^{\infty} (2^n - 1)x^n$ and the solution to the recurrent relation is indeed $a_n = 2^n - 1$. $\triangle$

Problem 2. *Find the general term of the sequence given recurrently by*

$$a_{n+1} = 2a_n + n, \quad (n \geqslant 0), \ a_0 = 1.$$

Solution. Let $\{a_n\}_0^\infty \overset{osr}{\leftrightarrow} A$. Then $\{a_{n+1}\}_0^\infty \overset{osr}{\leftrightarrow} \frac{A-1}{x}$. We also have that $xD\frac{1}{1-x} \overset{osr}{\leftrightarrow} \{n\cdot 1\}$. Since $xD\frac{1}{1-x} = x\frac{1}{(1-x)^2} = \frac{x}{(1-x)^2}$ the recurrent relation becomes

$$\frac{A-1}{x} = 2A + \frac{x}{(1-x)^2}.$$

From here we deduce

$$A = \frac{1-2x+2x^2}{(1-x)^2(1-2x)}.$$

Now we consider that we have *solved for the generating series*. If we wanted to show that the sequence is equal to some other sequence it would be enough to show that the functions are equal. However we need to find the terms explicitely. Let us try to represent A again in the form

$$\frac{1-2x+2x^2}{(1-x)^2(1-2x)} = \frac{P}{(1-x)^2} + \frac{Q}{1-x} + \frac{R}{1-2x}.$$

After multiplying both sides with $(1-x)^2(1-2x)$ we get

$$1-2x+2x^2 = P(1-2x) + Q(1-x)(1-2x) + R(1-x)^2,$$

or equivalently

$$1-2x+2x^2 = x^2(2Q+R) + x(-2P-3Q-2R) + (P+Q+R).$$

This implies $P=-1$, $Q=0$, and $R=2$. There was an easier way to get P, Q, and R. If we multiply both sides by $(1-x)^2$ and set $x=1$ we get $P=-1$. Similarly if we multiply everything by $1-2x$ and plug $x=\frac{1}{2}$ we get $R=2$. Now substituting P and R and setting $x=0$ we get $Q=0$.

Thus we have

$$A = \frac{-1}{(1-x)^2} + \frac{2}{1-2x}.$$

Since $\frac{2}{1-2x} \overset{osr}{\leftrightarrow} \{2^{n+1}\}$ and $\frac{1}{(1-x)^2} = D\frac{1}{1-x} \overset{osr}{\leftrightarrow} \{n+1\}$ we get $a_n = 2^{n+1} - n - 1$. $\triangle$

In previous two examples the term of the sequence was depending only on the previous term. We can use generating functions to solve recurrent relations of order greater than 1.

Problem 3 (Fibonacci's sequence). *$F_0 = 0$, $F_1 = 1$, and for $n \geqslant 1$, $F_{n+1} = F_n + F_{n-1}$. Find the general term of the sequence.*

Solution. Let F be the generating function of the series $\{F_n\}$. If we multiply both sides by x^n and add them all, the left-hand side becomes $\{F_{n+1}\} \overset{osr}{\leftrightarrow} \frac{F-x}{x}$, while the right-hand side becomes $F + xF$. Therefore

$$F = \frac{x}{1-x-x^2}.$$

Now we want to expand this function into power series. First we want to represent the function as a sum of two fractions. Let

$$-x^2 - x + 1 = (1-\alpha x)(1-\beta x).$$

Then $\alpha = (1+\sqrt{5})/2$, $\beta = (1-\sqrt{5})/2$, and $\alpha - \beta = \sqrt{5}$. We further have

$$\begin{aligned} \frac{x}{1-x-x^2} &= \frac{x}{(1-x\alpha)(1-x\beta)} = \frac{1}{\alpha-\beta}\left(\frac{1}{1-x\alpha} - \frac{1}{1-x\beta}\right) \\ &= \frac{1}{\sqrt{5}}\left\{\sum_{n=0}^{\infty}\alpha^n x^n - \sum_{n=0}^{\infty}\beta^n x^n\right\}. \end{aligned}$$

It is easy to obtain

$$F_n = \frac{1}{\sqrt{5}}(\alpha^n - \beta^n). \ \triangle$$

Remark: From here we can immediately get the approximate formula for F_n. Since $|\beta| < 1$ we have $\lim\limits_{n\to\infty} \beta^n = 0$ and

$$F_n \approx \frac{1}{\sqrt{5}} \left(\frac{1+\sqrt{5}}{2} \right)^n.$$

Now we will consider the case with the sequence of two variables.

Problem 4. *Find the number of k-element subsets of an n-element set.*

Solution. We know that the result is $\binom{n}{k}$, but we want to obtain this using the generating functions. Assume that the required number is equal to $c(n,k)$. Let $A = \{a_1, \ldots, a_n\}$ be an n-element set. There are two types of k-element subsets – those which contain a_n and those that don't. There are exactly $c(n-1,k-1)$ subsets containing a_n. Indeed they are all formed by taking $k-1$-element subsets of $\{a_1, \ldots, a_{n-1}\}$ and adding a_n to each of them. On the other hand there are exaclty $c(n-1,k)$ subsets not contianing a_n. Hence

$$c(n,k) = c(n-1,k) + c(n-1,k-1).$$

We also have $c(n,0) = 1$. Now we will define the generating function of the sequence $c(n,k)$ for a fixed n. Assume that

$$C_n(x) = \sum_k c(n,k)x^k.$$

If we multiply the recurrent relation by x^k and add for all $k \geqslant 1$ we get

$$C_n(x) - 1 = (C_{n-1}(x) - 1) + xC_{n-1}(x), \ \text{ for each } n \geqslant 0$$

and $C_0(x) = 1$. Now we have for $n \geqslant 1$:

$$C_n(x) = (1+x)C_{n-1}(x).$$

We finally have $C_n(x) = (1+x)^n$. Hence, $c(n,k)$ is the coefficient near x^k in the expansion of $(1+x)^n$, and that is exactly $\binom{n}{k}$. $\triangle$

Someone might think that this was a cheating. We have used binomial formula, and that is obtianed using a combinatorial technique which uses the result we wanted to prove. Fortunately, there is a proof of binomial formula involving Taylor expainsion.

We can also make a generating function of the sequece $C_n(x)$:

$$\sum_n C_n(x)y^n = \sum_n \sum_k \binom{n}{k} x^k y^n = \sum_n (1+x)^n y^n = \frac{1}{1-y(1+x)}.$$

In such a way we have $\binom{n}{k} = [x^k y^n](1-y(1+x))^{-1}$. Now we can calculate the sum $\sum\limits_n \binom{n}{k} y^n$:

$$\begin{aligned} [x^k] \sum_n \sum_k \binom{n}{k} x^k y^n &= [x^k] \frac{1}{1-y(1+x)} = \frac{1}{1-y} [x^k] \frac{1}{1-\frac{y}{1-y}x} \\ &= \frac{1}{1-y} \left(\frac{y}{1-y} \right)^k = \frac{y^k}{(1-y)^{k+1}}. \end{aligned}$$

Hence we have the identities

$$\sum_k \binom{n}{k} x^k = (1+x)^n; \qquad \sum_n \binom{n}{k} y^n = \frac{y^k}{(1-y)^{k+1}}.$$

Remark: For $n < k$ we define $\binom{n}{k} = 0$.

Problem 5. *Find the general term of the sequence* $a_{n+3} = 6a_{n+2} - 11a_{n+1} + 6a_n$, $n \geqslant 0$ *with the initial conditions* $a_0 = 2$, $a_1 = 0$, $a_2 = -2$.

Solution. If A is the generating function of the corresponding sequence then:

$$\frac{A-2-0\cdot x-(-2)x^2}{x^3} = 6\frac{A-2-0\cdot x}{x^2} - 11\frac{A-2}{x} + 6A,$$

from where we easily get

$$A = \frac{20x^2-12x+2}{1-6x+11x^2-6x^3} = \frac{20x^2-12x+2}{(1-x)(1-2x)(1-3x)}.$$

We want to find the real coefficients B, C, and D such that

$$\frac{20x^2-12x+2}{(1-x)(1-2x)(1-3x)} = \frac{B}{1-x} + \frac{C}{1-2x} + \frac{D}{1-3x}.$$

We will multiply both sides by $(1-x)$ and set $x=1$ to obtain $B = \frac{20-12+2}{(-1)\cdot(-2)} = 5$. Multiplying by $(1-2x)$ and setting $x = 1/2$ we further get $C = \frac{5-6+2}{-\frac{1}{4}} = -4$. If we now substitute the found values for B and C and put $x = 0$ we get $B+C+D = 2$ from where we deduce $D = 1$. We finally have

$$A = \frac{5}{1-x} - \frac{4}{1-2x} + \frac{1}{1-3x} = \sum_{n=0}^{\infty}(5-4\cdot 2^n+3^n)x^n$$

implying $a_n = 5-2^{n+2}+3^n$. $\triangle$

The following example will show that sometimes we can have troubles in finding the explicite formula for the elements of the sequence.

Problem 6. *Let the sequence be given by* $a_0 = 0$, $a_1 = 2$, *and for* $n \leqslant 0$*:*

$$a_{n+2} = -4a_{n+1} - 8a_n.$$

Find the general term of the sequence.

Solution. Let A be the generating function of the sequence. The recurrent relation can be written in the form

$$\frac{A-0-2x}{x^2} = -4\frac{A-0}{x} - 8A$$

implying

$$A = \frac{2x}{1+4x+8x^2}.$$

The roots $r_1 = -2+2i$ and $r_2 = -2-2i$ of the equation x^2+4x+8 are not real. However this should interfere too much with our intention for finding B and C. Pretending that nothing wierd is going on we get

$$\frac{2x}{1+4x+8x^2} = \frac{B}{1-r_1x} + \frac{C}{1-r_2x}.$$

Using the trick learned above we get $B = \frac{-i}{2}$ and $C = \frac{i}{2}$.

Did you read everything carefully? Why did we consider the roots of the polynomial x^2+4x+8 when the denumerator of A is $8x^2+4x+1$?! Well if we had considered the roots of the real denumerator we would get the fractions of the form $\frac{B}{r_1-x}$ which could give us a trouble if we wanted to use power series. However we can express the denominator as $x^2(8+4\frac{1}{x}+\frac{1}{x^2})$ and consider this as a polynomial in $\frac{1}{x}$! Then the denumerator becomes $x^2\left(\frac{1}{x}-r_1\right)\cdot\left(\frac{1}{x}-r_2\right)$.

Now we can proceed with solving the problem. We get

$$A = \frac{-i/2}{1-(-2+2i)x} + \frac{i/2}{1-(-2-2i)x}.$$

From here we get

$$A = \frac{-i}{2}\sum_{n=0}^{\infty}(-2+2i)^n x^n + \frac{i}{2}\sum_{n=0}^{\infty}(-2-2i)^n x^n,$$

implying

$$a_n = \frac{-i}{2}(-2+2i)^n + \frac{i}{2}(-2-2i)^n.$$

But the terms of the sequence are real, not complex numbers! We can now simplify the expression for a_n. Since

$$-2 \pm 2i = 2\sqrt{2}e^{\frac{\pm 3\pi i}{4}},$$

we get

$$a_n = \frac{i}{2}(2\sqrt{2})^n \left((\cos\frac{3n\pi}{4} - i\sin\frac{3n\pi}{4}) - (\cos\frac{3n\pi}{4} + i\sin\frac{3n\pi}{4}) \right),$$

hence $a_n = (2\sqrt{2})^n \sin\frac{3n\pi}{4}$. Written in another way we get

$$a_n = \begin{cases} 0, & n = 8k \\ (2\sqrt{2})^n, & n = 8k+6 \\ -(2\sqrt{2})^n, & n = 8k+2 \\ \frac{1}{\sqrt{2}}(2\sqrt{2})^n, & n = 8k+1 \text{ ili } n = 8k+3 \\ -\frac{1}{\sqrt{2}}(2\sqrt{2})^n, & n = 8k+5 \text{ ili } n = 8k+7. \end{cases} \triangle$$

Now we will consider on more complex recurrent equation.

Problem 7. *Find the general term of the sequence x_n given by:*

$$x_0 = x_1 = 0, \qquad x_{n+2} - 6x_{n+1} + 9x_n = 2^n + n \quad za\ n \geqslant 0.$$

Solution. Let $X(t)$ be the generating function of our sequence. Using the same methods as in the examples above we can see that the following holds:

$$\frac{X}{t^2} - 6\frac{X}{t} + 9X = \frac{1}{1-2t} + \frac{t}{(1-t)^2}.$$

Simplifying the expression we get

$$X(t) = \frac{t^2 - t^3 - t^4}{(1-t)^2(1-2t)(1-3t)^2},$$

hence

$$X(t) = \frac{1}{4(1-x)^2} + \frac{1}{1-2x} - \frac{5}{3(1-3x)} + \frac{5}{12(1-3x)^2}.$$

The sequence corresponding to the first summand is $\frac{n+1}{4}$, while the sequences for the second, third, and fourth are 2^n, $5\cdot 3^{n-1}$, and $\frac{5(n+1)3^{n+1}}{12}$ respectively. Now we have

$$x_n = \frac{2^{n+2} + n + 1 + 5(n-3)3^{n-1}}{4}. \ \triangle$$

Problem 8. *Let $f_1 = 1$, $f_{2n} = f_n$, and $f_{2n+1} = f_n + f_{n+1}$. Find the general term of the sequence.*

Solution. We see that the sequence is well define because each term is defined using the terms already defined. Let the generating function F be given by

$$F(x) = \sum_{n\geqslant 1} f_n x^{n-1}.$$

Multiplying the first given relation by x^{2n-1}, the second by x^{2n}, and adding all of them for $n \geq 1$ we get:

$$f_1 + \sum_{n\geqslant 1} f_{2n}x^{2n-1} + \sum_{n\geqslant 1} f_{2n+1}x^{2n} = 1 + \sum_{n\geqslant 1} f_n x^{2n-1} + \sum_{n\geqslant 1} f_n x^{2n} + \sum_{n\geqslant 1} f_{n+1}x^{2n}$$

or equivalently

$$\sum_{n\geqslant 1} f_n x^{n-1} = 1 + \sum_{n\geqslant 1} f_n x^{2n-1} + \sum_{n\geqslant 1} f_n x^{2n} + \sum_{n\geqslant 1} f_{n+1}x^{2n}.$$

This exactly means that $F(x) = x^2F(x^2) + xF(x^2) + F(x^2)$ i.e.

$$F(x) = (1+x+x^2)F(x^2).$$

Moreover we have

$$F(x) = \prod_{i=0}^{\infty}\left(1 + x^{2^i} + x^{2^{i+1}}\right).$$

We can show that the sequence defined by the previous formula has an interesting property. For every positive integer n we perform the following procedure: Write n in a binary expansion, discard the last "block" of zeroes (if it exists), and group the remaining digits in as few blocks as possible such that each block contains the digits of the same type. If for two numbers m and n the corresponding sets of blocks coincide the we have $f_m = f_n$. For example the binary expansion of 22 is 10110 hence the set of corresponding blocks is $\{1,0,11\}$, while the number 13 is represented as 1101 and has the very same set of blocks $\{11,0,1\}$, so we should have $f(22) = f(13)$. Easy verification gives us $f(22) = f(13) = 5$. From the last expression we conclude that f_n is the number of representations of n as a sum of powers of two, such that no two powers of two are taken from the same set of a collection $\{1,2\}$, $\{2,4\}$, $\{4,8\}$.

4 The Method of the Snake Oil

The method of the snake oil is very useful tool in evaluating various, frequently huge combinatorial sums, and in proving combinatorial identities.

The method is used to calculate many sums and as such it is not universal. Thus we will use several examples to give the flavor and illustration of the method.

The general principle is as follows: Suppose we want to calculate the sum S. First we wnat to identify the free variable on which S depends. Assume that n is such a variable and let $S = f(n)$. After that we have to obtain $F(x)$, the generating function of the sequence $f(n)$. We will multiply S by x^n and summ over all n. At this moment we have (at least) a double summation external in n and internal in S. Then we interchange the order of summation and get the value of internal sum in terms of n. In such a way we get certain coefficients of the generating function which are in fact the values of S in dependence of n.

In solving problems of this type we usually encounter several sums. Here we will first list some of these sums.

The identity involving $\sum_n \binom{m}{n} x^n$ is known from before:

$$(1+x)^m = \sum_n \binom{m}{n} x^n.$$

Sometimes we will use the identity for $\sum_n \binom{n}{k} x^n$ which is already mentioned in the list of generating functions:

$$\frac{1}{(1-x)^{k+1}} = \sum_n \binom{n+k}{k} x^n.$$

Among the common sums we will encounter those involving only even (or odd) indeces. For example we have $(1+x)^m = \sum_n \binom{m}{n} x^n$, hence $(1-x)^m = \sum_n \binom{m}{n} (-x)^n$. Adding and subtracting yields:

$$\sum_n \binom{m}{2n} x^{2n} = \frac{((1+x)^m + (1-x)^m)}{2},$$

$$\sum_n \binom{m}{2n+1} x^{2n+1} = \frac{((1+x)^m - (1-x)^m)}{2}.$$

In a similar fashion we prove:

$$\sum_n \binom{2n}{m} x^{2n} = \frac{x^m}{2}\left(\frac{1}{(1-x)^{m+1}} + \frac{(-1)^m}{(1-x)^{m+1}}\right), \text{ and}$$

$$\sum_n \binom{2n+1}{m} x^{2n+1} = \frac{x^m}{2}\left(\frac{1}{(1-x)^{m+1}} - \frac{(-1)^m}{(1-x)^{m+1}}\right).$$

The following identity is also used quite frequently:

$$\sum_n \frac{1}{n+1}\binom{2n}{n} x^n = \frac{1}{2x}(1-\sqrt{1-4x}).$$

Problem 9. *Evaluate the sum*

$$\sum_k \binom{k}{n-k}.$$

Solution. Let n be the free variable and denote the sum by

$$f(n) = \sum_k \binom{k}{n-k}.$$

Let $F(x)$ be the generating function of the sequence $f(n)$, i.e.

$$F(x) = \sum_n x^n f(n) = \sum_n x^n \sum_k \binom{k}{n-k} = \sum_n \sum_k \binom{k}{n-k} x^n.$$

We can rewrite the previous equation as

$$F(x) = \sum_k \sum_n \binom{k}{n-k} x^n = \sum_k x^k \sum_n \binom{k}{n-k} x^{n-k},$$

which gives

$$F(x) = \sum_k x^k (1+x)^k = \sum_k (x+x^2)^k = \frac{1}{1-(x-x^2)} = \frac{1}{1-x-x^2}.$$

However this is very similar to the generating function of a Fibonacci's sequence, i.e. $f(n) = F_{n+1}$ and we arrive to

$$\sum_k \binom{k}{n-k} = F_{n+1}. \ \triangle$$

Problem 10. *Evaluate the sum*

$$\sum_{k=m}^{n}(-1)^k\binom{n}{k}\binom{k}{m}.$$

Solution. If n is a fixed number, then m is a free variable on which the sum depends. Let $f(m)=\sum_{k=m}^{n}(-1)^k\binom{n}{k}\binom{k}{m}$ and let $F(x)$ be the generating function of the sequence $f(m)$, i.e. $F(x)=\sum_m f(m)x^m$. Then we have

$$F(x)=\sum_m f(m)x^m=\sum_m x^m\sum_{k=m}^{n}(-1)^k\binom{n}{k}\binom{k}{m}=$$

$$=\sum_{k\leqslant n}(-1)^k\binom{n}{k}\sum_{m\leqslant k}\binom{k}{m}x^m=\sum_{k\leqslant n}\binom{n}{k}(1+x)^k.$$

Here we have used $\sum_{m\leqslant k}\binom{k}{m}x^m=(1+x)^k$. Dalje je

$$F(x)=(-1)^n\sum_{k\leqslant n}\binom{n}{k}(-1)^{n-k}(1+x)^k=(-1)^n\Big((1+x)-1\Big)^n=(-1)^nx^n$$

Therefore we obtained $F(x)=(-1)^nx^n$ and since this is a generating function of the sequence $f(m)$ we have

$$f(m)=\begin{cases}(-1)^n, & n=m\\ 0, & m<n\,.\end{cases}\ \triangle$$

Problem 11. *Evaluate the sum* $\sum_{k=m}^{n}\binom{n}{k}\binom{k}{m}$.

Solution. Let $f(m)=\sum_{k=m}^{n}\binom{n}{k}\binom{k}{m}$ and $F(x)=\sum_m x^m f(m)$. Then we have

$$F(x)=\sum_m x^m f(m)=\sum_m x^m\sum_{k=m}^{n}\binom{n}{k}\binom{k}{m}=\sum_{k\leqslant n}\binom{n}{k}\sum_{m\leqslant k}\binom{k}{m}x^m=\sum_{k\leqslant n}\binom{n}{k}(1+x)^k,$$

implying $F(x)=(2+x)^n$. Since

$$(2+x)^n=\sum_m\binom{n}{m}2^{n-m}x^m,$$

the value of the required sum is $f(m)=\binom{n}{m}2^{n-m}$. $\triangle$

Problem 12. *Evaluate*

$$\sum_k\binom{n}{\left[\frac{k}{2}\right]}x^k.$$

Solution. We can divide this into two sums

$$\sum_k\binom{n}{\left[\frac{k}{2}\right]}x^k=\sum_{k=2k_1}\binom{n}{\left[\frac{2k_1}{2}\right]}x^{2k_1}+\sum_{k=2k_2+1}\binom{n}{\left[\frac{2k_2+1}{2}\right]}x^{2k_2+1}=$$

$$=\sum_{k_1}\binom{n}{k_1}(x^2)^{k_1}+x\sum_{k_2}\binom{n}{k_2}(x^2)^{k_2}=(1+x^2)^n+x(1+x^2)^n,$$

or equivalently

$$\sum_k\binom{n}{\left[\frac{k}{2}\right]}x^k=(1+x)(1+x^2)^n.\ \triangle$$

Problem 13. *Determine the elements of the sequence:*

$$f(m)=\sum_k \binom{n}{k}\binom{n-k}{\left[\frac{m-k}{2}\right]}y^k.$$

Solution. Let $F(x)=\sum_m x^m f(m)$. We then have

$$F(x)=\sum_m x^m \sum_k \binom{n}{k}\binom{n-k}{\left[\frac{m-k}{2}\right]}y^k=\sum_k \binom{n}{k}y^k\sum_m \binom{n-k}{\left[\frac{m-k}{2}\right]}x^m=$$

$$=\sum_k \binom{n}{k}y^k x^k\sum_m \binom{n-k}{\left[\frac{m-k}{2}\right]}x^{m-k}=\sum_k \binom{n}{k}y^k x^k(1+x)(1+x^2)^{n-k}.$$

Hence

$$F(x)=(1+x)\sum_k \binom{n}{k}(1+x^2)^{n-k}(xy)^k=(1+x)(1+x^2+xy)^n.$$

For $y=2$ we have that $F(x)=(1+x)^{2n+1}$, implying that $F(x)$ is the generating function of the sequence $\binom{2n+1}{m}$ and we get the following combinatorial identity:

$$\sum_k \binom{n}{k}\binom{n-k}{\left[\frac{m-k}{2}\right]}2^k=\binom{2n+1}{m}.$$

Setting $y=-2$ we get $F(x)=(1+x)(1-x)^{2n}=(1-x)^{2n}+x(1-x)^{2n}$ hence the coefficient near x^m equals $\binom{2n}{m}(-1)^m+\binom{2n}{m-1}(-1)^{m-1}=(-1)^m\left[\binom{2n}{m}-\binom{2n}{m-1}\right]$ which implies

$$\sum_k \binom{n}{k}\binom{n-k}{\left[\frac{m-k}{2}\right]}(-2)^k=(-1)^m\left[\binom{2n}{m}-\binom{2n}{m-1}\right].\ \triangle$$

Problem 14. *Prove that*

$$\sum_k \binom{n}{k}\binom{k}{j}x^k=\binom{n}{j}x^j(1+x)^{n-j}$$

for each $n\geqslant 0$

Solution. If we fix n and let j be the free variable and $f(j)=\sum_k \binom{n}{k}\binom{k}{j}x^k$, $g(j)=\binom{n}{j}x^j(1+x)^{n-j}$, then the corresponding generating functions are

$$F(y)=\sum_j y^j f(j),\qquad G(y)=\sum_j y^j g(j).$$

We want to prove that $F(y)=G(y)$. We have

$$F(y)=\sum_j y^j\sum_k \binom{n}{k}\binom{k}{j}x^k=\sum_k \binom{n}{k}x^k\sum_j \binom{k}{j}y^j=\sum_k \binom{n}{k}x^k(1+y)^k,$$

hence $F(y)=(1+x+xy)^n$. On the other hand we have

$$G(y)=\sum_j y^j\binom{n}{j}x^j(1+x)^{n-j}=\sum_j \binom{n}{j}(1+x)^{n-j}(xy)^j=(1+x+xy)^n,$$

hence $F(y)=G(y)$. $\triangle$

The real power of the generating functions method can be seen in the following example.

Problem 15. *Evaluate the sum*

$$\sum_k \binom{n+k}{m+2k}\binom{2k}{k}\frac{(-1)^k}{k+1}$$

for $m,n \geqslant 0$.

Solution. Since there are quite a lot of variables elementary combinatorial methods doesn't offer an effective way to treat the sum. Since n appears on only one place in the sum, it is natural to consider the sum as a function on n Let $F(x)$ be the generating series of such functions. Then

$$F(x)=\sum_n x^n\sum_k \binom{n+k}{m+2k}\binom{2k}{k}\frac{(-1)^k}{k+1}=\sum_k \binom{2k}{k}\frac{(-1)^k}{k+1}x^{-k}\sum_n \binom{n+k}{m+2k}x^{n+k}=$$

$$=\sum_k \binom{2k}{k}\frac{(-1)^k}{k+1}x^{-k}\frac{x^{m+2k}}{(1-x)^{m+2k+1}}=\frac{x^{m+2k}}{(1-x)^{m+2k+1}}\sum_k \binom{2k}{k}\frac{1}{k+1}\left\{\frac{-x}{(1-x)^2}\right\}^k=$$

$$=\frac{-x^{m-1}}{2(1-x)^{m-1}}\left\{1-\sqrt{1+\frac{4x}{(1-x)^2}}\right\}=\frac{-x^{m-1}}{2(1-x)^{m-1}}\left\{1-\frac{1+x}{1-x}\right\}=\frac{x^m}{(1-x)^m}.$$

This is a generating function of the sequence $\binom{n-1}{m-1}$ which establishes

$$\sum_k \binom{n+k}{m+2k}\binom{2k}{k}\frac{(-1)^k}{k+1}=\binom{n-1}{m-1}. \ \triangle$$

Problem 16. *Prove the identity*

$$\sum_k \binom{2n+1}{k}\binom{m+k}{2n}=\binom{2m+1}{2n}.$$

Solution. Let $F(x)=\sum_m x^m\sum_k \binom{2n+1}{k}\binom{m+k}{2n}$ and $G(x)=\sum_m x^m\binom{2m+1}{2n}$ the generating functions of the expressions on the left and right side of the required equality. We will prove that $F(x)=G(x)$. We have

$$F(x)=\sum_m x^m\sum_k \binom{2n+1}{k}\binom{m+k}{2n}=\sum_k \binom{2n+1}{2k}\sum_m \binom{m+k}{2n}=$$

$$=\sum_k \binom{2n+1}{2k}\sum_m \binom{m+k}{2n}x^m=\sum_k \binom{2n+1}{2k}x^{-k}\sum_m \binom{m+k}{2n}x^{m+k}=$$

$$=\sum_k \binom{2n+1}{2k}x^{-k}\frac{x^{2n}}{(1-x)^{2n+1}}=\frac{x^{2n}}{(1-x)^{2n+1}}\sum_k \binom{2n+1}{2k}\left(x^{-\frac{1}{2}}\right)^{2k}.$$

We already know that $\sum_k \binom{2n+1}{2k}\left(x^{-\frac{1}{2}}\right)^{2k}=\frac{1}{2}\left(\left(1+\frac{1}{\sqrt{x}}\right)^{2n+1}+\left(1-\frac{1}{\sqrt{x}}\right)^{2n+1}\right)$ so

$$F(x)=\frac{1}{2}(\sqrt{x})^{2n-1}\left(\frac{1}{(1-\sqrt{x})^{2n+1}}-\frac{1}{(1+\sqrt{x})^{2n+1}}\right).$$

On the other hand

$$G(x)=\sum_m \binom{2m+1}{2n}x^m=\left(x^{-1/2}\right)\sum_m \binom{2m+1}{2n}\left(x^{1/2}\right)^{2m+1},$$

implying

$$G(x)=\left(x^{-1/2}\right)\left[\frac{(x^{1/2})^{2n}}{2}\left(\frac{1}{(1-x^{1/2})^{2n+1}}-(-1)^{2n}\frac{1}{(1+x^{1/2})^{2n+1}}\right)\right],$$

or

$$G(x)=\frac{1}{2}(\sqrt{x})^{2n-1}\left(\frac{1}{(1-\sqrt{x})^{2n+1}}-\frac{1}{(1+\sqrt{x})^{2n+1}}\right).\ \triangle$$

Problem 17. *Prove that*

$$\sum_{k=0}^{n}\binom{2n}{2k}\binom{2k}{k}2^{2n-2k}=\binom{4n}{2n}.$$

Let n be the free variable on the left and right side of $F(x)$ and $G(x)$. We want to prove the equality of these generating functions.

$$F(x)=\sum_n x^n\sum_{0\leqslant k\leqslant n}\binom{2n}{2k}\binom{2k}{k}2^{2n-2k}=\sum_{0\leqslant k}\binom{2k}{k}2^{-2k}\sum_n\binom{2n}{2k}x^n2^{2n},$$

$$F(x)=\sum_{0\leqslant k}\binom{2k}{k}2^{-2k}\sum_n\binom{2n}{2k}(2\sqrt{x})^{2n}.$$

Now we use the formula for summation of even powers and get

$$\sum_n\binom{2n}{2k}(2\sqrt{x})^{2n}=\frac{1}{2}(2\sqrt{x})^{2k}\left(\frac{1}{(1-2\sqrt{x})^{2k+1}}+\frac{1}{(1+2\sqrt{x})^{2k+1}}\right),$$

and we further get

$$F(x)=\frac{1}{2(1-2\sqrt{x})}\sum_k\binom{2k}{k}\left(\frac{x}{(1-2\sqrt{x})^2}\right)^k+\frac{1}{2(1+2\sqrt{x})}\sum_k\binom{2k}{k}\left(\frac{x}{(1+2\sqrt{x})^2}\right)^k.$$

Since $\sum_n\binom{2n}{n}x^n=\frac{1}{\sqrt{1-4x}}$ we get

$$F(x)=\frac{1}{2(1-2\sqrt{x})}\cdot\frac{1}{\sqrt{1-4\frac{x}{(1-2\sqrt{x})^2}}}+\frac{1}{2(1+2\sqrt{x})}\cdot\frac{1}{\sqrt{1-4\frac{x}{(1+2\sqrt{x})^2}}},$$

which implies

$$F(x)=\frac{1}{2\sqrt{1-4\sqrt{x}}}+\frac{1}{2\sqrt{1+4\sqrt{x}}}.$$

On the other hand for $G(x)$ we would like to get the sum $\sum_n\binom{4n}{2n}x^n$. Since $\sum_n\binom{2n}{n}x^n=\frac{1}{\sqrt{1-4x}}$ we have $\sum_n\binom{2n}{n}(-x)^n=\frac{1}{\sqrt{1+4x}}$ hence

$$G(x)=\frac{1}{2}\left(\frac{1}{\sqrt{1-4\sqrt{x}}}+\frac{1}{\sqrt{1+4\sqrt{x}}}\right)$$

and $F(x)=G(x)$. $\triangle$

The followng problem is slightly harder because the standard idea of snake oil doesn't lead to a solution.

Problem 18 (Moriati). *For given n and p evaluate*

$$\sum_k \binom{2n+1}{2p+2k+1}\binom{p+k}{k}.$$

Solution. In order to have shorter formulas let us introduce $r=p+k$. If we assume that n is the free variable then the required sum is equal to

$$f(n)=\sum_r \binom{2n+1}{2r+1}\binom{r}{p}.$$

Take $F(x)=\sum_n x^{2n+1}f(n)$. This is somehow natural since the binomial coefficient contains the term $2n+1$. Now we have

$$F(x)=\sum_n x^{2n+1}\sum_r \binom{2n+1}{2r+1}\binom{r}{p}=\sum_r \binom{r}{p}\sum_n \binom{2n+1}{2r+1}x^{2n+1}.$$

Since

$$\sum_n \binom{2n+1}{2r+1}x^{2n+1}=\frac{x^{2r+1}}{2}\left(\frac{1}{(1-x)^{2r+2}}+\frac{1}{(1+x)^{2r+2}}\right),$$

we get

$$F(x)=\frac{1}{2}\cdot\frac{x}{(1-x)^2}\sum_r \binom{r}{p}\left(\frac{x^2}{(1-x)^2}\right)^r+\frac{1}{2}\cdot\frac{x}{(1+x)^2}\sum_r \binom{r}{p}\left(\frac{x^2}{(1+x)^2}\right)^r,$$

$$F(x)=\frac{1}{2}\frac{x}{(1-x)^2}\frac{\left(\frac{x^2}{(1-x)^2}\right)^p}{\left(1-\frac{x^2}{(1-x)^2}\right)^{p+1}}+\frac{1}{2}\frac{x}{(1+x)^2}\frac{\left(\frac{x^2}{(1+x)^2}\right)^p}{\left(1-\frac{x^2}{(1+x)^2}\right)^{p+1}},$$

$$F(x)=\frac{1}{2}\frac{x^{2p+1}}{(1-2x)^{p+1}}+\frac{1}{2}\frac{x^{2p+1}}{(1+2x)^{p+1}}=\frac{x^{2p+1}}{2}((1+2x)^{-p-1}+(1-2x)^{-p-1}),$$

implying

$$f(n)=\frac{1}{2}\left(\binom{-p-1}{2n-2p}2^{2n-2p}+\binom{-p-1}{2n-2p}2^{2n-2p}\right),$$

and after simplification

$$f(n)=\binom{2n-p}{2n-2p}2^{2n-2p}.\ \triangle$$

We notice that for most of the problems we didn't make a substantial deviation from the method and we used only a handful of identities. This method can also be used in writing computer algorithms for symbolic evaluation of number of sums with binomial coefficients.

5 Problems

1. Prove that for the sequence of Fibonacci numbers we have

$$F_0+F_1+\cdots+F_n=F_{n+2}+1.$$

2. Given a positive integer n, let A denote the number of ways in which n can be partitioned as a sum of odd integers. Let B be the number of ways in which n can be partitioned as a sum of different integers. Prove that $A=B$.

3. Find the number of permutations without fixed points of the set $\{1,2,\ldots,n\}$.

4. Evaluate $\sum_k (-1)^k \binom{n}{3k}$.

5. Let $n \in \mathbb{N}$ and assume that

$$x+2y=n \quad \text{has } R_1 \text{ solutions in } \mathbb{N}_0^2$$
$$2x+3y=n-1 \quad \text{has } R_2 \text{ solutions in } \mathbb{N}_0^2$$
$$\vdots$$
$$nx+(n+1)y=1 \quad \text{has } R_n \text{ solutions in } \mathbb{N}_0^2$$
$$(n+1)x+(n+2)y=0 \quad \text{has } R_{n+1} \text{ solutions in } \mathbb{N}_0^2$$

Prove that $\sum_k R_k = n+1$.

6. A polynomial $f(x_1,x_2,\ldots,x_n)$ is called a *symmetric* if each permutation $\sigma \in S_n$ we have $f(x_{\sigma(1)},\ldots,x_{\sigma(n)}) = f(x_1,\ldots,x_n)$. We will consider several classes of symmetric polynomials. The first class consists of the polynomials of the form:

$$\sigma_k(x_1,\ldots,x_n) = \sum_{i_1<\cdots<i_k} x_{i_1}\cdot\cdots\cdot x_{i_k}$$

for $1 \leqslant k \leqslant n$, $\sigma_0 = 1$, and $\sigma_k = 0$ for $k > n$. Another class of symmetric polynomials are the polynomials of the form

$$p_k(x_1,\ldots,x_n) = \sum_{i_1+\cdots+i_n=k} x_1^{i_1}\cdot\cdots\cdot x_n^{i_n}, \qquad \text{where } i_1,\cdots,i_n \in \mathbb{N}_0.$$

The third class consists of the polynomials of the form:

$$s_k(x_1,\ldots,x_n) = x_1^k+\cdots+x_n^k.$$

Prove the following relations between the polynomials introduced above:

$$\sum_{r=0}^{n}(-1)^r\sigma_r p_{n-r} = 0,\ np_n = \sum_{r=1}^{n} s_r p_{n-r},\ \text{and } n\sigma_n = \sum_{r=1}^{n}(-1)^{r-1}s_r\sigma_{n-r}.$$

7. Assume that for some $n \in \mathbb{N}$ there are sequences of positive numbers a_1, a_2, …, a_n and b_1, b_2, …, b_n such that the sums

$$a_1+a_2,\ a_1+a_3,\ \ldots,\ a_{n-1}+a_n$$

and

$$b_1+b_2,\ b_1+b_3,\ \ldots,\ b_{n-1}+b_n$$

the same up to permutation. Prove that n is a power of two.

8. (Leo Moser, Joe Lambek, 1959.) Prove that there is a unique way to partition the set of natural numbers in two sets A and B such that: For very non-negative integer n (including 0) the number of ways in which n can be written as a_1+a_2, $a_1,a_2 \in A$, $a_1 \neq a_2$ is at least 1 and is equal to the number of ways in which it can be represented as b_1+b_2, $b_1,b_2 \in B$, $b_1 \neq b_2$.

9. Given several (at least two, but finitely many) arithmetic progressions, if each natural number belongs to exactly one of them, prove there are two progressions whose common differences are equal.

10. (This problem was posed in the journal *American Mathematical Monthly*) Prove that in the contemporary calendar the 13*th* in a month is most likely to be Friday.

Remark: The contemporary calendar has a period of 400 years. Every fourth year has 366 days except those divisible by 100 and not by 400.

6 Solultions

1. According to the Theorem 7 the generating function of the sum of first n terms of the sequence (i.e. the left-hand side) is equal to $F/(1-x)$, where $F = x/(1-x-x^2)$ (such F is the generatring function of the Fibonacci sequence). On the right-hand side we have

$$\frac{F-x}{x} - \frac{1}{1-x},$$

and after some obvious calculation we arrive to the required identity.

2. We will first prove that the generating function of the number of odd partitions is equal to

$$(1+x+x^2+\cdots)\cdot(1+x^3+x^6+\cdots)\cdot(1+x^5+x^{10}+\cdots)\cdots = \prod_{k\geqslant 1}\frac{1}{1-x^{2k+1}}.$$

Indeed, to each partition in which i occurs a_i times corresponds exactly one term with coefficient 1 in the product. That term is equal to $x^{1\cdot a_1+3\cdot a_3+5\cdot a_5+\cdots}$.

The generating function to the number of partitions in different summands is equal to

$$(1+x)\cdot(1+x^2)\cdot(1+x^3)\cdots = \prod_{k\geqslant 1}(1+x^k),$$

because from each factor we may or may not take a power of x, which exactly correpsonds to taking or not taking the corresponding summand of a partition. By some elementary transformations we get

$$\prod_{k\geqslant 1}(1+x^k) = \prod_{k\geqslant 1}\frac{1-x^{2k}}{1-x^k} = \frac{(1-x^2)(1-x^4)\cdots}{(1-x)(1-x^2)(1-x^3)(1-x^4)\cdots} = \prod_{k\geqslant 1}\frac{1}{1-x^{2k+1}}$$

which proves the statement.

3. This example illustrates the usefullness of the exponential generating functions. This problem is known as *derangement problem* or "le Problème des Rencontres" posed by Pierre R. de Montmort (1678-1719).

Assume that the required number is D_n and let $D(x) \overset{esr}{\leftrightarrow} D_n$. The number of permutations having exactly k given fixed points is equal to D_{n-k}, hence the total number of permutations with exactly k fixed points is equal to $\binom{n}{k}D_{n-k}$, because we can choose k fixed points in $\binom{n}{k}$ ways. Since the total number of permutations is equal to $n!$, then

$$n! = \sum_k \binom{n}{k} D_{n-k}$$

and the Theorem 10 gives

$$\frac{1}{1-x} = e^x D(x)$$

implying $D(x) = e^{-x}/(1-x)$. Since e^{-x} is the generating function of the sequence $\frac{(-1)^n}{n!}$, we get

$$\frac{D_n}{n!} = 1-1+\frac{1}{2!}-\frac{1}{3!}+\cdots+(-1)^n\frac{1}{n!},$$

$$D_n = n!\cdot\left(\frac{1}{2!}-\frac{1}{3!}+\cdots+(-1)^n\frac{1}{n!}\right).$$

4. The idea here is to consider the generating function

$$F(x)=\sum_k \binom{n}{3k}x^{3k}.$$

The required sum is equal to $f(-1)$. The question now is how to make binomial formula to skip all terms except those of order $3k$. We will use the following identy for the sum of roots of unity in the complex plane

$$\sum_{\varepsilon^r=1}\varepsilon^n=\begin{cases} r, & r|n \\ 0, & \text{otherwise.}\end{cases}$$

Let $C(x)=(1+x)^n$ and let 1, ε, and ε^2 be the cube roots of 1. Then we have

$$F(x)=\frac{C(x)+C(\varepsilon x)+C(\varepsilon^2 x)}{3}$$

which for $x=-1$ gives

$$F(-1)=\frac{1}{3}\left\{\left(\frac{3-i\sqrt{3}}{2}\right)^n+\left(\frac{3+i\sqrt{3}}{2}\right)^n\right\}$$

and after simplification

$$\sum_k(-1)^k\binom{n}{3k}=2\cdot 3^{n/2-1}\cos\left(\frac{n\pi}{6}\right)$$

5. The number of solutions of $x+2y=n$ in $\mathbb{N}_0^2$ is the coefficient near t^n in

$$(1+t+t^2+\cdots)\cdot(1+t^2+t^4+\cdots)=\frac{1}{1-t}\frac{1}{1-t^2}$$

The reason is that each pair (x,y) that satisfies the condition of the problem increases the coefficient near t^n by 1 because it appears as a summand of the form $t^x t^{2y}=t^{x+2y}$. More generally, the number of solutions of $kx+(k+1)y=n+1-k$ is the coefficient near t^{n+1-k} in $\frac{1}{1-t^k}\frac{1}{1-t^{k+1}}$, i.e. the coefficient near t^n in $\frac{x^{k-1}}{(1-t^k)(1-t^{k+1})}$. Hence, $\sum_{k=1}^{n}R_k$ is the coefficient near t^n in $\sum_k\frac{t^{k-1}}{(1-t^k)(1-t^{k+1})}=\sum_k\frac{1}{t-t^2}\left(\frac{1}{1-t^{k+2}}-\frac{1}{1-t^{k+1}}\right)=\frac{1}{(1-t)^2}$. Now it is easy to see that $\sum_k R_k=n+1$.

6. The generating function of the symmetric polynomials $\sigma_k(x_1,\ldots,x_n)$ is

$$\Sigma(t)=\sum_{k=0}^{\infty}\sigma_k t^k=\prod_{i=1}^{n}(1+tx_i).$$

The generating function of the polynomials $p_k(x_1,\ldots,x_n)$ is:

$$P(t)=\sum_{k=0}^{\infty}p_k t^k=\prod\frac{1}{1-tx_i},$$

and the generating function of the polynomials s_k is:

$$S(t)=\sum_{k=0}^{\infty}s_k t^{k-1}=\sum_{i=1}^{n}\frac{x_i}{1-tx_i}.$$

The functions $\Sigma(t)$ and $P(t)$ satisfy the following condition $\Sigma(t)p(-t)=1$. If we calculate the coefficient of this product near t^n, $n \geqslant 1$ we get the relation

$$\sum_{r=0}^{n}(-1)^r \sigma_r p_{n-r} = 0.$$

Notice that

$$\log P(t) = \sum_{i=1}^{n} \log \frac{1}{1-tx_i} \qquad \text{and} \qquad \log \Sigma(t) = \sum_{i=1}^{n} \log(1+tx_i).$$

Now we can express $S(t)$ in terms of $P(t)$ and $\Sigma(t)$ by:

$$S(t) = \frac{d}{dt}\log P(t) = \frac{P'(t)}{P(t)}$$

and

$$S(-t) = -\frac{d}{dt}\log \Sigma(t) = -\frac{\Sigma'(t)}{\Sigma(t)}.$$

From the first formula we get $S(t)P(t) = P'(t)$, and from the second $S(-t)\Sigma(t) = -\Sigma'(t)$. Comparing the coefficients near t^{n+1} we get

$$np_n = \sum_{r=1}^{n} s_r p_{n-r} \qquad \text{and} \qquad n\sigma_n = \sum_{r=1}^{n}(-1)^{r-1} s_r \sigma_{n-r}.$$

7. Let F and G be polynomials generated by the given sequence: $F(x) = x^{a_1} + x^{a_2} + \cdots + x^{a_n}$ and $G(x) = x^{b_1} + x^{b_2} + \cdots + x^{b_n}$. Then

$$\begin{aligned} F^2(x) - G^2(x) &= \left(\sum_{i=1}^{n} x^{2a_i} + 2\sum_{1\leqslant i\leqslant j\leqslant n} x^{a_i+a_j}\right) - \left(\sum_{i=1}^{n} x^{2b_i} + 2\sum_{1\leqslant i\leqslant j\leqslant n} x^{b_i+b_j}\right) \\ &= F(x^2) - G(x^2). \end{aligned}$$

Since $F(1) = G(1) = n$, we have that 1 is zero of the order $k, (k \geqslant 1)$ of the polynomial $F(x) - G(x)$. Then we have $F(x) - G(x) = (x-1)^k H(x)$, hence

$$F(x) + G(x) = \frac{F^2(x) - G^2(x)}{F(x) - G(x)} = \frac{F(x^2) - G(x^2)}{F(x) - G(x)} = \frac{(x^2-1)^k H(x^2)}{(x-1)^k H(x)} = (x+1)^k \frac{H(x^2)}{H(x)}$$

Now for $x = 1$ we have:

$$2n = F(1) + G(1) = (1+1)^k \frac{H(x^2)}{H(x)} = 2^k,$$

implying that $n = 2^{k-1}$.

8. Consider the polynomials generated by the numbers from different sets:

$$A(x) = \sum_{a\in A} x^a, \quad B(x) = \sum_{b\in B} x^b.$$

The condition that A and B partition the whole $\mathbb{N}$ without intersection is equivalent to

$$A(x) + B(x) = \frac{1}{1-x}.$$

The number of ways in which some number can be represented as $a_1+a_2,\ \ a_1,a_2\in A,\ \ a_1\neq a_2$ has the generating function:

$$\sum_{a_i,a_j\in A,a_i\neq a_j} x^{a_i+a_j}=\frac{1}{2}\left(A^2(x)-A(x^2)\right).$$

Now the second condition can be expressed as

$$\left(A^2(x)-A(x^2)\right)=\left(B^2(x)-B(x^2)\right).$$

We further have

$$(A(x)-B(x))\frac{1}{1-x}=A(x^2)-B(x^2)$$

or equivalently

$$(A(x)-B(x))=(1-x)(A(x^2)-B(x^2)).$$

Changing x by $x^2,x^4,\ldots,x^{2^{n-1}}$ we get

$$A(x)-B(x)=(A(x^{2^n})-B(x^{2^n}))\prod_{i=0}^{n-1}(1-x^{2^i}),$$

implying

$$A(x)-B(x)=\prod_{i=0}^{\infty}(1-x^{2^i}).$$

The last product is series whose coefficients are ±1 hence A and B are uniquely determined (since their coefficients are 1). It is not difficult to notice that positive coefficients (i.e. coefficients originating from A) are precisely those corresponding to the terms x^n for which n can be represented as a sum of even numbers of 2s. This means that the binary partition of n has an even number of 1s. The other numbers form B.

Remark: The sequence representing the parity of the number of ones in the binary representation of n is called *Morse* sequence.

9. This problem is posed by Erdösz (in slightly different form), and was solved by Mirsky and Newman after many years. This is their original proof:

   Assume that k arithmetic progressionss $\{a_i+nb_i\}\ \ (i=1,2,\ldots,k)$ cover the entire set of positive integers. Then $\dfrac{z^a}{1-z^b}=\sum_{i=0}^{\infty}z^{a+ib}$, hence

   $$\frac{z}{1-z}=\frac{z^{a_1}}{1-z^{b_1}}+\frac{z^{a_2}}{1-z^{b_2}}+\cdots+\frac{z^{a_k}}{1-z^{b_k}}.$$

   Let $|z|\leqslant 1$. We will prove that the biggest number among b_i can't be unique. Assume the contrary, that b_1 is the greatest among the numbers $b_1,b_2,\ldots,b_n$ and set $\varepsilon=e^{2i\pi/b_1}$. Assume that z approaches ε in such a way that $|z|\leqslant 1$. Here we can choose ε such that $\varepsilon^{b_1}=1$, $\varepsilon\neq 1$, and $\varepsilon^{b_i}\neq 1,\ \ 1<i\leqslant k$. All terms except the first one converge to certain number while the first converges to ∞, which is impossible.

10. Friday the 13th corresponds to Sunday the 1st. Denote the days by numbers $1,2,3,\ldots$ and let t^i corresponds to the day i. Hence, $Jan.1st2001$ is denoted by 1 (or t), $Jan.4th2001$ by t^4 etc. Let A be the set of all days (i.e. corresponding numbers) which happen to be the first in a month. For instance, $1\in A$, $2\in A$, etc. $A=\{1,32,60,\ldots\}$. Let $f_A(t)=\sum_{n\in A}t^n$. If we replace t^{7k} by 1, t^{7k+1} by t, t^{7k+2} by t^2 etc. in the polynomial f_A we get another polynomial – denote it by $g_A(t)=\sum_{i=0}^{6}a_it^i$. Now the number a_i represents how many times the day (of a week)

denoted by i has appeared as the first in a month. Since $Jan1,2001$ was Monday, a_1 is the number of Mondays, a_2- the number of Tuesdays, ..., a_0- the number of Sundays. We will consider now f_A modulus t^7-1. The polynoimal $f_A(t)-g_A(t)$ is divisible by t^7-1. Since we only want to find which of the numbers $a_0,a_1,\ldots,a_6$ is the biggest, it is enough to consider the polynomial modulus $q(t)=1+t+t^2+\cdots+t^6$ which is a factor of t^7-1. Let $f_1(t)$ be the polynomial that represents the first days of months in 2001. Since the first day of January is Monday, Thursday– the first day of February, ..., Saturday the first day of December, we get

$$f_1(t)=t+t^4+t^4+1+t^2+t^5+1+t^3+t^6+t+t^4+t^6=$$

$$=2+2t+t^2+t^3+3t^4+t^5+2t^6\equiv 1+t+2t^4+t^6 \pmod{q(t)}.$$

Since the common year has $365\equiv 1 \pmod 7$ days, polynomials $f_2(t)$ and $f_3(t)$ corresponding to 2002. and 2003., satisfy

$$f_2(t)\equiv t\,f_1(t)\equiv t g_1(t)$$

and

$$f_3(t)\equiv t\,f_2(t)\equiv t^2 g_1(t),$$

where the congruences are modulus $q(t)$. Using plain counting we easily verify that $f_4(t)$ for leap 2004 is

$$f_4(t)=2+2t+t^2+2t^3+3t^4+t^5+t^6\equiv 1+t+t^3+2t^4=g_4(t).$$

We will introduce a new polynomial that will count the first days for the period $2001-2004$ $h_1(t)=g_1(t)(1+t+t^2)+g_4(t)$. Also, after each common year the days are shifted by one place, and after each leap year by 2 places, hence after the period of 4 years all days are shifted by 5 places. In such a way we get a polynomial that counts the numbers of first days of months between 2001 and 2100. It is:

$$p_1(t)=h_1(t)(1+t^5+t^{10}+\cdots+t^{115})+t^{120}g_1(t)(1+t+t^2+t^3).$$

Here we had to write the last for years in the form $g_1(t)(1+t+t^2+t^3)$ because 2100 is not leap, and we can't replace it by $h_1(t)$. The period of 100 years shifts the calendar for 100 days (common years) and additional 24 days (leap) which is congruent to 5 modulus 7. Now we get

$$g_A(t)\equiv p_1(t)(1+t^5+t^{10})+t^{15}h_1(t)(1+t^5+\cdots+t^{120}).$$

Similarly as before the last 100 are counted by last summands because 2400 is leap. Now we will use that $t^{5a}+t^{5(a+1)}+\cdots+t^{5(a+6)}\equiv 0$. Thus $1+t^5+\cdots+t^{23\cdot 5}\equiv 1+t^5+t^{2\cdot 5}\equiv 1+t^3+t^5$ and $1+t^5+\cdots+t^{25\cdot 5}\equiv 1+t^5+t^{2\cdot 5}+t^{4\cdot 5}\equiv 1+t+t^3+t^5$. We further have that

$$p_1(t)\equiv h_1(t)(1+t^3+t^5)+tg_1(t)(1+t+t^2+t^3)\equiv$$

$$g_1(t)[(1+t+t^2)(1+t^3+t^5)+t(1+t+t^2+t^3)]+g_4(t)(1+t^3+t^5)\equiv$$

$$g_1(t)(2+2t+2t^2+2t^3+2t^4+2t^5+t^6)+g_4(t)(1+t^3+t^5)\equiv -g_1(t)t^6+g_4(t)(1+t^3+t^5).$$

If we now put this into formula for $g_A(t)$ we get

$$\begin{aligned}
g_A(t) &\equiv p_1(t)(1+t^3+t^5)+th_1(t)(1+t+t^3+t^5)\\
&\equiv -g_1(t)t^6(1+t^3+t^5)+g_4(t)(1+t^3+t^5)^2\\
&\quad +tg_1(t)(1+t+t^2)(1+t+t^3+t^5)+tg_4(t)(1+t+t^3+t^5)\\
&\equiv g_1(t)(t+t^3)+g_4(t)(2t+2t^3+t^5+t^6)\\
&\equiv (1+t+2t^4+t^6)(t+t^3)+(1+t+t^3+2t^4)(2t+2t^3+t^5+t^6)\\
&\equiv 8+4t+7t^2+5t^3+5t^4+7t^5+4t^6\equiv 4+3t^2+t^3+t^4+3t^5.
\end{aligned}$$

This means that the most probable day for the first in a month is Sunday (because a_0 is the biggest).

We can precisely determine the probability. If we use the fact that there are 4800 months in a period of 400, we can easily get the Sunday is the first exactly 688 times, Monday – 684, Tuesday – 687, Wednesday – 685, Thursday – 685, Friday – 687, and Saturday – 684.

References

[1] H.S. Wilf, 1994, *generatingfunctionology* , University of Pennsylvania, Philadelphia, USA
Electronic version: http://www.math.upenn.edu/∼wilf/gfologyLinked.pdf

[2] Д. Стевановић, М. Милошевић, В. Балтић, 2004, *Дискретна математика, основе комбинаторике и теорије графова - збирка решених задатака*, Друштво математичара Србије, Београд (у припреми)
Electronic version: http://www.pmf.ni.ac.yu/people/dragance/zbirka.pdf

[3] В. Дренски, ,,Пораждащи функции“ у *Подготовка за олимпиади под редакцията на Сава Гроздев*, 2002, Съюз на математиците в България, София

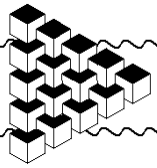
[4] Electronic materials: http://www.math.uvic.ca/faculty/gmacgill/guide/GenFuncs.pdf

[5] Lerma, A.M., 2003., *Generating functions* ,
http://www.math.northwestern.edu/ mlerma/problem_solving/results/gen_func.pdf

[6] Electronic materials: http://www.cs.brandeis.edu/∼ira/47a/gf.pdf

[7] П. Младеновић, *Комбинаторика*, Друштво математичара Србије, Београд, 2001

terug naar echt bestand



©2007 The Author(s) and The IMO Compendium Group

Polynomial Equations

Dušan Djukić

Contents

1 Introduction

The title refers to determining polynomials in one or more variables (e.g. with real or complex coefficients) which satisfy some given relation(s).

The following example illustrates some basic methods:

1. Determine the polynomials P for which $16P(x^2) = P(2x)^2$.

- *First method: evaluating at certain points and reducing degree.*

  Plugging $x = 0$ in the given relation yields $16P(0) = P(0)^2$, i.e. $P(0) = 0$ or 16.

  (i) Suppose that $P(0) = 0$. Then $P(x) = xQ(x)$ for some polynomial Q and $16x^2Q(x^2) = 4x^2Q(2x)^2$, which reduces to $4Q(x^2) = Q(2x)^2$. Now setting $4Q(x) = R(x)$ gives us $16R(x^2) = R(2x)^2$. Hence, $P(x) = \frac{1}{4}xR(x)$, with R satifying the same relation as P.

  (ii) Suppose that $P(0) = 16$. Putting $P(x) = xQ(x) + 16$ in the given relation we obtain $4xQ(x^2) = xQ(2x)^2 + 16Q(2x)$; hence $Q(0) = 0$, i.e. $Q(x) = xQ_1(x)$ for some polynomial Q_1. Furthermore, $x^2Q_1(x^2) = x^2Q_1(2x)^2 + 8Q_1(2x)$, implying that $Q_1(0) = 0$, so Q_1 too is divisible by x. Thus $Q(x) = x^2Q_1(x)$. Now suppose that x^n is the highest degree of x dividing Q, and $Q(x) = x^nR(x)$, where $R(0) \neq 0$. Then R satisfies $4x^{n+1}R(x^2) = 2^{2n}x^{n+1}R(2x)^2 + 2^{n+4}R(2x)$, which implies that $R(0) = 0$, a contradiction. It follows that $Q \equiv 0$ and $P(x) \equiv 16$.

  We conclude that $P(x) = 16\left(\frac{1}{4}x\right)^n$ for some $n \in \mathbb{N}_0$.

- *Second method: investigating coefficients.*

  We start by proving the following lemma (to be used frequently):

  Lemma 1. *If $P(x)^2$ is a polynomial in x^2, then so is either $P(x)$ or $P(x)/x$.*

  Proof. Let $P(x) = a_nx^n + a_{n-1}x^{n-1} + \cdots + a_0$, $a_n \neq 0$. The coefficient at x^{2n-1} is $2a_na_{n-1}$, from which we get $a_{n-1} = 0$. Now the coefficient at x^{2n-3} equals $2a_na_{n-3}$; hence $a_{n-3} = 0$, and so on. Continuing in this manner we conclude that $a_{n-2k-1} = 0$ for $k = 0, 1, 2, \ldots$, i.e. $P(x) = a_nx^n + a_{n-2}x^{n-2} + a_{n-4}x^{n-4} + \cdots$. $\triangle$

  Since $P(x)^2 = 16P(x^2/4)$ is a polynomial in x^2, we have $P(x) = Q(x^2)$ or $P(x) = xQ(x^2)$. In the former case we get $16Q(x^4) = Q(4x^2)^2$ and therefore $16Q(x^2) = Q(4x)^2$; in the latter case

we similarly get $4Q(x^2) = Q(4x)^2$. In either case, $Q(x) = R(x^2)$ or $Q(x) = xR(x^2)$ for some polynomial R, so $P(x) = x^i R(x^4)$ for some $i \in \{0,1,2,3\}$. Proceeding in this way we find that $P(x) = x^i S(x^{2^k})$ for each $k \in \mathbb{N}$ and some $i \in \{0,1,\ldots,2^k\}$. Now it is enough to take k with $2^k > \deg P$ and to conclude that S must be constant. Thus $P(x) = cx^i$ for some $c \in \mathbb{R}$. A simple verification gives us the general solution $P(x) = 16\left(\frac{1}{4}x\right)^n$ for $n \in \mathbb{N}_0$.

Investigating zeroes of the unknown polynomial is also counted under the first method.

A majority of problems of this type can be solved by one of the above two methods (although some cannot, making math more interesting!).

2 Problems with Solutions

1. Find all polynomials P such that $P(x)^2 + P(\frac{1}{x})^2 = P(x^2)P(\frac{1}{x^2})$.

   Solution. By the introducing lemma there exists a polynomial Q such that $P(x) = Q(x^2)$ or $P(x) = xQ(x^2)$. In the former case $Q(x^2)^2 + Q(\frac{1}{x^2}) = Q(x^4)Q(\frac{1}{x^4})$ and therefore $Q(x)^2 + Q(\frac{1}{x}) = Q(x^2)Q(\frac{1}{x^2})$ (which is precisely the relation fulfilled by P), whereas in the latter case (similarly) $xQ(x)^2 + \frac{1}{x}Q(\frac{1}{x})^2 = Q(x^2)Q(\frac{1}{x^2})$ which is impossible for the left and right hand side have odd and even degrees, respectively. We conclude that $P(x) = Q(x^2)$, where Q is also a solution of the considered polynomial equation. Considering the solution of the least degree we find that P must be constant.

2. Do there exist non-linear polynomials P and Q such that $P(Q(x)) = (x-1)(x-2)\cdots(x-15)$?

   Solution. Suppose there exist such polynomials. Then $\deg P \cdot \deg Q = 15$, so $\deg P = k \in \{3,5\}$. Putting $P(x) = c(x-a_1)\cdots(x-a_k)$ we get $c(Q(x)-a_1)\cdots(Q(x)-a_k) = (x-1)(x-2)\cdots(x-15)$. Thus the roots of polynomial $Q(x) - a_i$ are distinct and comprise the set $\{1,2,\ldots,15\}$. All these polynomials mutually differ at the last coefficient only. Now, investigating parity of the remaining (three or five) coefficients we conclude that each of them has the equally many odd roots. This is impossible, since the total number of odd roots is 8, not divisible by 3 or 5.

3. Determine all polynomials P for which $P(x)^2 - 2 = 2P(2x^2-1)$.

   Solution. Denote $P(1) = a$. We have $a^2 - 2a - 2 = 0$. Substituting $P(x) = (x-1)P_1(x) + a$ in the initial relation and simplifying yields $(x-1)P_1(x)^2 + 2aP_1(x) = 4(x+1)P_1(2x^2-1)$. For $x = 1$ we have $2aP_1(1) = 8P_1(1)$, which (since $a \neq 4$) gives us $P_1(1) = 0$, i.e. $P_1(x) = (x-1)P_2(x)$, so $P(x) = (x-1)^2P_2(x) + a$. Suppose that $P(x) = (x-1)^nQ(x) + a$, where $Q(1) \neq 0$. Substituting in the initial relation and simplifying yields $(x-1)^nQ(x)^2 + 2aQ(x) = 2(2x+2)^nQ(2x^2-1)$, giving us $Q(1) = 0$, a contradiction. It follows that $P(x) = a$.

4. Determine all polynomials P for which $P(x)^2 - 1 = 4P(x^2 - 4x + 1)$.

   Solution. Suppose that P is not constant. Fixing $\deg P = n$ and comparing coefficients of both sides we deduce that the coefficients of polynomial P must be rational. On the other hand, setting $x = a$ with $a = a^2 - 4a + 1$, that is, $a = \frac{5\pm\sqrt{21}}{2}$, we obtain $P(a) = b$, where $b^2 - 4b - 1 = 0$, i.e. $b = 2 \pm \sqrt{5}$. However, this is impossible because $P(a)$ must be of the form $p + q\sqrt{21}$ for some rational p, q for the coefficients of P are rational. It follows that $P(x)$ is constant.

5. For which real values of a does there exist a rational function $f(x)$ that satisfies $f(x^2) = f(x)^2 - a$?

   Solution. Write f as $f = P/Q$ with P and Q coprime polynomials and Q monic. By comparing leading coefficients we obtain that P too is monic. The condition of the problem became $P(x^2)/Q(x^2) = P(x)^2/Q(x)^2 - a$. Since $P(x^2)$ and $Q(x^2)$ are coprime (if, to the contrary, they

had a zero in common, then so do P and Q), it follows that $Q(x^2) = Q(x)^2$. Therefore $Q(x) = x^n$ for some $n \in \mathbb{N}$. Now we have $P(x^2) = P(x)^2 - ax^{2n}$.

Let $P(x) = a_0 + a_1x + \cdots + a_{m-1}x^{m-1} + x^m$. Comparing coefficients of $P(x)^2$ and $P(x^2)$ gives us $a_{n-1} = \cdots = a_{2m-n+1} = 0$, $a_{2m-n} = a/2$, $a_1 = \cdots = a_{m-1} = 0$ and $a_0 = 1$. This is only possible if $a = 2$ and $2m - n = 0$, or $a = 0$.

6. Find all polynomials P satisfying $P(x^2+1) = P(x)^2 + 1$ for all x.

   Solution. By the introducing lemma, there is a polynomial Q such that $P(x) = Q(x^2+1)$ or $P(x) = xQ(x^2+1)$. Then $Q((x^2+1)^2+1) = Q(x^2+1)^2 - 1$ or $(x^2+1)Q((x^2+1)^2+1) = x^2Q(x^2+1)^2 + 1$, respectively. Substituting $x^2+1 = y$ yields $Q(y^2+1) = Q(y)^2 + 1$ and $yQ(y^2+1) = (y-1)Q(y)^2 + 1$, respectively.

   Suppose that $yQ(y^2+1) = (y-1)Q(y)^2 + 1$. Setting $y = 1$ we obtain that $Q(2) = 1$. Note that, if $a \neq 0$ and $Q(a) = 1$, then also $aQ(a^2+1) = (a-1) + 1$ and hence $Q(a^2+1) = 1$. We thus obtain an infinite sequence of points at which Q takes value 1, namely the sequence given by $a_0 = 2$ and $a_{n+1} = a_n^2 + 1$. Therefore $Q \equiv 1$.

   It follows that if $Q \not\equiv 1$, then $P(x) = Q(x^2+1)$. Now we can easily list all solutions: these are the polynomials of the form $T(T(\cdots(T(x))\cdots))$, where $T(x) = x^2 + 1$.

7. If a polynomial P with real coefficients satisfies for all x

$$P(\cos x) = P(\sin x),$$

   prove that there exists a polynomial Q such that for all x, $P(x) = Q(x^4 - x^2)$.

   Solution. It follows from the condition of the problem that $P(-\sin x) = P(\sin x)$, so $P(-t) = P(t)$ for infinitely many t; hence the polynomials $P(x)$ and $P(-x)$ coincide. Therefore $P(x) = S(x^2)$ for some polynomial S. Now $S(\cos^2 x) = S(\sin^2 x)$ for all x, so $S(1-t) = S(t)$ for infinitely many t, which gives us $S(x) \equiv S(1-x)$. This is equivalent to $R(x - \frac{1}{2}) = R(\frac{1}{2} - x)$, i.e. $R(y) \equiv R(-y)$, where R is the polynomial such that $S(x) = R(x - \frac{1}{2})$. Now $R(x) = T(x^2)$ for some polynomial T, and finally, $P(x) = S(x^2) = R(x^2 - \frac{1}{2}) = T(x^4 - x^2 + \frac{1}{4}) = Q(x^4 - x^2)$ for some polynomial Q.

8. Find all quadruples of polynomials (P_1, P_2, P_3, P_4) such that, whenever natural numbers x, y, z, t satisfy $xy - zt = 1$, it holds that $P_1(x)P_2(y) - P_3(z)P_4(t) = 1$.

   Solution. Clearly $P_1(x)P_2(y) = P_2(x)P_1(y)$ for all natural numbers x and y. This implies that $P_2(x)/P_1(x)$ does not depend on x. Hence $P_2 = cP_1$ for some constant c. Analogously, $P_4 = dP_3$ for some constant d. Now we have $cP_1(x)P_1(y) - dP_3(z)P_3(t) = 1$ for all natural x, y, z, t with $xy - zt = 1$. Moreover, we see that $P_1(x)P_1(y)$ depends only on xy, i.e. $f(x) = P_1(x)P_1(n/x)$ is the same for all positive divisors x of natural number n. Since $f(x)$ is a rational function and the number of divisors x of n can be arbitrarily large, it follows that f is constant in x, i.e. a polynomial in n. It is easily verified that this is possible only when $P_1(x) = x^n$ for some n. Similarly, $P_3(x) = x^m$ for some m and $c(xy)^n - d(zt)^m = 1$. Therefore $m = n$ and $c = d = 1$, and finally $m = n = 1$. So, $P_1(x) = P_2(x) = P_3(x) = P_4(x) = x$.

9. Find all polynomials $P(x)$ with real coefficients that satisfy the equality

$$P(a-b) + P(b-c) + P(c-a) = 2P(a+b+c)$$

   for all triples a, b, c of real numbers such that $ab + bc + ca = 0$. *(IMO 2004.2)*

   Solution. Let $P(x) = a_0 + a_1x + \cdots + a_nx^n$. For every $x \in \mathbb{R}$ the triple $(a, b, c) = (6x, 3x, -2x)$ satisfies the condition $ab + bc + ca = 0$. Then the condition on P gives us $P(3x) + P(5x) + P(-8x) = 2P(7x)$ for all x, implying that for all $i = 0, 1, 2, \ldots, n$ the following equality holds:

$$\left(3^i + 5^i + (-8)^i - 2 \cdot 7^i\right) a_i = 0.$$

Suppose that $a_i \neq 0$. Then $K(i) = 3^i + 5^i + (-8)^i - 2 \cdot 7^i = 0$. But $K(i)$ is negative for i odd and positive for $i = 0$ or $i \geq 6$ even. Only for $i = 2$ and $i = 4$ do we have $K(i) = 0$. It follows that $P(x) = a_2x^2 + a_4x^4$ for some real numbers a_2, a_4. It is easily verified that all such $P(x)$ satisfy the required condition.

10. (a) If a real polynomial $P(x)$ satisfies $P(x) \geq 0$ for all x, show that there exist real polynomials $A(x)$ and $B(x)$ such that $P(x) = A(x)^2 + B(x)^2$.

    (b) If a real polynomial $P(x)$ satisfies $P(x) \geq 0$ for all $x \geq 0$, show that there exist real polynomials $A(x)$ and $B(x)$ such that $P(x) = A(x)^2 + xB(x)^2$.

    Solution. Polynomial $P(x)$ can be written in the form

$$P(x) = (x-a_1)^{\alpha_1} \cdots (x-a_k)^{\alpha_k} \cdot (x^2 - b_1x + c_1) \cdots (x^2 - b_mx + c_m), \qquad (*)$$

    where a_i, b_j, c_j are real numbers such that a_i are distinct and the polynomials $x^2 - b_ix + c_i$ have no real roots.

    It follows from the condition $P(x) \geq 0$ for all x that all the α_i are even, and from the condition $P(x) \geq 0$ for all $x \geq 0$ that $(\forall i)$ either α_i is even or $a_i < 0$. This ensures that each linear or quadratic factor in $(*)$ can be written in the required form $A^2 + B^2$ and/or $A^2 + xB^2$. The well-known formula $(a^2 + \gamma b^2)(c^2 + \gamma d^2) = (ac + \gamma bd)^2 + \gamma(ad - bc)^2$ now gives a required representation for their product $P(x)$.

11. Prove that if the polynomials P and Q have a real root each and

$$P(1 + x + Q(x)^2) = Q(1 + x + P(x)^2),$$

    then $P \equiv Q$.

    Solution. Note that there exists $x = a$ for which $P(a)^2 = Q(a)^2$. This follows from the fact that, if p and q are the respective real roots of P and Q, then $P(p)^2 - Q(p)^2 \leq 0 \leq P(q)^2 - Q(q)^2$, and moreover $P^2 - Q^2$ is continuous. Now $P(b) = Q(b)$ for $b = 1 + a + P(a)^2$. Taking a to be the largest real number for which $P(a) = Q(a)$ leads to an immediate contradiction.

12. If P and Q are monic polynomials with $P(P(x)) = Q(Q(x))$, prove that $P \equiv Q$.

    Solution. Suppose that $R = P - Q \neq 0$ and that $0 < k \leq n - 1$ is the degree of $R(x)$. Then

$$P(P(x)) - Q(Q(x)) = [Q(P(x)) - Q(Q(x))] + R(P(x)).$$

    Putting $Q(x) = x^n + \cdots + a_1x + a_0$ we have $Q(P(x)) - Q(Q(x)) = [P(x)^n - Q(x)^n] + \cdots + a_1[P(x) - Q(x)]$, where all summands but the first have a degree at most $n^2 - n$, while the first summand equals $R(x) \cdot \left(P(x)^{n-1} + P(x)^{n-2}Q(x) + \cdots + Q(x)^{n-1}\right)$ and therefore has the degree $n^2 - n + k$ with the leading coefficient n. Hence the degree of $Q(P(x)) - Q(Q(x))$ is $n^2 - n + k$. The degree of $R(P(x))$ is equal to $kn < n^2 - n + k$, from what we conclude that the degree of the difference $P(P(x)) - Q(Q(x))$ is $n^2 - n + k$, a contradiction.

    In the remaining case when $R \equiv c$ is constant, the condition $P(P(x)) = Q(Q(x))$ gives us $Q(Q(x) + c) = Q(Q(x)) - c$, so the equality $Q(y + c) = Q(y) - c$ holds for infinitely many y, implying $Q(y + c) \equiv Q(y) - c$. But this is only possible for $c = 0$.

13. Assume that there exist complex polynomials P, Q, R such that

$$P^a + Q^b = R^c,$$

    where a, b, c are natural numbers. Show that $\frac{1}{a} + \frac{1}{b} + \frac{1}{c} > 1$.

    Solution. We use the following auxilliary statement.

Lemma 2. *If A, B and C are pairwise coprime polynomials with $A+B=C$, then the degree of each of them is less than the number of different zeroes of the polynomial ABC.*

Proof. Let

$$A(x)=\prod_{i=1}^{k}(x-p_i)^{a_i},\quad B(x)=\prod_{i=1}^{l}(x-q_i)^{b_i},\quad C(x)=\prod_{i=1}^{m}(x-r_i)^{c_i}.$$

Writing the condition $A+B=C$ as $A(x)C(x)^{-1}+B(x)C(x)^{-1}=1$ and differentiating it with respect to x gives us

$$A(x)C(x)^{-1}\left(\sum_{i=1}^{k}\frac{a_i}{x-p_i}-\sum_{i=1}^{m}\frac{c_i}{x-r_i}\right)=-B(x)C(x)^{-1}\left(\sum_{i=1}^{l}\frac{b_i}{x-q_i}-\sum_{i=1}^{m}\frac{c_i}{x-r_i}\right),$$

from which we see that $A(x)/B(x)$ can be written as a quotient of two polynomials od degrees not exceeding $k+l+m-1$. Our statement now follows from the fact that A and B are coprime. Apply this statement on polynomials P^a, Q^b, R^c. Each of their degrees $a\deg P$, $b\deg Q$, $c\deg R$ is less than $\deg P+\deg Q+\deg R$ and hence $\frac{1}{a}>\frac{\deg P}{\deg P+\deg Q+\deg R}$, etc. Summing up yields the desired inequality.

Corollary. "The Last Fermat's theorem" for polynomials.

14. The lateral surface of a cylinder is divided by $n-1$ planes parallel to the base and m meridians into mn cells ($n\geq 1$, $m\geq 3$). Two cells are called neighbors if they have a common side. Prove that it is possible to write real numbers in the cells, not all zero, so that the number in each cell equals the sum of the numbers in the neighboring cells, if and only if there exist k, l with $n+1\nmid k$ such that $\cos\dfrac{2l\pi}{m}+\cos\dfrac{k\pi}{n+1}=\dfrac{1}{2}$.

Solution. Denote by a_{ij} the number in the intersection of i-th parallel and j-th meridian. We assign to the i-th parallel the polynomial $p_i(x)=a_{i1}+a_{i2}x+\cdots+a_{im}x^{m-1}$ and define $p_0(x)=p_{n+1}(x)=0$. The property that each number equals the sum of its neighbors can be written as $p_i(x)=p_{i-1}(x)+p_{i+1}(x)+(x^{m-1}+x)p_i(x)$ modulo x^m-1, i.e.

$$p_{i+1}(x)=(1-x-x^{m-1})p_i(x)-p_{i-1}(x)\pmod{x^m-1}.$$

This sequence of polynomials is entirely determined by term $p_1(x)$. The numbers a_{ij} can be written in the required way if and only if a polynomial $p_1(x)\neq 0$ of degree less than m can be chosen so that $p_{n+1}(x)=0$.

Consider the sequence of polynomials $r_i(x)$ given by $r_0=0$, $r_1=1$ and $r_{i+1}=(1-x-x^{m-1})r_i-r_{i-1}$. Clearly, $p_{n+1}(x)\equiv r_{n+1}(x)p_1(x)\pmod{x^m-1}$. Polynomial $p_1\neq 0$ of degree $<m$ for which $p_{n+1}=0$ exists if and only if $r_{n+1}(x)$ and x^m-1 are not coprime, i.e. if and only if there exists ε such that $\varepsilon^m=1$ and $r_{n+1}(\varepsilon)=0$. Now consider the sequence (x_i) given by $x_0=0$, $x_1=1$ and $x_{i+1}=(1-\varepsilon-\varepsilon^{m-1})x_i-x_{i-1}$. Let us write $c=1-\varepsilon-\varepsilon^{m-1}$ and denote by u_1, u_2 the zeroes of polynomial x^2-cx+1. The general term of the above recurrent sequence is $x_i=\dfrac{u_1^i-u_2^i}{u_1-u_2}$ if $u_1\neq u_2$ and $x_i=iu_1^i$ if $u_1=u_2$. The latter case is clearly impossible. In the former case ($u_1\neq u_2$) equality $x_{n+1}=0$ is equivalent to $u_1^{n+1}=u_2^{n+1}$ and hence to $\omega^{n+1}=1$, where $u_1=u_2\omega$, which holds if and only if $(\exists u_2)\ u_2^2\omega=1$ and $u_2(1+\omega)=c$. Therefore $(1+\omega)^2=c^2\omega$, so

$$2+\omega+\bar{\omega}=(1-\varepsilon-\bar{\varepsilon})^2.$$

Now if $\omega=\cos\frac{2k\pi}{n+1}+i\sin\frac{2k\pi}{n+1}$ and $\varepsilon=\cos\frac{2l\pi}{m}+i\sin\frac{2l\pi}{m}$, the above equality becomes the desired one.

terug naar echt bestand

Het Duivenhokprincipe

Arne Smeets

1 Duivenhok, zn., onz. (-ken), hok waarin men duiven houdt

Dit is alles wat de "Dikke Van Dale" ons weet te vertellen over het woord *duivenhok*. Een saai begrip, nee? Zeker niet interessant genoeg om een hele lesbrief te vullen... maar wanneer we het hebben over het *duivenhokprincipe*, dan worden de zaken al direct veel interessanter, en dan weet de Van Dale plots van toeten noch blazen. Deze lesbrief gaat dus over het *duivenhokprincipe*, ook wel het *ladenprincipe van Dirichlet* genoemd. In het Frans spreekt men over *le principe des tiroirs* en in het Engels heeft men het over *the pigeonhole principle*, dat we zullen afkorten als "PHP". Met behulp van een drietal eenvoudige stellingen (die we PHP1, PHP2 en PHP3 zullen noemen) kunnen we een aantal zeer leuke en uiteenlopende problemen oplossen: hoofdzakelijk opgaven uit de combinatoriek, maar soms ook algebraïsche, getaltheoretische en zelfs meetkundige problemen. Vaak zijn dit moeilijke vragen: het duivenhokprincipe mag dan wel zeer eenvoudig zijn, in veel gevallen is het verre van evident om het op de juiste manier toe te passen. Je bent gewaarschuwd... Succes!

2 Lang geleden waren er eens n duiven en k duivenhokken...

Veel theorie valt er hier niet te bespreken: ik presenteer hier drie eenvoudige stellingen, die door sommigen misschien zelfs als "vanzelfsprekend" en logisch zullen worden beschouwd. Men gebruikt dikwijls *duiven* en *duivenhokken* (of *laden* en *voorwerpen*) om deze stellingen te formuleren, maar in feite handelen deze stellingen over verzamelingen en functies. Na elke stelling vermeld ik dus ook de "formele"(wiskundig correcte) versie van de stelling. We zullen in dit hoofdstukje het aantal elementen van een verzameling S voorstellen door $|S|$. Met $[x]$ bedoel ik natuurlijk de *entier* van x, met andere woorden, het grootste geheel getal, kleiner dan of gelijk aan x.

Stelling. (*PHP1*) Zij $n \in \mathbb{N}_0$. Als men meer dan n duiven over n duivenhokken verdeelt, dan bestaat er duivenhok dat minstens twee duiven bevat.

Bewijs. Als elk duivenhok ten hoogste één duif zou bevatten, dan zou het totaal aantal duiven (in alle n duivenhokken samen) niet groter zijn dan n. Dit is een contradictie, dus er bestaat een duivenhok dat minstens twee duiven bevat. □

We herformuleren deze stelling: als A en B eindige verzamelingen zodat $|A| > |B|$, en als $f : A \to B$ een afbeelding is, dan kan f niet injectief zijn. Zie je waarom dit precies dezelfde stelling is?

Stelling. (*PHP2*) Als men oneindig veel duiven over een eindig aantal duivenhokken verdeelt, dan bestaat er een duivenhok dat oneindig veel duiven bevat.

Bewijs. Als elk duivenhok een eindig aantal duiven zou bevatten, dan zou het aantal duiven in alle duivenhokken samen eindig zijn (omdat er slechts een eindig aantal duivenhokken is). Dit is een contradictie, dus een van de duivenhokken bevat oneindig veel duiven. □

Als je "professioneel" wil klinken dan kan je de stelling ook als volgt formuleren: als A een oneindige verzameling is, als B een eindige verzameling is en als $f : A \to B$ een afbeelding is, dan bestaat er een oneindige verzameling $C \subset A$ zodat alle elementen van C hetzelfde beeld hebben onder f.

De meest algemene vorm van het duivenhokprincipe (en de meeste krachtige) is de volgende:

Stelling. (*PHP3*) Zijn $n, k \in \mathbb{N}_0$. Als men n duiven verdeelt over k duivenhokken, dan bestaat er een duivenhok dat minstens $\left[\frac{n-1}{k}\right] + 1$ duiven bevat.

(Of nog: als m en n natuurlijk getallen zijn, en men verdeelt meer dan mn duiven over n duivenhokken, dan bestaat er een duivenhok dat minstens $m + 1$ duiven bevat.)

Bewijs. Als elk duivenhok minder dan $\left[\frac{n-1}{k}\right] + 1$ duiven zou bevatten, dan zou het totaal aantal duiven in alle duivenhokken samen niet groter zijn dan $k\left[\frac{n-1}{k}\right]$. Nu is het duidelijk dat $[x] \leq x$ voor

alle reële getallen x. Bijgevolg is $k \cdot \left[\frac{n-1}{k}\right] \leq n-1$. Het totaal aantal duiven zou dus niet groter zijn dan $n-1$, maar dat is onmogelijk aangezien er n duiven zijn. Bijgevolg bestaat er een duivenhok dat minstens $\left[\frac{n-1}{k}\right]+1$ duiven bevat. □

De herformulering van de vraag laat ik deze keer aan jou over.

3 Minstens twee Vlamingen hebben evenveel hoofdharen!

In dit hoofdstukje presenteer ik een aantal mooie (en soms moeilijke) toepassingen van het duivenhokprincipe. Het eerste voorbeeldje is echter verre van moeilijk...

Voorbeeld 0. Er bestaan twee Vlamingen die evenveel hoofdharen hebben.
Bewijs. Een mens heeft niet niet meer dan 200.000 hoofdharen (een bekend biologisch feit). Omdat er veel meer dan 200.000 niet-kale Vlamingen zijn, moeten minstens twee Vlamingen evenveel haren op hun hoofd hebben, volgens PHP1. □

Voilà, tijd voor de serieuze voorbeelden...

Voorbeeld 1. (*VWO 1989*) Bewijs dat elke deelverzameling met 55 elementen van de verzameling $S = \{1, 2, 3, ..., 100\}$ twee getallen bevat waarvan het (positieve) verschil gelijk is aan 9.
Oplossing. Zij A een deelverzameling van S met $|A| = 55$ en zijn $a_1 < a_2 < ... < a_{55}$ de elementen van A. Definieer $b_i = a_i + 9$ voor $i = 1, 2, 3, ..., 55$. Dan geldt $b_{55} = a_{55} + 9 \leq 109$. We hebben dus 110 natuurlijke getallen a_i, b_i, allen kleiner dan 110 en verschillend van 0. Volgens PHP1 moeten twee van deze getallen gelijk zijn, zodat $a_i = b_j$ voor zekere indices i en j, en $a_i - a_j = 9$. □
Er bestaan verschillende alternatieve oplossingen voor deze vraag; je kan bijvoorbeeld restklassen modulo 9 beschouwen. Volgens PHP3 moeten minstens 7 elementen van A tot dezelfde restklasse behoren; noem deze getallen $b_1 < b_2 < ... < b_7$. We veronderstellen dat er geen indices i en j bestaan zodat $a_i - a_j = 9$. Omdat de zeven getallen b_i tot dezelfde restklasse behoren (mod 9) zal $b_j - b_i \geq 18$, waarbij $1 \leq i < j \leq 7$. Dan is $b_7 - b_1 = (b_7 - b_6) + (b_6 - b_5) + ... + (b_2 - b_1) \geq 6 \cdot 18 = 108$. Dat is natuurlijk onmogelijk voor twee getallen die tot S behoren; de veronderstelling was dus foutief. □

Soms kan het duivenhokprincipe in een meetkundige context worden toegepast:

Voorbeeld 2. In een vierkant waarvan de zijde lengte 1 heeft liggen 51 punten. Bewijs dat er een cirkelschijf met straal $\frac{1}{7}$ bestaat die minstens 3 van de gegeven punten bedekt.
Oplossing. Verdeel het vierkant in 25 congruente vierkantjes met zijden van lengte $\frac{1}{5}$. Volgens PHP3 bestaan er dan 3 punten die in hetzelfde vierkant liggen. Noem dit vierkant V en noem O het middelpunt van dit vierkant. De omgeschreven cirkel van V heeft middelpunt O en straal $\frac{1}{5\sqrt{2}} < \frac{1}{7}$. De cirkel met middelpunt O en straal $\frac{1}{7}$ zal het vierkant V bijgevolg volledig bedekken, en bijgevolg zullen ook alle punten binnen V (dit zijn er minstens 3) bedekt worden door deze cirkel. □

Het volgende voorbeeldje is een schitterende toepassing van het duivenhokprincipe:

Voorbeeld 3. Bewijs dat elke verzameling van 13 reële getallen twee getallen a en b bevat zodat

$$0 \leq \frac{a-b}{1+ab} \leq 2 - \sqrt{3}.$$

Oplossing. Zij $S = \{s_1, s_2, ..., s_{13}\}$ de gegeven verzameling en stel $t_i = \text{bgtan}\, s_i$ voor $i = 1, 2, ..., 13$. Dan geldt voor $i = 1, 2, ..., 13$ dat $t_i \in [-\pi/2, \pi/2]$. We verdelen het interval $[-\pi/2, \pi/2]$ in twaalf deelintervallen van gelijke lengte, namelijk de intervallen $[k\pi/12, (k+1)\pi/12]$ voor $k = -6, -5, ..., 5$. Omdat er 13 getallen t_i gegeven zijn zullen twee getallen tot hetzelfde interval behoren (PHP1), met andere woorden, er bestaan indices p en q zodat $0 \leq t_p - t_q \leq \pi/12$. Omdat de tangensfunctie stijgend is op het interval $]-\pi/2, \pi/2[$ geldt er dat $0 \leq \tan(t_p - t_q) \leq \tan(\pi/12)$. Merk nu op dat $\tan(\pi/12) = 2 - \sqrt{3}$ en dat

$$\tan(t_p - t_q) = \frac{\tan t_p - \tan t_q}{1 + \tan t_p \cdot \tan t_q} = \frac{s_p - s_q}{1 + s_p s_q}.$$

Stel $s_p = a$ en $s_q = b$: we zijn nu klaar. □

Soms moet je een opgave veralgemenen om die te kunnen oplossen:

Voorbeeld 4. (*Servië-Montenegro 2002*) Toon aan dat er een natuurlijk getal $k \neq 0$ bestaat zodat de cijfers 3, 4, 5 en 6 niet voorkomen in de decimale voorstelling van het getal $k \cdot 2002!$.
Oplossing. We bewijzen een veel mooiere stelling: alle natuurlijke getallen n hebben een veelvoud van de vorm $11\ldots100\ldots0$. (In dit geval is natuurlijk $n = 2002!$) Zij $a_k = 11\ldots11$ het getal dat uit k cijfers 1 bestaat. Natuurlijk bestaan er oneindig veel dergelijke getallen, maar er zijn slechts eindig veel restklassen modulo n. Volgens PHP2 bestaan er dus indices p en q zodat $a_p \equiv a_q \pmod{n}$. Bijgevolg is $n|a_p - a_q$, en $a_p - a_q$ is van de vorm $11\ldots100\ldots0$, dus we zijn klaar. □

De volgende drie voorbeelden zijn zeer moeilijke opgaven. Je zal zien dat de toepassing van het duivenhokprincipe vaak slechts een klein stukje is van de volledige oplossing van een opgave.

Voorbeeld 5. Gegeven is een rechthoekig rooster van punten met 13 rijen en 13 kolommen. Men kleurt 53 van de 169 gegeven punten rood. Bewijs dat er een rechthoek bestaat waarvan de zijden evenwijdig zijn aan de randen van het rooster en waarvan alle hoekpunten rode roosterpunten zijn.
Oplossing. Noem a_1, a_2, ..., a_{13} het aantal rode punten in de eerste, tweede, ..., dertiende rij respectievelijk. Natuurlijk is $a_1 + a_2 + ... + a_{13} = 53$. Na enig nadenken zien we dan dat er een dergelijke rechthoek bestaat als

$$\sum_{k=1}^{13} \binom{a_k}{2} > \binom{13}{2} = 78. \ (*)$$

Inderdaad, voor rij i zijn er $\binom{a_i}{2}$ mogelijke koppels van rode punten in die rij. Beschouw nu de 13 kolommen van het rooster. Er zijn $\binom{13}{2} = 78$ mogelijke combinaties van twee kolommen. Als nu $(*)$ geldt, dan zal een zekere combinatie van twee kolommen minstens twee maal bereikt worden door twee paren van punten (PHP1), waarbij de twee punten van elk paar in dezelfde rij liggen en waarbij de twee paren onderling in verschillende rijen gelegen zijn. (Dit klinkt moeilijk maar dat is het niet.) Het volstaat dus om te bewijzen dat $(*)$ geldt. Welnu,

$$\sum_{k=1}^{13} \binom{a_k}{2} = \sum_{k=1}^{13} \frac{a_k(a_k-1)}{2} = \frac{1}{2} \cdot \sum_{k=1}^{13} a_k^2 - \frac{1}{2} \cdot \sum_{k=1}^{13} a_k.$$

Nu is $a_1 + a_2 + ... + a_{13} = 53$ dus geldt volgens de ongelijkheid van Cauchy dat

$$\left(a_1^2 + a_2^2 + ... + a_{13}^2\right)\left(1^2 + 1^2 + ... + 1^2\right) \geq \left(a_1 + a_2 + ... + a_{13}\right)^2 \Rightarrow \sum_{k=1}^{13} a_k^2 \geq \frac{53^2}{13}.$$

Bijgevolg geldt inderdaad dat

$$\frac{1}{2} \cdot \sum_{k=1}^{13} a_k^2 - \frac{1}{2} \cdot \sum_{k=1}^{13} a_k \geq \frac{1}{2} \cdot \left(\frac{53^2}{13} - 53\right) > 78$$

dus moet er een dergelijke rechthoek bestaan. □

Voorbeeld 6. (*IMO 1987 Vraag 3*) Zijn x_1, x_2, ..., x_n reële getallen zodat $x_1^2 + x_2^2 + ... + x_n^2 = 1$. Bewijs dat, $\forall k \in \mathbb{N}$ zodat $k \geq 2$, er gehele getallen a_1, a_2, ..., a_n bestaan, niet allen gelijk aan 0, zodat $|a_i| \leq k - 1$ voor alle i en zodat

$$|a_1x_1 + a_2x_2 + ... + a_nx_n| \leq \frac{(k-1)\sqrt{n}}{k^n - 1}.$$

Oplossing. Zonder verlies van de algemeenheid mogen we veronderstellen dat $x_1 \geq x_2 \geq ... \geq x_n$. Zij m de unieke index waarvoor geldt dat x_1, x_2, ..., $x_m \geq 0$ en x_{m+1}, ..., $x_n < 0$. (Als alle getallen x_i strikt negatief zijn, dan stellen we $m = 0$; als alle getallen x_i positief zijn dan stellen we $m = n$.) Zij $\mathcal{C}$ de verzameling van alle vectoren $(c_1, c_2, ..., c_n) \in \mathbb{R}^n$ zodat $c_i \in \{0, 1, ..., k-1\}$. Beschouw alle

mogelijke waarden van de som $S = c_1x_1 + c_2x_2 + ... + c_nx_n$ voor $(c_1, c_2, ..., c_n) \in \mathcal{C}$. De kleinst en grootst mogelijke waarden van S worden respectievelijk bereikt voor

$$(c_1, c_2, ..., c_n) = (\underbrace{0, ..., 0}_{m}, \underbrace{k-1, ..., k-1}_{n-m});$$

$$(c_1, c_2, ..., c_n) = (\underbrace{k-1, ..., k-1}_{m}, \underbrace{0, ..., 0}_{n-m}).$$

Noem deze extreme waarden A en B respectievelijk, dan geldt er dat

$$B - A = (k-1)(|x_1| + |x_2| + ... + |x_n|).$$

Omdat $x_1^2 + x_2^2 + ... + x_n^2 = 1$ geldt volgens de ongelijkheid van Cauchy dat

$$(x_1^2 + x_2^2 + ... + x_n^2)(1 + 1 + ... + 1) \geq (|x_1| + |x_2| + ... + |x_n|)^2 \Rightarrow |x_1| + |x_2| + ... + |x_n| \leq \sqrt{n}.$$

Bijgevolg is $B - A \leq (k-1)\sqrt{n}$. Nu liggen alle waarden die S kan aannemen in het interval $[A, B]$. We verdelen dit interval in $N = k^n - 1$ deelintervallen van gelijke lengte, namelijk van lengte $(B - A)/N$. Omdat de verzameling $\mathcal{C}$ precies k^n vectoren bevat, bestaan er volgens PHP1 twee vectoren $(c_1', c_2', ..., c_n')$ en $(c_1'', c_2'', ..., c_n'')$ waarvoor de corresponderende sommen S' en S'' in hetzelfde deelinterval liggen, waarbij

$$S' = c_1'x_1 + c_2'x_2 + ... + c_n'x_n, \;\; S'' = c_1''x_1 + c_2''x_2 + ... = c_n''x_n.$$

Dan geldt dus $|S' - S''| \leq (B - A)/N$. Neem nu $a_1 = c_i' - c_i''$ voor $i = 1, 2, ..., n$. Op die manier verkrijgen we een vector $(a_1, a_2, ..., a_n)$, verschillend van de nulvector in $\mathbb{R}^n$, waarvoor geldt dat $|a_i| \leq k - 1$ voor $i = 1, 2, ..., n$ en

$$|a_1x_1 + a_2x_2 + ... + a_nx_n| = |S' - S''| \leq \frac{B - A}{N} \leq \frac{(k-1)\sqrt{n}}{k^n - 1}$$

dus we hebben gehele getallen $a_1, a_2, ..., a_n$ gevonden die voldoen aan het te bewijzen. □

Voorbeeld 7. (*IMO 2001 Vraag 3*) Aan een wiskundecompetitie namen 21 jongens en 21 meisjes deel. Achteraf bleek dat geen enkele deelnemer meer dan 6 problemen oploste, en dat er voor elke jongen en voor elk meisje minstens één probleem bestaat dat door zowel die jongen als door dat meisje opgelost werd. Bewijs dat er een probleem bestaat dat opgelost werd door minstens 3 jongens en minstens 3 meisjes.

Oplossing. Om te beginnen voeren we een paar notaties in. Noem J de verzameling van de jongens, M de verzameling van de meisjes, P de verzameling van de problemen. Dan is $|M| = |J| = 21$. Noem $P(m)$ de verzameling van de problemen die opgelost werden door een meisje $m \in M$ en $P(j)$ de verzameling van de problemen die opgelost werden door $j \in J$. Tenslotte noemen we $J(p)$ en $M(p)$ respectievelijk de verzamelingen van de jongens en meisjes die het probleem $p \in P$ oplossen. Nu kunnen we aan de slag. We willen bewijzen dat er een probleem $p \in P$ bestaat waarvoor geldt dat $|J(p)| \geq 3$ en $|M(p)| \geq 3$. We veronderstellen dat dit niet het geval is. We zullen op twee verschillende manieren het aantal elementen tellen van de verzameling $T = \{(p, m, j) \mid p \in P(m) \cap P(j)\}$. Natuurlijk is $|T| \geq 21^2 = 441$, omdat er 21^2 koppels (j, m) bestaan van een jongen en een meisje en omdat er voor elk koppel (j, m) een probleem bestaat dat zowel door j als door m opgelost werd. Veronderstel dus dat er geen $p \in P$ bestaat zodat $|M(p)| \geq 3$ en $|J(p)| \geq 3$. We merken op dat

$$\sum_{m \in M} |P(m)| = \sum_{p \in P} |M(p)| \leq 6\,|M| = 126, \; \sum_{j \in J} |P(j)| = \sum_{p \in P} |J(p)| \leq 6\,|J| = 126$$

omdat geen enkele deelnemer meer dan 6 problemen oplost. (De gelijkheden hierboven kan men eenvoudig aantonen en zijn eigenlijk niet meer dan logisch.) We definiëren nu

$$P_+ = \{p \in P \mid |M(p)| \geq 3\}; \; P_- = \{p \in P \mid |M(p)| \leq 2\}.$$

We zullen bewijzen dat

$$\sum_{p\in P_-} |M(p)| \geq |M|\,,\ \sum_{p\in P_+} |J(p)| \geq |J|\,. \ (*)$$

Zij $m \in M$ een willekeurig meisje. Volgens het duivenhokprincipe lost m een probleem p op dat door minstens $\left[\frac{21-1}{6}\right] + 1 = 4$ jongens opgelost wordt. Wegens onze veronderstelling volgt uit $|J(p)| \geq 4$ dat $p \in P_-$, dus elk meisje lost minstens één probleem uit P_- op. Op analoge wijze toont men aan dat elke jongen minstens één probleem uit P_+ oplost. Daarmee is $(*)$ bewezen. Dan geldt ook

$$\sum_{p\in P_+} |M(p)| = \sum_{p\in P} |M(p)| - \sum_{p\in P_-} |M(p)| \leq 5\,|M|\,,\ \sum_{p\in P_-} |J(p)| \leq 5\,|J|\,.$$

Nu is

$$|T| = \sum_{p\in P} |M(p)|\cdot|J(p)| = \sum_{p\in P_+} |M(p)|\cdot|J(p)| + \sum_{p\in P_-} |M(p)|\cdot|J(p)|$$

en bijgevolg geldt volgens onze veronderstelling dat

$$|T| \leq 2\sum_{p\in P_+} |M(p)| + 2\sum_{p\in P_-} |J(p)| \leq 10\,|M| + 10\,|J| = 420.$$

We bewezen echter dat $|T| \geq 441$. Onze veronderstelling was dus foutief; we zijn dus klaar. □

4 Al doende leert men!

De eerste opgaven zijn zeer eenvoudig, de laatste opgaven zijn aan de moeilijke kant (maar natuurlijk niet zo moeilijk als Voorbeeld 7.) De laatste opgave is vraag 6 van IMO 2005: normaalgezien is zo'n vraag onoplosbaar voor een Belg, maar in 2005 was vraag 6 iets gemakkelijker dan gewoonlijk, en maar liefst 2 van de 3 Vlaamse IMO-deelnemers losten deze vraag op! Succes!

1. Bewijs dat er een getal N van de vorm 20042004...2004 bestaat waarvoor geldt:
   - N is deelbaar door 2003, en
   - N heeft niet meer dan 10.000 cijfers in decimale voorstelling.
2. (*IMO 1972 Vraag 1*) Bewijs: elke verzameling van 10 natuurlijke getallen kleiner dan 100, heeft twee disjuncte deelverzamelingen waarvan de sommen van de elementen gelijk zijn.
3. Bewijs: elk veelvlak heeft twee zijvlakken die begrensd worden door een gelijk aantal zijden.
4. (*British Mathematical Olympiad 2000*) Bestaat er een verzameling van elf gehele getallen zodat geen zes van deze gehele getallen een zesvoud als som hebben?
5. Een basketbalteam speelde 45 wedstrijden in een maand met 30 dagen. Elke dag speelde het team minstens één wedstrijd. Bewijs dat er een periode van een bepaald aantal dagen bestaat zodat het team gedurende die periode precies 14 wedstrijden speelde.
6. (*Rusland 1961*) Men plaatst 120 vierkantjes met zijden van lengte 1 binnen een rechthoek met afmetingen 20×25, en dat op willekeurige wijze. Bewijs dat men een cirkel met diameter 1 binnen deze rechthoek kan plaatsen zodat deze cirkel geen enkel vierkantje snijdt.
7. (*Bulgarian Mathematical Olympiad 2003*) Bart en Ria spelen het volgende spel. Bart schrijft n verschillende natuurlijke getallen op een papier, met n een natuurlijk getal. Ria mag enkele van deze getallen wegstrepen (ze mag er ook geen enkel wegstrepen, maar ze mag ze zeker niet allemaal wegstrepen). Daarna mag Ria voor elk van de overblijvende getallen een $+$ of een $-$ zetten en de som bepalen van de getallen die ze op deze manier bekomt. Als deze som deelbaar is door 2003 dan wint Ria. In het andere geval wint Bart. Voor welke waarden van n heeft Bart een winnende strategie, en voor welke waarden van n heeft Ria een winnende strategie?

8. (*IMO 1988 Longlist*) Zijn $a_1, a_2, \ldots, a_{11} \in \mathbb{Z}$. Bewijs dat er $b_1, b_2, \ldots, b_{11} \in \{-1, 0, 1\}$ bestaan (niet allen gelijk aan 0) zodat $\sum_{k=1}^{11} a_k b_k$ deelbaar is door 2005.

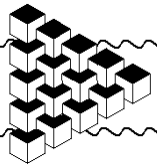
9. (*China 1990*) Bepaal het kleinste natuurlijk getal n zodat er voor elke verzameling $\{a_1, a_2, ..., a_n\}$ van n verschillende reële getallen, gekozen uit het interval $[1, 1000]$, indices i en j bestaan waarvoor geldt dat $0 < a_i - a_j < 1 + 3\sqrt[3]{a_i a_j}$.

10. Binnen een cirkel met straal 16 liggen 650 gegeven punten. Definieer een *ring* als het vlakdeel dat begrepen is tussen twee concentrische cirkels met stralen 2 en 3 respectievelijk. Bewijs dat men een ring kan plaatsen zodat minstens 10 van de 650 punten bedekt worden door deze ring.

11. (*IMO 2005 Vraag 6*) Bij een wiskundewedstrijd kregen de deelnemers 6 opgaven voorgelegd. Ieder tweetal van deze opgaven werd door meer dan $\frac{2}{5}$ van het aantal deelnemers opgelost. Niemand loste alle 6 de opgaven op. Laat zien dat minstens 2 deelnemers ieder precies 5 opgaven hebben opgelost.

∘ ∘ ∘

terug naar echt bestand



©2007 The Author(s) and The IMO Compendium Group

Polynomials in One Variable

Dušan Djukić

Contents

1 General Properties

A *Monomial* in variable x is an expression of the form cx^k, where c is a constant and k a nonnegative integer. Constant c can be e.g. an integer, rational, real or complex number.

A *Polynomial* in x is a sum of finitely many monomials in x. In other words, it is an expression of the form

$$P(x) = a_n x^n + a_{n-1} + \cdots + a_1 x + a_0. \qquad (*)$$

If only two or three of the above summands are nonzero, P is said to be a *binomial* and *trinomial*, respectively.

The constants $a_0, \ldots, a_n$ in $(*)$ are the *coefficients* of polynomial P. The set of polynomials with the coefficients in set A is denoted by $A[x]$ - for instance, $\mathbb{R}[x]$ is the set of polynomials with real coefficients.

We can assume in $(*)$ w.l.o.g. that $a_n \neq 0$ (if $a_n = 0$, the summand $a_n x^n$ can be erased without changing the polynomial). Then the exponent n is called the *degree* of polynomial P and denoted by $\deg P$. In particular, polynomials of degree one, two and three are called *linear*, *quadratic* and *cubic*. A nonzero constant polynomial has degree 0, while the zero-polynomial $P(x) \equiv 0$ is assigned the degree $-\infty$ for reasons soon to become clear.

Example 1. *$P(x) = x^3(x+1) + (1-x^2)^2 = 2x^4 + x^3 - 2x^2 + 1$ is a polynomial with integer coefficients of degree 4.*

$Q(x) = 0x^2 - \sqrt{2}x + 3$ is a linear polynomial with real coefficients.

$R(x) = \sqrt{x^2} = |x|$, $S(x) = \frac{1}{x}$ and $T(x) = \sqrt{2x+1}$ are not polynomials.

Polynomials can be added, subtracted or multiplied, and the result will be a polynomial too:

$$A(x) = a_0 + a_1 x + \cdots + a_n x^n, \quad B(x) = b_0 + b_1 x + \cdots + b_m x^m$$
$$A(x) \pm B(x) = (a_0 - b_0) + (a_1 - b_1)x + \cdots,$$
$$A(x)B(x) = a_0 b_0 + (a_0 b_1 + a_1 b_0)x + \cdots + a_n b_m x^{m+n}.$$

The behavior of the degrees of the polynomials under these operations is clear:

Theorem 1. *If A and B are two polynomials then:*

(i) $\deg(A \pm B) \le \max(\deg A, \deg B)$, *with the equality if* $\deg A \ne \deg B$.

(ii) $\deg(A \cdot B) = \deg A + \deg B$. □

The conventional equality $\deg 0 = -\infty$ actually arose from these properties of degrees, as else the equality (ii) would not be always true.

Unlike a sum, difference and product, a quotient of two polynomials is not necessarily a polynomial. Instead, like integers, they can be divided with a residue.

Theorem 2. *Given polynomials A and $B \ne 0$, there are unique polynomials Q (quotient) and R (residue) such that*

$$A = BQ + R \quad \textit{and} \quad \deg R < \deg B.$$

Proof. Let $A(x) = a_n x^n + \cdots + a_0$ and $B(x) = b_k x^k + \cdots + b_0$, where $a_n b_k \ne 0$. Assume k is fixed and use induction on n. For $n < k$ the statement is trivial. Suppose that $n = N \ge k$ and that the statement is true for $n < N$. Then $A_1(x) = A(x) - \frac{a_n}{b_k} x^{n-k} B(x)$ is a polynomial of degree less than n (for its coefficient at x^n iz zero); hence by the inductive assumption there are unique polynomials Q_1 and R such that $A_1 = BQ_1 + R$ and $\deg R$. But this also implies

$$A = BQ + R, \quad \text{where} \quad Q(x) = \frac{a_n}{b_k} x^{n-k} + Q_1(x)\,. \ \square$$

Example 2. *The quotient upon division of $A(x) = x^3 + x^2 - 1$ by $B(x) = x^2 - x - 3$ is $x + 2$ with the residue $5x + 5$, as*

$$\frac{x^3 + x^2 - 1}{x^2 - x - 3} = x + 2 + \frac{5x + 5}{x^2 - x - 3}\,.$$

We say that polynomial A is *divisible* by polynomial B if the remainder R when A is divided by B equal to 0, i.e. if there is a polynomial Q such that $A = BQ$.

Theorem 3 (Bezout's theorem). *Polynomial $P(x)$ is divisible by binomial $x - a$ if and only if $P(a) = 0$.*

Proof. There exist a polynomial Q and a constant c such that $P(x) = (x - a)Q(x) + c$. Here $P(a) = c$, making the statement obvious. □

Number a is a *zero (root)* of a given polynomial $P(x)$ if $P(a) = 0$, i.e. $(x - a) \mid P(x)$.

To determine a zero of a polynomial f means to solve the equation $f(x) = 0$. This is not always possible. For example, it is known that finding the exact values of zeros is impossible in general when f is of degree at least 5. Nevertheless, the zeros can always be computed with an arbitrary precision. Specifically, $f(a) < 0 < f(b)$ implies that f has a zero between a and b.

Example 3. *Polynomial $x^2 - 2x - 1$ has two real roots: $x_{1,2} = 1 \pm \sqrt{2}$.*

Polynomial $x^2 - 2x + 2$ has no real roots, but it has two complex roots: $x_{1,2} = 1 \pm i$.

Polynomial $x^5 - 5x + 1$ has a zero in the interval $[1.44, 1.441]$ which cannot be exactly computed.

More generally, the following simple statement holds.

Theorem 4. *If a polynomial P is divisible by a polynomial Q, then every zero of Q is also a zero of P.* □

The converse does not hold. Although every zero of x^2 is a zero of x, x^2 does not divide x.

Problem 1. *For which n is the polynomial $x^n + x - 1$ divisible by a) $x^2 - x + 1$, b) $x^3 - x + 1$?*

Solution. a) The zeros of polynomial x^2-x+1 are $\varepsilon_{1,2}=\frac{1\pm i\sqrt{3}}{2}$. If x^2-x+1 divides x^n+x-1, then $\varepsilon_{1,2}$ are zeros of polynomial x^n+x-1, so $\varepsilon_i^n=1-\varepsilon_i=\varepsilon_i^{-1}$. Since $\varepsilon^k=1$ if and only if $6\mid k$, the answer is $n=6i-1$.

b) If $f(x)=x^3-x+1$ divides x^n+x-1, then it also divides x^n+x^3. This means that every zero of $f(x)$ satisfies $x^{n-3}=-1$; in particular, each zero of f has modulus 1. However, $f(x)$ has a zero between -2 and -1 (for $f(-2)<0<f(-1)$) which is obviously not of modulus 1. Hence there is no such n. $\triangle$

Every nonconstant polynomial with complex coefficients has a complex root. We shall prove this statement later; until then we just believe.

The following statement is analogous to the unique factorization theorem in arithmetics.

Theorem 5. *Polynomial $P(x)$ of degree $n>0$ has a unique representation of the form*

$$P(x)=c(x-x_1)(x-x_2)\cdots(x-x_n),$$

not counting the ordering, where $c\neq 0$ and $x_1,\ldots,x_n$ are complex numbers, not necessarily distinct.

Therefore, $P(x)$ has at most $\deg P=n$ *different zeros.*

Proof. First we show the uniqueness. Suppose that

$$P(x)=c(x-x_1)(x-x_2)\cdots(x-x_n)=d(x-y_1)(x-y_2)\cdots(x-y_n).$$

Comparing the leading coefficients yields $c=d$. We may assume w.l.o.g. that there are no i,j for which $x_i=y_j$ (otherwise the factor $x-x_i$ can be canceled on both sides). Then $P(x_1)=0$. On the other hand, $P(x_1)=d(x_1-y_1)\cdots(x_1-y_n)\neq 0$, a contradiction.

The existence is shown by induction on n. The case $n=1$ is clear. Let $n>1$. The polynomial $P(x)$ has a complex root, say x_1. By Bezout's theorem, $P(x)=(x-x_1)P_1(x)$ for some polynomial P_1 of degree $n-1$. By the inductive assumption there exist complex numbers $x_2,\ldots,x_n$ for which $P_1(x)=c(x-x_2)\cdots(x-x_n)$, which also implies $P(x)=c(x-x_1)\cdots(x-x_n)$. $\square$

Corollary. If polynomials P and Q has degrees not exceeding n and coincide at $n+1$ different points, then they are equal.

Grouping equal factors yields the *canonical representation*:

$$P(x)=c(x-a_1)^{\alpha_1}(x-a_2)^{\alpha_2}\cdots(x-a_k)^{\alpha_k},$$

where α_i are natural numbers with $\alpha_1+\cdots+\alpha_k=n$. The exponent α_i is called the *multiplicity* of the root a_i. It is worth emphasizing that:

Theorem 6. *Polynomial of n-th degree has exactly n complex roots counted with their multiplicities.* $\square$

We say that two polynomials Q and R are *coprime* if they have no roots in common; Equivalently, there is no nonconstant polynomial dividing them both, in analogy with coprimeness of integers. The following statement is a direct consequence of the previous theorem:

Theorem 7. *If a polynomial P is divisible by two coprime polynomials Q and R, then it is divisible by $Q\cdot R$.* $\square$

Remark: This can be shown without using the existence of roots. By the Euclidean algorithm applied on polynomials there exist polynomials K and L such that $KQ+LR=1$. Now if $P=QS=RT$ for some polynomials R,S, then $R(KT-LS)=KQS-LRS=S$, and therefore $R\mid S$ and $QR\mid QS=P$.

If polynomial $P(x)=x^n+\cdots+a_1x+a_0$ with real coefficients has a complex zero ξ, then $P(\overline{\xi})=\overline{\xi^n}+\cdots+a_1\overline{\xi}+a_0=\overline{P(\xi)}=0$. Thus:

Theorem 8. *If ξ is a zero of a real polynomial $P(x)$, then so is $\overline{\xi}$.* $\square$

In the factorization of a real polynomial $P(x)$ into linear factors we can group conjugated complex zeros:

$$P(x) = (x-r_1)\cdots(x-r_k)(x-\xi_1)(x-\overline{\xi_1})\cdots(x-\xi_l)(x-\overline{\xi_l}),$$

where r_i are the real zeros, ξ complex, and $k+2l = n = \deg P$. Polynomial $(x-\xi)(x-\overline{\xi}) = x^2 - 2Re\xi + |\xi|^2 = x^2 - p_i x + q_i$ has real coefficients which satisfy $p_i^2 - 4q_i < 0$. This shows that:

Theorem 9. *A real polynomial $P(x)$ has a unique factorization (up to the order) of the form*

$$P(x) = (x-r_1)\cdots(x-r_k)(x^2-p_1x+q_1)\cdots(x^2-p_lx+q_l),$$

where r_i and p_j, q_j are real numbers with $p_i^2 < 4q_i$ and $k+2l = n$. □

It follows that a real polynomial of an odd degree always has an odd number of zeros (and at least one).

2 Zeros of Polynomials

In the first section we described some basic properties of polynomials. In this section we describe some further properties and at the end we prove that every complex polynomial actually has a root.

As we pointed out, in some cases the zeros of a given polynomial can be exactly determined. The case of polynomials of degree 2 has been known since the old age. The well-known formula gives the solutions of a quadratic equation $ax^2+bx+c = 0$ ($a \neq 0$) in the form

$$x_{1,2} = \frac{-b \pm \sqrt{b^2-4ac}}{2a}.$$

When f has degree 3 or 4, the (fairly impractical) formulas describing the solutions were given by the Italian mathematicians Tartaglia and Ferrari in the 16-th century. We show Tartaglia's method of solving a cubic equation.

At first, substituting $x = y - a/3$ reduces the cubic equation $x^3+ax^2+bx+c = 0$ with real coefficients to

$$y^3+py+q = 0, \quad \text{where} \quad p = b - \frac{a^2}{3}, \quad q = c - \frac{ab}{3} + \frac{2a^3}{27}.$$

Putting $y = u+v$ transforms this equation into $u^3+v^3+(3uv+p)y+q = 0$. But, since u and v are variable, we are allowed to bind them by the condition $3uv+p = 0$. Thus the above equation becomes the system

$$uv = -\frac{p}{3}, \quad u^3+v^3 = -q$$

which is easily solved: u^3 and v^3 are the solutions of the quadratic equation $t^2+qt-\frac{p^3}{27} = 0$ and $uv = -p/3$ must be real. Thus we come to the solutions:

Theorem 10 (Cardano's formula). *The solutions of the equation $y^3+py+q = 0$ with $p,q \in \mathbb{R}$ are*

$$y_i = \varepsilon^j \sqrt[3]{-\frac{q}{2}+\sqrt{\frac{q^2}{4}+\frac{p^3}{27}}} + \varepsilon^{-j}\sqrt[3]{-\frac{q}{2}-\sqrt{\frac{q^2}{4}+\frac{p^3}{27}}}, \quad j = 0,1,2,$$

where ε is a primitive cubic root of unity. □

A polynomial $f(x) = a_nx^n+\cdots+a_1x+a_0$ is *symmetric* if $a_{n-i} = a_i$ for all i. If $\deg f = n$ is odd, then -1 is a zero of f and the polynomial $f(x)/(x+1)$ is symmetric. If $n = 2k$ is even, then

$$f(x)/x^k = a_0(x^k+x^{-k})+\cdots+a_{k-1}(x+x^{-1})+a_k$$

is a polynomial in $y = x+x^{-1}$, for so is each of the expressions x^i+x^{-i} (see problem 3 in section 7). In particular, $x^2+x^{-2} = y^2-2$, $x^3+x^{-3} = y^3-3y$, etc. This reduces the equation $f(x) = 0$ to an equation of degree $n/2$.

Problem 2. *Show that the polynomial* $f(x)=x^6-2x^5+x^4-2x^3+x^2-2x+1$ *has exactly four zeros of modulus 1.*

Solution. Set $y=x+x^{-1}$. Then

$$\frac{f(x)}{x^3}=g(y)=y^3-2y^2-2y+2.$$

Observe that x is of modulus 1 if and only if $x=\cos t+i\sin t$ for some t, in which case $y=2\cos t$; conversely, $y=2\cos t$ implies that $x=\cos t\pm i\sin t$. In other words, $|x|=1$ if and only if y is real with $-2\le y\le 2$, where to each such y correspond two values of x if $y\neq\pm 2$. Therefore it remains to show that $g(y)$ has exactly two real roots in the interval $(-2,2)$. To see this, it is enough to note that $g(-2)=-10$, $g(0)=2$, $g(2)=-2$, and that therefore g has a zero in each of the intervals $(-2,0)$, $(0,2)$ and $(2,\infty)$. $\triangle$

How are the roots of a polynomial related to its coefficients? Consider a monic polynomial

$$P(x)=x^n+a_1x^{n-1}+\cdots+a_{n-1}x+a_n=(x-x_1)(x-x_2)\cdots(x-x_n)$$

of degree $n>0$. For example, comparing coefficients at x^{n-1} on both sides gives us $x_1+x_2+\cdots+x_n=-a_1$. Similarly, comparing the constant terms gives us $x_1x_2\cdots x_n=(-1)^na_n$. The general relations are given by the Vieta formulas below.

Definition 1. Elementary symmetric polynomials *in* $x_1,\ldots,x_n$ *are the polynomials* $\sigma_1,\sigma_2,\ldots,\sigma_n$, *where*

$$\sigma_k=\sigma_k(x_1,x_2,\ldots,x_n)=\sum x_{i_1}x_{i_2}\ldots x_{i_k},$$

the sum being over all k-element subsets $\{i_1,\ldots,i_k\}$ *of* $\{1,2,\ldots,n\}$.

In particular, $\sigma_1=x_1+x_2+\cdots+x_n$ and $\sigma_n=x_1x_2\cdots x_n$. Also, we usually set $\sigma_0=1$ and $\sigma_k=0$ for $k>n$.

Theorem 11 (Vieta's formulas). *If* $\alpha_1,\alpha_2,\ldots,\alpha_n$ *are the zeros of polynomial* $P(x)=x^n+a_1x^{n-1}+a_2x^{n-2}+\cdots+a_n$, *then* $a_k=(-1)^k\sigma_k(\alpha_1,\ldots,\alpha_n)$ *for* $k=1,2,\ldots,n$.

Proof. Induction on n. The case $n=1$ is trivial. Assume that $n>1$ and write $P(x)=(x-x_n)Q(x)$, where $Q(x)=(x-x_1)\cdots(x-x_{n-1})$. Let us compute the coefficient a_k of $P(x)$ at x^k. Since the coefficients of $Q(x)$ at x^{k-1} and x^k are $a'_{k-1}=(-1)^{k-1}\sigma_{k-1}(x_1,\ldots,x_{n-1})$ and $a'_k=(-1)^k\sigma_k(x_1,\ldots,x_{n-1})$ respectively, we have

$$a_k=-x_na'_{k-1}+a'_k=\sigma_k(x_1,\ldots,x_n).\ \square$$

Example 4. *The roots* x_1,x_2,x_3 *of polynomial* $P(x)=x^3-ax^2+bx-c$ *satisfy* $a=x_1+x_2+x_3$, $b=x_1x_2+x_2x_3+x_3x_1$ *and* $c=x_1x_2x_3$.

Problem 3. *Prove that not all zeros of a polynomial of the form* $x^n+2nx^{n-1}+2n^2x^{n-2}+\cdots$ *can be real.*

Solution. Suppose that all its zeros $x_1,x_2,\ldots,x_n$ are real. They satisfy

$$\sum_i x_i=-2n,\quad \sum_{i<j}x_ix_j=2n^2.$$

However, by the mean inequality we have

$$\sum_{i<j}x_ix_j=\frac{1}{2}\left(\sum_i x_i\right)^2-\frac{1}{2}\sum_i x_i^2\le\frac{n-1}{2n}\left(\sum_i x_i\right)^2=2n(n-1),$$

a contradiction. $\triangle$

Problem 4. *Find all polynomials of the form $a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$ with $a_j \in \{-1, 1\}$ $(j = 0, 1, \ldots, n)$, whose all roots are real.*

Solution. Let $x_1, \ldots, x_n$ be the roots of the given polynomial. Then

$$x_1^2 + x_2^2 + \cdots + x_n^2 = (\sum_i x_i)^2 - 2(\sum_{i<j} x_i x_j) = a_{n-1}^2 - 2a_{n-2} \leq 3;$$
$$x_1^2 x_2^2 \cdots x_n^2 = 1.$$

By the mean inequality, the second equality implies $x_1^2 + \cdots + x_n^2 \geq n$; hence $n \leq 3$. The case $n = 3$ is only possible if $x_1, x_2, x_3 = \pm 1$. Now we can easily find all solutions: $x \pm 1$, $x^2 \pm x - 1$, $x^3 - x \pm (x^2 - 1)$. $\triangle$

One contradiction is enough to show that not all zeros of a given polynomial are real. On the other hand, if the task is to show that all zeros of a polynomial *are* real, but not all are computable, the situation often gets more complicated.

Problem 5. *Show that all zeros of a polynomial $f(x) = x(x-2)(x-4)(x-6) + (x-1)(x-3)(x-5)(x-7)$ are real.*

Solution. Since $f(-\infty) = f(\infty) = +\infty$, $f(1) < 0$, $f(3) > 0$ and $f(5) < 0$, polynomial f has a real zero in each of the intervals $(-\infty, 1)$, $(1, 3)$, $(3, 5)$, $(5, \infty)$, that is four in total. $\triangle$

We now give the announced proof of the fact that every polynomial has a complex root. This fundamental theorem has many different proofs. The proof we present is, although more difficult than all the previous ones, still next to elementary. All imperfections in the proof are made on purpose.

Theorem 12 (The Fundamental Theorem of Algebra). *Every nonconstant complex polynomial $P(x)$ has a complex zero.*

Proof. Write $P(x) = x^n + a_{n-1}x^{n-1} + \cdots + a_0$. Suppose that $P(0) = a_0 \neq 0$. For each $r > 0$, let C_r be the circle in the complex plane with the center at point 0 and radius r. Consider the continuous curve $\gamma_r = P(C_r) = \{P(x) \mid |x| = r\}$. The curve described by the monomial x^n, i.e. $\{x^n \mid x \in C_r\}$ rounds point 0 n times. If r is large enough, for example $r > 1 + |a_0| + \cdots + |a_{n-1}|$, we have $|x^n| > |a_{n-1}x^{n-1} + \cdots + a_0| = |P(x) - x^n|$, which means that the rest $P(x) - x^n$ in the expression of $P(x)$ can not "reach" point 0. Thus for such r the curve γ_r also rounds point 0 n times; hence, it contains point 0 in its interior.

For very small r the curve γ_r is close to point $P(0) = a_0$ and leaves point 0 in its exterior. Thus, there exists a minimum $r = r_0$ for which point 0 is *not* in the exterior of γ_r. Since the curve γ_r changes continuously as a function of r, it cannot jump over the point 0, so point 0 must lie on the curve γ_{r_0}. Therefore, there is a zero of polynomial $P(x)$ of modulus r_0. $\square$

3 Polynomials with Integer Coefficients

Consider a polynomial $P(x) = a_n x^n + \cdots + a_1 x + a_0$ with integer coefficients. The difference $P(x) - P(y)$ can be written in the form

$$a_n(x^n - y^n) + \cdots + a_2(x^2 - y^2) + a_1(x - y),$$

in which all summands are multiples of polynomial $x - y$. This leads to the simple though important arithmetic property of polynomials from $\mathbb{Z}[x]$:

Theorem 13. *If P is a polynomial with integer coefficients, then $P(a) - P(b)$ is divisible by $a - b$ for any distinct integers a and b.*

In particular, all integer roots of P divide $P(0)$. $\square$

There is a similar statement about rational roots of polynomial $P(x) \in \mathbb{Z}[x]$.

Theorem 14. *If a rational number p/q ($p,q \in \mathbb{Z}$, $q \neq 0$, $nzd(p,q) = 1$) is a root of polynomial $P(x) = a_n x^n + \cdots + a_0$ with integer coefficients, then $p \mid a_0$ and $q \mid a_n$.*

Proof. We have

$$q^n P\left(\frac{p}{q}\right) = a_n p^n + a_{n-1} p^{n-1} q + \cdots + a_0 q^n.$$

All summands but possibly the first are multiples of q, and all but possibly the last are multiples of p. Hence $q \mid a_n p^n$ and $p \mid a_0 q^n$ and the claim follows. □

Problem 6. *Polynomial $P(x) \in \mathbb{Z}[x]$ takes values ± 1 at three different integer points. Prove that it has no integer zeros.*

Solution. Suppose to the contrary, that a,b,c,d are integers with $P(a)$, $P(b), P(c) \in \{-1,1\}$ and $P(d) = 0$. Then by the previous statement the integers $a-d, b-d$ and $c-d$ all divide 1, a contradiction. △

Problem 7. *Let $P(x)$ be a polynomial with integer coefficients. Prove that if $P(P(\cdots P(x)\cdots)) = x$ for some integer x (where P is iterated n times), then $P(P(x)) = x$.*

Solution. Consider the sequence given by $x_0 = x$ and $x_{k+1} = P(x_k)$ for $k \geq 0$. Assume $x_k = x_0$. We know that

$$d_i = x_{i+1} - x_i \mid P(x_{i+1}) - P(x_i) = x_{i+2} - x_{i+1} = d_{i+1}$$

for all i, which together with $d_k = d_0$ implies $|d_0| = |d_1| = \cdots = |d_k|$.

Suppose that $d_1 = d_0 = d \neq 0$. Then $d_2 = d$ (otherwise $x_3 = x_1$ and x_0 will never occur in the sequence again). Similarly, $d_3 = d$ etc, and hence $x_k = x_0 + kd \neq x_0$ for all k, a contradiction. It follows that $d_1 = -d_0$, so $x_2 = x_0$. △

Note that a polynomial that takes integer values at all integer points does not necessarily have integer coefficients, as seen on the polynomial $\frac{x(x-1)}{2}$.

Theorem 15. *If the value of the polynomial $P(x)$ is integral for every integer x, then there exist integers $c_0, \ldots, c_n$ such that*

$$P(x) = c_n \binom{x}{n} + c_{n-1} \binom{x}{n-1} + \cdots + c_0 \binom{x}{0}.$$

The converse is true, also.

Proof. We use induction on n. The case $n = 1$ is trivial; Now assume that $n > 1$. Polynomial $Q(x) = P(x+1) - P(x)$ is of degree $n-1$ and takes integer values at all integer points, so by the inductive hypothesis there exist $a_0, \ldots, a_{n-1} \in \mathbb{Z}$ such that

$$Q(x) = a_{n-1} \binom{x}{n-1} + \cdots + a_0 \binom{x}{0}.$$

For every integer $x > 0$ we have $P(x) = P(0) + Q(0) + Q(1) + \cdots + Q(x-1)$. Using the identity $\binom{0}{k} + \binom{1}{k} + \cdots + \binom{x-1}{k} = \binom{x}{k+1}$ for every integer k we obtain the desired representation of $P(x)$:

$$P(x) = a_{n-1} \binom{x}{n} + \cdots + a_0 \binom{x}{1} + P(0). \quad \square$$

Problem 8. *Suppose that a natural number m and a real polynomial $R(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_0$ are such that $R(x)$ is an integer divisible by m whenever x is an integer. Prove that $n! a_n$ is divisible by m.*

Solution. Apply the previous theorem on polynomial $\frac{1}{m} R(x)$ (with the same notation). The leading coefficient of this polynomial equals $c_n + n c_{n-1} + \cdots + n! c_0$, and the statement follows immediately. △

4 Irreducibility

Polynomial $P(x)$ with integer coefficients is said to be *irreducible* over $\mathbb{Z}[x]$ if it cannot be written as a product of two nonconstant polynomials with integer coefficients.

Example 5. *Every quadratic or cubic polynomial with no rational roots is irreducible over $\mathbb{Z}$. Such are e.g. x^2-x-1 and $2x^3-4x+1$.*

One analogously defines (ir)reducibility over the sets of polynomials with e.g. rational, real or complex coefficients. However, of the mentioned, only reducibility over $\mathbb{Z}[x]$ is of interest. Gauss' Lemma below claims that the reducibility over $\mathbb{Q}[x]$ is equivalent to the reducibility over $\mathbb{Z}[x]$. In addition, we have already shown that a real polynomial is always reducible into linear and quadratic factors over $\mathbb{R}[x]$, while a complex polynomial is always reducible into linear factors over $\mathbb{C}[x]$.

Theorem 16 (Gauss' Lema). *If a polynomial $P(x)$ with integer coefficients is reducible over $\mathbb{Q}[x]$, then it is reducible over $\mathbb{Z}[x]$, also.*

Proof. Suppose that $P(x)=a_nx^n+\cdots+a_0=Q(x)R(x)\in\mathbb{Z}[x]$, where $Q(x)$ and $R(x)$ nonconstant polynomials with rational coefficients. Let q and r be the smallest natural numbers such that the polynomials $qQ(x)=q_kx^k+\cdots+q_0$ and $rR(x)=r_mx^m+\cdots+r_0$ have integer coefficients. Then $qrP(x)=qQ(x)\cdot rR(x)$ is a factorization of the polynomial $qrP(x)$ into two polynomials from $\mathbb{Z}[x]$. Based on this, we shall construct such a factorization for $P(x)$.

Let p be an arbitrary prime divisor of q. All coefficients of $P(x)$ are divisible by p. Let i be such that $p\mid q_0,q_1,\ldots,q_{i-1}$, but $p\nmid q_i$. We have $p\mid a_i=q_0r_i+\cdots+q_ir_0\equiv q_ir_0 \pmod p$, which implies that $p\mid r_0$. Furthermore, $p\mid a_{i+1}=q_0r_{i+1}+\cdots+q_ir_1+q_{i+1}r_0\equiv q_ir_1 \pmod p$, so $p\mid r_1$. Continuing in this way, we deduce that $p\mid r_j$ for all j. Hence $rR(x)/p$ has integer coefficients. We have thus obtained a factorization of $\frac{rq}{p}P(x)$ into two polynomials from $\mathbb{Z}[x]$. Continuing this procedure and taking other values for p we shall eventually end up with a factorization of $P(x)$ itself. □

From now on, unless otherwise specified, by "irreducibility" we mean irreducibility over $\mathbb{Z}[x]$.

Problem 9. *If a_1, a_2, …, a_n are integers, prove that the polynomial $P(x)=(x-a_1)(x-a_2)\cdots(x-a_n)-1$ is irreducible.*

Solution. Suppose that $P(x)=Q(x)R(x)$ for some nonconstant polynomials $Q,R\in\mathbb{Z}[x]$. Since $Q(a_i)R(a_i)=-1$ for $i=1,\ldots,n$, we have $Q(a_i)=1$ and $R(a_i)=-1$ or $Q(a_i)=-1$ and $R(a_i)=1$; either way, we have $Q(a_i)+R(a_i)=0$. It follows that the polynomial $Q(x)+R(x)$ (which is obviously nonzero) has n zeros $a_1,\ldots,a_n$ which is impossible for its degree is less than n. △

Theorem 17 (Extended Eisenstein's Criterion). *Let $P(x)=a_nx^n+\cdots+a_1x+a_0$ be a polynomial with integer coefficients. If there exist a prime number p and an integer $k\in\{0,1,\ldots,n-1\}$ such that*

$$p\mid a_0,a_1,\ldots,a_k,\ \ p\nmid a_{k+1}\ \text{ and }\ p^2\nmid a_0,$$

then $P(x)$ has an irreducible factor of a degree greater than k.

In particular, if p can be taken so that $k=n-1$, then $P(x)$ is irreducible.

Proof. Like in the proof of Gauss's lemma, suppose that $P(x)=Q(x)R(x)$, where $Q(x)=q_kx^k+\cdots+q_0$ and $R(x)=r_mx^m+\cdots+r_0$ are polynomials from $\mathbb{Z}[x]$. Since $a_0=q_0r_0$ is divisible by p and not by p^2, exactly one of q_0,r_0 is a multiple of p. Assume that $p\mid q_0$ and $p\nmid r_0$. Further, $p\mid a_1=q_0r_1+q_1r_0$, implying that $p\mid q_1r_0$, i.e. $p\mid q_1$, and so on. We conclude that all coefficients $q_0,q_1,\ldots,q_k$ are divisible by p, but $p\nmid q_{k+1}$. It follows that $\deg Q\geq k+1$. □

Problem 10. *Given an integer $n>1$, consider the polynomial $f(x)=x^n+5x^{n-1}+3$. Prove that there are no nonconstant polynomials $g(x),h(x)$ with integer coefficients such that $f(x)=g(x)h(x)$. (IMO93-1)*

Solution. By the (extended) Eisenstein criterion, f has an irreducible factor of degree at least $n-1$. Since f has no integer zeros, it must be irreducible. $\triangle$

Problem 11. *If p is a prime number, prove that the polynomial $\Phi_p(x) = x^{p-1} + \cdots + x + 1$ is irreducible.*

Solution. Instead of $\Phi_p(x)$, we shall consider $\Phi_p(x+1)$ and show that it is irreducible, which will clearly imply that so is Φ_p. We have

$$\Phi_p(x+1) = \frac{(x+1)^p - 1}{x} = x^{p-1} + \binom{p}{p-1}x^{p-2} + \cdots + \binom{p}{2}x + p.$$

This polynomial satisfies all the assumptions of Eisenstein's criterion, based on which it is irreducible. $\triangle$

In investigating reducibility of a polynomial, it can be useful to investigate its zeros and their modules. The following problems provide us an illustration.

Problem 12. *Prove that the polynomial $P(x) = x^n + 4$ is irreducible over $\mathbb{Z}[x]$ if and only if n is a multiple of 4.*

Solution. All zeros of polynomial P have the modulus equal to $2^{2/n}$. If Q and R are polynomials from $\mathbb{Z}[x]$ and $\deg Q = k$, then $|Q(0)|$ is the product of the modules of the zeros of Q and equals $2^{2k/n}$; since this should be an integer, we deduce that $n = 2k$.

If k is odd, polynomial Q has a real zero, which is impossible since $P(x)$ has none. Therefore, $2 \mid k$ and $4 \mid n$. $\triangle$

If the zeros cannot be exactly determined, one should find a good enough bound. Estimating complex zeros of a polynomial is not always simple. Our main tool is the triangle inequality for complex numbers:

$$|x| - |y| \le |x+y| \le |x| + |y|.$$

Consider a polynomial $P(x) = a_n x^n + a_{n-k} x n - k + \cdots + a_1 x + a_0$ with complex coefficients ($a_n \neq 0$). Let α be its zero. If M is a real number such that $|a_i| < M|a_n|$ for all i, it holds that

$$0 = |P(\alpha)| \ge |a_n||\alpha|^n - M|a_n|(|\alpha|^{n-k} + \cdots + |\alpha| + 1) > |a_n||\alpha|^n \left(1 - \frac{M}{|\alpha|^{k-1}(|\alpha|-1)}\right),$$

which yields $|\alpha|^{k-1}(|\alpha|-1) < M$. We thus come to the following estimate:

Theorem 18. *Let $P(x) = a_n x^n + \cdots + a_0$ be a complex polynomial with $a_n \neq 0$ and $M = \max\limits_{0 \le k < n} \left|\frac{a_k}{a_n}\right|$. If $a_{n-1} = \cdots = a_{n-k+1} = 0$, then all roots of the polynomial P are less than $1 + \sqrt[k]{M}$ in modulus.*

In particular, for $k = 1$, each zero of $P(x)$ is of modulus less than $M + 1$. $\square$

Problem 13. *If $\overline{a_n \ldots a_1 a_0}$ is a decimal representation of a prime number and $a_n > 1$, prove that the polynomial $P(x) = a_n x^n + \cdots + a_1 x + a_0$ is irreducible. (*BMO 1989.2*)*

Solution. Suppose that Q and R are nonconstant polynomials from $\mathbb{Z}[x]$ with $Q(x)R(x) = P(x)$. Let $x_1, \ldots, x_k$ be the zeros of Q and $x_{k+1}, \ldots, x_n$ be the zeros of R. The condition of the problem means that $P(10) = Q(10)R(10)$ is a prime, so we can assume w.l.o.g. that

$$|Q(10)| = (10 - x_1)(10 - x_2)\cdots(10 - x_k) = 1.$$

On the other hand, by the estimate in 18, each zero x_i has a modulus less than $1 + 9/2 = 11/2 < 9$; hence $|10 - x_i| > 1$ for all i, contradicting the above inequality. $\triangle$

Problem 14. *Let $p > 2$ be a prime number and $P(x) = x^p - x + p$.*

1. *Prove that all zeros of polynomial P are less than $p^{\frac{1}{p-1}}$ in modulus.*

2. *Prove that the polynomial $P(x)$ is irreducible.*

Solution.

1. Let y be a zero of P. Then $|y|^p - |y| \leq |y^p - y| = p$. If we assume that $|y| \geq p^{\frac{1}{p-1}}$, we obtain

$$|y|^p - |y| \geq (p-1)p^{\frac{1}{p-1}} > p,$$

a contradiction. Here we used the inequality $p^{\frac{1}{p-1}} > \frac{p}{p-1}$ which follows for example from the binomial expansion of $p^{p-1} = ((p-1)+1)^{p-1}$.

2. Suppose that $P(x)$ is the product of two nonconstant polynomials $Q(x)$ and $R(x)$ with integer coefficients. One of these two polynomials, say Q, has the constant term equal to $\pm p$. On the other hand, the zeros $x_1, \ldots, x_k$ of Q satisfy $|x_1|, \ldots, |x_k| < p^{\frac{1}{p-1}}$ by part (a), and $x_1 \cdots x_k = \pm p$, so we conclude that $k \geq p$, which is impossible. $\triangle$

5 Interpolating polynomials

A polynomial of n-th degree is uniquely determined, given its values at $n+1$ points. So, suppose that P is an n-th degree polynomial and that $P(x_i) = y_i$ in different points $x_0, x_1, \ldots, x_n$. There exist unique polynomials $E_0, E_1, \ldots, E_n$ of n-th degree such that $E_i(x_i) = 1$ and $E_i(x_j) = 0$ for $j \neq i$. Then the polynomial

$$P(x) = y_0 E_0(x) + y_1 E_1(x) + \cdots + y_n E_n(x)$$

has the desired properties: indeed, $P(x_i) = \sum_j y_j E_j(x_i) = y_i E_i(x_i) = y_i$. It remains to find the polynomials $E_0, \ldots, E_n$. A polynomial that vanishes at the n points x_j, $j \neq i$, is divisible by $\prod_{j \neq i}(x - x_j)$, from which we easily obtain $E_i(x) = \prod_{j \neq i} \frac{(x-x_j)}{(x_i - x_j)}$. This shows that:

Theorem 19 (Newton's interpolating polynomial). *For given numbers $y_0, \ldots, y_n$ and distinct x_0, $\ldots$, x_n there is a unique polynomial $P(x)$ of n-th degree such that $P(x_i) = y_i$ for $i = 0, 1, \ldots, n$. This polynomial is given by the formula*

$$P(x) = \sum_{i=0}^{n} y_i \prod_{j \neq i} \frac{(x - x_j)}{(x_i - x_j)}. \ \square$$

Example 6. *Find the cubic polynomial Q such that $Q(i) = 2^i$ for $i = 0, 1, 2, 3$.*

Solution. $Q(x) = \frac{(x-1)(x-2)(x-3)}{-6} + \frac{2x(x-2)(x-3)}{2} + \frac{4x(x-1)(x-3)}{-2} + \frac{8x(x-1)(x-2)}{6} = \frac{x^3+5x+6}{6}$. $\triangle$

In order to compute the value of a polynomial given in this way in some point, sometimes we do not need to determine its Newton's polynomial. In fact, Newton's polynomial has an unpleasant property of giving the answer in a complicated form.

Example 7. *If the polynomial P of n-th degree takes the value 1 in points $0, 2, 4, \ldots, 2n$, compute $P(-1)$.*

Solution. $P(x)$ is of course identically equal to 1, so $P(-1) = 1$. But if we apply the Newton polynomial, here is what we get:

$$P(1) = \sum_{i=0}^{n} \prod_{j \neq i} \frac{1-2i}{(2j-2i)} = \sum_{i=0}^{n} \prod_{j \neq i} \frac{-1-2j}{(2i-2j)} = \frac{(2n+1)!!}{2^n} \sum_{i=1}^{n+1} \frac{(-1)^{n-i}}{(2i+1)i!(n-i)!}. \ \triangle$$

Instead, it is often useful to consider the *finite difference* of polynomial P, defined by $P^{[1]}(x) = P(x+1) - P(x)$, which has the degree by 1 less than that of P. Further, we define the k-th finite difference, $P^{[k]} = (P^{[k-1]})^{[1]}$, which is of degree $n-k$ (where $\deg P = n$). A simple induction gives a general formula

$$P^{[k]} = \sum_{i=0}^{k} (-1)^{k-i} \binom{k}{i} P(x+i).$$

In particular, $P^{[n]}$ is constant and $P^{[n+1]} = 0$, which leads to

Theorem 20. $P(x+n+1) = \sum_{i=0}^{n} (-1)^{n-i} \binom{n+1}{i} P(x+i)$. □

Problem 15. *Polynomial P of degree n satisfies $P(i) = \binom{n+1}{i}^{-1}$ for $i = 0, 1, \ldots, n$. Evaluate $P(n+1)$.*

Solution. We have

$$0 = \sum_{i=0}^{n+1} (-1)^i \binom{n+1}{i} P(i) = (-1)^{n+1} P(n+1) + \begin{cases} 1, & 2 \mid n; \\ 0, & 2 \nmid n. \end{cases}$$

It follows that $P(n+1) = \begin{cases} 1, & 2 \mid n; \\ 0, & 2 \nmid n. \end{cases}$ △

Problem 16. *If $P(x)$ is a polynomial of an even degree n with $P(0) = 1$ and $P(i) = 2^{i-1}$ for $i = 1, \ldots, n$, prove that $P(n+2) = 2P(n+1) - 1$.*

Solution. We observe that $P^{[1]}(0) = 0$ i $P^{[1]}(i) = 2^{i-1}$ for $i = 1, \ldots, n-1$; furthermore, $P^{[2]}(0) = 1$ i $P^{[2]}(i) = 2^{i-1}$ for $i = 1, \ldots, n-2$, etc. In general, it is easily seen that $P^{[k]}(i) = 2^{i-1}$ for $i = 1, \ldots, n-k$, and $P^{[k]}(0)$ is 0 for k odd and 1 for k even. Now

$$P(n+1) = P(n) + P^{[1]}(n) = \cdots = P(n) + P^{[1]}(n-1) + \cdots + P^{[n]}(0) = \begin{cases} 2^n, & 2 \mid n; \\ 2^n - 1, & 2 \nmid n. \end{cases}$$

Similarly, $P(n+2) = 2^{2n+1} - 1$. △

6 Applications of Calculus

The derivative of a polynomial $P(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0$ is given by

$$P'(x) = n a_n x^{n-1} + (n-1) a_{n-1} x^{n-2} + \cdots + a_1.$$

The inverse operation, the indefinite integral, is given by

$$\int P(x)dx = \frac{a_n}{n+1} x^{n+1} + \frac{a_{n-1}}{n} x^n + \cdots + a_0 x + C.$$

If the polynomial P is not given by its coefficients but rather by its canonical factorization, as $P(x) = (x - x_1)^{k_1} \cdots (x - x_n)^{k_n}$, a more suitable expression for the derivative is obtained by using the logarithmic derivative rule or product rule:

$$P'(x) = P(x) \left(\frac{k_1}{x - x_1} + \cdots + \frac{k_n}{x - x_n} \right).$$

A similar formula can be obtained for the second derivative.

Problem 17. *Suppose that real numbers* $0 = x_0 < x_1 < \cdots < x_n < x_{n+1} = 1$ *satisfy*

$$\sum_{j=0,\, j\neq i}^{n+1} \frac{1}{x_i - x_j} = 0 \quad za\ i = 1, 2, \ldots, n. \tag{1}$$

Prove that $x_{n+1-i} = 1 - x_i$ *for* $i = 1, 2, \ldots, n$.

Solution. Let $P(x) = (x - x_0)(x - x_1)\cdots(x - x_n)(x - x_{n+1})$. We have

$$P'(x) = \sum_{j=0}^{n+1} \frac{P(x)}{x - x_j} \quad \text{i} \quad P''(x) = \sum_{j=0}^{n+1} \sum_{k\neq j} \frac{P(x)}{(x - x_j)(x - x_k)}\ .$$

Therefore

$$P''(x_i) = 2P'(x_i) \sum_{j\neq i} \frac{1}{(x_i - x_j)}$$

for $i = 0, 1, \ldots, n+1$. Thus the condition of the problem is equivalent to $P''(x_i) = 0$ for $i = 1, 2, \ldots, n$. Therefore

$$x(x-1)P''(x) = (n+2)(n+1)P(x).$$

It is easy to see that there is a unique monic polynomial of degree $n+2$ satisfying the above differential equation. On the other hand, the monic polynomial $Q(x) = (-1)^n P(1-x)$ satisfies the same equation and has degree $n+2$, so we must have $(-1)^n P(1-x) = P(x)$, which implies the statement. $\triangle$

What makes derivatives of polynomials especially suitable is their property of preserving multiple zeros.

Theorem 21. *If* $(x - \alpha)^k \mid P(x)$, *then* $(x - \alpha)^{k-1} \mid P'(x)$.

Proof. If $P(x) = (x - \alpha)^k Q(x)$, then $P'(x) = (x - \alpha)^k Q'(x) + k(x - \alpha)^{k-1} Q(x)$. $\square$

Problem 18. *Determine a real polynomial* $P(x)$ *of degree at most 5 which leaves remainders* -1 *and 1 upon division by* $(x-1)^3$ *and* $(x+1)^3$, *respectively.*

Solution. If $P(x) + 1$ has a triple zero at point 1, then its derivative $P'(x)$ has a double zero at that point. Similarly, $P'(x)$ has a double zero at point -1 too. It follows that $P'(x)$ is divisible by the polynomial $(x-1)^2(x+1)^2$. Since $P'(x)$ is of degree at most 4, it follows that

$$P'(x) = c(x-1)^2(x+1)^2 = c(x^4 - 2x^2 + 1)$$

for some constant c. Now $P(x) = c(\frac{1}{5}x^5 - \frac{2}{3}x^3 + x) + d$ for some real numbers c and d. The conditions $P(-1) = 1$ and $P(1) = -1$ now give us $c = -15/8$, $d = 0$ and

$$P(x) = -\frac{3}{8}x^5 + \frac{5}{4}x^3 - \frac{15}{8}x.\ \triangle$$

Problem 19. *For polynomials* $P(x)$ *and* $Q(x)$ *and an arbitrary* $k \in \mathbb{C}$, *denote*

$$P_k = \{z \in \mathbb{C} \mid P(z) = k\} \quad \textit{and} \quad Q_k = \{z \in \mathbb{C} \mid Q(z) = k\}.$$

Prove that $P_0 = Q_0$ *and* $P_1 = Q_1$ *imply that* $P(x) = Q(x)$.

Solution. Let us assume w.l.o.g. that $n = \deg P \geq \deg Q$. Let $P_0 = \{z_1, z_2, \ldots, z_k\}$ and $P_1 = \{z_{k+1}, z_{k+2}, \ldots, z_{k+m}\}$. Polynomials P and Q coincide at $k+m$ points $z_1, z_2, \ldots, z_{k+m}$. The result will follow if we show that $k + m > n$.

We have

$$P(x) = (x-z_1)^{\alpha_1}\cdots(x-z_k)^{\alpha_k} = (x-z_{k+1})^{\alpha_{k+1}}\cdots(x-z_{k+m})^{\alpha_{k+m}} + 1$$

for some natural numbers $\alpha_1,\ldots,\alpha_{k+m}$. Let us consider $P'(x)$. We know that it is divisible by $(x-z_i)^{\alpha_i-1}$ for $i=1,2,\ldots,k+m$; hence,

$$\prod_{i=1}^{k+m}(x-z_i)^{\alpha_i-1} \mid P'(x).$$

Therefore, $2n-k-m = \deg\prod_{i=1}^{k+m}(x-z_i)^{\alpha_i-1} \le \deg P' = n-1$, i.e. $k+m \ge n+1$, as desired. △

Even if P has no multiple zeros, certain relations between zeros of P and P' still hold. For example, the following statement holds for all differentiable functions.

Theorem 22 (Rolle's Theorem). *Between every two zeros of a polynomial $P(x)$ there is a zero of $P'(x)$.*

Corollary. If all zeros of $P(x)$ are real, then so are all zeros of $P'(x)$.)

Proof. Let $a<b$ be two zeros of polynomial P. Assume w.l.o.g. that $P'(a)>0$ and consider the point c in the interval $[a,b]$ in which P attains a local maximum (such a point exists since the interval $[a,b]$ is compact). We know that $P(x) = P(c) + (x-c)[P'(c)+o(1)]$. If for example $P'(c)>0$ (the case $P'(c)<0$ leads to a similar contradiction), then $P(x)>P(c)$ would hold in a small neighborhood of c, a contradiction. It is only possible that $P'(c)=0$, so c is a root of $P'(x)$ between a and b. □

7 Symmetric polynomials

A symmetric polynomial in variables $x_1,\ldots,x_n$ is every polynomial that is not varied by permuting the indices of the variables. For instance, polynomial x_1^2 is symmetric as a polynomial in x_1 (no wonder), but is not symmetric as a polynomial in x_1,x_2 as changing places of the indices 1 and 2 changes it to the polynomial x_2^2.

Definition 2. *The polynomial $P(x_1,x_2,\ldots,x_n)$ is* symmetric *if, for every permutation π of $\{1, 2, \ldots, n\}$, $P(x_1,x_2,\ldots,x_n) \equiv P(x_{\pi(1)},x_{\pi(2)},\ldots,x_{\pi(n)})$.*

An obvious property of a symmetric polynomial is that its coefficients at two terms of the forms $x_1^{i_1}\cdots x_n^{i_n}$ and $x_1^{j_1}\cdots x_n^{j_n}$, where $(j_1,\ldots,j_n)$ is a permutation $(i_1,\ldots,i_n)$, always coincide. For example, if the expansion of a symmetric polynomial in x,y,z contains the terms x^2y, then it also contains x^2z, xy^2, etc, with the same coefficient.

Thus, the polynomials σ_k $(1\le k\le n)$ introduced in section 2 are symmetric. Also symmetric is e.g. the polynomial $x_1^2+x_2^2$.

A symmetric polynomial is said to be *homogenous* if all its terms are of the same degree. Equivalently, polynomial T is homogenous of degree d if $T(tx_1,\ldots,tx_n) = t^dT(x_1,\ldots,x_n)$ holds for all x and t. For instance, $x_1^2+x_2^2$ is homogenous of degree $d=2$, but $x_1^2+x_2^2+1$, although symmetric, is not homogenous.

Every symmetric polynomial in $x_1,\ldots,x_n$ can be written as a sum of homogenous polynomials. Moreover, it can also be represented as a linear combination of certain "bricks". These bricks are the polynomials

$$T_a = \sum x_1^{a_{i_1}}\cdots x_n^{a_{i_n}} \qquad (*)$$

for each n-tuple $a=(a_1,\ldots,a_n)$ of nonnegative integers with $a_1\ge\cdots\ge a_n$, where the summation goes over all permutations $(i_1,\ldots,i_n)$ of the indices $1,\ldots,n$. In the expression for T_a the same summand can occur more than once, so we define S_a as the sum of the *different* terms in $(*)$. The polynomial T_a is always an integral multiple of S_a. For instance,

$$T_{(2,2,0)} = 2(x_1^2x_2^2+x_2^2x_3^2+x_3^2x_1^2) = 2S_{(2,2,0)}.$$

All the n-tuples a of degree $d = a_1 + \cdots + a_n$ can be ordered in a lexicographic order so that

$$a > a' \quad \text{if} \quad s_1 = s_1', \ldots, s_k = s_k' \text{ and } s_{k+1} > s_{k+1}' \text{ for some } k \geq 1,$$

where $s_i = a_1 + \cdots + a_i$. In this ordering, the least n-tuple is $m = (x+1, \ldots, x+1, x, \ldots, x)$, where $x = [d/n]$ and $x+1$ occurs $d - n[d/n]$ times.

The polynomials T_a can be multiplies according to the following simple formula:

Theorem 23. *If $a = (a_1, \ldots, a_n)$ and $b = (b_1, \ldots, b_n)$ are n-tuples of nonnegative integers, it holds that*

$$T_a \cdot T_b = \sum_{\pi} T_{a+\pi(b)},$$

where the sum goes over all permutations $\pi(b)$ of the n-tuple b. (We define $(x_i)_{i=1}^n + (y_i)_{i=1}^n = (x_i + y_i)_{i=1}^n$.)

Proof. It suffices to observe that

$$x_1^{\pi_1(b)} \cdots x_n^{\pi_n(b)} T_a = \sum x_{i_1}^{a_1 + \pi_{i_1}(b)} \cdots x_{i_n}^{a_n + \pi_{i_n}(b)},$$

and to sum up over all permutations π. □

There are infinitely many mentioned bricks, and these are obviously not mutually independent. We need simpler elements which are independent and using which one can express every symmetric polynomial by basic operations. It turns out that these atoms are $\sigma_1, \ldots, \sigma_n$.

Example 8. *The following polynomials in x, y, z can be written in terms of $\sigma_1, \sigma_2, \sigma_3$:*
$xy + yz + zx + x + y + z = \sigma_2 + \sigma_1$;
$x^2y + x^2z + y^2x + y^2z + z^2x + z^2y = \sigma_1\sigma_2 - 3\sigma_3$;
$x^2y^2 + y^2z^2 + z^2x^2 = \sigma_2^2 - 2\sigma_1\sigma_3$.

Theorem 24. *Every symmetric polynomial in $x_1, \ldots, x_n$ can be represented in the form of a polynomial in $\sigma_1, \ldots, \sigma_n$. Moreover, a symmetric polynomial with integer coefficients is also a polynomial in $\sigma_1, \ldots, \sigma_n$ with integer coefficients.*

Proof. It is enough to prove the statement for the polynomials S_a of degree d (for each d). Assuming that it holds for the degrees less than d, we use induction on n-tuples a. The statement is true for the smallest n-tuple m: Indeed, $S_m = \sigma_n^q \sigma_r$, where $d = nq + r$, $0 \leq r < n$. Now suppose that the statement is true for all S_b with $b < a$; we show that it also holds for S_a.

Suppose that $a = (a_1, \ldots, a_n)$ with $a_1 = \cdots = a_k > a_{k+1}$ ($k \geq 1$). Consider the polynomial $S_a - \sigma_k S_{a'}$, where $a' = (a_1 - 1, \ldots, a_k - 1, a_{k+1}, \ldots, a_n)$. According to theorem 23 it is easy to see that this polynomial is of the form $\sum_{b<a} c_b S_b$, where c_b are integers, and is therefore by the inductive hypothesis representable in the form of a polynomial in σ_i with integer coefficients. □

The proof of the previous theorem also gives us an algorithm for expressing each symmetric polynomial in terms of the σ_i. Nevertheless, for some particular symmetric polynomials there are simpler formulas.

Theorem 25 (Newton's Theorem on Symmetric Polynomials). *If we denote $s_k = x_1^k + x_2^k + \cdots + x_n^k$, then:*

$$\begin{aligned} k\sigma_k &= s_1\sigma_{k-1} - s_2\sigma_{k-2} + \cdots + (-1)^k s_{k-1}\sigma_1 + (-1)^{k+1} s_k; \\ s_m &= \sigma_1 s_{m-1} - \sigma_2 s_{m-2} + \cdots + (-1)^{n-1}\sigma_n s_{m-n} \quad za\ m \geq n. \end{aligned}$$

(All the polynomials are in n variables.)

Proof. Direct, for example by using the formula 23. □

Problem 20. *Suppose that complex numbers $x_1, x_2, \ldots, x_k$ satisfy*

$$x_1^j + x_2^j + \cdots + x_k^j = n, \qquad \text{for } j = 1, 2, \ldots, k,$$

where n, k are given positive integers. Prove that

$$(x - x_1)(x - x_2)\ldots(x - x_k) = x^k - \binom{n}{1}x^{k-1} + \binom{n}{2}x^{k-2} - \cdots + (-1)^k\binom{n}{k}.$$

Solution. We are given $s_k = n$ for $k = 1, \ldots, n$. The Newton's theorem gives us $\sigma_1 = n$, $\sigma_2 = \frac{1}{2}(n\sigma_1 - n) = \binom{n}{2}$, $\sigma_3 = \frac{1}{3}(n\sigma_2 - n\sigma_1 + n) = \binom{n}{3}$, etc. We prove by induction on k that $\sigma_k = \binom{n}{k}$. If this holds for $1, \ldots, k-1$, we have

$$\sigma_k = \frac{n}{k}\left[\binom{n}{k-1} - \binom{n}{k-2} + \binom{n}{k-3} - \cdots\right].$$

Since $\binom{n}{i} = \binom{n-1}{i} + \binom{n-1}{i-1}$, the above equality telescopes to $\sigma_k = \frac{n}{k}\binom{n-1}{k-1}$, which is exactly equal to $\binom{n}{k}$. $\triangle$

8 Problems

1. A monic polynomial $f(x)$ of fourth degree satisfies $f(1) = 10$, $f(2) = 20$ and $f(3) = 30$. Determine $f(12) + f(-8)$.

2. Consider complex polynomials $P(x) = x^n + a_1x^{n-1} + \cdots + a_n$ with the zeros $x_1, \ldots, x_n$, and $Q(x) = x^n + b_1x^{n-1} + \cdots + b_n$ with the zeros $x_1^2, \ldots, x_n^2$. Prove that if $a_1 + a_3 + a_5 + \cdots$ and $a_2 + a_4 + a_6 + \cdots$ are real numbers, then $b_1 + b_2 + \cdots + b_n$ is also real.

3. If a polynomial P with real coefficients satisfies for all x

$$P(\cos x) = P(\sin x),$$

show that there exists a polynomial Q such that $P(x) = Q(x^4 - x^2)$ for each x.

4. (a) Prove that for each $n \in \mathbb{N}$ there is a polynomial T_n with integer coefficients and the leading coefficient 2^{n-1} such that $T_n(\cos x) = \cos nx$ for all x.

   (b) Prove that the polynomials T_n satisfy $T_{m+n} + T_{m-n} = 2T_mT_n$ for all $m, n \in \mathbb{N}$, $m \geq n$.

   (c) Prove that the polynomial U_n given by $U_n(2x) = 2T_n(x)$ also has integer coefficients and satisfies $U_n(x + x^{-1}) = x^n + x^{-n}$.

   The polynomials $T_n(x)$ are known as the *Chebyshev polynomials.*

5. Prove that if $\cos\frac{p}{q}\pi = a$ is a rational number for some $p, q \in \mathbb{Z}$, then $a \in \{0, \pm\frac{1}{2}, \pm 1\}$.

6. Prove that the maximum in absolute value of any monic real polynomial of n-th degree on $[-1, 1]$ is not less than $\frac{1}{2^{n-1}}$.

7. The polynomial P of n-th degree is such that, for each $i = 0, 1, \ldots, n$, $P(i)$ equals the remainder of i modulo 2. Evaluate $P(n+1)$.

8. A polynomial $P(x)$ of n-th degree satisfies $P(i) = \frac{1}{i}$ for $i = 1, 2, \ldots, n+1$. Find $P(n+2)$.

9. Let $P(x)$ be a real polynomial.

   (a) If $P(x) \geq 0$ for all x, show that there exist real polynomials $A(x)$ and $B(x)$ such that $P(x) = A(x)^2 + B(x)^2$.

(b) If $P(x) \geq 0$ for all $x \geq 0$, show that there exist real polynomials $A(x)$ and $B(x)$ such that $P(x) = A(x)^2 + xB(x)^2$.

10. Prove that if the equation $Q(x) = ax^2 + (c-b)x + (e-d) = 0$ has real roots greater than 1, where $a, b, c, d, e \in \mathbb{R}$, then the equation $P(x) = ax^4 + bx^3 + cx^2 + dx + e = 0$ has at least one real root.

11. A monic polynomial P with real coefficients satisfies $|P(i)| < 1$. Prove that there is a root $z = a + bi$ of P such that $(a^2 + b^2 + 1)^2 < 4b^2 + 1$.

12. For what real values of a does there exist a rational function $f(x)$ that satisfies $f(x^2) = f(x)^2 - a$? (A rational function is a quotient of two polynomials.)

13. Find all polynomials P satisfying $P(x^2 + 1) = P(x)^2 + 1$ for all x.

14. Find all P for which $P(x)^2 - 2 = 2P(2x^2 - 1)$.

15. If the polynomials P and Q each have a real root and

$$P(1 + x + Q(x)^2) = Q(1 + x + P(x)^2),$$

prove that $P \equiv Q$.

16. Find all polynomials $P(x)$ with real coefficients satisfying the equality

$$P(a-b) + P(b-c) + P(c-a) = 2P(a+b+c)$$

for all triples (a, b, c) of real numbers such that $ab + bc + ca = 0$. *(IMO04-2)*

17. A sequence of integers $(a_n)_{n=1}^{\infty}$ has the property that $m - n \mid a_m - a_n$ for any distinct $m, n \in \mathbb{N}$. Suppose that there is a polynomial $P(x)$ such that $|a_n| < P(n)$ for all n. Show that there exists a polynomial $Q(x)$ such that $a_n = Q(n)$ for all n.

18. Let $P(x)$ be a polynomial of degree $n > 1$ with integer coefficients and let k be a natural number. Consider the polynomial $Q(x) = P(P(\ldots P(P(x))\ldots))$, where P is applied k times. Prove that there exist at most n integers t such that $Q(t) = t$. *(IMO06-5)*

19. If P and Q are monic polynomials such that $P(P(x)) = Q(Q(x))$, prove that $P \equiv Q$.

20. Let m, n and a be natural numbers and $p < a - 1$ a prime number. Prove that the polynomial $f(x) = x^m(x-a)^n + p$ is irreducible.

21. Prove that the polynomial $F(x) = (x^2 + x)^{2^n} + 1$ is irreducible for all $n \in \mathbb{N}$.

22. A polynomial $P(x)$ has the property that for every $y \in \mathbb{Q}$ there exists $x \in \mathbb{Q}$ such that $P(x) = y$. Prove that P is a linear polynomial.

23. Let $P(x)$ be a monic polynomial of degree n whose zeros are $i-1, i-2, \ldots, i-n$ (where $i^2 = -1$) and let $R(x)$ and $S(x)$ be the real polynomials such that $P(x) = R(x) + iS(x)$. Prove that the polynomial $R(x)$ has n real zeros.

24. Let a, b, c be natural numbers. Prove that if there exist coprime polynomials P, Q, R with complex coefficients such that

$$P^a + Q^b = R^c,$$

then $\frac{1}{a} + \frac{1}{b} + \frac{1}{c} > 1$.

Corollary: The Last Fermat Theorem for polynomials.

25. Suppose that all zeros of a monic polynomial $P(x)$ with integer coefficients are of module 1. Prove that there are only finitely many such polynomials of any given degree; hence show that all its zeros are actually roots of unity, i.e. $P(x) \mid (x^n - 1)^k$ for some natural n, k.

9 Solutions

1. The polynomial $f(x)-10x$ vanishes at points $x=1,2,3$, so it is divisible by polynomial $(x-1)(x-2)(x-3)$. The monicity implies that $f(x)-10x=(x-1)(x-2)(x-3)(x-c)$ for some c. Now

$$f(12)+f(-8)=11\cdot 10\cdot 9\cdot(12-c)+120+(-9)(-10)(-11)(-8-c)-80=19840.$$

2. Note that $Q(x^2)=\prod(x^2-x_i^2)=\prod(x-x_i)\cdot\prod(x+x_i)=(-1)^nP(x)P(-x)$. We now have

$$b_1+b_2+\cdots+b_n=Q(1)-1=(-1)^nP(1)P(-1)-1=(-1)^n(1+B-A)(1+B+A),$$

where $A=a_1+a_3+a_5+\cdots$ and $B=a_2+a_4+\cdots$.

3. It follows from the conditions that $P(-\sin x)=P(\sin x)$, i.e. $P(-t)=P(t)$ for infinitely many t, so the polynomials $P(x)$ and $P(-x)$ coincide. Therefore, $P(x)=S(x^2)$ for some polynomial S. Now $S(\cos^2 x)=S(\sin^2 x)$ for all x, i.e. $S(1-t)=S(t)$ for infinitely many t, which implies $S(x)\equiv S(1-x)$. This is equivalent to $R(x-\frac{1}{2})=R(\frac{1}{2}-x)$, i.e. $R(y)\equiv R(-y)$, where R is a polynomial such that $S(x)=R(x-\frac{1}{2})$. Now $R(x)=T(x^2)$ for some polynomial T, and therefore $P(x)=S(x^2)=R(x^2-\frac{1}{2})=T(x^4-x^2+\frac{1}{4})=Q(x^4-x^2)$ for some polynomial Q.

4. (a) Clearly, $T_0(x)=1$ and $T_1(x)=x$ satisfy the requirements. For $n>1$ we use induction on n. Since $\cos(n+1)x=2\cos x\cos nx-\cos(n-1)x$, we can define $T_{n+1}=2T_1T_n-T_{n-1}$. Since T_1T_n and T_{n-1} are of degrees $n+1$ and $n-1$ respectively, T_{n+1} is of degree $n+1$ and has the leading coefficient $2\cdot 2^n=2^{n+1}$. It also follows from the construction that all its coefficients are integers.

   (b) The relation follows from the identity $\cos(m+n)x+\cos(m-n)x=2\cos mx\cos nx$.

   (c) The sequence of polynomials (U_n) satisfies $U_0(x)=2$, $U_1(x)=x$ and $U_{n+1}=U_1U_n-U_{n-1}$, implying that each U_n has integer coefficients. The equality $U_n(x+x^{-1})=x^n+x^{-n}$ holds for each $x=\cos t+i\sin t$, and therefore it holds for all x.

5. Suppose that $\cos\frac{p}{q}\pi=a$. It follows from the previous problem that $U_q(2a)=2\cos p\pi=\pm 2$, where U_q is monic with integer coefficients, so $2a$ is an integer by theorem 14.

6. Note that equality holds for a multiple of the n-th Chebyshev polynomial $T_n(x)$. The leading coefficient of T_n equals 2^{n-1}, so $C_n(x)=\frac{1}{2^{n-1}}T_n(x)$ is a monic polynomial and

$$|T_n(x)|=\frac{1}{2^{n-1}}\,|\cos(n\arccos x)|\le\frac{1}{2^{n-1}}\quad\text{za } x\in[-1,1].$$

   Moreover, the values of T_n at points $1,\cos\frac{\pi}{n},\cos\frac{2\pi}{n},\cdots,\cos\frac{(n-1)\pi}{n},-1$ are alternately $\frac{1}{2^{n-1}}$ and $-\frac{1}{2^{n-1}}$.

   Now suppose that $P\neq T_n$ is a monic polynomial such that $\max_{-1\le x\le 1}|P(x)|<\frac{1}{2^{n-1}}$. Then $P(x)-C_n(x)$ at points $1,\cos\frac{\pi}{n},\cdots,\cos\frac{(n-1)\pi}{n},-1$ alternately takes positive and negative values. Therefore the polynomial $P-C_n$ has at least n zeros, namely, at least one in every interval between two adjacent points. However, $P-C_n$ is a polynomial of degree $n-1$ as the monomial x^n is canceled, so we have arrived at a contradiction.

7. Since $P^{[i]}(x)=(-2)^{i-1}(-1)^x$ for $x=0,1,\ldots,n-i$, we have

$$P(n+1)=P(n)+P^{[1]}(n-1)+\cdots+P^{[n]}(0)=\begin{cases}2^n, & 2\nmid n;\\ 1-2^n, & 2\mid n.\end{cases}$$

8. By theorem 20 we have

$$P(n+2)=\sum_{i=0}^{n}(-1)^{n-i}\frac{1}{i+1}\binom{n+1}{i}=\frac{1}{n+2}\sum_{i=0}^{n}(-1)^{n-i}\binom{n+2}{i+1}=\begin{cases}0, & 2\nmid n;\\ \frac{2}{n+2}, & 2\mid n.\end{cases}$$

9. By theorem 9, the polynomial $P(x)$ can be factorized as

$$P(x)=(x-a_1)^{\alpha_1}\cdots(x-a_k)^{\alpha_k}\cdot(x^2-b_1x+c_1)\cdots(x^2-b_mx+c_m), \qquad (*)$$

where a_i, b_j, c_j are real numbers such that the a_i are different and the polynomials $x^2-b_ix+c_i$ has no real zeros.

The condition $P(x)\geq 0$ for all x implies that the α_i are even, whereas the condition $P(x)\geq 0$ for $x\geq 0$ implies that $(\forall i)$ α_i is even or $a_i<0$. It is now easy to write each factor in $(*)$ in the form A^2+B^2, respectively A^2+xB^2, so by the known formula $(a^2+\gamma b^2)(c^2+\gamma d^2)=(ac+\gamma bd)^2+\gamma(ad-bc)^2$ one can express their product $P(x)$ in the desired form.

10. Write

$$P(-x)=ax^4+(c-b)x^2+(e-d)-b(x^3-x^2)-d(x-1).$$

If r is a root of the polynomial Q, we have $P(\sqrt{r})=-(\sqrt{r}-1)(br+d)$ and $P(-\sqrt{r})=(\sqrt{r}+1)(br+d)$. Note that one of the two numbers $P(\pm\sqrt{r})$ positive and the other is negative (or both are zero). Hence there must be a zero of P between $-\sqrt{r}$ and $\sqrt{r}$.

11. Let us write $P(x)=(x-x_1)\cdots(x-x_m)(x^2-p_1x+q_1)\cdots(x^2-p_nx+q_n)$, where the polynomials $x^2-p_kx+q_k$ have no real zeros. We have

$$1>|P(i)|=\prod_{j=1}^{m}|i-x_j|\prod_{k=1}^{n}|-1-p_ki+q_k|,$$

and since $|i-x_j|^2=1+x_j^2>1$ for all j, we must have $|-1-p_ki+q_k|<1$ for some k, i.e.

$$p_k^2+(q_k-1)^2<1. \qquad (*)$$

Let $a\pm bi$ be the zeros of the polynomial $x^2-p_kx+q_k$ (and also of the polynomial P). Then $p_k=2a$ and $q_k=a^2+b^2$, so the inequality $(*)$ becomes $4a^2+(a^2+b^2-1)^2<1$, which is equivalent to the desired inequality.

12. Write f in the form $f=P/Q$, where P and Q are coprime polynomials and Q is monic. Comparing the leading coefficients we conclude that P is also monic. The condition of the problem becomes $P(x^2)/Q(x^2)=P(x)^2/Q(x)^2-a$. Since $P(x^2)$ and $Q(x^2)$ are coprime (if they have a common zero, so do P and Q), it follows that $Q(x^2)=Q(x)^2$ and hence $Q(x)=x^n$ for some $n\in\mathbb{N}$. Therefore, $P(x^2)=P(x)^2-ax^{2n}$.

Let $P(x)=a_0+a_1x+\cdots+a_{m-1}x^{m-1}+x^m$. Comparing the coefficients of $P(x)^2$ and $P(x^2)$ we find that $a_{n-1}=\cdots=a_{2m-n+1}=0$, $a_{2m-n}=a/2$, $a_1=\cdots=a_{m-1}=0$ and $a_0=1$. Clearly, this is only possible if $a=0$, or $a=2$ and $2m-n=0$.

13. Since P is symmetric with respect to point 0, it is easy to show that P is also a polynomial in x^2, so there is a polynomial Q such that $P(x)=Q(x^2+1)$ or $P(x)=xQ(x^2+1)$. Then $Q((x^2+1)^2+1)=Q(x^2+1)^2-1$, respectively $(x^2+1)Q((x^2+1)^2+1)=x^2Q(x^2+1)^2+1$. The substitution $x^2+1=y$ yields $Q(y^2+1)=Q(y)^2+1$, resp. $yQ(y^2+1)=(y-1)Q(y)^2+1$.

Suppose that $yQ(y^2+1)=(y-1)Q(y)^2+1$. Setting $y=1$ gives us $Q(2)=1$. Note that if $a\neq 0$ and $Q(a)=1$ then $aQ(a^2+1)=(a-1)+1$, so $Q(a^2+1)=1$ as well. This leads to an infinite sequence (a_n) of points at which Q takes the value 1, given by $a_0=2$ and $a_{n+1}=a_n^2+1$. We conclude that $Q\equiv 1$.

We have shown that if $Q\not\equiv 1$, then $P(x)=Q(x^2+1)$. Now we easily come to all solutions: these are the polynomials of the form $T(T(\cdots(T(x))\cdots))$, where $T(x)=x^2+1$.

14. Let us denote $P(1)=a$. We have $a^2-2a-2=0$. Since $P(x)=(x-1)P_1(x)+a$, substituting in the original equation and simplifying yields $(x-1)P_1(x)^2+2aP_1(x)=4(x+1)P_1(2x^2-1)$. For $x=1$ we have $2aP_1(1)=8P_1(1)$, which together with $a\neq 4$ implies $P_1(1)=0$, i.e. $P_1(x)=(x-1)P_2(x)$, so $P(x)=(x-1)^2P_2(x)+a$. Assume that $P(x)=(x-1)^nQ(x)+a$, where $Q(1)\neq 0$. Again substituting in the original equation and simplifying yields $(x-1)^nQ(x)^2+2aQ(x)=2(2x+2)^nQ(2x^2-1)$, which implies that $Q(1)=0$, a contradiction. We conclude that $P(x)=a$.

15. At first, note that there exists $x=a$ for which $P(a)^2=Q(a)^2$. This follows from the fact that, if p and q are real roots of P and Q respectively, then $P(p)^2-Q(p)^2\leq 0\leq P(q)^2-Q(q)^2$, whereby P^2-Q^2 is a continuous function. Then we also have $P(b)=Q(b)$ for $b=1+a+P(a)^2$. Assuming that a is the largest real number with $P(a)=Q(a)$, we come to an immediate contradiction.

16. Let $P(x)=a_0+a_1x+\cdots+a_nx^n$. For every x the triple $(a,b,c)=(6x,3x,-2x)$ satisfies the condition $ab+bc+ca=0$. The condition in P gives us $P(3x)+P(5x)+P(-8x)=2P(7x)$ for all x, so by comparing the coefficients on both sides we obtain $K(i)=\left(3^i+5^i+(-8)^i-2\cdot 7^i\right)=0$ whenever $a_i\neq 0$. Since $K(i)$ is negative for odd i and positive for $i=0$ and even $i\geq 6$, $a_i=0$ is only possible for $i=2$ and $i=4$. Therefore, $P(x)=a_2x^2+a_4x^4$ for some real numbers a_2,a_4. It is easily verified that all such $P(x)$ satisfy the conditions.

17. Let d be the degree of P. There is a unique polynomial Q of degree at most d such that $Q(k)=a_k$ for $k=1,2,\ldots,d+1$. Let us show that $Q(n)=a_n$ for all n.

    Let $n>d+1$. Polynomial Q might not have integral coefficients, so we cannot deduce that $n-m\mid Q(n)-Q(m)$, but it certainly has rational coefficients, i.e. there is a natural number M for which $R(x)=MQ(x)$ has integral coefficients. By the condition of the problem, $M(a_n-Q(n))=M(a_n-a_k)-(R(n)-R(k))$ is divisible by $n-k$ for each $k=1,2,\ldots,d+1$. Therefore, for each n we either have $a_n=Q(n)$ or

    $$L_n=\operatorname{lcm}(n-1,n-2,\ldots,n-d-1)\leq M(a_n-Q(n))<Cn^d$$

    for some constant C independent of n.

    Suppose that $a_n\neq Q(n)$ for some n. note that L_n is not less than the product $(n-1)\cdots(n-d-1)$ divided by the product P of numbers $\gcd(n-i,n-j)$ over all pairs (i,j) of different numbers from $\{1,2,\ldots,d+1\}$. Since $\gcd(n-i,n-j)\leq i-j$, we have $P\leq 1^d2^{d-1}\cdots d$. It follows that

    $$(n-1)(n-2)\cdots(n-d-1)\leq PL_n<CPn^d,$$

    which is false for large enough n as the left hand side is of degree $d+1$. Thus, $a_n=Q(n)$ for each sufficiently large n, say $n>N$.

    What happens for $n\leq N$? By the condition of the problem, $M(a_n-Q(n))=M(a_n-a_k)-(R(n)-R(k))$ is divisible by $m-n$ for every $m>N$, so it must be equal to zero. Hence $a_n=Q(n)$ for all n.

18. We have shown in 7 from the text that every such t satisfies $P(P(t))=t$. If every such t also satisfies $P(t)=t$, the number of solutions is clearly at most $\deg P=n$. Suppose that $P(t_1)=t_2$, $P(t_2)=t_1$, $P(t_3)=t_4$ i $P(t_4)=t_3$, where $t_1\neq t_{2,3,4}$. By theorem 10, t_1-t_3 divides t_2-t_4 and vice versa, from which we deduce that $t_1-t_3=\pm(t_2-t_4)$. Assume that $t_1-t_3=t_2-t_4$, i.e. $t_1-t_2=t_3-t_4=k\neq 0$. Since the relation $t_1-t_4=\pm(t_2-t_3)$ similarly holds, we obtain $t_1-t_3+k=\pm(t_1-t_3-k)$ which is impossible. Therefore, we must have $t_1-t_3=t_4-t_2$, which gives us $P(t_1)+t_1=P(t_3)+t_3=c$ for some c. It follows that all integral solutions t of the equation $P(P(t))=t$ satisfy $P(t)+t=c$, and hence their number does not exceed n.

19. Suppose that $R = P - Q \neq 0$ and that $0 < k \leq n-1$ is the degree of $R(x)$. Then

$$P(P(x)) - Q(Q(x)) = [Q(P(x)) - Q(Q(x))] + R(P(x)).$$

Writing $Q(x) = x^n + \cdots + a_1 x + a_0$ yields

$$Q(P(x)) - Q(Q(x)) = [P(x)^n - Q(x)^n] + \cdots + a_1[P(x) - Q(x)],$$

where all the summands but the first have a degree at most $n^2 - n$, while the first summand equals $R(x) \cdot \left(P(x)^{n-1} + P(x)^{n-2}Q(x) + \cdots + Q(x)^{n-1}\right)$ and has the degree $n^2 - n + k$ with the leading coefficient n. Therefore the degree of $Q(P(x)) - Q(Q(x))$ is $n^2 - n + k$. On the other hand, the degree of the polynomial $R(P(x))$ equals $kn < n^2 - n + k$, from which we conclude that the difference $P(P(x)) - Q(Q(x))$ has the degree $n^2 - n + k$, a contradiction.

It remains to check the case of a constant $R \equiv c$. Then the condition $P(P(x)) = Q(Q(x))$ yields $Q(Q(x) + c) = Q(Q(x)) - c$, so the equality $Q(y + c) = Q(y) - c$ holds for infinitely many values of y; hence $Q(y + c) \equiv Q(y) - c$ which is only possible for $c = 0$ (to see this, just compare the coefficients).

20. Suppose that $f(x) = g(x)h(x)$ for some nonconstant polynomials with integer coefficients. Since $|f(0)| = p$, either $|g(0)| = 1$ or $|h(0)| = 1$ holds. Assume w.l.o.g. that $|g(0)| = 1$. Write $g(x) = (x - \alpha_1) \cdots (x - \alpha_k)$. Then $|\alpha_1 \cdots \alpha_k| = 1$. Since $f(\alpha_i) - p = \alpha_i^m(\alpha_i - a)^n = -p$, taking the product over $i = 1, 2, \ldots, k$ yields $|g(a)|^n = |(\alpha_1 - a) \cdots (\alpha_k - a)|^n = p^k$. Since $g(a)$ divides $|g(a)h(a)| = p$, we must have $|g(a)| = p$ and $n = k$. However, a must divide $|g(a) - g(0)| = p \pm 1$, which is impossible.

21. Suppose that $F = G \cdot H$ for some polynomials G, H with integer coefficients. Let us consider this equality modulo 2. Since $(x^2 + x + 1)^{2^n} \equiv F(x) \pmod 2$, we obtain $(x^2 + x + 1)^{2^n} = g(x)h(x)$, where $g \equiv G$ and $h \equiv H$ are polynomials over $\mathbb{Z}_2$. The polynomial $x^2 + x + 1$ is irreducible over $\mathbb{Z}_2[x]$, so there exists a natural number k for which $g(x) = (x^2 + x + 1)^k$ and $h(x) = (x^2 + x + 1)^{2^n - k}$; of course, these equalities hold in $\mathbb{Z}_2[x]$ only.

Back in $\mathbb{Z}[x]$, these equalities become $H(x) = (x^2 + x + 1)^{2^n - k} + 2V(x)$ and $G(x) = (x^2 + x + 1)^k + 2U(x)$ for some polynomials U and V with integer coefficients. Thus,

$$[(x^2 + x + 1)^k + 2U(x)][(x^2 + x + 1)^{2^n - k} + 2V(x)] = F(x).$$

Now if we set $x = \varepsilon = \frac{-1 + i\sqrt{3}}{2}$ in this equality, we obtain $U(\varepsilon)V(\varepsilon) = \frac{1}{4}F(\varepsilon) = \frac{1}{2}$. However, this is impossible as the polynomial $U(x)V(x)$ has integer coefficients, so $U(\varepsilon)V(\varepsilon)$ must be of the form $a + b\varepsilon$ for some $a, b \in \mathbb{Z}$ (since $\varepsilon^2 = -1 - \varepsilon$), which is not the case with $\frac{1}{2}$.

22. It is clear, for example by theorem 16, that P must have rational coefficients. For some $m \in \mathbb{N}$ the coefficients of the polynomial $mP(x)$ are integral. Let p be a prime number not dividing m. We claim that, if P is not linear, there is no rational number x for which $P(x) = \frac{1}{mp}$. Namely, such an x would also satisfy $Q(x) = mpP(x) - 1 = 0$. On the other hand, the polynomial $Q(x)$ is irreducible because so is the polynomial $x^n Q(1/x)$ by the Eisenstain criterion; indeed, all the coefficients of $x^n Q(1/x)$ but the first are divisible by p and the constant term is not divisible by p^2. This proves our claim.

23. Denote $P(x) = P_n(x) = R_n(x) + iS_n(x)$. We prove by induction on n that all zeros of P_n are real; moreover, if $x_1 > x_2 > \cdots > x_n$ are the zeros of R_n and $y_1 > y_2 > \cdots > y_{n-1}$ the zeros of R_{n-1}, then

$$x_1 > y_1 > x_2 > y_2 > \cdots > x_{n-1} > y_{n-1} > x_n.$$

This statement is trivially true for $n = 1$. Suppose that it is true for $n - 1$.

Since $R_n + iS_n = (x-i+n)(R_{n-1}+iS_{n-1})$, the polynomials R_n and S_n satisfy the recurrent relations $R_n = (x+n)R_{n-1} + S_{n-1}$ and $S_n = (x+n)S_{n-1} - R_{n-1}$. This gives us

$$R_n - (2x+2n-1)R_{n-1} + [(x+n-1)^2+1]R_{n-2} = 0.$$

If $z_1 > \cdots > z_{n-2}$ are the (real) zeros R_{n-2}, by the inductive hypothesis we have $z_{i-1} > y_i > z_i$. Since the value of R_{n-2} is alternately positive and negative on the intervals $(z_1, +\infty)$, (z_2, z_1), etc, it follows that $\operatorname{sgn} R_{n-2}(y_i) = (-1)^{i-1}$. Now we conclude from the relation $R_n(y_i) = -[(x+n-1)^2+1]R_{n-2}(y_i)$ that

$$\operatorname{sgn} R_n(y_i) = (-1)^i,$$

which means that the polynomial R_n has a zero on each of the n intervals $(y_1, +\infty)$, (y_2, y_1), $\ldots$, $(-\infty, y_{n-1})$. This finishes the induction.

24. We first prove the following auxiliary statement.

Lemma. If A, B and C are coprime polynomials with $A + B = C$, then the degree of each of the polynomials A, B, C is less than the number of different zeros of the polynomial ABC.

Proof. Let

$$A(x) = \prod_{i=1}^{k} (x-p_i)^{a_i}, \quad B(x) = \prod_{i=1}^{l} (x-q_i)^{b_i}, \quad C(x) = \prod_{i=1}^{m} (x-r_i)^{c_i}.$$

Let us rewrite the given equality as $A(x)/C(x) + B(x)/C(x) = 1$ and differentiate it with respect to x. We obtain

$$\frac{A(x)}{C(x)} \left(\sum_{i=1}^{k} \frac{a_i}{x-p_i} - \sum_{i=1}^{m} \frac{c_i}{x-r_i} \right) = -\frac{B(x)}{C(x)} \left(\sum_{i=1}^{l} \frac{b_i}{x-q_i} - \sum_{i=1}^{m} \frac{c_i}{x-r_i} \right),$$

from which we see that $A(x)/B(x)$ can be expressed as a quotient of two polynomials of degree not exceeding $k+l+m-1$. The statement follows from the coprimeness of A and B.

Now we apply the Lemma on the polynomials P^a, Q^b, R^c. We obtain that each of the numbers $a \deg P$, $b \deg Q$, $c \deg R$ is less than $\deg P + \deg Q + \deg R$, and therefore

$$\frac{1}{a} > \frac{\deg P}{\deg P + \deg Q + \deg R},$$

etc. Adding these yields the desired inequality.

25. Let us fix $\deg P = n$. Let $P(x) = (x-z_1)\cdots(x-z_n) = x^n + a_{n-1}x^{n-1} + \cdots + a_0$, where $|z_i| = 1$ for $i = 1, \ldots, n$. By the Vieta formulas, $a_{n-i} = \pm\sigma_i(z_1, \ldots, z_n)$, which is a sum of $\binom{n}{i}$ summands of modulus 1, and hence $|a_{n-i}| \le \binom{n}{i}$. Therefore, there are at most $2\binom{m}{i} + 1$ possible values of the coefficient of $P(x)$ at x^{n-i} for each i. Thus the number of possible polynomials P of degree n is finite.

Now consider the polynomial $P_r(x) = (x-z_1^r)\cdots(x-z_n^r)$ for each natural number r. All coefficients of polynomial P_r are symmetric polynomials in z_i with integral coefficients, so by the theorem 24 they must be integers. Therefore, every polynomial P_r satisfies the conditions of the problem, but there are infinitely many r's and only finitely many such polynomials. We conclude that $P_r(x) = P_s(x)$ for some distinct $r, s \in \mathbb{N}$, and the main statement of the problem follows.

terug naar echt bestand

1
Graphs

A *graph* G consists of a set V (or $V(G)$) of *vertices*, a set E (or $E(G)$) of *edges*, and a mapping associating to each edge $e \in E(G)$ an unordered pair x, y of vertices called the *endpoints* (or simply the *ends*) of e. We say an edge is *incident* with its ends, and that it *joins* its ends. We allow $x = y$, in which case the edge is called a *loop*. A vertex is *isolated* when it is incident with no edges.

It is common to represent a graph by a *drawing* where we represent each vertex by a point in the plane, and represent edges by line segments or arcs joining some of the pairs of points. One can think e.g. of a network of roads between cities. A graph is called *planar* if it can be drawn in the plane such that no two edges (that is, the line segments or arcs representing the edges) cross. The topic of planarity will be dealt with in Chapter 33; we wish to deal with graphs more purely combinatorially for the present.

edge	ends
a	x, z
b	y, w
c	x, z
d	z, w
e	z, w
f	x, y
g	z, w

Figure 1.1

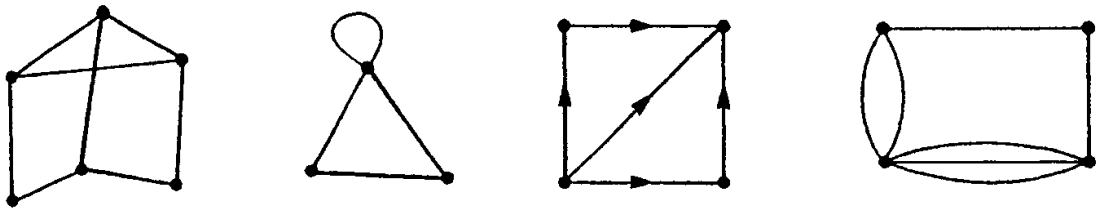
Thus a graph is described by a table such as the one in Fig. 1.1 that lists the ends of each edge. Here the graph we are describing

has vertex set $V = \{x, y, z, w\}$ and edge set $E = \{a, b, c, d, e, f, g\}$; a drawing of this graph may be found as Fig. 1.2(iv).

A graph is *simple* when it has no loops and no two distinct edges have exactly the same pair of ends. Two nonloops are *parallel* when they have the same ends; graphs that contain them are called *multigraphs* by some authors, or are said to have 'multiple edges'.

If an *ordered* pair of vertices is associated to each edge, we have a *directed graph* or *digraph*. In a drawing of a digraph, we use an arrowhead to point from the first vertex (the *tail*) towards the second vertex (the *head*) incident with an edge. For a *simple* digraph, we disallow loops and require that no two distinct edges have the same ordered pair of ends.

When dealing with simple graphs, it is often convenient to identify the edges with the unordered pairs of vertices they join; thus an edge joining x and y can be called $\{x, y\}$. Similarly, the edges of a simple digraph can be identified with ordered pairs (x, y) of distinct vertices.



(i) graph (ii) graph with loop (iii) digraph (iv) multiple edges

Figure 1.2

There are several ways to draw the same graph. For example, the two graphs of Fig. 1.3 are essentially the same.

We make this more precise, but to avoid unnecessarily technical definitions at this point, let us assume that all graphs are undirected and simple for the next two definitions.

We say two graphs are *isomorphic* if there is a one-to-one correspondence between the vertex sets such that if two vertices are joined by an edge in one graph, then the corresponding vertices are joined by an edge in the other graph. To show that the two graphs in Fig. 1.3 are the same, find a suitable numbering of the vertices

in both graphs (using $1, 2, 3, 4, 5, 6$) and observe that the edge sets are the same sets of unordered pairs.

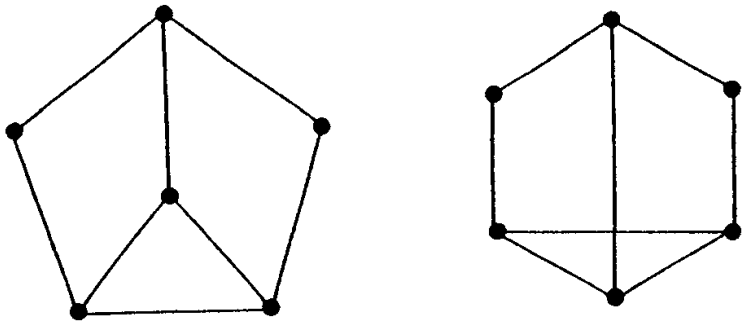


Figure 1.3

A permutation σ of the vertex set of a graph G with the property that $\{a, b\}$ is an edge if and only if $\{\sigma(a), \sigma(b)\}$ is an edge, is called an *automorphism* of G.

Problem 1A. (i) Show that the drawings in Fig. 1.4 represent the same graph (or isomorphic graphs).

(ii) Find the group of automorphisms of the graph in Fig. 1.4. Remark: There is no quick or easy way to do this unless you are lucky; you will have to experiment and try things.

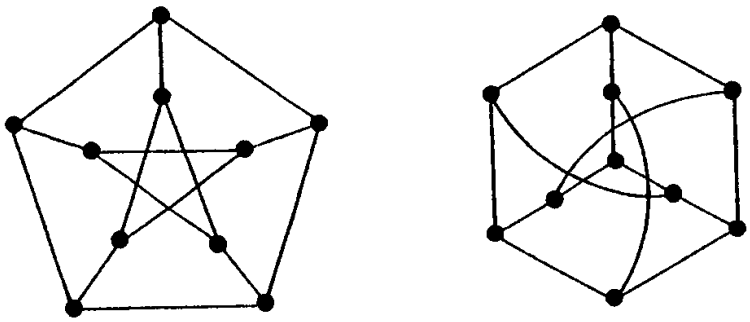


Figure 1.4

The *complete* graph K_n on n vertices is the simple graph that has all $\binom{n}{2}$ possible edges.

Two vertices a and b of a graph G are called *adjacent* if they are distinct and joined by an edge. We will use $\Gamma(x)$ to denote the set of all vertices adjacent to a given vertex x; these vertices are also called the *neighbors* of x.

The number of edges incident with a vertex x is called the *degree* or the *valency* of x. Loops are considered to contribute 2 to the valency, as the pictures we draw suggest. If all the vertices of a graph have the same degree, then the graph is called *regular*.

One of the important tools in combinatorics is the method of *counting* certain objects in two different ways. It is a well known fact that if one makes no mistakes, then the two answers are the same. We give a first elementary example. A graph is *finite* when both $E(G)$ and $V(G)$ are finite sets. We will be primarily concerned with finite graphs, so much so that it is possible we have occasionally forgotten to specify this condition as a hypothesis in some assertions.

Theorem 1.1. *A finite graph G has an even number of vertices with odd valency.*

PROOF: Consider a table listing the ends of the edges, as in Fig. 1.1. The number of entries in the right column of the table is twice the number of edges. On the other hand, the degree of a vertex x is, by definition, the number of times it occurs in the table. So the number of entries in the right column is

$$\sum_{x \in V(G)} \deg(x) = 2|E(G)|. \tag{1.1}$$

The assertion follows immediately. □

The equation (1.1) is simple but important. It might be called the 'first theorem of graph theory', and our Theorem 1.1 is its first corollary.

A *subgraph* of a graph G is a graph H such that $V(H) \subseteq V(G)$, $E(H) \subseteq E(G)$, and the ends of an edge $e \in E(H)$ are the same as its ends in G. H is a *spanning* subgraph when $V(H) = V(G)$. The subgraph of G *induced* by a subset S of vertices of G is the subgraph whose vertex set is S and whose edges are *all* the edges of G with both ends in S.

A *walk* in a graph G consists of an alternating sequence

$$x_0, e_1, x_1, e_2, x_2, \ldots, x_{k-1}, e_k, x_k$$

of vertices x_i, not necessarily distinct, and edges e_i so that the ends of e_i are exactly x_{i-1} and x_i, $i = 1, 2, \ldots, k$. Such a walk has *length* k. If the graph is simple, a walk is determined by its sequence of vertices, any two successive elements of which are adjacent.

If the edge terms $e_1, \ldots, e_k$ are distinct, then the walk is called a *path* from x_0 to x_k. If $x_0 = x_k$, then a walk (or path) is called *closed*. A *simple* path is one in which the vertex terms $x_0, x_1, \ldots, x_k$ are also distinct, although we say we have a *simple closed path* when $k \geq 1$ and all vertex terms are distinct except $x_0 = x_k$.

If a path from x to y exists for every pair of vertices x, y of G, then G is called *connected*. Otherwise G consists of a number of connected *components* (maximal connected subgraphs). It will be convenient to agree that the null graph with no vertices and no edges is not connected.

Problem 1B. Suppose G is a simple graph on 10 vertices that is not connected. Prove that G has at most 36 edges. Can equality occur?

The length of the shortest walk from a to b, if such walks exist, is called the *distance* $d(a, b)$ between these vertices. Such a shortest walk is necessarily a simple path.

Example 1.1. A well known graph has the mathematicians of the world as vertices. Two vertices are adjacent if and only if they have published a joint paper. The distance in this graph from some mathematician to the vertex P. Erdős is known as his or her Erdős-number.

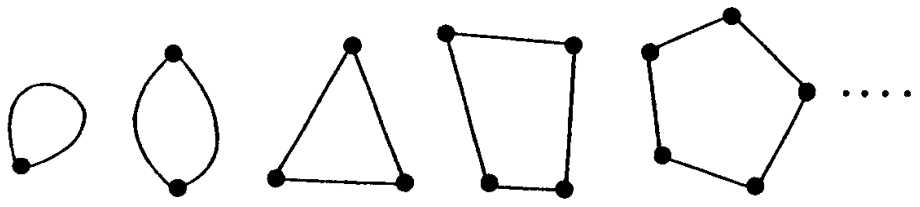


Figure 1.5

A *polygon* is the 'graph of' a simple closed path, but more precisely it can be defined as a finite connected graph that is regular of degree 2. There is, up to isomorphism, exactly one polygon P_n

with n vertices (often called the *n-gon*) for each positive integer n. The sequence of polygons is shown in Fig. 1.5.

A connected graph that contains no simple closed paths, i.e. that has no polygons as subgraphs, is called a *tree*.

Problem 1C. Show that a connected graph on n vertices is a tree if and only if it has $n-1$ edges.

Problem 1D. The *complete bipartite graph* $K_{n,m}$ has $n+m$ vertices $a_1, \ldots, a_n$ and $b_1, \ldots, b_m$, and as edges all mn pairs $\{a_i, b_j\}$. Show that $K_{3,3}$ is not planar.

No introduction to graph theory can omit the problem of the bridges of Königsberg (formerly a city in Prussia). The river Pregel flowed through this city and split into two parts. In the river was the island Kneiphof. There were seven bridges connecting different parts of the city as shown in the diagram of Fig. 1.6.

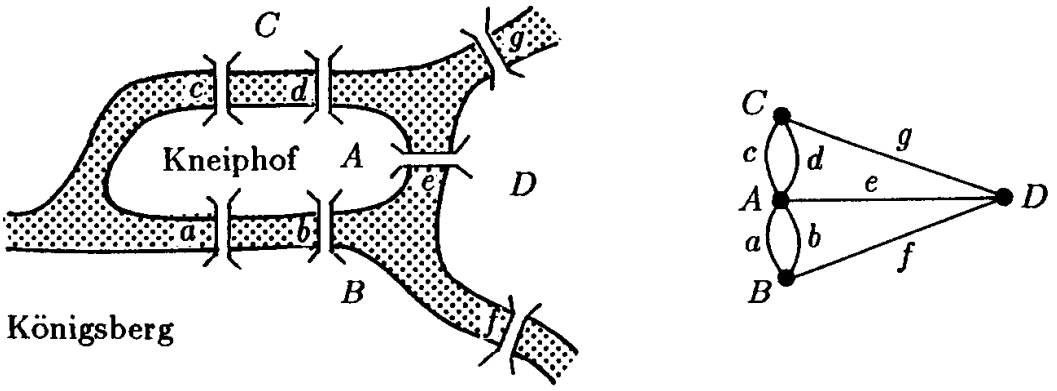


Figure 1.6

In a paper written in 1736 by L. Euler (considered the first paper on graph theory) the author claims that the following question was considered difficult: Is it possible to make a walk through the city, returning to the starting point and crossing each bridge exactly once? This paper has led to the following definition. A closed path through a graph using every edge once is called an *Eulerian circuit* and a graph that has such a path is called an *Eulerian graph*.

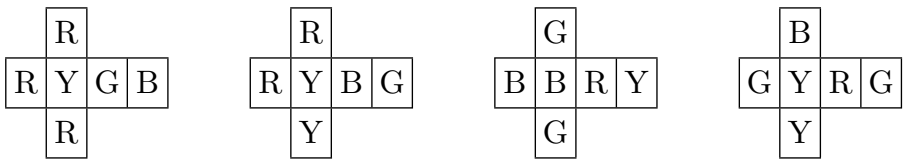
Theorem 1.2. *A finite graph G with no isolated vertices (but possibly with multiple edges) is Eulerian if and only if it is connected and every vertex has even degree.*

PROOF: That G must be connected is obvious. Since the path enters a vertex through some edge and leaves by another edge, it is clear that all degrees must be even. To show that the conditions are sufficient, we start in a vertex x and begin making a path. We keep going, never using the same edge twice, until we cannot go further. Since every vertex has even degree, this can only happen when we return to x and all edges from x have been used. If there are unused edges, then we consider the subgraph formed by these edges. We use the same procedure on a component of this subgraph, producing a second closed path. If we start this second path in a point occurring in the first path, then the two paths can be combined to a longer closed path from x to x. Therefore the longest of these paths uses all the edges. □

The problem of the bridges of Königsberg is described by the graph in Fig. 1.6. No vertex has even degree, so there is no Eulerian circuit.

One can consider a similar problem for digraphs. The necessary and sufficient condition for a directed Eulerian circuit is that the graph is connected and that each vertex has the same 'in-degree' as 'out-degree'.

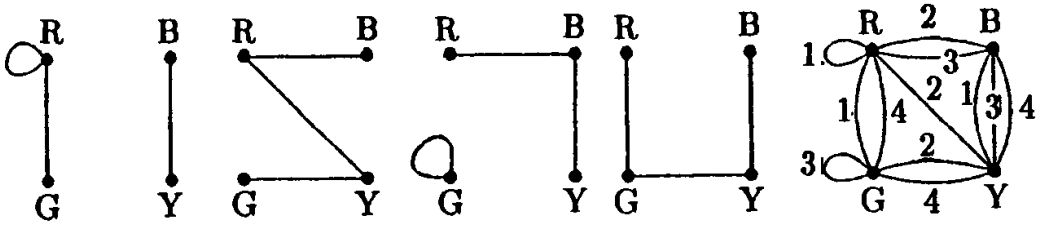
Example 1.2. A puzzle with the name *Instant Insanity* concerns four cubes with faces colored red, blue, green, and yellow, in such a way that each cube has at least one face of each color. The problem is to make a stack of these cubes so that all four colors appear on each of the four sides of the stack. In Fig. 1.7 we describe four possible cubes in flattened form.



cube 1 cube 2 cube 3 cube 4

Figure 1.7

It is not a very good idea to try all possibilities. A systematic approach is as follows. The essential information about the cubes is given by the four graphs in Fig. 1.8.



cube 1 cube 2 cube 3 cube 4 the graph G

Figure 1.8

An edge indicates that the two adjacent colors occur on opposite faces of the cube. We obtain a graph G by superposition of the four graphs and number the edges according to their origin. It is not difficult to see that we need to find in G two subgraphs that are regular of degree 2, with edges numbered $1, 2, 3, 4$ and such that they have no edge in common. One of the subgraphs tells us which pairs of colors to align on the left side and right side of the stack. The other graph describes the colors on front and back. Of course it is easy to rotate the cubes in such a way that the colors are where we wish them to be. The point of the example is that it takes only a minute to find two subgraphs as described above. In this example the solution is unique.

We mention a concept that seems similar to Eulerian circuits but that is in reality quite different. A *Hamiltonian circuit* in a graph G is a simple closed path that passes through each *vertex* exactly once (rather than each *edge*). So a graph admits a Hamiltonian circuit if and only if it has a polygon as a spanning subgraph. In the mid-19th century, Sir William Rowan Hamilton tried to popularize the exercise of finding such a closed path in the graph of the dodecahedron (Fig. 1.9).

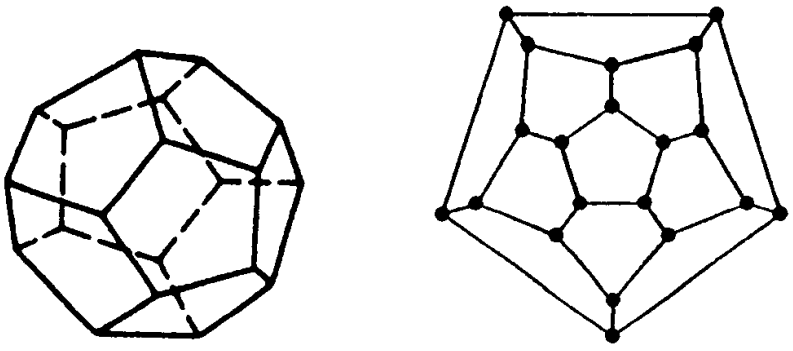


Figure 1.9

The graph in Fig. 1.4 is called the Petersen graph (cf. Chapter 21) and one of the reasons it is famous is that it is *not* 'Hamiltonian'; it contains n-gons only for $n = 5, 6, 8, 9$, and not when $n = 7$ or $n = 10$.

By Theorem 1.2, it is easy to decide whether a graph admits an Eulerian circuit. A computer can easily be programmed to check whether the degrees of a graph are even and whether the graph is connected, and even to produce an Eulerian circuit when one exists. In contrast to this, the problem of deciding whether an arbitrary graph admits a Hamiltonian circuit is likely 'intractable'. To be more precise, it has been proved to be *NP-complete*—see Garey and Johnson (1979).

Problem 1E. Let $A_1, \ldots, A_n$ be n distinct subsets of the n-set $N := \{1, \ldots, n\}$. Show that there is an element $x \in N$ such that the sets $A_i \backslash \{x\}$, $1 \leq i \leq n$, are all distinct. To do this, form a graph G on the vertices A_i with an edge with 'color' x between A_i and A_j if and only if the symmetric difference of the sets A_i and A_j is $\{x\}$. Consider the colors occurring on the edges of a polygon. Show that one can delete edges from G in such a way that *no* polygons are left and the number of different colors remains the same. Then use 1C. (This idea is due to J. A. Bondy (1972).)

Problem 1F. The *girth* of a graph is the length of the smallest polygon in the graph. Let G be a graph with girth 5 for which all vertices have degree $\geq d$. Show that G has at least $d^2 + 1$ vertices. Can equality hold?

Problem 1G. Show that a finite simple graph with more than one vertex has at least two vertices with the same degree.

Problem 1H. A graph on the vertex set $\{1, 2, \ldots, n\}$ is often described by a matrix A of size n, where a_{ij} and a_{ji} are equal to the number of edges with ends i and j. What is the combinatorial interpretation of the entries of the matrix A^2?

Problem 1I. Let $Q := \{1, 2, \ldots, q\}$. Let G be a graph with the elements of Q^n as vertices and an edge between $(a_1, a_2, \ldots, a_n)$ and $(b_1, b_2, \ldots, b_n)$ if and only if $a_i \neq b_i$ for exactly one value of i. Show that G is Hamiltonian.

Problem 1J. Let G be a simple graph on n vertices ($n > 3$) with no vertex of degree $n-1$. Suppose that for any two vertices of G, there is a *unique* vertex joined to both of them.

(i) If x and y are not adjacent, prove that they have the same degree.

(ii) Now show that G is a regular graph.

Notes.

Paul Erdős (1913–1996) (cf. Example 1.1) was probably the most prolific mathematician of the 20th century with well over 1400 papers having been published. His contributions to combinatorics, number theory, set theory, etc., include many important results. He collaborated with many mathematicians all over the world, all of them proud to have Erdős-number 1, among them the authors of this book; see J. W. Grossman (1997).

Leonhard Euler (1707–1783) was a Swiss mathematician who spent most of his life in St. Petersburg. He was probably the most productive mathematician of all times. Even after becoming blind in 1766, his work continued at the same pace. The celebration in 1986 of the 250th birthday of graph theory was based on Euler's paper on the Königsberg bridge problem. Königsberg is now the city of Kaliningrad in Russia.

For an elementary introduction to graph theory, we recommend R. J. Wilson (1979), and J. J. Watkins and R. J. Wilson (1990).

Sir William Rowan Hamilton (1805–1865) was an Irish mathematician. He was considered a genius. He knew 13 languages at the age of 12 and was appointed professor of astronomy at Trinity College Dublin at the age of 22 (before completing his degree). His most important work was in mathematical physics.

References.

M. Garey and D. S. Johnson (1979), *Computers and Intractability; A Guide to the Theory of NP-completeness*, W. H. Freeman and Co.

J. W. Grossman (1997), Paul Erdős: The Master of Collaboration, pp. 467–475 in *The Mathematics of Paul Erdős*, R. L. Graham and J. Nešetřil (eds.), Springer-Verlag.

J. J. Watkins and R. J. Wilson (1990), *Graphs (An Introductory Approach)*, J. Wiley & Sons.
R. J. Wilson (1979), *Introduction to Graph Theory*, Longman.

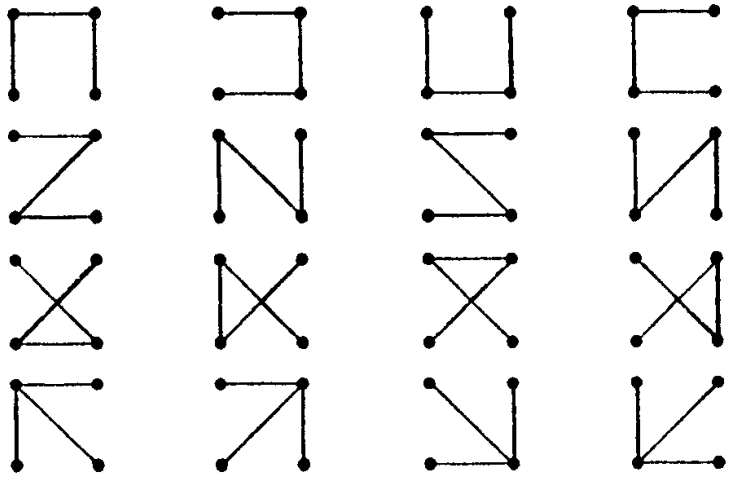
2
Trees

We come to the first not so easy theorem. It is due to A. Cayley (1889). We shall give three different proofs here. Two more proofs will occur in later chapters; see Example 14.14 and Example 38.2. The first two proofs illustrate a method that is used very often in combinatorics. In order to count certain objects that seem hard to count, one finds a one-to-one mapping onto a set of other objects whose number is easier to determine.

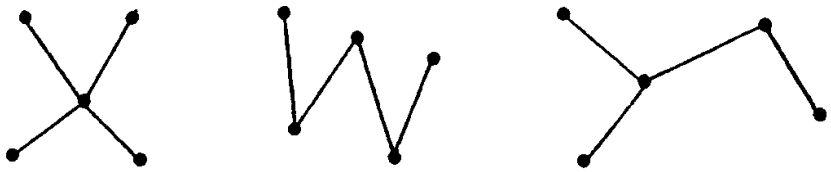
Theorem 2.1. *There are n^{n-2} different labeled trees on n vertices.*

The term *labeled* emphasizes that we are not identifying isomorphic graphs. We have fixed the set of vertices, and two trees are counted as the same if and only if exactly the same pairs of vertices are adjacent. A *spanning tree* of a connected graph G is a spanning subgraph of G that is a tree. The theorem could have been stated: the complete graph K_n has n^{n-2} spanning trees.

Example 2.1. Here are the 16 labeled trees on four vertices:



Example 2.2. There are three nonisomorphic trees on five vertices:



The number of spanning trees in K_5 isomorphic to a specific tree T on five vertices is 5! divided by the order of the automorphism group of T (why?). Thus there are $5!/4! = 5$ trees in K_5 isomorphic to the first tree above, and $5!/2 = 60$ trees isomorphic to either of the other two trees, for a total of 125 spanning trees.

Problem 2A. Find the six nonisomorphic trees on 6 vertices, and for each compute the number of distinct spanning trees in K_6 isomorphic to it.

Before starting the proofs, we make the following observations. (Probably the reader has already noticed these things in solving Problem 1C.) Firstly, every tree with $n \geq 2$ vertices has at least two monovalent vertices (vertices of degree 1). This is immediate, for example, from Problem 1C and equation (1.1): the sum of the degrees $d_1, d_2, \ldots, d_n$, all of which are at least 1, is $2n-2$. Secondly, if a monovalent vertex and its incident edge are deleted from a tree, the resulting graph is still a tree. Finally, given a tree T, if we introduce a new vertex x and a new edge joining x to *any* vertex of T, the new graph is again a tree.

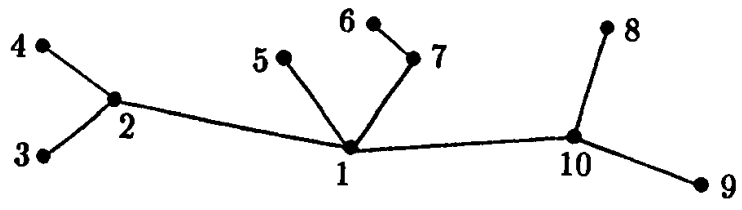


Figure 2.1

Proof 1: The first proof we present, due to H. Prüfer (1918), uses an algorithm that associates to any tree T a 'name' $\mathcal{P}(T)$ (called the *Prüfer code*) that characterizes the tree.

For the vertices of K_n, we take the ordered set $V = \{1, 2, 3, \ldots, n\}$. Given a spanning tree T in K_n, we let $T_1 = T$ and generate a sequence of trees $T_1, T_2, \ldots, T_{n-1}$ and two sequences of vertices as follows: Given the tree T_i with $n-i+1$ vertices, $i = 1, 2, \ldots, n-1$, let x_i be the least monovalent vertex of T_i and delete x_i and its incident edge $\{x_i, y_i\}$ from T_i to obtain a tree T_{i+1} on $n-i$ vertices. The *name* of T is to be

$$\mathcal{P}(T) = (y_1, y_2, \ldots, y_{n-2}).$$

We claim that the mapping $\mathcal{P}$, from the set of all spanning trees in K_n to the set V^{n-2} of all possible names, is one-to-one and onto (bijective). This will prove that the number of spanning trees in K_n is n^{n-2}.

For the tree in Fig. 2.1, where $n = 10$, we have $(x_1, y_1) = (3, 2)$, $(x_2, y_2) = (4, 2)$, $(x_3, y_3) = (2, 1)$,..., $(x_9, y_9) = (9, 10)$; these edges are the columns of the matrix below.

$$\begin{bmatrix} 3 & 4 & 2 & 5 & 6 & 7 & 1 & 8 & 9 \\ 2 & 2 & 1 & 1 & 7 & 1 & 10 & 10 & 10 \end{bmatrix}$$

So $\mathcal{P}(T) = (2, 2, 1, 1, 7, 1, 10, 10)$. Don't include $y_9 = 10$.

To understand why $\mathcal{P}$ is bijective, we first note some simple facts about the x_i's and y_i's. First, $y_{n-1} = n$, always. This is because every tree (with at least two vertices) has at least two monovalent vertices, so the vertex n will never be the least monovalent vertex. Second, $x_k, x_{k+1}, \ldots, x_{n-1}$ and n are the vertices of the tree T_k. Third, $\{x_i, y_i\}$, $k \le i \le n-1$, are exactly the edges of T_k, in some order.

The number of times a vertex v occurs among $y_1, y_2, \ldots, y_{n-2}$ is $\deg_T(v) - 1$. This is because v occurs $\deg_T(v)$ times among the edges $\{x_i, y_i\}$, $1 \le i \le n-1$, and exactly once in $x_1, x_2, \ldots, x_{n-1}, y_{n-1}$. Similarly, the number of times a vertex v of T_k occurs among $y_k, y_{k+1}, \ldots, y_{n-2}$ is its degree in the tree T_k less 1. In particular, *the monovalent vertices of T_k are those elements of V not in*

$$\{x_1, x_2, \ldots, x_{k-1}\} \cup \{y_k, y_{k+1}, \ldots, y_{n-1}\},$$

and this means that x_k, the least monovalent vertex of T_k, is the least element of $\{1, 2, \ldots, n\}$ not in the above set. In particular,

x_1 is the least element of V not in the name $\mathcal{P}(T)$, and we can uniquely determine x_k from $\mathcal{P}(T)$ and $x_1, \ldots, x_{k-1}$. □

Problem 2B. How many trees T are there on the set of vertices $\{1, 2, 3, 4, 5, 6, 7\}$ in which the vertices 2 and 3 have degree 3, vertex 5 has degree 2, and hence all others have degree 1? Do not just draw pictures but consider the possible Prüfer codes of these trees.

PROOF 2: We give another proof, again by a reversible algorithm. Consider any mapping f from $\{2, 3, \ldots, n-1\}$ to $\{1, 2, \ldots, n\}$. There are n^{n-2} such mappings f. Construct a digraph D on the vertices 1 to n by defining $(i, f(i))$, $i = 2, \ldots, n-1$, to be the edges. Fig. 2.2 shows an example with $n = 21$.

D consists of two trees 'rooted' at 1 and n and a number (say k) of circuits (directed polygons) to which trees are attached. (Directed trees with all the edges pointing in the direction of one vertex, called the root, are called *arborescences.*) These circuits are placed as in Fig. 2.2 where the rightmost vertex in the i-th component, denoted by r_i, is its minimal element (and l_i is the vertex on the left). The circuits are ordered by the condition $r_1 < r_2 < \cdots < r_k$. To D we adjoin the tree obtained by adding the edges $\{1, l_1\}$, $\{r_1, l_2\}$, …, $\{r_{k-1}, l_k\}$, $\{r_k, n\}$ and deleting the edges $\{r_i, l_i\}$ as in Fig. 2.3.

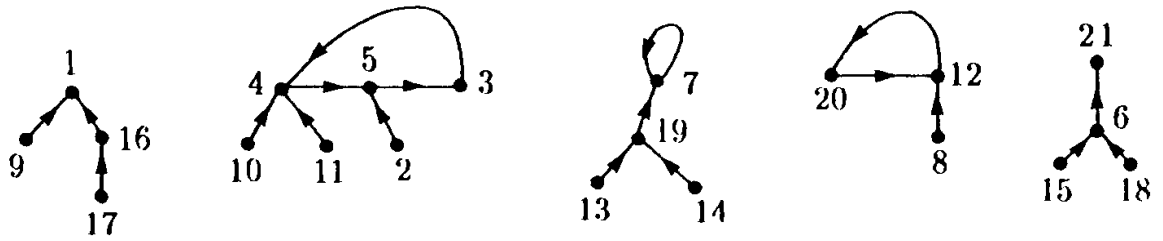


Figure 2.2

If the tree of Fig. 2.3 is given, consider the path from 1 to n (=21). Let $r_0 := 1$. Define r_1 to be the minimal number on this path (excluding $r_0 = 1$) and in general r_i as the minimal number on the path from r_{i-1} to n. It is easily seen that we recover the function f in this way. □

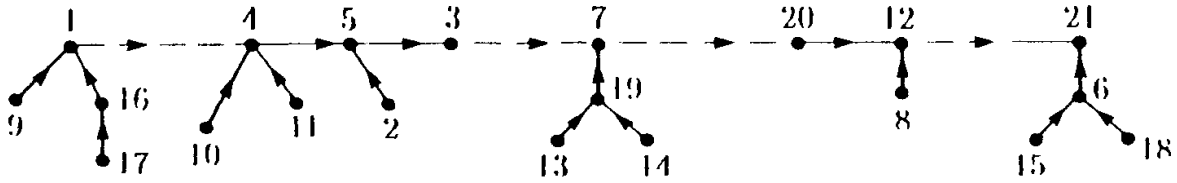


Figure 2.3

Generalizations of this proof may be found in Eğecioğlu and Remmel (1986).

Problem 2C. Let G be a directed graph with vertices $x_1, \ldots, x_n$ for which a (directed) Eulerian circuit exists. A *spanning arborescence* with root x_i is a spanning tree T of G, with root x_i, such that for all $j \neq i$ there is a directed path from x_j to x_i in T. Show that the number of spanning arborescences of G with root x_i does not depend on i. (This is difficult; see the hints.)

Proof 3: We now give a proof by counting since it is useful to have seen this method. We remind the reader of the definition of a *multinomial coefficient*. Let $r_1, r_2, \ldots, r_k$ be nonnegative integers with sum n. Then $\binom{n}{r_1,\ldots,r_k}$ is defined by

$$(x_1 + x_2 + \cdots + x_k)^n = \sum \binom{n}{r_1, \ldots, r_k} x_1^{r_1} x_2^{r_2} \ldots x_k^{r_k}, \tag{2.1}$$

where the sum is over all k-tuples $(r_1, \ldots, r_k)$ with sum n.

Since $(x_1 + \cdots + x_k)^n = (x_1 + \cdots + x_k)^{n-1}(x_1 + \cdots + x_k)$, we have

$$\binom{n}{r_1, \ldots, r_k} = \sum_{i=1}^{k} \binom{n-1}{r_1, \ldots, r_i - 1, \ldots, r_k}. \tag{2.2}$$

We denote the number of labeled trees with n vertices for which the degrees are $d_1, d_2, \ldots, d_n$ by $t(n; d_1, d_2, \ldots, d_n)$. Clearly this number is 0 if one of the d_i is 0. The value of $t(n; d_1, d_2, \ldots, d_n)$ depends only on the multiset of numbers d_i and not on their order. We may assume without loss of generality that $d_1 \geq d_2 \geq \cdots \geq d_n$, so $d_n = 1$. Take the vertex v_n corresponding to d_n. It is joined to some vertex v_i of degree $d_i \geq 2$, and any of the remaining vertices

is a candidate. Therefore

$$t(n; d_1, \ldots, d_n) = \sum_{i=1}^{n-1} t(n-1; d_1, \ldots, d_i - 1, \ldots, d_{n-1}). \tag{2.3}$$

It is trivial to check by hand that

$$t(n; d_1, \ldots, d_n) = \binom{n-2}{d_1 - 1, \ldots, d_n - 1} \tag{2.4}$$

for $n = 3$. Since the numbers on the left-hand side, respectively right-hand side, of (2.4) satisfy the same recurrence relation ((2.3), respectively (2.2)) it follows by induction that (2.4) is true for all n. In (2.1), we replace n by $n-2$, k by n, r_i by $d_i - 1$ and x_i by 1. We find

$$n^{n-2} = \sum t(n; d_1, d_2, \ldots, d_n).$$

□

Compare (2.4) with Problem 2B.

A spanning tree is easily constructed by starting at any vertex, taking the edges to vertices at distance 1, then one edge to each vertex at distance 2, etc. Several other constructions are possible (e.g. by starting with G and deleting suitable edges).

A graph with no polygons as subgraphs is called a *forest*. Each component $C_1, C_2, \ldots, C_k$ of a forest G is a tree, so if a forest with n vertices has k components, it has

$$(|V(C_1)| - 1) + (|V(C_2)| - 1) + \cdots + (|V(C_k)| - 1) = n - k$$

edges.

A *weighted graph* is a graph G together with a function associating a real number $c(e)$ (usually nonnegative) to each edge e, called its *length* or *cost* according to context. Let us use the term 'cost' here. Given a weighted connected graph G, define the *cost* of a spanning tree T of G as

$$c(T) := \sum_{e \in E(T)} c(e).$$

The graph may represent a network of cities where $c(\{x,y\})$ is the cost of erecting a telephone line joining cities x and y, and so it is clear that finding a *cheapest* spanning tree in G is a problem of practical importance.

The following method is often called the *greedy algorithm.* In fact, it is only one of a number of algorithms which can be called greedy algorithms, where one does not plan ahead but takes what seems to be the best alternative at each moment and does not look back. It is surprising that such a simple procedure actually produces a cheapest spanning tree, but this is proved in Theorem 2.2 below. Let us say that a set S of edges of a graph G is *independent* when the spanning subgraph with edge set S (denoted $G{:}S$) is a forest.

Greedy algorithm. Let G be a connected weighted graph with n vertices. At each point, we will have a set $\{e_1, e_2, \ldots, e_i\}$ of i independent edges ($i = 0$ to start), so that $G{:}\{e_1, e_2, \ldots, e_i\}$ has $n-i$ components. If $i < n-1$, let e_{i+1} be an edge with ends in different components of $G{:}\{e_1, e_2, \ldots, e_i\}$ and whose cost is minimum with respect to this property. Stop when we have chosen $n-1$ edges.

Theorem 2.2. *With $e_1, \ldots, e_{n-1}$ chosen as above, the spanning tree $T_0 := G{:}\{e_1, \ldots, e_{n-1}\}$ has the property that $c(T_0) \le c(T)$ for any spanning tree T.*

PROOF: Let $\{a_1, a_2, \ldots, a_{n-1}\}$ be the edge set of a tree T, numbered so that $c(a_1) \le c(a_2) \le \cdots \le c(a_{n-1})$. We claim something much stronger than $c(T_0) \le c(T)$; namely, we claim that $c(e_i) \le c(a_i)$ for each $i = 1, 2 \ldots, n-1$. If this is false, then

$$c(e_k) > c(a_k) \ge c(a_{k-1}) \ge \cdots \ge c(a_1)$$

for some k. Since none of $a_1, a_2, \ldots, a_k$ was chosen at the point when e_k was chosen, each of these k edges has both ends in the same component of $G{:}\{e_1, e_2, \ldots, e_{k-1}\}$. Then the number of components of $G{:}\{a_1, a_2, \ldots, a_k\}$ is at least the number $n-k+1$ of components of $G{:}\{e_1, e_2, \ldots, e_{k-1}\}$ and this contradicts the fact that $\{a_1, a_2, \ldots, a_k\}$ is independent. □

Problem 2D. Here is a variation on the above greedy algorithm. Let x_1 be any vertex of a weighted connected graph G with n

vertices and let T_1 be the subgraph with the one vertex x_1 and no edges. After a tree (subgraph) T_k, $k < n$, has been defined, let e_k be a cheapest edge among all edges with one end in $V(T_k)$ and the other end *not* in $V(T_k)$, and let T_{k+1} be the tree obtained by adding that edge and its other end to T_k. Prove that T_n is a cheapest spanning tree in G.

In many practical situations, it is necessary to search through a tree starting from a particular vertex. (A tree with a distinguished vertex—the *root*—is called a *rooted tree*.) There are two well known methods known as *depth-first search* and *breadth-first search*. We explain the terminology by the example of Fig. 2.4.

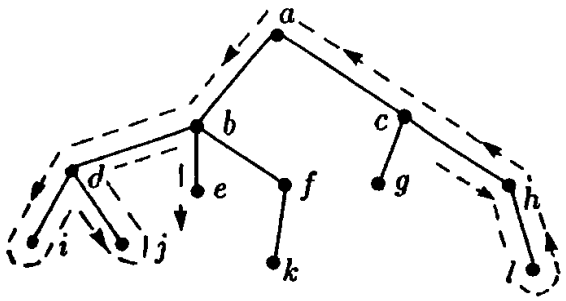


Figure 2.4

In a depth-first search starting at a, one essentially considers the tree as a fence and walks around it, keeping the fence on the left, i.e. along the walk $abdidjdbebfk \ldots lhca$. If one decides to number the vertices in accordance with the search, one finds the numbering $a = 1$, $b = 2$, $d = 3$, $i = 4$, $\ldots$, $l = 12$. In this description, we rely on a planar drawing of the tree; but see below.

In a breadth-first search, one proceeds as in the construction of a spanning tree mentioned above. The vertices are then numbered in the order of Fig. 2.4, i.e. alphabetically.

These ideas apply, more generally, to searching through the vertices of a connected graph.

Given a finite connected graph G, we can obtain a numbering of the vertices of G and a spanning tree T of G, called a 'depth-first search tree' for G, in the following manner. Pick a vertex v_0 and start with the tree T_0 with vertex v_0 and no edges. Proceed inductively: once vertices $v_0, v_1, v_2, \ldots, v_k$ and a tree T_k with exactly those vertices and some of the edges of G have been chosen, let ℓ

be the largest index $\leq k$ so that v_ℓ is adjacent to some vertex not in T_k. Such a vertex will exist if T_k is not yet a spanning tree; otherwise stop and let $T = T_k$. Call that new vertex v_{k+1} and add it and the edge $\{v_\ell, v_{k+1}\}$ to T_k to obtain a tree T_{k+1}. We consider T as a rooted tree with root v_0.

We give two properties of depth-first search trees and use them to give a simple constructive proof of a theorem on orientations of graphs.

Given a vertex x of a rooted tree with root v_0, the *ancestors* of x are the vertices traversed by the (unique) path from x to the root v_0. The first vertex other than x on that path is the *parent* of x. If x is an ancestor of y, we also say that y is a *descendant* of x. We will count x as a descendent and ancestor of itself.

Proposition 2.3. If vertices x and y are adjacent in G, then one of them is a descendant of the other in any depth-first search tree T of G.

PROOF: Suppose x is numbered with a smaller index than y in the depth-first search, say $x = v_k$.

At the stage when $v_0, v_1, v_2, \ldots, v_k$ have been chosen, the largest index ℓ so that v_ℓ is adjacent to an as-yet-unnumbered vertex is clearly $\ell = k$. So v_{k+1} is joined to v_k (and not to some v_i with $i < k$) in T. If $v_{k+1} = y$, we are done; y is a descendant of v_k. Otherwise, since v_k is still adjacent to an as-yet-unnumbered vertex, namely y, the choice of ℓ will be $\ell = k$ or $\ell = k+1$, and v_{k+2} is adjacent to either v_k or v_{k+1} in T. If $v_{k+2} = y$, we are done; y is a descendant of v_k.

Inductively, *as long as y remains unnumbered*, $v_k, v_{k+1}, \ldots, v_{k+j}$ will be descendants of v_k and the next choice of ℓ will be as one of $k, k+1, \ldots, k+j$. Then the vertex v_{k+j+1} numbered next will be adjacent to one of $v_k, v_{k+1}, \ldots, v_{k+j}$ and hence will be a descendant of v_k. Since the graph is finite, eventually this newly numbered vertex must be y. □

An *isthmus* of a connected graph G is an edge whose deletion results in a disconnected graph. (Some authors call this a *bridge*.)

Proposition 2.4. Let $\{x, y\}$ be an edge of T which is not an isthmus in G; say x is the parent of y. Then there is an edge in G

but not in T joining some descendant a of y and some ancestor b of x.

PROOF: Let D be the set of descendants of y. So $y \in D$ and $x \notin D$. Since G with $\{x, y\}$ deleted is still connected, there is some edge $\{a, b\} \neq \{x, y\}$ with one end $a \in D$ and the other end $b \notin D$. This edge $\{a, b\}$ is certainly not in T and by Proposition 2.3, b is an ancestor of a (since it cannot be a descendant of a because then it would be a descendant of y too and hence would be in D). The (unique) path in T from a to v_0 passes through y (an ancestor of a) and then x (the parent of y); but b must be on this path (since it is an ancestor of a), so b must be an ancestor of x too. □

Any directed graph obtained from an undirected graph G by assigning a direction to each edge of G is called an *orientation* of G. A walk in a digraph D may be called *strong* when each edge is traversed by the walk according to its direction, i.e. from its tail to its head, and the digraph D is *strongly connected* when for any two vertices x, y, there is a strong walk from x to y.

Theorem 2.5. Let G be a finite connected graph without isthmuses. Then G admits a strong orientation, i.e. an orientation that is a strongly connected digraph.

PROOF: We will construct a digraph D from G by choosing a direction for each edge. Find a depth-first search tree T and numbering $v_0, v_1, \ldots$ of the vertices of G. Let $\{v_i, v_j\}$ be an edge of G with $i < j$. If $\{v_i, v_j\}$ is in T, direct it from v_i to v_j, i.e. (v_i, v_j) is an edge of D. If $\{v_i, v_j\}$ is in not T, direct it from v_j to v_i, i.e. (v_j, v_i) is an edge of D.

It remains to show that D is strongly connected. There is a strong walk from v_0 to any vertex x of G (using only edges of the tree T), so it will suffice to show that we can find a strong walk from any vertex x to v_0.

Given a vertex x_k, $k > 0$, Proposition 2.4 says that some edge $\{a, b\}$ in G but not in T joins some descendant a of y to some ancestor $b = v_i$ of v_k. We get a strong walk in D from v_k to v_i by appending the directed edge (a, v_i) to a strong walk in T from v_k to its descendant a. Of course, $i < k$ since v_i is an ancestor of v_k. If $i = 0$ we are done. Otherwise, we repeat the argument to find

a strong walk from v_i to some v_j with $j < i$ and concatenate the walks to get a strong walk from v_k to v_j. Continue in this manner until you reach v_0. □

Problem 2E. A *graceful labeling* of a tree T on n vertices is a mapping $f : V(T) \to \{1, 2, \ldots, n\}$ so that the numbers $|f(x)-f(y)|$ computed across edges $\{x, y\}$ are all different. Show that the path-graphs (trees with exactly two monovalent vertices) admit graceful labelings. (It is conjectured that all trees admit graceful labelings.)

Problem 2F. Suppose a tree G has exactly one vertex of degree i for $2 \leq i \leq m$ and all other vertices have degree 1. How many vertices does G have?

Problem 2G. Let G be a graph with exactly one vertex of degree i for $2 \leq i \leq m$ and k other vertices, all of degree 1. Prove that $k \geq \lfloor \frac{m+3}{2} \rfloor$. Give a construction for such a graph.

Problem 2H. Consider labeled trivalent rooted trees T with $2n$ vertices, counting the root labeled $2n$; see Figure 14.3. The labels are chosen in such a way that the procedure leading to $\mathcal{P}(T)$ has $1, 2, 3, \ldots, 2n-1$ as first row. How many possible codes $\mathcal{P}(T)$ are there?

Notes.

A. Cayley (1821–1895), professor at Cambridge from 1863 until his death, was one of the great mathematicians of the 19th century. His work includes important contributions to the theory of elliptic functions, analytic geometry and algebra, e.g. the theory of invariants. His paper on trees appeared in 1889 but it did not contain what we would consider a proof. Of the many proofs (five of which are treated in this book) the one by Prüfer is the best known.

H. Prüfer (1896–1934) was one of I. Schur's many pupils. He was professor at Münster.

References.

A. Cayley (1889), A theorem on trees, *Quart. J. Pure and App. Math.* **23**, 376–378.

Ö. Eğecioğlu and J. B. Remmel (1986), Bijections for Cayley trees, spanning trees, and their q-analogues, *J. Combinatorial Theory* (A) **42**, 15–30.

H. Prüfer (1918), Neuer Beweis eines Satzes über Permutationen, *Archiv der Math. und Phys.* (3) **27**, 142–144.

3
Colorings of graphs and Ramsey's theorem

We shall first look at a few so-called coloring problems for graphs.

A *proper coloring* of a graph G is a function from the vertices to a set C of 'colors' (e.g. $C = \{1, 2, 3, 4\}$) such that the ends of every edge have distinct colors. (So a graph with a loop will admit no proper colorings.) If $|C| = k$, we say that G is *k-colored*.

The *chromatic number* $\chi(G)$ of a graph G is the minimal number of colors for which a proper coloring exists.

If $\chi(G) = 2$ (or $\chi(G) = 1$, which is the case when and only when G has no edges), then G is called *bipartite*. A graph with no odd polygons (equivalently, no closed paths of odd length) is bipartite as the reader should verify.

The famous 'Four Color Theorem' (K. Appel and W. Haken, 1977) states that if G is planar, then $\chi(G) \leq 4$.

Clearly $\chi(K_n) = n$. If k is odd then $\chi(P_k) = 3$. In the following theorem, we show that, with the exception of these examples, the chromatic number is at most equal to the maximum degree (R. L. Brooks, 1941).

Theorem 3.1. *Let $d \geq 3$ and let G be a graph in which all vertices have degree $\leq d$ and such that K_{d+1} is not a subgraph of G. Then $\chi(G) \leq d$.*

Proof 1: As is the case in many theorems in combinatorial analysis, one can prove the theorem by assuming that it is not true, then considering a *minimal* counterexample (in this case a graph with the minimal number of vertices) and arriving at a contradiction. We shall use the technique of recoloring: it is possible to change

the colors of certain vertices to go from one proper coloring to another. For example, let S be a subset of the set C of colors. On any connected component of the subgraph induced by the vertices with colors from S, we arbitrarily permute the colors (without changing those of the vertices with colors in $C\backslash S$). Clearly we again have a proper coloring.

So let G be a counterexample with the minimum number of vertices. Let $x \in G$ and let $\Gamma(x) = \{x_1, \ldots, x_l\}$, $l \leq d$. Since G is a minimal counterexample, the graph H, obtained by deleting x and the edges incident with x, has a d-coloring, say with colors $1, 2, \ldots, d$. If one of these colors is not used in the coloring of $\Gamma(x)$, then we can assign this color to x and obtain a d-coloring of G. It follows that $l = d$ and every d-coloring of H must use all the colors on the set $\Gamma(x)$. Let us assume that x_i has color i for $i = 1, 2, \ldots, d$.

Now consider x_i and x_j and the induced subgraph H_{ij} of H with colors i and j. If x_i and x_j were in different connected components of H_{ij}, then we could interchange the colors in one of these components, after which x_i and x_j would have the same color, which is impossible. So x_i and x_j are in the same component (say C_{ij}) of H_{ij}. We shall now show that this component is (the graph of) a *simple path* (with alternating colors i and j) from x_i to x_j. If two neighbors of x_i in H had color j, then the neighbors of x_i in H would have at most $d-2$ different colors. Then we could recolor x_i and that is impossible. Suppose y is the first vertex on a path from x_i to x_j in C_{ij} that has degree ≥ 3. The neighbors of y in H have at most $d-2$ colors, so we can recolor y to some color $\notin \{i, j\}$ and then x_i and x_j are no longer connected in H_{ij}, which we know to be impossible. So such a y does not exist, proving that C_{ij} is a path.

Suppose that z is a vertex $\neq x_i$ on C_{ij} and on C_{ik}. Then z has two neighbors with color j and two with color k. Again the neighbors of z in H have at most $d-2$ colors and z can be recolored to some color $\notin \{i, j, k\}$, again a contradiction. Hence $C_{ij} \cap C_{ik} = \{x_i\}$.

Our assumption that $K_{d+1} \nsubseteq G$ shows that there are two vertices in $\Gamma(x)$, say x_1 and x_2, that are not connected by an edge. We have the situation of Fig. 3.1. The vertex a is the neighbor of x_1 with color 2 on C_{12}.

We recolor H by interchanging the colors 1 and 3 on the subgraph C_{13}. For the new coloring, we have new paths that we call C'_{ij}. Clearly $a \in C'_{23}$ (since x_1 now has color 3). However, on C_{12} no point except x_1 has changed color, so $a \in C'_{12}$. Hence $C'_{12} \cap C'_{23} \neq \{x_2\}$, contradicting what we proved above. The contradiction shows that our assumption that a minimal counterexample exists is false. □

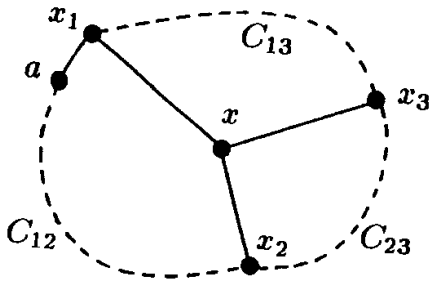


Figure 3.1

In preparation for a second proof of Brook's Theorem, the reader should do the following problem without applying the theorem.

Problem 3A. Fix an integer $d \geq 3$. Let H be a simple graph with all degrees $\leq d$ which cannot be d-colored and which is minimal (with the fewest vertices) subject to these properties. (We claim H is complete on $d+1$ vertices, but we don't know that yet.) (i) Show that H is nonseparable (this means that every graph obtained from H by deleting a vertex is connected). (ii) Then show that if the vertex set $V(H)$ is partitioned into sets X and Y with $|Y| \geq 3$, then there are at least three vertices $a, b, c \in Y$ each of which is adjacent to at least one vertex in X.

Proof 2: Let d and H be as in Problem 3A. If H is not complete, there are vertices x_1, x_{n-1}, x_n so that x_1 is adjacent to x_{n-1} and x_n but so that these last two vertices are not adjacent. We want to number the other $n-3$ vertices so that the sequence

$$x_1, x_2, \ldots, x_{n-1}, x_n$$

has the property that each x_k, $k \geq 2$, is adjacent to at least one of the vertices x_i preceding it, i.e. with $i < k$. This is simple: when $x_1, x_2, \ldots, x_k$ have been chosen, $k < n-2$, choose x_{k+1} to be any

vertex other than x_{n-1} or x_n adjacent to one of $x_1, x_2, \ldots, x_k$; since there are at least three such vertices, this is always possible.

Once this is done, we d-color the vertices starting at the end of the sequence. Assign x_{n-1} and x_n the same color. When the last k vertices $x_{k+1}, \ldots, x_{n-1}, x_n$ have been colored, $k \geq 2$, there is a color available for x_k since x_k is adjacent to at most $d-1$ of the already-colored vertices. Finally, there is a color available for x_1 since two of the vertices adjacent to x_1 have been given the same color. □

We now consider a coloring problem of a completely different nature. It serves as an introduction to a very important theorem of combinatorics, namely Ramsey's theorem. Before reading on, the reader should try the following problem.

Problem 3B. Let the edges of K_7 be colored with the colors red and blue. Show that there are at least four subgraphs K_3 with all three edges the same color (*monochromatic triangles*). Also show that equality can occur.

The example that is always used to introduce this subject is K_6 with its edges colored red or blue. We shall show that there is at least one monochromatic triangle. A proof is as follows. Let a be any vertex. Because a has degree 5, it is incident with at least three edges of the same color, say red edges to the vertices b, c, d. If one of the edges between these three vertices is red, then we have a red triangle; if not, then they form a blue triangle.

The idea of the proof is the same as for the more difficult situation of Ramsey's theorem. However, it does not show as much as the following counting argument.

Theorem 3.2. *If the edges of K_n are colored red or blue, and r_i, $i = 1, 2, \ldots, n$, denotes the number of red edges with vertex i as an endpoint, and if Δ denotes the number of monochromatic triangles, then*

$$\Delta = \binom{n}{3} - \frac{1}{2}\sum_{i=1}^{n} r_i(n-1-r_i). \tag{3.1}$$

Proof: Every triangle in K_n that is not monochromatic has exactly two vertices where a red and a blue edge meet. On the i-th

vertex, two such edges can be chosen in $r_i(n-1-r_i)$ ways. So the sum in (3.1) counts the bichromatic triangles twice. □

Corollary.

$$\Delta \geq \binom{n}{3} - \lfloor \frac{n}{2} \lfloor (\frac{n-1}{2})^2 \rfloor \rfloor. \tag{3.2}$$

PROOF: From (3.1) we see that Δ is minimized if $r_i = (n-1-r_i)$ for all i when n is odd, or if $r_i = \frac{n}{2}$ or $r_i = \frac{n}{2} - 1$ for all i in the case that n is even. Since Δ is an integer, the first situation cannot always arise. It is easy to show that (3.2) cannot be improved. □

Note that this argument shows that a red-blue coloring of K_6 must always have at least two monochromatic triangles.

We now treat Ramsey's theorem (Ramsey, 1930).

Theorem 3.3. *Let $r \geq 1$ and $q_i \geq r$, $i = 1, 2, \ldots, s$ be given. There exists a minimal positive integer $N(q_1, q_2, \ldots, q_s; r)$ with the following property. Let S be a set with n elements. Suppose that all $\binom{n}{r}$ r-subsets of S are divided into s mutually exclusive families $T_1, \ldots, T_s$ ('colors'). Then if $n \geq N(q_1, q_2, \ldots, q_s; r)$ there is an i, $1 \leq i \leq s$, and some q_i-subset of S for which every r-subset is in T_i.*

(The reader should compare this with our introductory example and show that $N(3, 3; 2) = 6$.)

PROOF: We give the proof only for $s = 2$. The general case only involves a little more bookkeeping.

(a) Trivially, the theorem is true for $r = 1$ and $N(p, q; 1) = p + q - 1$.

(b) For any r and $p \geq r$ it is also obvious that $N(p, r; r) = p$ and similarly $N(r, q; r) = q$ for $q \geq r$.

(c) We proceed by induction on r. So assume the theorem is true for $r - 1$. We now use induction on $p + q$, using (b). So we can define $p_1 = N(p-1, q; r)$, $q_1 = N(p, q-1; r)$. Let S be a set with n elements, where $n \geq 1 + N(p_1, q_1; r-1)$. Let the r-subsets of S be colored with two colors, say red and blue. As in the proof for K_6, we pick an arbitrary element a of S. We now define a coloring of the $(r-1)$-subsets of $S' := S \backslash \{a\}$ by giving $X \subseteq S'$ the same

color as $X \cup \{a\}$. By induction S' either contains a subset A of size p_1 such that all its $(r-1)$-subsets are red or a subset B of size q_1 such that all its $(r-1)$-subsets are colored blue. Without loss of generality the first situation occurs. Since A has $N(p-1,q;r)$ elements, there are two possibilities. The first is that A has a subset of q elements with all its r-subsets blue, in which case we are done. The other possibility is that A has a subset A' of $p-1$ elements with all its r-subsets red. The set $A' \cup \{a\}$ also has this property because $A' \subseteq A$. This proves the theorem and furthermore we have shown

$$N(p,q;r) \leq N(N(p-1,q;r), N(p,q-1;r); r-1) + 1. \tag{3.3}$$

□

A special case of (3.3) occurs when we go back to the coloring of edges ($r = 2$) of a graph with two colors. Using (a) from the proof, we find

$$N(p,q;2) \leq N(p-1,q;2) + N(p,q-1;2). \tag{3.4}$$

Problem 3C. Show that equality cannot hold in (3.4) if both terms on the right-hand side are even.

Theorem 3.4.

$$N(p,q;2) \leq \binom{p+q-2}{p-1}.$$

PROOF: Since $N(p,2;2) = p$, the result follows from (3.4) because binomial coefficients satisfy the same relation with equality. □

Let us look at what we now know about $N(p,q;2)$. By Problem 3C, we have $N(3,4;2) \leq 9$. To show that equality holds, we have to color K_8 such that there is no red triangle and no blue K_4. We do this as follows: number the vertices with the elements of $\mathbb{Z}_8$. Let the edge $\{i,j\}$ be red if and only if $i-j \equiv \pm 3$ or $i-j \equiv 4 \pmod 8$. One easily checks that this coloring does the job.

Problem 3D. Use the same method to show that $N(4,4;2) = 18$ and that $N(3,5;2) = 14$.

With a lot more work it has been shown that

$$N(3,6;2) = 18, \quad N(3,7;2) = 23, \quad N(3,8;2) = 28, \quad N(3,9;2) = 36,$$

$$N(4,5;2) = 25.$$

No other values of $N(p,q;2)$ are known.

One of the interesting problems in this area that has seen virtually no progress for 30 years is the asymptotic behavior of $N(p,p;2)$. We know by Theorem 3.4 that

$$N(p,p;2) \le \binom{2p-2}{p-1} \le 2^{2p-2}. \tag{3.5}$$

We now show that $N(p,p;2)$ grows exponentially, using a method that is used quite often in combinatorics. It is often referred to as 'probabilistic' since it estimates the probability that a random coloring has a monochromatic K_p. Consider a K_n. There are $2^{\binom{n}{2}}$ different ways of coloring the edges red or blue. Now fix a subgraph K_p. There are $2^{\binom{n}{2}-\binom{p}{2}+1}$ colorings for which that K_p is monochromatic. The number of colorings for which some K_p is monochromatic is at most $\binom{n}{p}$ times as large (because we may count some colorings more than once). If this number is less than the total number of colorings, then there exist colorings with no monochromatic K_p. Using the fact that $\binom{n}{p} < n^p/p!$, we find that such a coloring certainly exists if $n < 2^{p/2}$ (unless $p = 2$). This proves the following theorem.

Theorem 3.5. $N(p,p;2) \ge 2^{p/2}$.

From (3.5) and Theorem 3.5, we know that

$$\sqrt{2} \le \sqrt[p]{N(p,p;2)} \le 4 \qquad (p \ge 2).$$

It would be very nice if one could show that this p-th root has a limit for $p \to \infty$.

To give a considerable improvement of Theorem 3.5, we discuss a probabilistic method that is useful in many parts of combinatorics. We consider *events* $A_1, A_2, \ldots, A_n$ in a probability space. Denote by $Pr[A_i]$ the probability of the event A_i and, as usual, let $\overline{A_i}$ denote the complement of A_i, i.e. non-occurrence of A_i. We are interested in applications where the A_i denote situations we *do not* wish to occur and where we would like to assert that there is a

positive probability that *none* of the events A_i occurs. In some easy counting situations one can use

$$\sum_{i=1}^{n} Pr[A_i] < 1 \Rightarrow \cap \overline{A_i} \neq \emptyset.$$

Cf. Problem 5E. However, in general, dependence among the undesirable events leads to multiple counting which results in a sum that is much larger than 1. Of course, if the events A_i are independent, then it suffices that each has probability less than one to guarantee that non-occurrence of all of them has positive probability. The *Lovász Sieve* handles situations where there is indeed some dependence but simultaneously there are many obviously independent combinations of events.

We define what we shall call *a dependency graph* for the events $A_1, \ldots, A_n$. This is a graph G on the set of indices $\{1, 2, \ldots, n\}$ with the property that for *every* i the event A_i is independent of *every subset* of $\{A_j : \{i, j\} \notin E(G)\}$. Note that we require a lot more than that A_i is independent of each of the A_j in this subset.

Theorem 3.6. *Let G be some dependency graph for the events $A_1, \ldots, A_n$. Suppose that $Pr[A_i] \leq p$, $i = 1, \ldots, n$ and that every vertex in G has degree $\leq d$. If $4dp < 1$, then $\cap \overline{A_i} \neq \emptyset$.*

Proof: We first show that for every subset $\{i_1, i_2, \ldots, i_m\}$ of the index set,

$$Pr[A_{i_1} | \overline{A_{i_2}} \ldots \overline{A_{i_m}}] \leq \frac{1}{2d}. \tag{3.6}$$

The case $m = 1$ is trivial and for $m = 2$ we have

$$Pr[A_1 | \overline{A_2}] \leq \frac{p_1}{1 - p_2} \leq \frac{1}{4d - 1} < \frac{1}{2d},$$

where for convenience of notation we have taken $i_j = j$ and $p_i := Pr[A_i]$. We proceed by induction.

Suppose that in G, 1 is adjacent to $2, 3, \ldots, q$ and not adjacent to $q + 1, \ldots, m$. We have

$$Pr[A_1 | \overline{A_2} \ldots \overline{A_m}] = \frac{Pr[A_1 \overline{A_2} \ldots \overline{A_q} | \overline{A_{q+1}} \ldots \overline{A_m}]}{Pr[\overline{A_2} \ldots \overline{A_q} | \overline{A_{q+1}} \ldots \overline{A_m}]}.$$

The numerator is (by definition of G) at most

$$Pr[A_1|\overline{A_{q+1}}\ldots\overline{A_m}] = Pr[A_1] \leq \frac{1}{4d}.$$

Using the induction hypothesis, we find that the denominator is at least

$$1 - \sum_{i=2}^{q} Pr[A_1|\overline{A_{q+1}}\ldots\overline{A_m}] \geq 1 - \frac{q-1}{2d} \geq \frac{1}{2}.$$

This proves (3.6). We now have

$$Pr[\overline{A_1}\ldots\overline{A_n}] = \prod_{i=1}^{n} Pr[\overline{A_i}|\overline{A_1}\ldots\overline{A_{i-1}}] \geq (1 - \frac{1}{2d})^n > 0,$$

where we have used (3.6) for each term in the product. □

We apply this method to obtain a lower bound for $N(p,p;2)$.

Theorem 3.7. *$N(p,p;2) \geq c \cdot p \cdot 2^{p/2}$, where c is a constant.*

PROOF: Consider K_n and color the edges randomly with two colors. For each set S of k vertices let A_S be the event that the subgraph on S is colored monochromatically. We wish to assert that among the random colorings, there is at least one in which no monochromatic subgraph on k vertices occurs. We define a dependency graph by making S and T adjacent if and only if $|S \cap T| \geq 2$, i.e. the subgraphs on S and T have an edge in common. The degree d of G is clearly at most $\binom{k}{2}\binom{n}{k-2}$. The events A_S all have probability $2^{1-\binom{k}{2}}$. From Theorem 3.6, Stirling's formula and a little manipulation, we find the result (and if we wish an estimate for c). □

We have given some examples of an area of combinatorics known as Ramsey theory. We just mention one more example, namely a theorem due to B. L. van der Waerden (1927). It states that there exists a number $N(r)$ such that if $N \geq N(r)$ and the integers from 1 to N are colored red or blue, then there is a monochromatic arithmetic progression of length r in the set. For a short (but not easy) proof see Graham and Rothschild (1974). A general reference

for this area is the book *Ramsey Theory* by R. L. Graham, B. L. Rothschild and J. L. Spencer (1980).

An interesting application of Ramsey's theorem is the following theorem due to Erdős and Szekeres (1935).

Theorem 3.8. *For a given n, there is an integer $N(n)$ such that any collection of $N \geq N(n)$ points in the plane, no three on a line, has a subset of n points forming a convex n-gon.*

PROOF: (i) First we observe that if we have n points, no three on a line, then they form a convex n-gon if and only if every quadrilateral formed by taking four of the points is convex.

(ii) We now claim that $N(n) = N(n, n; 3)$ will do the job. Let S be a set of $N(n)$ points. Number the points and then color triangles red, respectively blue, if the path from the smallest number via the middle one to the largest number is clockwise, respectively counterclockwise. There is an n-subset with all its triangles the same color, say red. We shall show that this set cannot contain the configuration of Fig. 3.2.

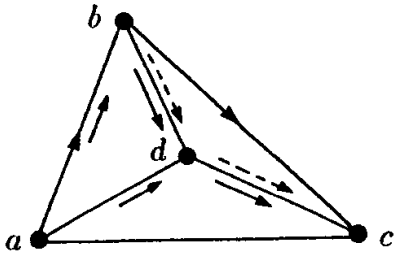


Figure 3.2

Without loss of generality $a < b < c$. From triangle adc, we see that $a < d < c$. Then from triangle abd it follows that $a < b < d$. But then triangle bcd is blue, a contradiction. So all quadrilaterals formed from the n-subset are convex and by (i), we are done. □

Problem 3E. A *tournament* on n vertices is an orientation of K_n. A *transitive tournament* is a tournament for which the vertices can be numbered in such a way that (i, j) is an edge if and only if $i < j$.

(a) Show that if $k \leq \log_2 n$, every tournament on n vertices has a transitive subtournament on k vertices.

(b) Show that if $k > 1 + 2\log_2 n$, there exists a tournament on n vertices with no transitive subtournament on k vertices.

Problem 3F. Prove that for all $r \in \mathbb{N}$ there is a minimal number $N(r)$ with the following property. If $n \geq N(r)$ and the integers in $\{1, 2, \ldots, n\}$ are colored with r colors, then there are three elements x, y, z (not necessarily distinct) with the same color and $x + y = z$. (A result due to I. Schur.) Determine $N(2)$. Show by an elementary argument that $N(3) > 13$.

Problem 3G. Let m be given. Show that if n is large enough, every $n \times n$ $(0,1)$-matrix has a principal submatrix of size m, in which all the elements below the diagonal are the same, and all the elements above the diagonal are the same.

Problem 3H. Show that if the edges of K_{17} are colored with three colors, there must be a monochromatic triangle.

Problem 3I. Let $\{1, \alpha, \alpha^2, \ldots, \alpha^{14}\}$ be the multiplicative group of $\mathbb{F}_{16}$. Number the vertices of K_{16} with the elements of $\mathbb{F}_{16}$. We will color the edges of K_{16} with three colors. Edge $\{i, j\}$ is to be colored with a color that depends only on ν, where $i - j = \alpha^\nu$. Do this in such a way that there is *no* monochromatic triangle. (In the notation of Theorem 3.3, this problem and the previous one show that $N(3, 3, 3; 2) = 17$.)

Problem 3J. The edges of K_n are colored red and blue in such a way that a red edge is in at most one red triangle. Show that there is a subgraph K_k with $k \geq \lfloor \sqrt{2n} \rfloor$ that contains *no* red triangle.

Problem 3K. Let G satisfy the conditions of Theorem 3.1. Show that by removing at most n/d edges we can find a subgraph G' with chromatic number $\leq d - 1$.

Notes.

The four color conjecture was considered one of the most famous open problems in combinatorics until 1976. Its solution by Appel and Haken has caused much debate because the proof depends on an extensive computer analysis of many cases. The validity of the argument depends on one's trust in computers and programs. (See Chapter 34 for two proofs of the Five Color Theorem.)

The theorem by Brooks, Theorem 3.1, which he discovered while an undergraduate at Cambridge, is a typical example of the inge-

nious arguments that are quite often necessary in that part of graph theory where algebraic methods do not apply.

F. P. Ramsey (1902–1928) died too young to produce the results that he probably would have. He was interested in decision procedures for logical systems and, strangely enough, this led to the theorem that in turn led to so-called Ramsey theory.

Theorem 3.5 is due to P. Erdős (1947). For more on probabilistic methods, see Erdős and Spencer (1974).

Values of and estimates for the numbers $N(p, q; 2)$ can be found in Radziszowski (1999). The value of $N(3, 9; 2)$ is from Grinstead and Roberts (1982).

The proof of Theorem 3.6 is not the original proof by Erdős and Szekeres that one usually finds in books. This proof was produced by a student in Haifa (M. Tarsy) during an examination! He had missed the class in which the proof had been presented. See Lewin (1976). A proof in a similar vein was given by Johnson (1986).

References.

K. Appel and W. Haken (1977), Every planar map is four-colorable, *Illinois J. Math.* **21**, 429–567.

R. L. Brooks (1941), On colouring the nodes of a network, *Cambridge Philos. Soc.* **37**, 194–197.

P. Erdős (1947), Some remarks on the theory of graphs, *Bull. Amer. Math. Soc.* **53**, 292–294.

P. Erdős and J. L. Spencer (1974), *Probabilistic Methods in Combinatorics*, Academic Press.

P. Erdős and G. Szekeres (1935), A combinatorial problem in geometry, *Compositio Math.* **2**, 463–470.

R. L. Graham and B. L. Rothschild (1974), A short proof of van der Waerden's theorem on arithmetic progressions, *Proc. Amer. Math. Soc.* **42**, 385–386.

R. L. Graham, B. L. Rothschild, and J. L. Spencer (1980), *Ramsey Theory*, Wiley.

C. M. Grinstead and S. M. Roberts (1982), On the Ramsey numbers $R(3, 8)$ and $R(3, 9)$, *J. Combinatorial Theory* (B) **33**, 27–51.

S. Johnson (1986), A new proof of the Erdős–Szekeres convex k-gon result, *J. Combinatorial Theory* (A) **42**, 318–319.

M. Lewin (1976), A new proof of a theorem of Erdős and Szekeres, *The Math. Gazette* **60**, 136–138, 298.

S. P. Radziszowski (1999), Small Ramsey Numbers, *The Electronic Journal of Combinatorics* **1** DS 1.

F. P. Ramsey (1930), On a problem of formal logic, *Proc. London Math. Soc.* (2) **30**, 264–286.

B. L. van der Waerden (1927), Beweis einer Baudetschen Vermutung, *Nieuw Archief voor Wiskunde* **15**, 212–216.

4
Turán's theorem and extremal graphs

As an introduction, we first ask the question how many edges a simple graph must have to guarantee that the graph contains a *triangle*. Since $K_{m,m}$ and $K_{m,m+1}$ do not contain triangles, we see that if the graph has n vertices, then $\lfloor n^2/4 \rfloor$ edges are not enough. We claim that if there are more edges, then the graph contains a triangle (W. Mantel, 1907). The following proof is surprising. Let G have n vertices, numbered from 1 to n, and no triangles. We give vertex i a weight $z_i \geq 0$ such that $\sum z_i = 1$ and we wish to maximize $S := \sum z_i z_j$, where the sum is taken over all edges $\{i, j\}$. Suppose that vertex k and vertex l are not joined. Let the neighbors of k have total weight x, and those of l total weight y, where $x \geq y$. Since $(z_k + \epsilon)x + (z_l - \epsilon)y \geq z_k x + z_l y$, we do not decrease the value of S if we shift some of the weight of vertex l to the vertex k. It follows that S is maximal if all of the weight is concentrated on some complete subgraph of G, i.e. on *one* edge! Therefore $S \leq \frac{1}{4}$. On the other hand, taking all z_i equal to n^{-1} would yield a value of $n^{-2}|E|$ for S. Therefore $|E| \leq \frac{1}{4}n^2$.

Note that Ramsey's theorem states that if a graph on n vertices has $n \geq N(p, q; 2)$, then the graph either has a complete subgraph on p vertices or a set of q vertices with no edges between them (called an *independent* set). We now ask the question whether some condition on the number of edges guarantees a K_p as a subgraph. We saw above what the answer is for $p = 3$. We also already have an idea of how to avoid a K_p. Divide the vertices into $p-1$ subsets $S_1, \ldots, S_{p-1}$ of almost equal size, i.e. r subsets of size $t + 1$ and $p - 1 - r$ subsets of size t, where $n = t(p - 1) + r$, $1 \leq r \leq p - 1$. Within each S_i there are no edges but every vertex in S_i is joined to

every vertex in S_j if $i \neq j$. (This is a *complete multipartite graph.*) The number of edges is

$$M(n,p) := \frac{p-2}{2(p-1)}n^2 - \frac{r(p-1-r)}{2(p-1)}.$$

Theorem 4.1. (*Turán, 1941*) *If a simple graph on n vertices has more than $M(n,p)$ edges, then it contains a K_p as a subgraph.*

PROOF: The proof is by induction on t. If $t = 0$, the theorem is obvious. Consider a graph G with n vertices, no K_p, and the maximum number of edges subject to those properties. Clearly G contains a K_{p-1} (otherwise adding an edge would not produce a K_p), say H. Each of the remaining vertices is joined to at most $p-2$ vertices of H. The remaining $n-p+1$ vertices do not contain a K_p as subgraph. Since $n-p+1 = (t-1)(p-1)+r$, we can apply the induction hypothesis to this set of points. So the number of edges of G is at most

$$M(n-p+1,p) + (n-p+1)(p-2) + \binom{p-1}{2}$$

and this number is equal to $M(n,p)$. □

Remark. The argument that was used for Mantel's theorem would show that if there is no K_p, then $|E| \leq \frac{p-2}{2(p-1)}n^2$.

Problem 4A. Let G be a simple graph with 10 vertices and 26 edges. Show that G has at least 5 triangles. Can equality occur?

Turán's paper on graph theory that contains Theorem 4.1 is considered the starting point of what is now known as extremal graph theory—see Bollobás (1978). A simple instance of an extremal problem will ask for the maximum number of edges a graph with a certain property may have. The graphs whose number of edges *is* maximum are called the *extremal graphs* with respect to the property.

The extremal graphs for Turán's problem are *only* the complete multipartite graphs described above. This follows from an analysis of the proof of Theorem 4.1; we ask the reader to do this at least in the case $p = 3$ in the problem below.

Problem 4B. Show that a simple graph on n vertices with $\lfloor n^2/4 \rfloor$ edges and no triangles is a complete bipartite graph $K_{k,k}$ if $n = 2k$, or $K_{k,k+1}$ if $n = 2k + 1$.

Problem 4C. If a simple graph on n vertices has e edges, then it has at least $\frac{e}{3n}(4e - n^2)$ triangles.

The *girth* of a graph G is the size of a smallest polygon P_n in G. (A forest has infinite girth.) By definition, a graph is simple if and only if it has girth ≥ 3. By Mantel's theorem, a graph with more than $n^2/4$ edges has girth ≤ 3.

Theorem 4.2. *If a graph G on n vertices has more than $\frac{1}{2}n\sqrt{n-1}$ edges, then G has girth ≤ 4. That is, G is not simple or contains a P_3 or a P_4 (a triangle or a quadrilateral).*

Proof: Suppose G has girth ≥ 5. Let $y_1, y_2, \ldots, y_d$ be the vertices adjacent to a vertex x, where $d := \deg(x)$. No two of these are adjacent since G has no triangles. Moreover, no vertex (other than x) can be adjacent to more than one of $y_1, \ldots, y_d$ since there are no quadrilaterals in G. Thus $(\deg(y_1)-1)+\cdots+(\deg(y_d)-1)+(d+1)$ cannot exceed the total number n of vertices. That is,

$$\sum_{y \text{ adjacent to } x} \deg(y) \leq n - 1.$$

Then

$$n(n-1) \geq \sum_x \sum_{y \text{ adjacent to } x} \deg(y) = \sum_y \deg(y)^2$$
$$\geq \frac{1}{n}\left(\sum_y \deg(y)\right)^2 = \frac{1}{n}(2|E(G)|)^2.$$

□

The number $\frac{1}{2}n\sqrt{n-1}$ in Theorem 4.2 is only a bound—it is not the exact answer for all n. Determination of the extremal graphs for this problem (maximum number of edges subject to girth ≥ 5) for all values of n is impossibly difficult; a determination of the graphs for which equality holds has, however, been almost possible.

Perhaps surprisingly, there are at most four graphs with $n > 2$ vertices, girth ≥ 5, and $\frac{1}{2}n\sqrt{n-1}$ edges: The pentagon ($n = 5$), the Petersen graph ($n = 10$), one with $n = 50$, and possibly one with $n = 3250$. See the notes and Chapter 21.

Problem 4D. Suppose G is regular of degree r and has girth g or greater. Find a lower bound for $|V(G)|$. (Consider the cases g even and odd separately.)

It is not so interesting to ask how many edges are required to force a Hamiltonian circuit. But we can ask what bound on the minimum degree will do the job.

Theorem 4.3. *If a simple graph G on n vertices has all vertices of degree at least $n/2$, then it contains a P_n as a subgraph, i.e. it has a Hamiltonian circuit.*

Proof: Suppose the theorem is not true and let G be a graph satisfying the hypothesis for some n but having no Hamiltonian circuits. We may take G to be such a counterexample with the maximum number of edges; then the addition of any edge to G (i.e. joining two nonadjacent vertices by an edge) creates a Hamiltonian circuit.

Let y and z be nonadjacent vertices. Since adding $\{y, z\}$ creates a Hamiltonian circuit, there exists a simple path from y to z with vertex terms, $y = x_1, x_2, \ldots, x_n = z$, say. The sets

$$\{i : y \text{ is adjacent to } x_{i+1}\}$$

and

$$\{i : z \text{ is adjacent to } x_i\}$$

each have cardinality $\geq n/2$ and are contained in $\{1, 2, 3, \ldots, n-1\}$, so they must meet; let i_0 belong to both. Then

$$y = x_1, x_2, \ldots, x_{i_0}, z = x_n, x_{n-1}, \ldots, x_{i_0+1}, x_1 = y$$

is the vertex sequence of a simple closed path of length n in G, contradicting our choice of G as a counterexample. □

Theorem 4.3 is due to G. A. Dirac and is best possible at least in the sense that it does not remain true if we replace $n/2$ by

$(n-1)/2$. For example, the complete bipartite graphs $K_{k,k+1}$ have no Hamiltonian circuits. But it does admit improvements and generalizations—see e.g. Lovász (1979), Problem 10.21.

Problem 4E. A $3\times3\times3$ cube of cheese is divided into 27 $1\times1\times1$ small cubes. A mouse eats one small cube each day and an *adjacent* small cube (sharing a face) the next day. Can the mouse eat the *center* small cube on the last day?

Problem 4F. Let G be a simple graph with n vertices. Prove: If each vertex of G has degree $\geq (n+1)/2$, then for any edge e, there exists a Hamiltonian circuit of G that passes through e.

Problem 4G. Prove the remark following Theorem 4.1.

Problem 4H. Show that a graph on n vertices that does not contain a circuit on four vertices has at most $\frac{n}{4}(1+\sqrt{4n-3})$ edges.

Notes.

P. Turán (1910–1976), one of the famous Hungarian mathematicians of the 20th century, is best known for his work in analytic number theory and real and complex analysis.

For every $r \geq 2$ and $g \geq 2$, there exists a graph that is regular of degree r and has girth $\geq g$. See Lovász (1979), Problem 10.12.

Analysis of the proof of Theorem 4.2 shows that a graph with $n > 2$ vertices, girth ≥ 5, and $\frac{1}{2}n\sqrt{n-1}$ edges is regular of degree $k := \sqrt{n-1}$ and also that any two vertices are joined by a (unique) path of length 2 if they are not adjacent. With the notation of Chapter 21, such a graph is an $srg(n,k,0,1)$ and the methods of that chapter show that $k = 2, 3, 7,$ or 57. This was first shown in Hoffman and Singleton (1960) where in addition an example with $k = 7$ and $n = 50$ was described (that is now known as the *Hoffman-Singleton graph*). It is not known at this time whether there exists an $srg(3250, 57, 0, 1)$.

References.

B. Bollobás (1978), *Extremal Graph Theory*, Academic Press.

A. J. Hoffman and R. R. Singleton (1960), On Moore graphs with diameters two and three, *IBM J. Res. Develop.* **4**, 497–504.

L. Lovász (1979), *Combinatorial Problems and Exercises*, North Holland.

W. Mantel (1907), Problem 28, *Wiskundige Opgaven* **10**, 60–61.

P. Turán (1941), An extremal problem in graph theory (in Hungarian), *Mat. Fiz. Lapok* **48**, 435–452.

5
Systems of distinct representatives

We first give two different formulations of a theorem known as P. Hall's *marriage theorem*. We give a constructive proof and an enumerative one. If A is a subset of the vertices of a graph, then denote by $\Gamma(A)$ the set $\bigcup_{a\in A}\Gamma(a)$. Consider a bipartite graph G with vertex set $X \cup Y$ (every edge has one endpoint in X and one in Y). A *matching* in G is a subset E_1 of the edge set such that no vertex is incident with more than one edge in E_1. A *complete* matching from X to Y is a matching such that every vertex in X is incident with an edge in E_1. If the vertices of X and Y are thought of as boys and girls, respectively, or vice versa, and an edge is present when the persons corresponding to its ends have amicable feelings towards one another, then a complete matching represents a possible assignment of marriage partners to the persons in X.

Theorem 5.1. *A necessary and sufficient condition for there to be a complete matching from X to Y in G is that $|\Gamma(A)| \geq |A|$ for every $A \subseteq X$.*

Proof: (i) It is obvious that the condition is necessary.

(ii) Assume that $|\Gamma(A)| \geq |A|$ for every $A \subseteq X$. Let $|X| = n$, $m < n$, and suppose we have a matching M with m edges. We shall show that a larger matching exists. (We mean larger in *cardinality*; we may not be able to find a complete matching containing these particular m edges.)

Call the edges of M red and all other edges blue. Let $x_0 \in X$ be a vertex not incident with an edge of the matching. We claim that there exists a simple path (of odd length) starting with x_0 and a blue edge, using red and blue edges alternately, and terminating

with a blue edge and a vertex y not incident with an edge of the matching. If we find such a path p, we are done because we obtain a matching with $m+1$ edges by deleting the red edges of p from M and replacing them with the blue edges of p. In other words, we switch the colors on the edges of p.

Since $|\Gamma(\{x_0\})| \geq 1$, there is a vertex y_1 adjacent to x_0 (obviously by a blue edge since x_0 is not incident with any red edges). If y_1 is also not incident with any red edges, we have the required path (of length one); if y_1 is incident with a red edge, let x_1 be the other end of that red edge. Recursively, define $x_0, x_1, \ldots$ and $y_1, y_2, \ldots$ as follows. If $x_0, x_1, \ldots, x_k$ and $y_1, \ldots, y_k$ have been defined, then since $|\Gamma(\{x_0, x_1, \ldots, x_k\})| \geq k+1$, there exists a vertex y_{k+1}, distinct from $y_1, \ldots, y_k$, that is adjacent to at least one vertex in $\{x_0, x_1, \ldots, x_k\}$. If y_{k+1} is not incident with a red edge, stop; otherwise, let x_{k+1} be the other end of that red edge.

When the procedure terminates, we construct the path p by starting with y_{k+1} and the blue edge joining it to, say, x_{i_1}, $i_1 < k+1$. Then add the red edge $\{x_{i_1}, y_{i_1}\}$. By construction, y_{i_1} is joined by an edge (necessarily blue) to some x_{i_2}, $i_2 < i_1$. Then add the red edge $\{x_{i_2}, y_{i_2}\}$. Continue in this way until x_0 is reached. □

Problem 5A. A *perfect* matching in a graph G (not necessarily bipartite) is a matching so that each vertex of G is incident with one edge of the matching. (i) Show that a finite *regular* bipartite graph (regular of degree $d > 0$) has a perfect matching. (ii) Find a trivalent (regular of degree 3) simple graph which does not have a perfect matching. (iii) Suppose G is bipartite with vertices $X \cup Y$ (every edge having one end in X and one in Y). Further assume that every vertex in X has the same degree $s > 0$ and every vertex in Y has the same degree t. (This condition is called *semiregularity.*) Prove: If $|X| \leq |Y|$ (equivalently, if $s \geq t$), then there is a complete matching M of X into Y.

Example 5.1. A parlor trick involving a standard deck of 52 cards is as follows. You are dealt five cards at random. You keep one and put the other four (in a specific order) into an envelope which is taken to your partner in another room. Your partner looks at these and announces the name of the fifth card, that you had retained.

Using the values of suits and ranks, it is possible to think of clever

or relatively simple ways to determine which card to keep and which to pass, and for the partner to determine the retained card; see the notes. Ignoring complexity of solution, however, the relevant general mathematical problem, stated for a deck of N cards, is does there exist an injective mapping f from the set X of the $\binom{N}{5}$ 5-element subsets of the N cards into the set Y of $N(N-1)(N-2)(N-3)$ ordered 4-tuples of distinct cards, subject to the condition that if $f(S) = (c_1, c_2, c_3, c_4)$, then $\{c_1, c_2, c_3, c_4\} \subseteq S$.

In terms of matchings, we consider the bipartite graph G whose vertices are $X \cup Y$ as defined above, and where there is an edge joining $S \in X$ to $\{c_1, c_2, c_3, c_4\} \in Y$ exactly when $\{c_1, c_2, c_3, c_4\} \subseteq S$. We require a complete matching M of X into Y. The reader may check that G is semiregular and that $|X| \leq |Y|$ if and only if $N \leq 124$. So by Problem 5A(iii), for $N \leq 124$, there exists such a matching.

We now reformulate Theorem 5.1 in terms of sets and not only prove the theorem but also give a lower bound for the number of matchings. We consider subsets $A_0, A_1, \ldots, A_{n-1}$ of a finite set S. We shall say that this collection has property H (Hall's condition) if (for all k) the union of any k-tuple of subsets A_i has at least k elements. If the union of some k-tuple of subsets contains exactly k elements ($0 < k < n$), then we call this k-tuple a *critical block.*

We define a *system of distinct representatives* (SDR) of the sets $A_0, \ldots, A_{n-1}$ to be a sequence of n *distinct* elements $a_0, \ldots, a_{n-1}$ with $a_i \in A_i$, $0 \leq i \leq n-1$.

Let $m_0 \leq m_1 \leq \cdots \leq m_{n-1}$. We define

$$F_n(m_0, m_1, \ldots, m_{n-1}) := \prod_{i=0}^{n-1} (m_i - i)_*,$$

where $(a)_* := \max\{1, a\}$.

From now on, we assume that the sequence $m_i := |A_i|$ is nondecreasing.

For the proof of the main theorem, we need a lemma.

Lemma 5.2. *For $n \geq 1$, let $f_n : \mathbb{Z}^n \to \mathbb{N}$ be defined by*

$$f_n(a_0, a_1, \ldots, a_{n-1}) := F_n(m_0, m_1, \ldots, m_{n-1})$$

if $(m_0, \ldots, m_{n-1})$ *is a nondecreasing rearrangement of the* n*-tuple* $(a_0, \ldots, a_{n-1})$*. Then* f_n *is nondecreasing with respect to each of the variables* a_i*.*

PROOF: Let

$$m_0 \leq \cdots \leq m_{k-1} \leq a_i = m_k \leq m_{k+1} \leq \cdots \leq m_l \leq m_{l+1} \leq \cdots \leq m_{n-1}$$

be a nondecreasing rearrangement of $(a_0, \ldots, a_{n-1})$. If $a_i' \geq a_i$ and

$$m_0 \leq \cdots \leq m_{k-1} \leq m_{k+1} \leq \cdots \leq m_l \leq a_i' \leq m_{l+1} \leq \cdots \leq m_{n-1}$$

is a nondecreasing rearrangement of $(a_0, \ldots, a_{i-1}, a_i', a_{i+1}, \ldots, a_{n-1})$ then

$$\frac{f_n(a_0, \ldots, a_{i-1}, a_i', a_{i+1}, \ldots, a_{n-1})}{f_n(a_0, \ldots, a_{n-1})} =$$

$$= \frac{(m_{k+1} - k)_*}{(a_i - k)_*} \cdot \frac{(a_i' - l)_*}{(m_l - l)_*} \prod_{j=k+1}^{l-1} \frac{(m_{j+1} - j)_*}{(m_j - j)_*}$$

and this is ≥ 1 since $a_i \leq m_{k+1}$, $a_i' \geq m_l$, and $m_{j+1} \geq m_j$ for $j = k+1, \ldots, l-1$. □

We now come to the second form of Hall's theorem. We denote by $N(A_0, \ldots, A_{n-1})$ the number of SDRs of $(A_0, \ldots, A_{n-1})$.

Theorem 5.3. *Let* $(A_0, \ldots, A_{n-1})$ *be a sequence of subsets of a set* S*. Let* $m_i := |A_i|$ $(i = 0, \ldots, n-1)$ *and let* $m_0 \leq m_1 \leq \cdots \leq m_{n-1}$*. If the sequence has property* H*, then*

$$N(A_0, \ldots, A_{n-1}) \geq F_n(m_0, \ldots, m_{n-1}).$$

PROOF: The proof is by induction. Clearly the theorem is true for $n = 1$. We distinguish two cases.

Case 1. There is no critical block. In this case, we choose any element a of A_0 as its representative and then remove a from all the other sets. This yields sets, that we call $A_1(a), \ldots, A_{n-1}(a)$, and

for these sets property H still holds. By the induction hypothesis and by the lemma, we find

$$\begin{aligned} N(A_0,\dots,A_{n-1}) &\geq \sum_{a\in A_0} f_{n-1}(|A_1(a)|,\dots,|A_{n-1}(a)|) \\ &\geq \sum_{a\in A_0} f_{n-1}(m_1-1,\dots,m_{n-1}-1) \\ &= m_0 f_{n-1}(m_1-1,\dots,m_{n-1}-1) \\ &= F_n(m_0,m_1,\dots,m_{n-1}). \end{aligned}$$

Case 2. There is a critical block $(A_{\nu_0},\dots,A_{\nu_{k-1}})$ with $\nu_0<\dots<\nu_{k-1}$ and $0<k<n$. In this case, we delete all elements of $A_{\nu_o}\cup\dots\cup A_{\nu_{k-1}}$ from all the other sets A_i which produces $A'_{\mu_0},\dots,A'_{\mu_{l-1}}$, where $\{\nu_0,\dots,\nu_{k-1},\mu_0,\dots,\mu_{l-1}\}=\{0,1,\dots,n-1\}, \quad k+l=n.$

Now both $(A_{\nu_0},\dots,A_{\nu_{k-1}})$ and $(A'_{\mu_0},\dots,A'_{\mu_{l-1}})$ satisfy property H and SDRs of the two sequences are always disjoint. Hence by the induction hypothesis and the lemma, we have

(5.1)

$$\begin{aligned} N(A_0,\dots,A_{n-1}) &= N(A_{\nu_0},\dots,A_{\nu_{k-1}})N(A'_{\mu_0},\dots,A'_{\mu_{l-1}}) \\ &\geq f_k(m_{\nu_0},\dots,m_{\nu_{k-1}})f_l(|A'_{\mu_0}|,\dots,|A'_{\mu_{l-1}}|) \\ &\geq f_k(m_{\nu_0},\dots,m_{\nu_{k-1}})f_l(m_{\mu_0}-k,\dots,m_{\mu_{l-1}}-k) \\ &\geq f_k(m_0,\dots,m_{k-1})f_l(m_{\mu_0}-k,\dots,m_{\mu_{l-1}}-k). \end{aligned}$$

Now we remark that

$$m_{\nu_{k-1}} \leq |A_{\nu_0}\cup\dots\cup A_{\nu_{k-1}}| = k,$$

and therefore we have

$$(m_r-r)_* = 1 \qquad \text{if } k\leq r\leq \nu_{k-1},$$

and

$$(m_{\mu_i}-k-i)_* = 1 \qquad \text{if } \mu_i\leq \nu_{k-1}.$$

This implies that

$$f_k(m_0,\dots,m_{k-1}) = \prod_{0\leq i\leq \nu_{k-1}} (m_i-i)_*,$$

$$f_l(m_{\mu_0}-k,\dots,m_{\mu_{l-1}}-k) = \prod_{\nu_{k-1}<j<n} (m_j-j)_*,$$

i.e. the product (5.1) is equal to $F_n(m_0, \ldots, m_{n-1})$, which proves the theorem. □

Problem 5B. Show that Theorem 5.3 gives the best lower bound for the number of SDRs of the sets A_i that only involves the numbers $|A_i|$.

We now come to a theorem known as König's theorem. It is equivalent (whatever that means) to Hall's theorem. In the theorem, A is a (0,1)-matrix with entries a_{ij}. By a *line*, we mean a row or a column of A.

Theorem 5.4. *The minimum number of lines of A that contain all the 1's of A is equal to the maximum number of 1's in A, no two on a line.*

PROOF: Let m be the minimum number of lines of A containing all the 1's of A and let M be the maximum number of 1's, no two on a line. Clearly $m \geq M$. Let the minimum covering by lines consist of r rows and s columns ($r + s = m$). Without loss of generality, these are the first r rows and the first s columns. We now define sets A_i, $1 \leq i \leq r$, by $A_i := \{j > s : a_{ij} = 1\}$. If some k-tuple of the A_i's contained less than k elements, then we could replace the corresponding k rows by $k - 1$ columns, still covering all the 1's. Since this is impossible, we see that the A_i's satisfy property H. So the A_i's have an SDR. This means that there are r 1's, no two on a line, in the first r rows and not in the first s columns. By the same argument there are s 1's, no two on a line, in the first s columns and not in the first r rows. This shows that $M \geq r + s = m$ and we are done. □

The following theorem of G. Birkhoff is an application of Hall's theorem.

Theorem 5.5. *Let $A = (a_{ij})$ be an $n \times n$ matrix with nonnegative integers as entries, such that every row and column of A has sum l. Then A is the sum of l permutation matrices.*

PROOF: Define A_i, $1 \leq i \leq n$, by $A_i := \{j : a_{ij} > 0\}$. For any k-tuple of the A_i's, the sum of the corresponding rows of A is kl. Since every column of A has sum l, the nonzero entries in the

chosen k rows must be in at least k columns. Hence the A_i's satisfy property H. An SDR of the A_i's corresponds to a permutation matrix $P = (p_{ij})$ such that $a_{ij} > 0$ if $p_{ij} = 1$. The theorem now follows by induction on l. □

Problem 5C. In the hypothesis of Theorem 5.5, we replace 'integers' by 'reals'. Show that in this case, A is a nonnegative linear combination of permutation matrices. (Equivalently, every doubly stochastic matrix—see Chapter 11—is a *convex* combination of permutation matrices.)

Problem 5D. Let S be the set $\{1, 2, \ldots, mn\}$. We partition S into m sets $A_1, \ldots, A_m$ of size n. Let a second partitioning into m sets of size n be $B_1, \ldots, B_m$. Show that the sets A_i can be renumbered in such a way that $A_i \cap B_i \neq \emptyset$.

Problem 5E. Let $A_i = \{i-1, i, i+1\} \cap \{1, 2, \ldots, n\}$, $i = 1, 2, \ldots, n$. Let S_n denote the number of SDR's of the collection $\{A_1, \ldots, A_n\}$. Determine S_n and $\lim_{n\to\infty} S_n^{1/n}$.

Let G be a bipartite graph, finite or infinite; say the vertex set is partitioned into sets X, Y of vertices so that every edge of G has one end in X and one end in Y. We say that a matching M in G *covers* a subset S of the vertices when every vertex in S is incident with one of the edges in M.

Theorem 5.6. *If there exists a matching M_1 that covers a subset X_0 of X and there exists a matching M_2 that covers a subset Y_0 of Y, then there exists a matching M_3 that covers $X_0 \cup Y_0$.*

Proof: Think of the edges of M_1 as 'red edges' and the edges of M_2 as 'blue edges'. If an edge belongs to both M_1 and M_2, it is 'purple'.

A connected graph all of whose vertices have degree at most two is easily seen to be one of: a finite path-graph (allowing the trivial case of length 0, when the component has one vertex and no edges), a finite polygon, an infinte 'one-sided' path-graph (with one monovalent vertex), or an infinite 'two-sided' path-graph. The graph H whose vertices are those of G and whose edges $M_1 \cup M_2$ has the property that every vertex has degree at most two, so its connected components are of the types enumerated above. The

edges of any of these components, other than the graphs consisting of a purple edge and its two ends, are alternately colored red and blue; in particular, all polygons have even length. Every vertex of $X_0 \cup Y_0$ is in one of these nontrivial components.

For the matching M_3, we will take all purple edges, and either all red edges or all blue edges from every other component of H. From the cycles and infinite two-sided paths, it doesn't matter; take all red or all blue edges, and they will cover all vertices of the component. From the paths of odd length and infinite one-sided paths, take all red or blue edges depending on whether the first edge is red or blue (for a path of odd length, the first and last edge have the same color, so it doesn't matter what side you start on). Again, the selected edges will cover all vertices of the component.

We have to think just a tiny bit harder for a component that is a finite path P of even length, with vertices $v_0, v_1, \ldots, v_k$, say. There is an odd number of vertices and they alternate between X and Y, so v_0 and v_k are both in X or both in Y. If they are both in X, take all red edges of P (those of M_1) and put them in M_3; if they are both in Y, take all blue edges of P and put them in M_3. Only one end of the path is not covered by the chosen edges.

Consider the case that both $v_0, v_k \in X$ (the case when they are in Y is completely analogous). If the first edge of P is red, then the last is blue and it follows that $v_k \notin X_0$ since no edge of M_1 covers v_k. Thus the red edges of P still cover all vertices of X_0 and Y_0 that were in P. Similarly, if the first edge of P is blue, then $v_0 \notin X_0$ and the red edges of P still cover all vertices of X_0 and Y_0 that were in P. □

For the case $X_0 = X$, $Y_0 = Y$, and when the graph G is complete bipartite, matchings that cover X_0 or Y_0 correspond to, or can be interpreted as, injective mappings $X \to Y$ or $Y \to X$, respectively. Theorem 5.6 says:

Corollary. *If X and Y are sets and there exist injective mappings $f : X \to Y$ and $g : Y \to X$, then there exists a bijective mapping from X to Y, or from Y to X, i.e. there is a one-to-one correspondence between the two sets.*

In terms of 'cardinality' of sets, this says that if $|X| \leq |Y|$ and

$|Y| \leq |X|$, then $|X| = |Y|$. This is the Schröder-Bernstein Theorem; see Section 22 of P. R. Halmos (1974). It is trivial for finite sets, of course.

Problem 5F. Let $A_1, A_2, \ldots, A_n$ be finite sets. Show that if

$$\sum_{1 \leq i < j \leq n} \frac{|A_i \cap A_j|}{|A_i| \cdot |A_j|} < 1,$$

then the sets $A_1, A_2, \ldots A_n$ have a system of distinct representatives.

Problem 5G. (i) From Problem 5A we know that a bipartite graph on $2n$ vertices that is regular of degree 3 has a perfect matching. How many different perfect matchings are there if $n = 4$? (ii) The same question for a bipartite graph on 10 vertices that is regular of degree 4.

Notes.

Philip Hall published his result in 1935 (with a rather difficult proof). The proof that we gave is a generalization of ideas of Halmos and Vaughan, Rado, and M. Hall. The proof is due to Ostrand (1970) and Hautus and Van Lint (1972). See Van Lint (1974). The problem of complete matchings is often referred to as the *marriage problem.*

D. König (1884–1944) was professor at Budapest. He wrote the first comprehensive treatise on graph theory (Theorie der endlichen und unendlichen Graphen, 1936). König (1916) contains the first proof of (one of the theorems called) König's theorem.

Just before Theorem 5.4, we referred to the 'equivalence'of these theorems. This expression is often used when each of two theorems is more or less an immediate consequence of the other.

The theorem by Birkhoff (1946), i.e. Theorem 5.5, is extremely useful and will be applied a number of times in later chapters.

For the card problem mentioned in Example 5.1, here is one solution. In a set of five cards, some suit must be represented at least twice. The first card you pass to your partner should be the 'larger' of two cards of the same suit, and you will retain the 'smaller', where we think of the ranks $2, 3, \ldots, 10, J, Q, K, A$ as arranged clockwise on a circle (modulo 13), and by 'smaller' we mean

the card from which we must travel the least distance clockwise to get to the other. For example, if $S = \{3\spadesuit, Q\diamondsuit, 6\clubsuit, 3\diamondsuit, 7\spadesuit\}$, you pass either the $7\spadesuit$ or $3\diamondsuit$ to your partner. This already tells your partner the suit of the retained card and limits that card to six possibilities in that suit. To determine how far to count back (counterclockwise) from the rank of the first card, you and your partner use the *order* of the remaining three cards (the 52 cards are ordered lexicographically), and agree on some correspondence between the six permutations of three objects and the integers from 1 to 6.

References.

G. Birkhoff (1946), Tres observaciones sobre el algebra lineal, *Univ. Nac. Tucumán, Rev.* Ser. A, **5**, 147–151.

P. Hall (1935), On representatives of subsets, *J. London Math. Soc.* **10**, 26–30.

P. R. Halmos (1974), Naive Set Theory, Springer-Verlag.

D. König (1916), Über Graphen und ihre Anwendung auf Determinantentheorie und Mengenlehre, *Math. Annalen* **77**, 453–465.

J. H. van Lint (1974), *Combinatorial Theory Seminar Eindhoven University of Technology*, Lecture Notes in Mathematics **382**, Springer-Verlag.

P. Ostrand (1970), Systems of distinct representatives, *J. of Math. Analysis and Applic.* **32**, 1–4.

6
Dilworth's theorem and extremal set theory

A *partially ordered set* (also *poset*) is a set S with a binary relation $\leq$ (sometimes $\subseteq$ is used) such that:
(i) $a \leq a$ for all $a \in S$ (reflexivity),
(ii) if $a \leq b$ and $b \leq c$ then $a \leq c$ (transitivity),
(iii) if $a \leq b$ and $b \leq a$ then $a = b$ (antisymmetry).

If for any a and b in S, either $a \leq b$ or $b \leq a$, then the partial order is called a *total* order, or a *linear* order. If $a \leq b$ and $a \neq b$, then we also write $a < b$. Examples of posets include the integers with the usual order or the subsets of a set, ordered by inclusion. If a subset of S is totally ordered, it is called a *chain*. An *antichain* is a set of elements that are pairwise incomparable.

The following theorem is due to R. Dilworth (1950). This proof is due to H. Tverberg (1967).

Theorem 6.1. *Let P be a partially ordered finite set. The minimum number m of disjoint chains which together contain all elements of P is equal to the maximum number M of elements in an antichain of P.*

Proof: (i) It is trivial that $m \geq M$.

(ii) We use induction on $|P|$. If $|P| = 0$, there is nothing to prove. Let C be a maximal chain in P. If every antichain in $P \backslash C$ contains at most $M - 1$ elements, we are done. So assume that $\{a_1, \ldots, a_M\}$ is an antichain in $P \backslash C$. Now define $S^- := \{x \in P : \exists_i [x \leq a_i]\}$, and define S^+ analogously. Since C is a maximal chain, the largest element in C is not in S^- and hence by the induction hypothesis, the theorem holds for S^-. Hence S^- is the union of M disjoint chains $S_1^-, \ldots, S_M^-$, where $a_i \in S_i^-$. Suppose $x \in S_i^-$ and

$x > a_i$. Since there is a j with $x \le a_j$, we would have $a_i < a_j$, a contradiction. This shows that a_i is the maximal element of the chain S_i^-, $i = 1, \ldots, m$. We do the same for S^+. By combining the chains the theorem follows. □

A 'dual' to Dilworth's theorem was given by Mirsky (1971).

Theorem 6.2. *Let P be a partially ordered set. If P possesses no chain of $m+1$ elements, then P is the union of m antichains.*

PROOF: For $m = 1$ the theorem is trivial. Let $m \ge 2$ and assume that the theorem is true for $m-1$. Let P be a partially ordered set that has no chain of $m+1$ elements. Let M be the set of maximal elements of P. M is an antichain. Suppose $x_1 < x_2 < \cdots < x_m$ were a chain in $P \backslash M$. Then this would also be a maximal chain in P and hence we would have $x_m \in M$, a contradiction. Hence $P \backslash M$ has no chain of m elements. By the induction hypothesis, $P \backslash M$ is the union of $m-1$ antichains. This proves the theorem. □

The following famous theorem due to Sperner (1928) is of a similar nature. This proof is due to Lubell (1966).

Theorem 6.3. *If $A_1, A_2, \ldots, A_m$ are subsets of $N := \{1, 2, \ldots, n\}$ such that A_i is not a subset of A_j if $i \ne j$, then $m \le \binom{n}{\lfloor n/2 \rfloor}$.*

PROOF: Consider the poset of subsets of N. $\mathcal{A} := \{A_1, \ldots, A_m\}$ is an antichain in this poset.

A maximal chain $\mathcal{C}$ in this poset will consist of one subset of each cardinality $0, 1, \ldots, n$, and is obtained by starting with the empty set, then any singleton set (n choices), then any 2-subset containing the singleton ($n-1$ choices), then any 3-subset containing the 2-subset ($n-2$ choices), etc. Thus there are $n!$ maximal chains. Similarly, there are exactly $k!(n-k)!$ maximal chains which contain a given k-subset A of N.

Now count the number of ordered pairs $(A, \mathcal{C})$ such that $A \in \mathcal{A}$, $\mathcal{C}$ is a maximal chain, and $A \in \mathcal{C}$. Since each maximal chain $\mathcal{C}$ contains at most one member of an antichain, this number is at most $n!$. If we let α_k denote the number of sets $A \in \mathcal{A}$ with $|A| = k$, then this number is $\sum_{k=0}^{n} \alpha_k k!(n-k)!$. Thus

$$\sum_{k=0}^{n} \alpha_k k!(n-k)! \le n!, \quad \text{or equivalently,} \quad \sum_{k=0}^{n} \frac{\alpha_k}{\binom{n}{k}} \le 1.$$

Since $\binom{n}{k}$ is maximal for $k = \lfloor n/2 \rfloor$ and $\sum \alpha_k = m$, the result follows. □

Equality holds in Theorem 6.3 if we take all $\lfloor n/2 \rfloor$-subsets of N as the antichain.

We now consider the poset B_n (with 2^n elements) of the subsets of the n-set N, ordered by inclusion. The set of i-subsets of N is denoted by $\mathcal{A}_i$. We define a *symmetric chain* in B_n to be a sequence $P_k, P_{k+1}, \ldots, P_{n-k}$ of vertices such that $P_i \in \mathcal{A}_i$ and $P_i \subseteq P_{i+1}$ for $i = k, k+1, \ldots, n-k-1$. We describe an algorithm due to De Bruijn, Van Ebbenhorst Tengbergen and Kruyswijk (1949), that splits B_n into (disjoint) symmetric chains.

Algorithm: Start with B_1. Proceed by induction. If B_n has been split into symmetric chains, then for each such symmetric chain $P_k, \ldots, P_{n-k}$ define two symmetric chains in B_{n+1}, namely $P_{k+1}, \ldots, P_{n-k}$ and $P_k, P_k \cup \{n+1\}, P_{k+1} \cup \{n+1\}, \ldots, P_{n-k} \cup \{n+1\}$.

It is easy to see that this algorithm does what we claim. Furthermore it provides a natural matching between k-subsets and $(n-k)$-subsets in B_n (cf. Theorem 5.1). Also, see Problem 6D below.

Problem 6A. Let $a_1, a_2, \ldots, a_{n^2+1}$ be a permutation of the integers $1, 2, \ldots, n^2+1$. Show that Dilworth's theorem implies that the sequence has a subsequence of length $n+1$ that is monotone.

A nice direct proof of the assertion of Problem 6A is as follows. Suppose there is no increasing subsequence of $n+1$ terms. Define b_i to be the length of the longest increasing subsequence that starts with the term a_i. Then by the pigeonhole principle, there are at least $n+1$ terms in the b_i-sequence that have the same value. Since $i < j$ and $b_i = b_j$ imply that $a_i > a_j$, we have a decreasing subsequence of $n+1$ terms.

To show a connection between Chapters 5 and 6, we now prove that Theorem 5.1 immediately follows from Theorem 6.1. We consider the bipartite graph G of Theorem 5.1. Let $|X| = n$, $|Y| = n' \geq n$. We introduce a partial order by defining $x_i < y_j$ if and only if there is an edge from vertex x_i to vertex y_j. Suppose that the largest antichain contains s elements. Let this antichain be $\{x_1, \ldots, x_h, y_1, \ldots, y_k\}$, where $h + k = s$. Since $\Gamma(\{x_1, \ldots, x_h\}) \subseteq$

$Y\backslash\{y_1,\ldots,y_k\}$, we have $h \le n'-k$. Hence $s \le n'$. The partially ordered set is the union of s disjoint chains. This will consist of a matching of size a, the remaining $n-a$ elements of X, and the remaining $n'-a$ elements of Y. Therefore $n+n'-a = s \le n'$, i.e. $a \ge n$, which means that we have a complete matching.

Theorem 6.3 is a (fairly easy) example of an area known as *extremal set theory* in which the problems are often quite difficult. We first give one more example as an easy exercise.

Problem 6B. Let the sets A_i, $1 \le i \le k$, be distinct subsets of $\{1,2,\ldots,n\}$. Suppose $A_i \cap A_j \ne \emptyset$ for all i and j. Show that $k \le 2^{n-1}$ and give an example where equality holds.

We now give one more example of the method that we used to prove Sperner's theorem. We prove the so-called Erdős–Ko–Rado theorem (1961).

Theorem 6.4. *Let $\mathcal{A} = \{A_1,\ldots,A_m\}$ be a collection of m distinct k-subsets of $\{1,2,\ldots,n\}$, where $k \le n/2$, with the property that any two of the subsets have a nonempty intersection. Then $m \le \binom{n-1}{k-1}$.*

Proof: Place the integers 1 to n on a circle and consider the family $\mathcal{F} := \{F_1,\ldots,F_n\}$ of all consecutive k-tuples on the circle, i.e. F_i denotes $\{i, i+1,\ldots,i+k-1\}$ where the integers should be taken mod n. We observe that $|\mathcal{A}\cap\mathcal{F}| \le k$ because if some F_i equals A_j, then at most one of the sets $\{l, l+1,\ldots,l+k-1\}$, $\{l-k,\ldots,l-1\}$ $(i < l < i+k)$ is in $\mathcal{A}$. The same assertion holds for the collection $\mathcal{F}^\pi$ obtained from $\mathcal{F}$ by applying a permutation π to $\{1,\ldots,n\}$. Therefore

$$\Sigma := \sum_{\pi\in S_n} |\mathcal{A}\cap\mathcal{F}^\pi| \le k\cdot n!.$$

We now count this sum by fixing $A_j \in \mathcal{A}$, $F_i \in \mathcal{F}$ and observing that there are $k!(n-k)!$ permutations π such that $F_i^\pi = A_j$. Hence $\Sigma = m\cdot n\cdot k!(n-k)!$. This proves the theorem. □

By a slight modification of the proof, one can show that the theorem also holds if the sets in $\mathcal{A}$ are assumed to have size at most k and they form an antichain. However we shall give a proof using Theorem 5.1.

Theorem 6.5. *Let $\mathcal{A} = \{A_1, \ldots, A_m\}$ be a collection of m subsets of $N := \{1, 2, \ldots, n\}$ such that $A_i \not\subseteq A_j$ and $A_i \cap A_j \neq \emptyset$ if $i \neq j$ and $|A_i| \leq k \leq n/2$ for all i. Then $m \leq \binom{n-1}{k-1}$.*

PROOF: (i) If all the subsets have size k, then we are done by Theorem 6.4.

(ii) Let $A_1, \ldots, A_s$ be the subsets with the smallest cardinality, say $l \leq \frac{n}{2} - 1$. Consider all the $(l+1)$-subsets B_j of N that contain one or more of the sets A_i, $1 \leq i \leq s$. Clearly none of these is in $\mathcal{A}$. Each of the sets A_i, $1 \leq i \leq s$, is in exactly $n - l$ of the B_j's and each B_j contains at most $l + 1 \leq n - l$ of the A_i's. So by Theorem 5.1, we can pick s distinct sets, say $B_1, \ldots, B_s$, such that $A_i \subseteq B_i$. If we replace $A_1, \ldots, A_s$ by $B_1, \ldots, B_s$, then the new collection $\mathcal{A}'$ satisfies the conditions of the theorem and the subsets of smallest cardinality now all have size $> l$. By induction, we can reduce to case (i). □

By replacing the counting argument of the proof of Theorem 6.4 by an argument in which the subsets are counted with weights, we can prove the following generalization due to B. Bollobás (1973).

Theorem 6.6. *Let $\mathcal{A} = \{A_1, \ldots, A_m\}$ be a collection of m distinct subsets of $\{1, 2, \ldots, n\}$, where $|A_i| \leq n/2$ for $i = 1, \ldots, m$, with the property that any two of the subsets have a nonempty intersection. Then*

$$\sum_{i=1}^{m} \frac{1}{\binom{n-1}{|A_i|-1}} \leq 1.$$

PROOF: Let π be a permutation of $1, 2, \ldots, n$ placed on a circle and let us say that $A_i \in \pi$ if the elements of A_i occur consecutively somewhere on that circle. By the same argument as in the proof of Theorem 6.4 we see that if $A_i \in \pi$, then $A_j \in \pi$ for at most $|A_i|$ values of j.

Now define

$$f(\pi, i) := \begin{cases} \frac{1}{|A_i|}, & \text{if } A_i \in \pi \\ 0, & \text{otherwise.} \end{cases}$$

By the argument above $\sum_{\pi \in S_n} \sum_{i=1}^{m} f(\pi, i) \leq n!$. Changing the order of summation we have to count, for a fixed A_i, the number of permutations π placed on a circle such that $A_i \in \pi$. This number

(by the same argument as in Theorem 6.4) is $n \cdot |A_i|!(n-|A_i|)!$. So we have

$$\sum_{i=1}^{m} \frac{1}{|A_i|} \cdot n \cdot |A_i|!(n-|A_i|)! \leq n!,$$

which yields the result. □

Problem 6C. Let $\mathcal{A} = \{A_1, \ldots, A_m\}$ be a collection of m distinct subsets of $N := \{1, 2, \ldots, n\}$ such that if $i \neq j$ then $A_i \not\subseteq A_j$, $A_i \cap A_j \neq \emptyset$, $A_i \cup A_j \neq N$. Prove that

$$m \leq \binom{n-1}{\lfloor \frac{n}{2} \rfloor - 1}.$$

Problem 6D. Consider the decomposition of B_n into symmetric chains as described above. Show that Theorem 6.3 is an immediate consequence of this decomposition. Show that Theorem 6.5 reduces to Theorem 6.4 via this decomposition. How many of the chains have their smallest element in $\mathcal{A}_i$?

Problem 6E. Here is an algorithm to construct a symmetric chain in the poset B_n which contains a given element S (a subset of $\{1, 2, \ldots, n\}$). Consider the characteristic vector x of S; for example, if $n = 7$ and $S = \{3, 4, 7\}$, then $x = 0011001$. Mark all consecutive pairs 10, temporarily delete these pairs and again mark all consecutive pairs 10, and repeat until only a string of the form $00\cdots01\cdots11$ remains. In our example, we obtain $00\dot{1}\dot{1}\dot{0}\dot{0}1$, where the i-th coordinates are marked for $i = 3, 4, 5, 6$; when these are deleted, the string 001 remains. The characteristic vectors of the subsets in the chain are obtained by fixing all marked coordinates and letting the remaining coordinates range over the strings $0\cdots000$, $0\cdots001$, $0\cdots011$, $\ldots$, $1\cdots111$. In our example, these characteristic vectors are

$$00\dot{1}\dot{1}\dot{0}\dot{0}0,$$
$$00\dot{1}\dot{1}\dot{0}\dot{0}1,$$
$$01\dot{1}\dot{1}\dot{0}\dot{0}1,$$
$$11\dot{1}\dot{1}\dot{0}\dot{0}1,$$

which correspond to the subsets

$$\{3,4\}, \quad \{3,4,7\}, \quad \{2,3,4,7\}, \quad \{1,2,3,4,7\}.$$

Show that this algorithm produces exactly the same symmetric chain containing S as is produced by the inductive algorithm of De Bruijn *et al.* described above.

Notes.

We shall return to partially ordered sets in Chapters 23 and 25.

E. Sperner (1905–1980) is best known for a lemma in combinatorial topology known as 'Sperner's lemma', which occurred in his thesis (1928). It was used to give a proof of Brouwer's fixed point theorem. (Another connection to combinatorics: his first professorship was in Königsberg!) He was one of the pioneers of the famous Oberwolfach research institute.

For a survey of extremal set theory, we refer to Frankl (1988).

The short proof of the Erdős–Ko–Rado theorem is due to Katona (1974). Theorem 6.5 is due to Kleitman and Spencer (1973) and Schönheim (1971). The proof of Theorem 6.6 is due to Greene, Katona and Kleitman (1976).

References.

B. Bollobás (1973), Sperner systems consisting of pairs of complementary subsets, *J. Combinatorial Theory* (A) **15**, 363–366.

N. G. de Bruijn, C. van Ebbenhorst Tengbergen and D. Kruyswijk (1949), On the set of divisors of a number, *Nieuw Archief v. Wisk.* (2) **23**, 191–193.

R. P. Dilworth (1950), A decomposition theorem for partially ordered sets, *Annals of Math.* (2) **51**, 161–166.

P. Erdős, Chao Ko, and R. Rado (1961), Extremal problems among subsets of a set, *Quart. J. Math. Oxford* Ser. (2) **12**, 313–318.

P. Frankl (1988), Old and new problems on finite sets, Proc. Nineteenth S. E. Conf. on Combinatorics, Graph Th. and Computing, Baton Rouge, 1988.

C. Greene, G. Katona, and D. J. Kleitman (1976), Extensions of the Erdős–Ko–Rado theorem, *Stud. Appl. Math.* **55**, 1–8.

G. O. H. Katona (1974), Extremal problems for hypergraphs, in *Combinatorics* (edited by M. Hall, Jr. and J. H. van Lint), Reidel.

D. J. Kleitman and J. Spencer (1973), Families of k-independent sets, *Discrete Math.* **6**, 255–262.

D. Lubell (1966), A short proof of Sperner's lemma, *J. Combinatorial Theory* **1**, 299.

L. Mirsky (1971), A dual of Dilworth's decomposition theorem, *Amer. Math. Monthly* **78**, 876–877.

J. Schönheim (1971), A generalization of results of P. Erdős, G. Katona, and D. J. Kleitman concerning Sperner's theorem, *J. Combinatorial Theory* (A) **11**, 111–117.

E. Sperner (1928), Ein Satz über Untermengen einer endlichen Menge, *Math. Zeitschrift* **27**, 544–548.

H. Tverberg (1967), On Dilworth's decomposition theorem for partially ordered sets, *J. Combinatorial Theory* **3**, 305–306.

7
Flows in networks

By a *transportation network*, we will mean a finite directed graph D together with two distinguished vertices s and t called the *source* and the *sink*, respectively, and which is provided with a function c associating to each edge e a nonnegative real number $c(e)$ called its *capacity*. We may further assume that there are no loops, no multiple edges, and that no edges enter the source s or leave the sink t (although there would be no harm in admitting any of these types of edges other than our having to be more careful in a definition or two).

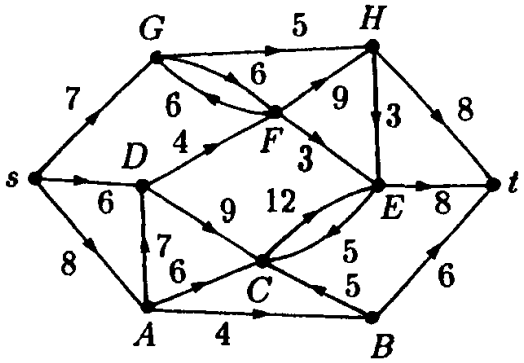


Figure 7.1

In Fig. 7.1 we give an example. We could think of a network of pipes through which some liquid could flow in the direction of the arrows. The capacity would indicate the maximal possible flow (per time unit) in that section of pipe.

A *flow* in a transportation network is a function f assigning a real number $f(e)$ to each edge e such that:

(a) $0 \leq f(e) \leq c(e)$ for all edges e (the flow is *feasible*);

(b) for each vertex x (not the source or the sink) the sum of the

values of f on incoming edges equals the sum of the values of f on outgoing edges (*conservation of flow*).

The sum of the values of a flow f on the edges leaving the source is called the *strength* of the flow (denoted by $|f|$). It seems obvious that the strength of the flow is also equal to the sum of the values of f on edges entering the sink; the reader might try to verify this formally before reading further.

One of our objectives will be to find a method for constructing a *maximum flow*, that is, a flow with maximum strength. Before we begin, it will be good to have a goal or an upper bound for the strength of a flow; for example, the sum of the capacities of all edges leaving the source is clearly such an upper bound. More generally, by a *cut separating s and t* (or simply a *cut*), we mean here a pair (X, Y) of subsets of the vertex set $V := V(D)$ which partition V and such that $s \in X$ and $t \in Y$. We define the *capacity* $c(X, Y)$ of the cut to be the sum of the capacities of the edges directed from X to Y (that is, edges $e = (x, y)$ with $x \in X$ and $y \in Y$). We claim that the capacity of any cut is an upper bound for the strength of any flow. More strongly, we claim that the conservation law implies (see below) that the strength of a flow f can be computed as

$$|f| = f(X, Y) - f(Y, X), \tag{7.1}$$

where $f(A, B)$ denotes the sum of the values of f on all edges directed from A to B; then the feasibility of f immediately implies that $|f| \le c(X, Y)$. Thus the minimum capacity of all cuts in a network (e.g. in Fig. 7.1 the minimum cut capacity is 20) is an upper bound for the strength of a flow in that network.

To establish (7.1), we introduce the function ϕ by defining for each pair (x, e), where x is a vertex incident with the edge e, $\phi(x, e) := -1$ if the edge is incoming, and $\phi(x, e) := +1$ if the edge is outgoing; $\phi(x, e)$ is to be 0 if x is not incident with e. (We remark that ϕ is essentially the *incidence matrix* of the directed graph—see Chapter 36.) The conservation law is equivalent to $\sum_{e \in E} \phi(x, e) f(e) = 0$ for $x \neq s, t$. Notice that $\sum_{x \in X} \phi(x, e)$ is $+1$ if e is directed from X to Y, -1 if e is directed from Y to X, and

0 if e has both endpoints in X or both in Y. Then

$$|f| = \sum_{e \in E} \phi(s,e) f(e) = \sum_{x \in X} \sum_{e \in E} \phi(x,e) f(e)$$
$$= \sum_{e \in E} f(e) \sum_{x \in X} \phi(x,e) = f(X,Y) - f(Y,X).$$

(In the first double sum above, the inner sum is 0 for all terms x other than s.)

A special instance of (7.1) is $|f| = f(V \backslash \{t\}, \{t\})$, the assertion which we invited the reader to reflect on earlier.

We now construct flows. Fix a flow f, possibly the 0-flow. We shall say that the sequence $x_0, x_1, \ldots, x_{k-1}, x_k$ of distinct vertices is a *special path* from x_0 to x_k if for each i, $1 \leq i \leq k$, either

(i) $e = (x_{i-1}, x_i)$ is an edge with $c(e) - f(e) > 0$, or
(ii) $e = (x_i, x_{i-1})$ is an edge with $f(e) > 0$.

Edges e with $f(e) = c(e)$ are said to be *saturated* and conditions (i) and (ii) can be stated in words as requiring that 'forward' edges of the path are unsaturated while 'backward' edges are positive—all with respect to a given flow f. Suppose there exists such a special path from s to t. Define α_i as $c(e) - f(e)$ in the first case and as $f(e)$ in the second case (picking one of the edges to use if both cases hold) and let α be the minimum of these positive numbers α_i. On each edge of type (i) *increase* the flow value by α, and on each edge of type (ii) *decrease* the flow by α. It is easy to check that the two conditions for a flow (feasibility and conservation of flow) are still satisfied. Clearly the new flow has strength $|f| + \alpha$.

This idea for obtaining a stronger flow becomes an algorithm when we iterate it (starting with the 0-flow) and incorporate a systematic procedure for searching for special paths from s to t with respect to the current flow. We make brief remarks concerning termination in the notes to this chapter. But what happens when we can go no further?

Suppose that *no special path from source to sink exists* with respect to some flow f_0. Let X_0 be the set of vertices x which can be reached from s by a special path, Y_0 the set of remaining vertices.

In this way we produce a cut. If $x \in X_0$, $y \in Y_0$ and $e = (x, y)$ is an edge, then e must be saturated or we could adjoin y to a special path from s to x to get a special path from s to y, contradicting the definitions of X_0 and Y_0. If, on the other hand, $e = (y, x)$ is an edge, then, for a similar reason, $f(e)$ must be 0. In view of (7.1), we have then

$$|f_0| = f_0(X_0, Y_0) - f_0(Y_0, X_0) = c(X_0, Y_0).$$

Now it is clear that not only can no stronger flow be obtained by our method of special paths, but that no stronger flows exist at all because $|f| \leq c(X_0, Y_0)$ for any flow f.

If f_0 is chosen to be a maximum flow (which exists by continuity reasons in case one is unsure of the termination of the algorithm), then surely no special paths from s to t exist. Note that the constructed cut (X_0, Y_0) is a minimum cut (i.e. a cut of minimum capacity), since $c(X, Y) \geq |f_0|$ for any cut (X, Y). Our observations have combined to prove the following famous theorem of Ford and Fulkerson (1956).

Theorem 7.1. *In a transportation network, the maximum value of $|f|$ over all flows f is equal to the minimum value of $c(X, Y)$ over all cuts (X, Y).*

This theorem is usually referred to as the 'maxflow-mincut' theorem. The procedure for increasing the strength of a flow that we used above shows somewhat more.

Theorem 7.2. *If all the capacities in a transportation network are integers, then there is a maximum strength flow f for which all values $f(e)$ are integers.*

PROOF: Start with the 0-flow. The argument above provides a way to increase the strength until a maximum flow is reached. At each step α is an integer, so the next flow is integer valued too. □

Problem 7A. Construct a maximum flow for the transportation network of Fig. 7.1.

Problem 7B. An *elementary flow* in a transportation network is a flow f which is obtained by assigning a constant positive value

α to the set of edges traversed by a simple (directed) path from s to t, and 0 to all other edges. Show that every flow is the sum of elementary flows and perhaps a flow of strength zero. (This means we can arrive at a maxflow by starting from the 0-flow and using only special paths with 'forward' edges.) Give an example of a network and a flow which is not maximum, but with respect to which there are no special paths using only 'forward' edges.

Problem 7C. Let (X_1, Y_1) and (X_2, Y_2) be minimum cuts (i.e. cuts of minimum capacity) in a transportation network. Show that $(X_1 \cup X_2, Y_1 \cap Y_2)$ is also a minimum cut. (This can be done either from first principles, or with an argument involving maximum flows.)

Problem 7D. Prove P. Hall's marriage theorem, Theorem 5.1, from Theorems 7.1 and 7.2.

It should be clear that the topic of this chapter is of great practical importance. Routing schemes for all kinds of products depend on algorithms that produce optimal flows through transportation networks. We do not go into the algorithmic aspect of this area. Instead, we shall show a beautiful application of Theorem 7.2 to a problem related to Birkhoff's theorem, Theorem 5.5. Before giving the theorem and its proof, we observe that several attempts were made to prove it by reducing it to Theorem 5.5 but with no success. The proof below is due to A. Schrijver. (If $b = v$ in Theorem 7.3, then we have the situation of Theorem 5.5.)

Theorem 7.3. *Let A be a $b \times v$ $(0,1)$-matrix with k ones per row and r ones per column (so $bk = vr$). Let α be a rational number, $0 < \alpha < 1$, such that $k' = \alpha k$ and $r' = \alpha r$ are integers. Then there is a $(0,1)$-matrix A' of size $b \times v$ with k' ones per row and r' ones per column such that entries a'_{ij} of A' are 1 only if the corresponding entries of A are 1, i.e. A' can be obtained from A by changing some ones into zeros.*

PROOF: We construct a transportation network with vertices s (the source), $x_1, \ldots, x_b$ (corresponding to the rows of A), $y_1, \ldots, y_v$ (corresponding to the columns of A), and t (the sink). Edges are (s, x_i) with capacity k, $1 \le i \le b$, (x_i, y_j) with capacity 1 if and only if $a_{ij} = 1$, and (y_j, t) with capacity r, $1 \le j \le v$. The definition

ensures that there is a maximum flow with all edges saturated. We now change the capacities of the edges from the source to k' and those of the edges to the sink to r'. Again, all the capacities are integers and clearly a maximum flow exists for which the flows $f((x_i, y_j))$ are equal to α. By Theorem 7.2 there is also a maximum flow f^* for which all the flows are integers, i.e. $f^*((x_i, y_j)) = 0$ or 1. From this flow, we immediately find the required matrix A'. □

The theorem above can be generalized in several ways with essentially the same proof idea, but see below for a slightly different approach.

For some combinatorial applications, it is convenient to use the following theorem, which does not require the introduction of capacities or the concept of strength. It can be derived from Theorem 7.2—see Ford and Fulkerson (1956)—but we choose to give a direct proof.

A *circulation* on a digraph D is a mapping f from $E(D)$ to the reals satisfying conservation of flow at every vertex. We do not require nonnegativity. Circulations may be identified with vectors in the null space of the incidence matrix of the digraph.

Theorem 7.4. *Let f be a circulation on a finite digraph D. Then there exists an integral circulation g such that for every edge e, $g(e)$ is equal to one of $\lfloor f(e) \rfloor$ or $\lceil f(e) \rceil$.*

We may say that the values of g are those of f 'rounded up or down'. Of course, if $f(e)$ is already an integer, then $g(e) = f(e)$.

Proof: Given a circulation f, consider a circulation g that satisfies

$$\lfloor f(e) \rfloor \leq g(e) \leq \lceil f(e) \rceil \tag{7.2}$$

and for which the number of edges e with $g(e)$ an integer is as large as possible subject to (7.2).

Let H be the spanning subgraph of D with edge set consisting of those edges of D for which $g(e)$ is not an integer, i.e. for which strict inequality holds both times in (7.2). Conservation of flow implies that no vertex can have degree 1 in H, so if g is not integral, then H contains a polygon.

Let P be a polygon in H and traverse P with a simple closed path; let A be the set of edges of P that are forward edges of the

path in D, and B the set of edges of P that are backward edges in this path. For any constant c, we obtain a new circulation g' by

$$g'(e) := \begin{cases} g(e) + c & \text{if } e \in A, \\ g(e) - c & \text{if } e \in B, \\ g(e) & \text{if } e \notin E(P). \end{cases}$$

If c is small, (7.2) will still hold with g replaced by g'. Now choose

$$c := \min \Big\{ \min_{e \in A} \Big(\lceil f(e) \rceil - g(e) \Big), \min_{e \in B} \Big(g(e) - \lfloor f(e) \rfloor \Big) \Big\}.$$

Then g' still satisfies (7.2), yet $g'(e)$ is an integer for at least one more edge (any edge for the which term in the expression above achieves the minimum). This would contradict the choice of g, were g not integral. □

Corollary. *Let f be an integral circulation on a finite digraph D and d any positive integer. Then f can be written as the sum $g_1 + g_2 + \cdots + g_d$ of integral circulations such that for each index j and each edge e,*

$$\lfloor f(e)/d \rfloor \le g_j(e) \le \lceil f(e)/d \rceil. \tag{7.3}$$

Proof: By induction on d. For $d = 1$, there is nothing to prove.

Given $d \ge 2$, apply Theorem 7.4 to f/d to find an integral circulation g_1 satisfying (7.3) for $j = 1$. Apply the induction hypothesis to find

$$f - g_1 = g_2 + g_3 + \cdots + g_d$$

where for each $j = 2, 3, \ldots, d$, g_j is an integral circulation satisfying

$$\lfloor (f(e) - g_1(e))/(d-1) \rfloor \le g_j(e) \le \lceil (f(e) - g_1(e))/(d-1) \rceil.$$

An easy exercise is that if a is an integer and b is either $\lfloor a/d \rfloor$ or $\lceil a/d \rceil$, then

$$\lfloor \frac{a}{d} \rfloor \le \lfloor \frac{a-b}{d-1} \rfloor \qquad \text{and} \qquad \lceil \frac{a-b}{d-1} \rceil \le \lceil \frac{a}{d} \rceil,$$

so that the above inequalities imply (7.3) for $j = 2, 3, \ldots, d$. □

From an $m \times n$ matrix A of real numbers a_{ij}, not necessarily nonnegative or integers, we obtain a circulation on a digraph with $m + n + 2$ vertices and $mn + m + n + 1$ edges. The digraph is similar to the one used in the proof of Theorem 7.3. There are vertices $x_1, \ldots, x_m$ corresponding to the rows, vertices $y_1, \ldots, y_n$ corresponding to the columns, and two others called s and t. There is an edge from x_i to y_j with circulation value a_{ij}, an edge from s to x_i with circulation value equal to the i-th row-sum r_i, an edge from y_j to t with circulation value equal to the j-th column-sum k_j ($i = 1, \ldots, m$, $j = 1, \ldots, n$), and an edge from t to s with circulation value equal to the sum of all entries of M. If we multiply this circulation f by any scalar α, apply Theorem 7.4 to αf, and reinterpret the resulting integral circulation as a matrix, we obtain part (i) of the following theorem. Part (ii) follows from the corollary.

Theorem 7.5. (i) *Given a matrix A and a real number α, there is an integral matrix B so that the entries of B, the row-sums of B, the column-sums of B, and the sum of all entries of B, are the corresponding values for αA rounded up or down.* (ii) *If A is an integral matrix and d any positive integer, then*

$$A = B_1 + B_2 + \cdots + B_d$$

where each B_i is an integral matrix whose entries, row-sums, column-sums, and sum of all entries, are those of $(1/d)A$, rounded up or down.

Problem 7E. Show that the following results are quick consequences of Theorem 7.5: (i) Problem 5A(iii); (ii) Theorem 5.5; (iii) Theorem 7.3; (iv) A finite graph all of whose vertices have even degree has a *balanced orientation*, where the in-degree and out-degree of each vertex are equal; (v) If a bipartite graph has minimum degree $\underline{d}$ and maximum degree $\overline{d}$, then its edges may be colored with $\overline{d}$ colors so that the colors that appear at every vertex are distinct, and with $\underline{d}$ colors so that all colors appear at each vertex.

Problem 7F. Show that the dimension of the vector space of all circulations on a connected digraph D is $|E(D)| - |V(D)| + 1$.

Notes.

The term *augmenting path* is often used instead of *special path.*

If the capacities of a transportation network are integers, the special path method for constructing maximum flows will terminate after finitely many iterations, since the strength increases by at least one each time. But Ford and Fulkerson (1962) give an example with irrational capacities where certain contrived choices of special paths lead to an infinite sequence of flows whose strengths converge—but only to one-fourth of the actual maximum flow strength! If one is careful to pick *shortest* special paths, however, then it can be shown that a maximum flow is reached after at most $O(n^3)$ iterations, where n is the number of vertices. See Edmonds and Karp (1972).

The problem of finding a maximum flow is an example of a linear programming problem and can be solved e.g. by the simplex algorithm. The network flow problem is special in that its matrix is totally unimodular, and this is one way of explaining why Theorem 7.2 holds. See the references below for more discussion of linear and integer programming. Graphical methods are usually faster than the simplex algorithm, and add insight.

Circulations on a digraph are called 1-*cycles* in algebraic topology. An analogue of Theorem 7.4 holds for vectors f in the null space of any totally unimodular matrix.

Theorems 7.1, 7.2, 7.4, and the algorithm have many further combinatorial applications, since certain combinatorial problems can be phrased in terms of transportation networks. For example, finding a maximum matching in a bipartite graph is equivalent to finding a maximum (integer valued) flow in a certain associated network—see the references—and thus a good algorithm exists to find a maximum matching. We give further applications of these theorems in Chapter 16 to an existence problem on (0,1)-matrices, and in Chapter 38 to a problem on partitions of sets.

References.

J. Edmonds and R. M. Karp (1972), Theoretical improvements in algorithm efficiency for network flow problems, *J. Assn. for Computing Machinery* **19**, 248–264.

L. R. Ford, Jr. and D. R. Fulkerson (1962), *Flows in Networks*, Princeton University Press.
T. C. Hu (1969), *Integer Programming and Network Flows*, Addison-Wesley.
V. Chvátal (1983), *Linear Programming*, W. H. Freeman.

8
De Bruijn sequences

The following problem has a practical origin: the so-called *rotating drum problem*. Consider a rotating drum as in Fig. 8.1.

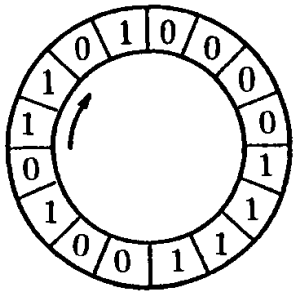


Figure 8.1

Each of the segments is of one of two types, denoted by 0 and 1. We require that any four consecutive segments uniquely determine the position of the drum. This means that the 16 possible quadruples of consecutive 0's and 1's on the drum should be the binary representations of the integers 0 to 15. Can this be done and, if yes, in how many different ways? The first question is easy to answer. Both questions were treated by N. G. de Bruijn (1946) and for this reason the graphs described below and the corresponding circular sequences of 0's and 1's are often called *De Bruijn graphs* and *De Bruijn sequences*, respectively.

We consider a digraph (later to be called G_4) by taking all 3-tuples of 0's and 1's (i.e. 3-bit binary words) as vertices and joining the vertex $x_1x_2x_3$ by a directed edge (arc) to x_2x_30 and x_2x_31. The arc $(x_1x_2x_3, x_2x_3x_4)$ is numbered e_j, where $x_1x_2x_3x_4$ is the binary representation of the integer j. The graph has a loop at 000 and at 111. As we saw before, the graph has an Eulerian circuit because every vertex has in-degree 2 and out-degree 2. Such a

closed path produces the required 16-bit sequence for the drum. Such a (circular) sequence is called a De Bruijn sequence. For example the path $000 \to 000 \to 001 \to 011 \to 111 \to 111 \to 110 \to 100 \to 001 \to 010 \to 101 \to 011 \to 110 \to 101 \to 010 \to 100 \to 000$ corresponds to 0000111100101101 (to be read circularly). We call such a path a *complete cycle*.

We define the graph G_n to be the directed graph on $(n-1)$-tuples of 0's and 1's in a similar way as above. (So G_n has 2^n edges.)

The graph G_4 is given in Fig. 8.2. In this chapter, we shall call a digraph with in-degree 2 and out-degree 2 for every vertex, a '2-in 2-out graph'. For such a graph G we define the 'doubled' graph G^* as follows:

(i) to each edge of G there corresponds a vertex of G^*;

(ii) if a and b are vertices of G^*, then there is an edge from a to b if and only if the edge of G corresponding to a has as terminal end (head) the initial end (tail) of the edge of G corresponding to b.

Clearly $G_n^* = G_{n+1}$.

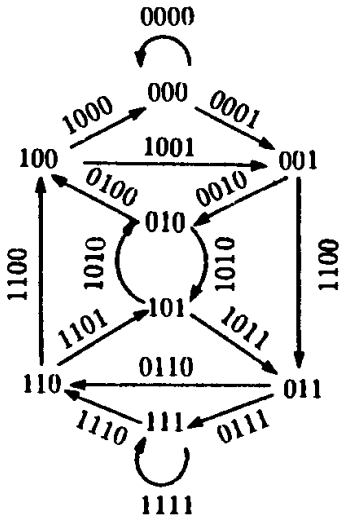


Figure 8.2

Theorem 8.1. *Let G be a 2-in 2-out graph on m vertices with M complete cycles. Then G^* has $2^{m-1}M$ complete cycles.*

PROOF: The proof is by induction on m.

(a) If $m = 1$ then G has one vertex p and two loops from p to p. Then $G^* = G_2$ which has one complete cycle.

(b) We may assume that G is connected. If G has m vertices and there is a loop at *every* vertex, then, besides these loops, G is a circuit $p_1 \to p_2 \to \cdots \to p_m \to p_1$. Let A_i be the loop $p_i \to p_i$ and B_i the arc $p_i \to p_{i+1}$. We shall always denote the corresponding vertices in G^* by lower case letters. The situation in G^* is as in Fig. 8.3.

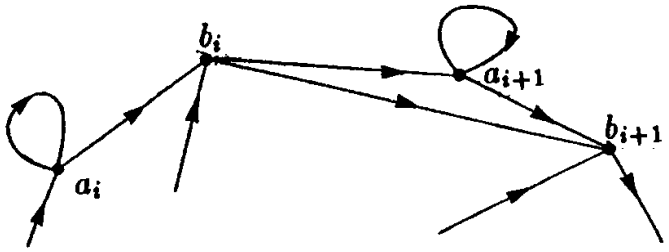


Figure 8.3

Clearly a cycle in G^* has two ways of going from b_i to b_{i+1}. So G^* has 2^{m-1} complete cycles, whereas G has only one.

(c) We now assume that G has a vertex x that does not have a loop on it. The situation is as in Fig. 8.4, where P, Q, R, S are different edges of G (although some of the vertices a, b, c, d may coincide).

From G we form a new 2-in 2-out graph with one vertex less by deleting the vertex x. This can be done in two ways: G_1 is obtained by the identification $P = R$, $Q = S$, and G_2 is obtained by $P = S$, $Q = R$. By the induction hypothesis, the theorem applies to G_1 and to G_2.

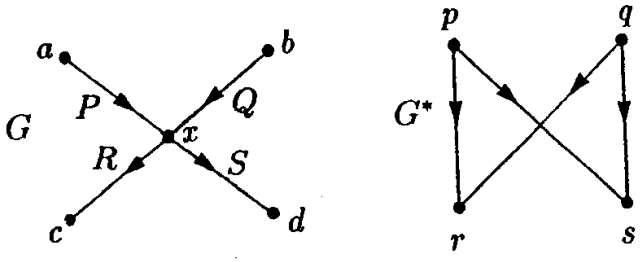


Figure 8.4

There are three different types of complete cycle in G^*, depending on whether the two paths leaving r and returning to p, respectively q, both go to p, both to q, or one to p and one to q. We treat one

case; the other two are similar and left to the reader. In Fig. 8.5 we show the situation where path 1 goes from r to p, path 2 from s to q, path 3 from s to p, and path 4 from r to q.

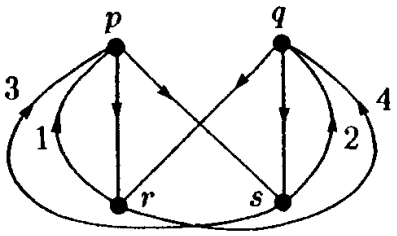


Figure 8.5

These yield the following four complete cycles in G^*:

$$\begin{array}{cccccccc} 1, & pr, & 4, & qs, & 3, & ps, & 2, & qr \\ 1, & ps, & 2, & qr, & 4, & qs, & 3, & pr \\ 1, & ps, & 3, & pr, & 4, & qs, & 2, & qr \\ 1, & ps, & 2, & qs, & 3, & pr, & 4, & qr \end{array}$$

In G_1^* and G_2^* the situation reduces to Fig. 8.6.

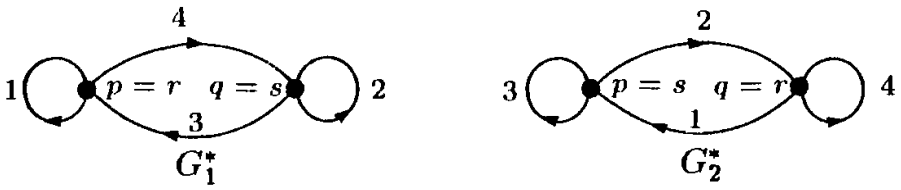


Figure 8.6

In each of G_1^* and G_2^* *one* complete cycle using the paths $1, 2, 3, 4$ is possible. In the remaining two cases, we also find two complete cycles in G_1^* and G_2^* corresponding to four complete cycles in G^*. Therefore the number of complete cycles in G^* is twice the sum of the numbers for G_1^* and G_2^*. On the other hand, the number of complete cycles in G is clearly equal to the sum of the corresponding numbers for G_1 and G_2. The theorem then follows from the induction hypothesis. □

We are now able to answer the question how many complete cycles there are in a De Bruijn graph.

Theorem 8.2. *G_n has exactly $2^{2^{n-1}-n}$ complete cycles.*

PROOF: The theorem is true for $n = 1$. Since $G_n^* = G_{n+1}$, the result follows by induction from Theorem 8.1. □

For a second proof, see Chapter 36.

Problem 8A. Let α be a primitive element in $\mathbb{F}_{2^n}$. For $1 \le i \le m := 2^n - 1$, let

$$\alpha^i = \sum_{j=0}^{n-1} c_{ij}\alpha^j.$$

Show that the sequence

$$0, c_{10}, c_{20}, \ldots, c_{m0}$$

is a De Bruijn sequence.

Problem 8B. Find a circular ternary sequence (with symbols $0, 1, 2$) of length 27 so that each possible ternary ordered triple occurs as three (circularly) consecutive positions of the sequence. First sketch a certain directed graph on 9 vertices so that Eulerian circuits in the graph correspond to such sequences.

Problem 8C. We wish to construct a circular sequence $a_0, \ldots, a_7$ (indices mod 8) in such a way that a sliding window a_i, a_{i+1}, a_{i+3} $(i = 0, 1, \ldots 7)$ will contain every possible three-tuple once. Show (not just by trial and error) that this is impossible.

Problem 8D. Let $m := 2^n - 1$. An algorithm to construct a De Bruijn sequence $a_0, a_1, \ldots, a_m$ works as follows. Start with $a_0 = a_1 = \cdots = a_{n-1} = 0$. For $k > n$, we define a_k to be the maximal value in $\{0, 1\}$ such that the sequence $(a_{k-n+1}, \ldots, a_{k-1}, a_k)$ has not occurred in $(a_0, \ldots, a_{k-1})$ as a (consecutive) subsequence. The resulting sequence is known as a *Ford sequence*. Prove that this algorithm indeed produces a De Bruijn sequence.

Notes.

Although the graphs of this chapter are commonly called De Bruijn graphs, Theorem 8.1 was proved in 1894 by C. Flye Sainte-Marie. This went unnoticed for a long time. We refer to De Bruijn (1975).

N. G. de Bruijn (1918–), one of the best-known Dutch mathematicians, worked in many different areas such as analysis, number theory, combinatorics, and also computing science and crystalography.

We mention a peculiarity concerning the spelling of some Dutch names. When omitting the initials of N. G. de Bruijn, one should capitalize the word 'de' and furthermore the name should be listed under B. Similarly Van der Waerden is correct when the initials are omitted and he should be listed under W.

For a proof of Theorem 8.1 using algebraic methods, we refer to Chapter 36.

References.

N. G. de Bruijn (1946), A combinatorial problem, *Proc. Kon. Ned. Akad. v. Wetensch.* **49**, 758–764.

N. G. de Bruijn (1975), Acknowledgement of priority to C. Flye Sainte-Marie on the counting of circular arrangements of 2^n zeros and ones that show each n-letter word exactly once, T. H. report 75-WSK-06, Eindhoven University of Technology.

C. Flye Sainte-Marie (1894), Solution to question nr. 48, *Intermédiaire des Mathématiciens* **1**, 107–110.

9
Two (0,1,*) problems: addressing for graphs and a hash-coding scheme

The following problem originated in communication theory. For a telephone network, a connection between terminals A and B is established before messages flow in either direction. For a network of computers it is desirable to be able to send a message from A to B without B knowing that a message is on its way. The idea is to let the message be preceded by some 'address' of B such that at each node of the network a decision can be made concerning the direction in which the message should proceed.

A natural thing to try is to give each vertex of a graph G a binary address, say in $\{0,1\}^k$, in such a way that the distance of two vertices in the graph is equal to the so-called *Hamming distance* of the addresses, i.e. the number of places where the addresses differ. This is equivalent to regarding G as an *induced subgraph* of the *hypercube* H_k, which has $V(H_k) := \{0,1\}^k$ and where k-tuples are adjacent when they differ in exactly one coordinate. The example $G = K_3$ already shows that this is impossible. We now introduce a new alphabet $\{0,1,*\}$ and form addresses by taking n-tuples from this alphabet. The distance between two addresses is defined to be the number of places where one has a 0 and the other a 1 (so stars do not contribute to the distance). For an addressing of a graph G, we require that the distance of any two vertices in G is equal to the distance of their addresses. It is trivial to show that this can be done if n is large enough. We denote by $N(G)$ the minimum value of n for which there exists an addressing of G with length n.

For a tree we can do without the stars as follows. We use induction. For a tree with two vertices, we have a trivial addressing with length 1. Suppose that we can address trees with k vertices. If $x_0, x_1, \ldots, x_k$ are the vertices of the tree T and x_0 is a monovalent vertex, then consider an addressing for the tree obtained by removing x_0. Let $\mathbf{x}_i$ be the address of x_i and suppose x_0 is joined to x_1. We change all addresses to $(0, \mathbf{x}_i)$, $1 \leq i \leq k$, and give x_0 the address $(1, \mathbf{x}_1)$. Clearly this is now an addressing for T. So for a tree, we have $N(T) \leq |V(T)| - 1$.

As a second example, consider K_m. In the identity matrix of size $m-1$, we replace the zeros above the diagonal by stars and add a row of zeros. Any two rows now have distance 1 and hence $N(K_m) \leq m - 1$.

As a third example, we consider the graph of Fig. 9.1.

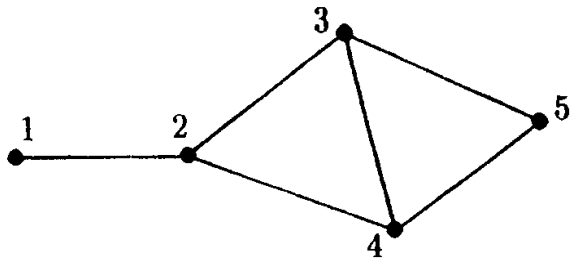


Figure 9.1

A possible (though not optimal) addressing is

1	1	1	1	$*$	$*$
2	1	0	$*$	1	$*$
3	$*$	0	0	0	1
4	0	0	1	$*$	$*$
5	0	0	0	0	0

We now show a correspondence between addressings of a graph and *quadratic forms* (an idea of Graham and Pollak, 1971). Consider the graph G of Fig. 9.1 and the addressing given above. To the first column of the addressing, we associate the product $(x_1+x_2)(x_4+x_5)$. Here x_i is in the first, respectively second, factor if the address of i has a 1, respectively a 0, in the first column. If we do the same thing for each column and then add the terms, we obtain a quadratic form $\sum d_{ij}x_ix_j$, where d_{ij} is the distance of

the vertices i and j in G. Thus an addressing of G corresponds to writing the quadratic form $\sum d_{ij}x_ix_j$ as a sum of n products

$$(x_{i_1} + \cdots + x_{i_k})(x_{j_1} + \cdots + x_{j_l})$$

such that no x_i occurs in both of the factors. The number of variables is $|V(G)|$.

Theorem 9.1. *Let n_+, respectively n_-, be the number of positive, respectively negative, eigenvalues of the distance matrix (d_{ij}) of the graph G. Then $N(G) \geq \max\{n_+, n_-\}$.*

Proof: Each of the quadratic forms mentioned above can be represented as $\frac{1}{2}\mathbf{x}^\top A\mathbf{x}$, where $\mathbf{x} := (x_1, x_2, \ldots, x_n)$ and A has entry $a_{ij} = 1$ if the term x_ix_j occurs in the quadratic form and 0 otherwise. Such a matrix has rank 2 and trace 0. Therefore it has one positive and one negative eigenvalue. Since (d_{ij}) is the sum of the matrices corresponding to the quadratic forms, it can have at most n positive (respectively negative) eigenvalues. □

Theorem 9.2. $N(K_m) = m - 1$.

Proof: We have already seen that $N(K_m) \leq m - 1$. Since $J - I$, of size m, is the distance matrix of K_m and the eigenvalues of $J - I$ are $m - 1$, with multiplicity 1, and -1, with multiplicity $m - 1$, the result follows from Theorem 9.1. □

With slightly more work, we shall now show that the shortest addressing for a tree T has length $|V(T)| - 1$.

Theorem 9.3. *If T is a tree on n vertices, then $N(T) = n - 1$.*

Proof: We first calculate the determinant of the distance matrix (d_{ij}) of T. We number the vertices $p_1, \ldots, p_n$ in such a way that p_n is an endpoint adjacent to p_{n-1}. In the distance matrix, we subtract row $n - 1$ from row n, and similarly for the columns. Then all the entries in the new last row and column are 1 except for the diagonal element which is equal to -2. Now renumber the vertices $p_1, \ldots, p_{n-1}$ in such a way that the new vertex p_{n-1} is an endpoint

of $T\backslash\{p_n\}$ adjacent to p_{n-2}. Repeat the procedure for the rows and columns with numbers $n-1$ and $n-2$. After $n-1$ steps, we have the determinant

$$\begin{vmatrix} 0 & 1 & 1 & \dots & 1 \\ 1 & -2 & 0 & \dots & 0 \\ 1 & 0 & -2 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 0 & 0 & \dots & -2 \end{vmatrix}.$$

From this we find the remarkable result that the determinant D_n of the distance matrix of a tree on n vertices satisfies

$$D_n = (-1)^{n-1}(n-1)2^{n-2},$$

i.e. it depends only on $|V(T)|$. If we number the vertices according to the procedure described above, then the $k \times k$ principal minor in the upper left-hand corner of the distance matrix is the distance matrix of a subtree on k vertices. Therefore the sequence $1, D_1, D_2, \dots, D_n$, where D_k is the determinant of the $k \times k$ minor, is equal to

$$1, 0, -1, 4, -12, \dots, (-1)^{n-1}(n-1)2^{n-2}.$$

If we consider the sign of 0 to be positive, then this sequence has only one occurrence of two consecutive terms of the same sign. By an elementary theorem on quadratic forms this implies that the corresponding quadratic form has index 1, and hence (d_{ij}) has one positive eigenvalue; see B. W. Jones (1950), Theorem 4. Now the result follows from Theorem 9.1. □

The conjecture that in fact $N(G) \leq |V(G)|-1$ for all (connected) graphs G was proved by P. Winkler in 1983. The proof is constructive. In order to describe the addressing, we need some preparation. Consider the graph of Fig. 9.2.

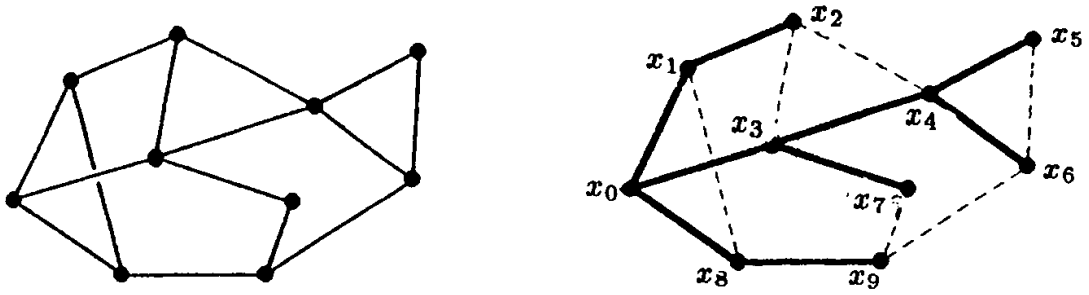


Figure 9.2

We pick a vertex x_0, then construct a spanning tree T by a breadth-first search, and then number the vertices by a depth-first search. The result is shown on the right-hand side of Fig. 9.2, where edges of $E(G)\backslash E(T)$ are dashed.

Let $n := |V(G)| - 1$. We need several definitions.
For $i \leq n$, we define

$$P(i) := \{j : x_j \text{ is on a path from } x_0 \text{ to } x_i \text{ in } T\}.$$

For example, $P(6) = \{0, 3, 4, 6\}$. Let

$$i \triangle j := \max(P(i) \cap P(j)).$$

We describe the general situation in Fig. 9.3.

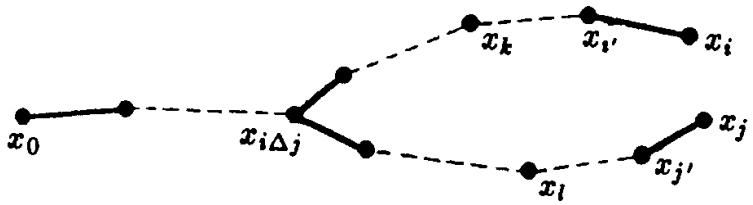


Figure 9.3

Note that in Fig. 9.3, we have $i < j$ if and only if $k < l$.
For $i \leq n$, we define

$$i' := \max(P(i)\backslash\{i\}).$$

For example, $7' = 3$ in Fig. 9.2. Define

$$i \sim j \Leftrightarrow P(i) \subseteq P(j) \text{ or } P(j) \subseteq P(i).$$

We denote distances in G, respectively T, by d_G, respectively d_T. The *discrepancy function* $c(i,j)$ is now defined by

$$c(i,j) := d_T(x_i,x_j) - d_G(x_i,x_j).$$

For example, in Fig. 9.2, $c(6,9) = 4$.

Lemma 9.4.
(i) $c(i,j) = c(j,i) \geq 0$;
(ii) *if* $i \sim j$, *then* $c(i,j) = 0$;
(iii) *if* $i \not\sim j$, *then* $c(i,j') \leq c(i,j) \leq c(i,j') + 2$.

PROOF: (i) is trivial; (ii) follows from the definition of T since

$$d_G(x_i,x_j) \geq |d_G(x_j,x_0) - d_G(x_i,x_0)| = d_T(x_i,x_j);$$

(iii) follows from the fact that $|d_G(x_i,x_j) - d_G(x_i,x_{j'})| \leq 1$ and that $d_T(x_i,x_j) = 1 + d_T(x_i,x_{j'})$. □

Now we can define the addressing. For $0 \leq i \leq n$ the vertex x_i is given the address $\mathbf{a}_i \in \{0,1,*\}^n$, where

$$\mathbf{a}_i = (a_i(1), a_i(2), \ldots, a_i(n))$$

and

$$a_i(j) := \begin{cases} 1 & \text{if } j \in P(i), \\ * & \text{if} \begin{cases} c(i,j) - c(i,j') = 2, \text{ or} \\ c(i,j) - c(i,j') = 1,\ i < j,\ c(i,j) \text{ even, or} \\ c(i,j) - c(i,j') = 1,\ i > j,\ c(i,j) \text{ odd,} \end{cases} \\ 0 & \text{otherwise.} \end{cases}$$

Theorem 9.5. $d(\mathbf{a}_i, \mathbf{a}_k) = d_G(x_i, x_k)$.

PROOF: We may assume $i < k$.

(i) Suppose $i \sim k$. Then $d_G(x_i,x_k) = |P(k) \backslash P(i)|$. The values of j such that $j \in P(k) \backslash P(i)$ are exactly the positions where $a_k(j) = 1$, $a_i(j) \neq 1$. For these values of j we see that $c(i,j) = 0$, hence $a_i(j) = 0$ and we are done.

(ii) The hard case is when $i \not\sim k$. The key observation is the following. Let $n_1 \leq n_2 \leq \cdots \leq n_l$ be a nondecreasing sequence of

integers such that $|n_{i+1} - n_i| \le 2$ for all i. If m is an even integer between n_1 and n_l that does *not* occur in the sequence, then there is an i such that $n_i = m - 1$, $n_{i+1} = m + 1$. Now consider the sequence

$$c(i,k) \ge c(i,k') \ge c(i,k'') \ge \cdots \ge c(i,i\triangle k) = 0.$$

By the definition of $a_i(j)$ and the observation above, $a_i(j) = *$ and $a_k(j) = 1$ exactly as many times as there are even integers between $c(i,i\triangle k)$ and $c(i,k)$. Similarly $a_k(j) = *$ and $a_i(j) = 1$ as many times as there are odd integers between $c(i,i\triangle k)$ and $c(i,k)$. So

$$\begin{aligned} d(\mathbf{a}_i,\mathbf{a}_k) &= |P(k)\backslash P(i)| + |P(i)\backslash P(k)| - c(i,k) \\ &= d_T(x_i,x_k) - c(i,k) = d_G(x_i,x_k). \end{aligned}$$

□

Therefore we have proved the following theorem.

Theorem 9.6. $N(G) \le |V(G)| - 1$.

Problem 9A. If we use the scheme defined above, what are the addresses of x_2 and x_6 in Fig. 9.2?

Problem 9B. Let G be a cycle (polygon) on $2n$ vertices. Determine $N(G)$.

Problem 9C. Let G be a cycle (polygon) on $2n+1$ vertices. Prove that $N(G) = 2n$. Hint: if C_k is the permutation matrix with entries $c_{ij} = 1$ if and only if $j - i \equiv 1 \pmod{k}$ and $\zeta^k = 1$, then $(1, \zeta, \zeta^2, \ldots, \zeta^{k-1})$ is an eigenvector of C_k.

* * *

We now look at a second problem involving k-tuples from the alphabet $\{0, 1, \star\}$. The objects we shall study were introduced by Rivest (1974) and given the (unfortunate) name *associative block design*; cf. Chapter 19 for block designs. An $ABD(k,w)$ is a set of $b := 2^w$ elements of $\{0,1,\star\}^k$ with the following properties: if the elements are the rows of a $b \times k$ matrix C, then

(i) each row of C has $k-w$ stars;
(ii) each column of C has $b(k-w)/k$ stars;
(iii) any two distinct rows have distance at least 1.

Note that the definition implies that each vector in $\mathbb{F}_2^k$ has distance 0 to exactly one row of C.

The origin of the problem is the following. Consider a file of k-bit binary words. Each sequence in $\{0,1,\star\}^k$ is called a *partial match query*. The partial match retrieval problem is to retrieve from the file all words agreeing with the query in those positions where the query specifies a bit. So-called *hash-coding* schemes divide a file into b disjoint lists $L_1, L_2, \ldots, L_b$. A record x will be stored in the list with index $h(x)$, where h is the 'hash-function' mapping $\{0,1\}^k$ onto $\{1,2,\ldots,b\}$. For a given partial match query, some of the lists must be searched. An analysis of the worst-case number of lists to be searched led to the concept of ABD. In this case $h(x)$ is the index of the unique row of C which has distance 0 to x.

Example 9.1. The following matrix is an $ABD(4,3)$:

$$\begin{pmatrix} \star & 0 & 0 & 0 \\ 0 & \star & 1 & 0 \\ 0 & 0 & \star & 1 \\ 0 & 1 & 0 & \star \\ \star & 1 & 1 & 1 \\ 1 & \star & 0 & 1 \\ 1 & 1 & \star & 0 \\ 1 & 0 & 1 & \star \end{pmatrix}.$$

We first prove some elementary properties of an ABD.

Theorem 9.7. *If an $ABD(k,w)$ exists, then:*

(1) *it has exactly $bw/(2k)$ zeros and $bw/(2k)$ ones in each column;*
(2) *for each $\mathbf{x}$ in $\mathbb{F}_2^k$ it has exactly $\binom{w}{u}$ rows which agree with $\mathbf{x}$ in u positions;*
(3) *the parameters satisfy*

$$w^2 \geq 2k\left(1-\frac{1}{b}\right);$$

(4) *for any row, the number of rows with stars in the same positions is even.*

PROOF: Let C be the $ABD(k, w)$.

(1) A row of C with a star, respectively a zero, in column j represents (i. e. has distance 0 to) 2^{k-w-1}, respectively 2^{k-w}, elements of $\mathbb{F}_2^k$. From (i) and (ii) of the definition, it follows that column j must contain $bw/(2k)$ zeros.

(2) Let $\mathbf{x} \in \mathbb{F}_2^k$. Denote by n_i the number of rows of C which agree with $\mathbf{x}$ in i positions. There are $\binom{k}{l}$ vectors in $\mathbb{F}_2^k$ which agree with $\mathbf{x}$ in exactly l positions. Therefore $\binom{k}{l} = \sum n_i \binom{k-w}{l-i}$, i. e.

$$(1+z)^k = (1+z)^{k-w} \cdot \sum n_i z^i.$$

This proves that $n_i = \binom{w}{i}$.

(3) The sum of the distances between pairs of rows of C is $k(\frac{bw}{2k})^2$ by (1). Since any two rows have distance at least 1, this sum is at least $\binom{b}{2}$.

(4) Consider a row of C. Count vectors in $\mathbb{F}_2^k$ which have zeros in the positions where the row has stars. Each row with a different star pattern represents an even number of such vectors whereas a row with the same star pattern represents exactly one such vector. □

Note that property (1) in Theorem 9.7(3) implies that a necessary condition for the existence of an $ABD(k, w)$ is that k divides $w \cdot 2^{w-1}$.

The following strengthening of Theorem 9.7(3) is due to A. E. Brouwer (1999).

Theorem 9.8. *Let C be an $ABD(k, w)$ with $w > 3$.*

(1) *If two rows of C agree in all but one position, then*

$$\binom{w}{2} \geq k;$$

(2) *otherwise $w^2 > 2k$.*

PROOF: Suppose $\mathbf{c}_1$ and $\mathbf{c}_2$ are two rows of C which differ only in position one. Then all the other rows of C must differ from $\mathbf{c}_1$ in some other position. So, by (i) of the definition and Theorem 9.7(3), we find

$$b - 2 \leq (w-1) \cdot \frac{bw}{2k}.$$

To prove the assertion, we must show that the right-hand side of the inequality cannot be equal to $b-2$ or $b-1$. In both cases, equality would imply $2^{w-1} | k$ which contradicts Theorem 9.7 unless $w = 4$, which is excluded by substitution.

(ii) Consider two rows of C which have the same star pattern. By hypothesis, they differ in more than one position. Again, count the sum of the distances of all the rows from one of this pair. This sum is at least $2 + (b-2) = b$ and, by Theorem 9.7.1, it is equal to $w \cdot (bw)/(2k)$. So $w^2 \geq 2k$. We must show that equality cannot hold. By the argument above, equality would imply that rows with the same star pattern occur in pairs which have distance 2, and furthermore all the other rows have distance 1 to each row of such a pair. Without loss of generality, such a pair would be

$$(\star\star\cdots\star 00\ldots000) \text{ and } (\star\star\cdots\star 00\ldots011).$$

The $bw/(2k) - 1$ other rows ending in a 1 would have to end in 01, for otherwise they would have distance 0 to the second row or distance > 1 to the first row. Similarly, there would be $bw/(2k) - 1$ rows ending in 10. Since we now have rows with distance 2, we find that necessarily $bw/(2k) - 1 = 1$. Therefore $2^w = 2w$, which is impossible if $w \geq 3$. □

Corollary. *An $ABD(8, 4)$ does not exist.*

Using these results, it is easy to find all $ABD(k, w)$ with $w \leq 4$. Of course, $w = 0$ is trivial. For $w = 1, 2$, or 4, we must have $k = w$ (no stars). If $w = 3$, then either $k = 3$ (no stars) or $k = 4$. In that case there are two types of ABD, one given in Example 9.1.

Problem 9D. Construct an $ABD(4, 3)$ that has the same first four rows as Example 9.1 but differs in the others.

In 1987 La Poutré and Van Lint proved that an $ABD(10, 5)$ does not exist but the smaller possibility $ABD(8, 5)$ turned out to

be quite difficult to handle. In 1999 D. E. Knuth asked Brouwer whether any progress had been made in this area since the early results and this made Brouwer decide it was time the question was settled. An example can be found in Brouwer (1999). It does not seem to have any structure. Now the smallest open case is the question whether an $ABD(12, 6)$ exists.

We shall now describe some construction methods. Some of the ideas will be used in other chapters.

Theorem 9.9. *If an $ABD(k_i, w_i)$ exists for $i = 1, 2$, then an $ABD(k_1k_2, w_1w_2)$ exists.*

PROOF: We can assume $w_2 > 0$. Partition the rows of $ABD(k_2, w_2)$ into two classes R_0 and R_1 of equal size. In $ABD(k_1, w_1)$ we replace each star by a row of k_2 stars, each 0 by a row from R_0 and each 1 by a row from R_1 in all possible ways. A trivial calculation shows that the resulting matrix is an $ABD(k_1k_2, w_1w_2)$. □

Corollary. *An $ABD(4^t, 3^t)$ exists.*

For the proof of the next theorem, we introduce a new symbol, namely $-$. A k-tuple consisting of the symbols $0, 1, \star$, and $-$ represents all possible words with only $0, 1, \star$ that can be obtained by replacing each $-$ by a 0 or a 1 in all possible ways.

Theorem 9.10. *Let $w > 0$. Suppose an $ABD(k, w)$ exists, where $k = k_0 \cdot 2^l$, k_0 odd. Then an $ABD(k, w + ik_0)$ exists for $0 \leq i \leq (k - w)/k_0$.*

PROOF: It is sufficient to consider $i = 1$. Let C be the $ABD(k, w)$. Define a matrix A of the same size by requiring $a_{ij} = 1$ if $C_{ij} = \star$ and $a_{ij} = 0$ otherwise. By Theorem 7.3, A is the sum of two matrices A_1 and A_2, where A_1 has k_0 ones in each row and 2^{w-l} ones in each column. In a row of C, replace stars by $-$ if the star occurs in a position where A_1 has a one. This produces the required $ABD(k, w + k_0)$. □

Theorem 9.11. *If $ABD(k, w)$ exists and $\alpha \geq 1$ is a number such that αk and αw are integers, then an $ABD(\alpha k, \alpha w)$ exists.*

PROOF: It is sufficient to show that $ABD(k + l, w + m)$ exists for $(k + l)/(w + m) = k/w$ and $(l, m) = 1$. Let $k = k_0 \cdot 2^e$, k_0 odd.

From (ii) of the definition we know that $k_0|w$. Therefore $wl = mk$ and $(l, m) = 1$ imply that l is a power of 2. Consider the $l \times l$ circulant matrix with a row of $l - m$ stars and m minus signs as first row. Since l divides b we can adjoin a column of b/l copies of this circulant to the matrix C of the $ABD(k, w)$. It is easy to check that this larger matrix is an $ABD(k + l, w + m)$. □

Example 9.2. From the corollary to Theorem 9.9 we have an $ABD(64, 27)$. Theorem 9.10 then shows that an $ABD(64, w)$ exists for $27 \leq w \leq 64$. In particular, there is an $ABD(64, 32)$. Then Theorem 9.11 implies that an $ABD(2w, w)$ exists for all $w \geq 32$. As mentioned before, nonexistence has been shown for $w = 4$ and $w = 5$ and the case $w = 6$ is still open.

Notes.

The first problem considered in this chapter was introduced by J. R. Pierce at Bell Laboratories as the *loop switching problem.* Several people (including one of the present authors) tried in vain to solve it. Shortly after R. L. Graham raised the reward for the solution to $200, it was solved by P. Winkler. It is worth noting that Winkler stated that the idea of numbering the vertices as was done in the proof was a regular habit due to his background in computer science. Going over the proof, one sees that this numbering indeed played a crucial role.

References.

A. E. Brouwer (1999), An Associative Block Design $ABD(8, 5)$, *SIAM J. Comput.* **28**, 1970–1971.

R. L. Graham and H. O. Pollak (1971), On the addressing problem for loop switching, *Bell System Tech. J.* **50**, 2495–2519.

B. W. Jones (1950), *The Theory of Quadratic Forms*, Carus Math. Monogr. **10**, Math. Assoc. of America.

J. A. La Poutré and J. H. van Lint (1985), An associative block design $ABD(10, 5)$ does not exist, *Utilitas Math.* **31**, 219–225.

P. Winkler (1983), Proof of the squashed cube conjecture, *Combinatorica* **3**, 135–139.

10
The principle of inclusion and exclusion; inversion formulae

As we have seen in several of the previous chapters, many problems of combinatorial analysis involve the *counting* of certain objects. We now treat one of the most useful methods for counting. It is known as the *principle of inclusion and exclusion*. The idea is as follows. If A and B are subsets of S and we wish to count the elements of $S\backslash\{A \cup B\}$, then the answer is not $|S| - |A| - |B|$ because the elements of $A \cap B$ have been subtracted twice. However $|S| - |A| - |B| + |A \cap B|$ is correct. The following theorem generalizes this idea.

Theorem 10.1. *Let S be an N-set; $E_1, \ldots, E_r$ not necessarily distinct subsets of S. For any subset M of $\{1, \ldots, r\}$, we define $N(M)$ to be the number of elements of S in $\bigcap_{i \in M} E_i$ and for $0 \leq j \leq r$, we define $N_j := \sum_{|M|=j} N(M)$. Then the number of elements of S not in any of the subsets E_i, $1 \leq i \leq r$, is*

$$N - N_1 + N_2 - N_3 + \cdots + (-1)^r N_r. \tag{10.1}$$

Proof: (i) If $x \in S$ and x is in none of the E_i, then x contributes 1 to the expression (10.1).

(ii) If $x \in S$ and x is in exactly k of the sets E_i, then the contribution to (10.1) equals

$$1 - \binom{k}{1} + \binom{k}{2} - \cdots + (-1)^k \binom{k}{k} = (1-1)^k = 0.$$

□

Remark. If we truncate the sum in (10.1) after a positive (respectively, negative) term, then we have an upper (respectively, lower) bound for the number of elements of S not in any of the E_i.

Because this method is of great importance, we shall give several examples as illustration.

Example 10.1. Let d_n denote the number of permutations π of $1, 2, \ldots, n$ such that $\pi(i) \neq i$ for all i (these are called *derangements*). Let $S := S_n$, and let E_i be the subset of those permutations π with $\pi(i) = i$. By (10.1) we find

$$d_n = \sum_{i=0}^{n} (-1)^i \binom{n}{i} (n-i)! = n! \sum_{i=0}^{n} \frac{(-1)^i}{i!}. \tag{10.2}$$

From this formula, we see that for large values of n the probability that a permutation is a derangement is nearly e^{-1}. From (10.2) for n and $n-1$, we find a recursion formula for d_n:

$$d_n = n d_{n-1} + (-1)^n. \tag{10.3}$$

The formula (10.2) can also be obtained by inversion as follows. Consider the power series $D(x) := \sum_{n=0}^{\infty} d_n \frac{x^n}{n!}$ ($d_0 = 1$). Now if $F(x) := e^x D(x)$, then

$$F(x) = \sum_{m=0}^{\infty} \left(\sum_{r=0}^{m} \binom{m}{r} d_{m-r} \right) \frac{x^m}{m!}$$

and since $\sum_{r=0}^{m} \binom{m}{r} d_{m-r} = m!$, we find $F(x) = (1-x)^{-1}$. It follows that $D(x) = e^{-x}(1-x)^{-1}$ and by multiplying the power series for the two factors, we find (10.2) again.

Example 10.2. Let X be an n-set, $Y = \{y_1, \ldots, y_k\}$ a k-set. We count the surjections of X to Y. Let S be the set of all mappings from X to Y, E_i the subset of mappings for which y_i is not in the image of X. By (10.1) we find the number of surjections to be $\sum_{i=0}^{k} (-1)^i \binom{k}{i} (k-i)^n$. Now this number is trivially 0 if $k > n$ and clearly $n!$ if $k = n$. So we have proved:

$$\sum_{i=0}^{k} (-1)^i \binom{k}{i} (k-i)^n = \begin{cases} n! & \text{if } k = n, \\ 0 & \text{if } k > n. \end{cases} \tag{10.4}$$

There are many formulae like (10.4) that are often quite hard to prove directly. The occurrence of $(-1)^i$ is usually a sign that counting the right kind of objects using the principle of inclusion and exclusion can produce the formula, as in this example. Nevertheless it is useful in this case to see another proof.

Let $P(x)$ be a polynomial of degree n, with highest coefficient a_n. We denote the sequence of values $P(0), P(1), \dots$ by $\mathbf{P}$. We now consider the sequence of differences $P(1) - P(0), P(2) - P(1), \dots$. This is $\mathbf{Q}_1$, where $Q_1(x) := P(x+1) - P(x)$, a polynomial of degree $n-1$ with highest coefficient na_n. By repeating this procedure a number of times, we find a sequence $\mathbf{Q}_k$ whose terms are $\sum_{i=0}^{k}(-1)^i\binom{k}{i}P(x+k-i)$, corresponding to the polynomial $Q_k(x)$ of degree $n-k$ with highest coefficient $n(n-1)\dots(n-k+1)a_n$. If $k = n$, then all the terms of $\mathbf{Q}_k$ are $n!a_n$ and if $k > n$, then they are all 0. Take $P(x) = x^n$. We again find (10.4).

Example 10.3. The following identity is a well known relation between binomial coefficients:

$$\sum_{i=0}^{n}(-1)^i\binom{n}{i}\binom{m+n-i}{k-i} = \begin{cases}\binom{m}{k} & \text{if } m \geq k,\\ 0 & \text{if } m < k.\end{cases} \tag{10.5}$$

We see that if we wish to prove this using inclusion-exclusion, then the sets E_i that we wish to exclude involve choosing from an n-set, and after choosing i of them, we must choose $k-i$ elements from some set of size $m+n-i$. This shows us that the following combinatorial problem will lead us to the result (10.5). Consider a set $Z = X \cup Y$, where $X = \{x_1, \dots, x_n\}$ is an n-set of blue points and Y is an m-set of red points. How many k-subsets consist of red points only? The answer is trivially the right-hand side of (10.5). If we take S to be all the k-subsets of Z and E_i those k-subsets that contain x_i, then (10.1) gives us the left-hand side of (10.5).

Again we can ask whether this result can be proved directly. The answer is yes. To do this, we use the following expansion:

$$\sum_{j=0}^{\infty}\binom{a+j}{j}x^j = (1-x)^{-a-1}. \tag{10.6}$$

Note that $(-1)^i\binom{n}{i}$ is the coefficient of x^i in the expansion of $(1-x)^n$. From (10.6) we find that $\binom{m+n-i}{k-i}$ is the coefficient of x^{k-i} in the expansion of $(1-x)^{k-m-n-1}$. So the left-hand side of (10.5) is the coefficient of x^k in the expansion of $(1-x)^{k-m-1}$. If $m \leq k-1$, this is obviously 0 and if $m \geq k$, it is $\binom{m}{k}$, again by (10.6).

Example 10.4. (The *Euler function*) Let $n = p_1^{a_1}p_2^{a_2}\dots p_r^{a_r}$ be a positive integer. We denote by $\phi(n)$ the number of integers k with $1 \leq k \leq n$ such that the g.c.d. $(n,k) = 1$. We apply Theorem 10.1 with $S := \{1,2,\dots,n\}$ and E_i the set of integers divisible by p_i, $1 \leq i \leq r$. Then (10.1) yields

$$\phi(n) = n - \sum_{i=1}^{r}\frac{n}{p_i} + \sum_{1\leq i<j\leq r}\frac{n}{p_ip_j} - \dots = n\prod_{i=1}^{r}(1-\frac{1}{p_i}). \tag{10.7}$$

The next theorem is used quite often.

Theorem 10.2. $\sum_{d|n}\phi(d) = n$.

Proof: Consider $\{1,2,\dots,n\} = N$. For each $m \in N$, we have $(m,n)|n$. The number of integers m with $(m,n) = d$, i.e. $m = m_1d$, $n = n_1d$ and $(m_1,n_1) = 1$ clearly equals $\phi(n_1) = \phi(n/d)$. So $n = \sum_{d|n}\phi(n/d)$ which is equivalent to the assertion. □

At this point, it is useful to introduce the so-called *Möbius function*:

$$\mu(d) := \begin{cases} 1 & \text{if } d = \text{product of an even number of distinct primes,} \\ -1 & \text{if } d = \text{product of an odd number of distinct primes,} \\ 0 & \text{otherwise, i.e. } d \text{ not squarefree.} \end{cases} \tag{10.8}$$

Theorem 10.3.

$$\sum_{d|n}\mu(d) = \begin{cases} 1 & \textit{if } n = 1, \\ 0 & \textit{otherwise.} \end{cases}$$

Proof: If $n = 1$, there is nothing to prove. If $n = p_1^{a_1}\dots p_r^{a_r}$, then

by (10.8) we have

$$\sum_{d|n} \mu(d) = \sum_{i=0}^{r} \binom{r}{i} (-1)^i = (1-1)^r = 0.$$

□

Note how similar the proofs of Theorems 10.1 and 10.3 are.

Using the Möbius function, we can reformulate (10.7) as follows:

$$\frac{\phi(n)}{n} = \sum_{d|n} \frac{\mu(d)}{d}. \tag{10.9}$$

Problem 10A. How many positive integers less than 1000 have no factor between 1 and 10?

Problem 10B. How many monic polynomials of degree n are there in $\mathbb{F}_p[x]$ that do not take on the value 0 for $x \in \mathbb{F}_p$?

Problem 10C. Determine $\sum_{n \le x} \mu(n) \lfloor \frac{x}{n} \rfloor$.

Problem 10D. One of the most famous functions in complex analysis is the so-called *Riemann ζ-function* $\zeta(s) := \sum_{n=1}^{\infty} n^{-s}$, defined in the complex plane for $\mathrm{Re}(s) > 1$. Prove that $1/\zeta(s) = \sum_{n=1}^{\infty} \mu(n) n^{-s}$.

Problem 10E. Let $f_n(z)$ be the function that has as its zeros all numbers η for which $\eta^n = 1$ but $\eta^k \neq 1$ for $1 \le k < n$. Prove that

$$f_n(z) = \prod_{k|n} (z^k - 1)^{\mu(n/k)}.$$

Theorem 10.3 makes it possible to derive a very useful inversion formula known as the *Möbius inversion formula.*

Theorem 10.4. *Let $f(n)$ and $g(n)$ be functions defined for every positive integer n satisfying*

$$f(n) = \sum_{d|n} g(d). \tag{10.10}$$

Then g satisfies

$$g(n) = \sum_{d|n} \mu(d) f(\frac{n}{d}). \tag{10.11}$$

PROOF: By (10.10) we have

$$\begin{aligned}\sum_{d|n} \mu(d) f(\frac{n}{d}) &= \sum_{d|n} \mu(\frac{n}{d}) f(d) \\ &= \sum_{d|n} \mu(\frac{n}{d}) \sum_{d'|d} g(d') = \sum_{d'|n} g(d') \sum_{m|(n/d')} \mu(m).\end{aligned}$$

By Theorem 10.3 the inner sum on the right-hand side is 0 unless $d' = n$. □

Remark. The equation (10.11) also implies (10.10).

Example 10.5. We shall count the number N_n of circular sequences of 0's and 1's, where two sequences obtained by a rotation are considered the same. Let $M(d)$ be the number of circular sequences of length d that are not periodic. Then $N_n = \sum_{d|n} M(d)$. We observe that $\sum_{d|n} dM(d) = 2^n$ since this counts all possible circular sequences. By Theorem 10.4 we find from this equation that $nM(n) = \sum_{d|n} \mu(d) 2^{n/d}$ and therefore

$$\begin{aligned}N_n = \sum_{d|n} M(d) &= \sum_{d|n} \frac{1}{d} \sum_{l|d} \mu(\frac{d}{l}) 2^l \\ &= \sum_{l|n} \frac{2^l}{l} \sum_{k|\frac{n}{l}} \frac{\mu(k)}{k} = \frac{1}{n} \sum_{l|n} \phi(\frac{n}{l}) 2^l.\end{aligned} \tag{10.12}$$

The final expression has the advantage that all the terms are positive. This raises the question whether we could have obtained that expression by some other counting technique. We shall see that the following theorem, known as *Burnside's lemma* (although the theorem is actually due to Cauchy and Frobenius; see the notes), provides the answer.

Theorem 10.5. *Let G be a permutation group acting on a set X. For $g \in G$ let $\psi(g)$ denote the number of points of X fixed by g. Then the number of orbits of G is equal to $\frac{1}{|G|}\sum_{g\in G}\psi(g)$.*

PROOF: Count pairs (g, x), where $g \in G$, $x \in X$, $x^g = x$. Starting with g, we find $\sum_{g\in G}\psi(g)$. For each $x \in X$ there are $|G|/|O_x|$ such pairs, where O_x is the orbit of x. So the total number equals $|G|\sum_{x\in X} 1/|O_x|$. The orbits of G partition X, and if we sum the terms $1/|O_x|$ over all x in a particular orbit, we obtain 1. Thus $\sum_{x\in X} 1/|O_x|$ is the number of orbits. □

Example 10.5 (continued). *Let G be the cyclic group of order n, i.e. the group of rotations of a circular sequence of 0's and 1's. If $d|n$ there are $\phi(n/d)$ integers g such that $(n, g) = d$ and for each such g there are 2^d circular sequences that are fixed by the rotation over g positions. So Theorem 10.5 immediately yields the result (10.12).*

Example 10.6. The following problem, introduced by Lucas in 1891, is known as the *'problème des ménages'*. We wish to seat n couples at a circular table so that men and women are in alternate places and no husband will sit on either side of his wife. In how many ways can this be done? We assume that the women have been seated at alternate places. Call the ladies 1 to n and the corresponding men also 1 to n. The problem amounts to placing the integers 1 to n on a circle with positions numbered 1 to n such that for all i the integer i is not in position i or position $i + 1 \pmod n$. Let E_i be the set of seatings in which husband i is sitting next to his wife. We now wish to use inclusion-exclusion and we must therefore calculate in how many ways it is possible to seat r husbands incorrectly. Call this number A_r. We do this as follows. Consider a circular sequence of $2n$ positions. Put a 1 in position $2i - 1$ if husband i is sitting to the right of his wife; put a 1 in position $2i$ if he is sitting to the left of his wife. Put zeros in the remaining positions. The configurations that we wish to count therefore are circular sequences of $2n$ zeros and ones, with exactly r ones, no two adjacent. Let A'_r be the number of sequences starting with a 1 (followed by a 0). By considering 10 as one symbol, we see that we must choose $r - 1$ out of $2n - r - 1$ positions. To count

the number A''_r of sequences starting with a 0, we place the 0 at the end, and then it amounts to choosing r out of $2n-r$ places. Hence

$$A_r = A'_r + A''_r = \binom{2n-r-1}{r-1} + \binom{2n-r}{r} = \frac{2n}{2n-r}\binom{2n-r}{r}.$$

By (10.1) we find that the number of ways to seat the men is

$$\sum_{r=0}^{n}(-1)^r(n-r)!\binom{2n-r}{r}\frac{2n}{2n-r}. \tag{10.13}$$

Problem 10F. We color the integers 1 to $2n$ red or blue in such a way that if i is red then $i-1$ is not blue. Prove that

$$\sum_{k=0}^{n}(-1)^k\binom{2n-k}{k}2^{2n-2k} = 2n+1.$$

Can you prove this identity directly?

Problem 10G. Count the number of permutations $x_1, x_2, \ldots, x_{2n}$ of the integers 1 to $2n$ such that $x_i + x_{i+1} \neq 2n+1$ for $i = 1, 2, \ldots, 2n-1$.

Problem 10H. Prove that for $0 \leq k \leq n$

$$\sum_{i=0}^{k}\binom{k}{i}D_{n-i} = \sum_{j=0}^{n-k}(-1)^j\binom{n-k}{j}(n-j)!.$$

Notes.

The principle of inclusion and exclusion occurred as early as 1854 in a paper by Da Silva and later in a paper by Sylvester in 1883. For this reason (10.1) and similar formulae are sometimes called the formula of Da Silva, respectively Sylvester. A better name that is also often used is 'sieve formula'. The formula is indeed an example of a principle that is used extensively in number theory, referred to as 'sieve methods'. An example that is probably familiar to most readers is the sieve of Eratosthenes: to find the primes $\leq n^2$, take the integers $\leq n^2$ and sieve out all the multiples of primes $\leq n$.

The derangements treated in Example 10.1 occur again in Example 14.1 and Example 14.10. The first occurrence of this question is in one of the early books on games of chance: *Essai d'analyse sur les jeux de hazard* by P. R. de Montmort (1678–1719). It is still often referred to by the name that he gave it: 'problème des rencontres'. Formula (10.2) is sometimes stated as follows. If n persons check their umbrellas (a typical Dutch example; it's always raining in Holland) and subsequently pick one at random in the dark after a power failure, then the probability that nobody gets his own umbrella is roughly e^{-1} (if n is large).

The second proof in Example 10.2 is an example of the use of 'calculus of finite differences', used extensively in numerical analysis.

A. F. Möbius (1790–1868) was an astronomer (and before that an assistant to Gauss) who made important contributions to geometry and topology (e.g. the Möbius band).

G. F. B. Riemann (1826–1866) was professor in Göttingen, where he also obtained his doctorate under Gauss. He is famous for many of his ideas, which include the Riemann integral, Riemann surfaces and manifolds, and of course the so-called *Riemann hypothesis* on the location of the zeros of the ζ-function. One wonders what he would have left us if he had not died so young.

In most books in which it occurs, Theorem 10.5 is called Burnside's lemma. This is just one of many examples of theorems, etc. attributed to the wrong person. For a history of this misnomer, we refer to Neumann (1979).

F. E. A. Lucas (1842–1891) was a French number theorist. He is known for his books on number theory and mathematical recreations. The former book contained the problem of Example 10.6. The *Fibonacci numbers* were given this name by Lucas. See the notes to Chapter 14.

References.

F. E. A. Lucas (1891), *Théorie des nombres*, Gauthier-Villars, Paris.

P. M. Neumann (1979), A lemma that is not Burnside's, *Math. Scientist*, **4**, 133–141.

eindelijk terug naar echt bestand
(en voor zij die dit uitgebreider willen, kunnen ons contacteren)

The IMO Compendium Group
Olympiad Training Materials
www.imomath.com

©2007 The Author(s) and The IMO Compendium Group

Projective Geometry

Milivoje Lukić

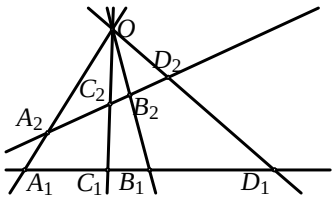
Contents

1 Cross Ratio. Harmonic Conjugates. Perspectivity. Projectivity

Definition 1. *Let A, B, C, and D be colinear points. The* cross ratio *of the pairs of points* (A,B) *and* (C,D) *is*

$$\mathscr{R}(A,B;C,D) = \frac{\overrightarrow{AC}}{\overrightarrow{CB}} : \frac{\overrightarrow{AD}}{\overrightarrow{DB}}. \tag{1}$$

Let a, b, c, d be four concurrent lines. For the given lines p_1 and p_2 let us denote $A_i = a \cap p_i$, $B_i = b \cap p_i$, $C_i = c \cap p_i$, $D_i = d \cap p_i$, for $i = 1, 2$. Then

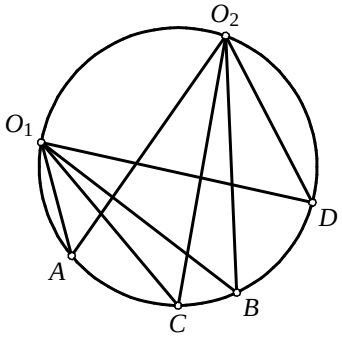


$$\mathscr{R}(A_1,B_1;C_1,D_1) = \mathscr{R}(A_2,B_2;C_2,D_2). \tag{2}$$

Thus it is meaningful to define the cross ratio of the pairs of concurrent points as

$$\mathscr{R}(a,b;c,d) = \mathscr{R}(A_1,B_1;C_1,D_1). \tag{3}$$

Assume that points O_1, O_2, A, B, C, D belong to a circle. Then



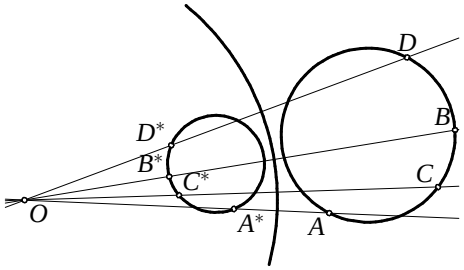
$$\begin{aligned} & \mathscr{R}(O_1A, O_1B; O_1C, O_1D) \\ = \; & \mathscr{R}(O_2A, O_2B; O_2C, O_2D). \end{aligned} \tag{4}$$

Hence it is meaningful to define the cross-ratio for cocyclic points as

$$\mathscr{R}(A,B;C,D) = \mathscr{R}(O_1A, O_1B; O_1C, O_1D). \tag{5}$$

Assume that the points A, B, C, D are colinear or cocyclic. Let an inversion with center O maps A, B, C, D into A^*, B^*, C^*, D^*. Then

$$\mathscr{R}(A,B;C,D)=\mathscr{R}(A^*,B^*;C^*,D^*). \qquad (6)$$



Definition 2. *Assume that A, B, C, and D are cocyclic or colinear points. Pairs of points (A,B) and (C,D) are* harmonic conjugates *if $\mathscr{R}(A,B;C,D)=-1$. We also write $\mathscr{H}(A,B;C,D)$ when we want to say that (A,B) and (C,D) are harmonic conjugates to each other.*

Definition 3. *Let each of l_1 and l_2 be either line or circle.* Perspectivity *with respect to the point S $\overset{S}{\curlywedge}$, is the mapping of $l_1 \to l_2$, such that*

(i) If either l_1 or l_2 is a circle than it contains S;

(ii) every point $A_1 \in l_1$ is mapped to the point $A_2 = OA_1 \cap l_2$.

According to the previous statements perspectivity preserves the cross ratio and hence the harmonic conjugates.

Definition 4. *Let each of l_1 and l_2 be either line or circle.* Projectivity *is any mapping from l_1 to l_2 that can be represented as a finite composition of perspectivities.*

Theorem 1. *Assume that the points A, B, C, D_1, and D_2 are either colinear or cocyclic. If the equation $\mathscr{R}(A,B;C,D_1)=\mathscr{R}(A,B;C,D_2)$ is satisfied, then $D_1=D_2$. In other words, a projectivity with three fixed points is the identity.*

Theorem 2. *If the points A, B, C, D are mutually discjoint and $\mathscr{R}(A,B;C,D)=\mathscr{R}(B,A;C,D)$ then $\mathscr{H}(A,B;C,D)$.*

2 Desargue's Theorem

The triangles $A_1B_1C_1$ and $A_2B_2C_2$ are *perspective with respect to a center* if the lines A_1A_2, B_1B_2, and C_1C_2 are concurrent. They are *perspective with respect to an axis* if the points $K=B_1C_1\cap B_2C_2$, $L=A_1C_1\cap A_2C_2$, $M=A_1B_1\cap A_2B_2$ are colinear.

Theorem 3 (Desargue). *Two triangles are perspective with respect to a center if and only if they are perspective with respect to a point.*

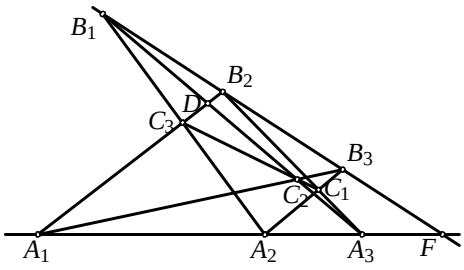
3 Theorems of Pappus and Pascal

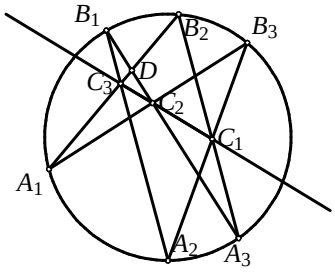
Theorem 4 (Pappus). *The points A_1, A_2, A_3 belong to the line a, and the points B_1, B_2, B_3 belong to the line b. Assume that $A_1B_2\cap A_2B_1=C_3$, $A_1B_3\cap A_3B_1=C_2$, $A_2B_3\cap A_3B_2=C_1$. Then C_1, C_2, C_3 are colinear.*

Proof. Denote $C_2'=C_1C_3\cap A_3B_1$, $D=A_1B_2\cap A_3B_1$, $E=A_2B_1\cap A_3B_2$, $F=a\cap b$. Our goal is to prove that the points C_2 and C_2' are identical. Consider the sequence of projectivities:

$$A_3B_1DC_2 \overset{A_1}{\curlywedge} FB_1B_2B_3 \overset{A_2}{\curlywedge} A_3EB_2C_1 \overset{C_3}{\curlywedge} A_3B_1DC_2'.$$

We have got the projective transformation of the line A_3B_1 that fixes the points A_3, B_1, D, and maps C_2 to C_2'. Since the projective mapping with three fixed points is the identity we have $C_2=C_2'$. $\square$





Theorem 5 (Pascal). *Assume that the points A_1, A_2, A_3, B_1, B_2, B_3 belong to a circle. The point in intersections of A_1B_2 with A_2B_1, A_1B_3 with A_3B_1, A_2B_3 with A_3B_2 lie on a line.*

Proof. The points C_2', D, and E as in the proof of the Pappus theorem. Consider the sequence of perspectivities

$$A_3B_1DC_2 \overset{A_1}{\wedge} A_3B_1B_2B_3 \overset{A_2}{\wedge} A_3EB_2C_1 \overset{C_3}{\wedge} A_3B_1DC_2'.$$

In the same way as above we conclude that $C_2 = C_2'$. □

4 Pole. Polar. Theorems of Brianchon and Brokard

Definition 5. *Given a circle $k(O,r)$, let A^* be the image of the point $A \neq O$ under the inversion with respect to k. The line a passing through A^* and perpendicular to OA is called the* polar *of A with respect to k. Conversely A is called the* pole *of a with respect to k.*

Theorem 6. *Given a circle $k(O,r)$, let and a and b be the polars of A and B with respect to k. The $A \in b$ if and only if $B \in a$.*

Proof. $A \in b$ if and only if $\angle AB^*O = 90^\circ$. Analogously $B \in a$ if and only if $\angle BA^*O = 90^\circ$, and it reamins to notice that according to the basic properties of inversion we have $\angle AB^*O = \angle BA^*O$. □

Definition 6. *Points A and B are called* conjugated *with respect to the circle k if one of them lies on a polar of the other.*

Theorem 7. *If the line determined by two conjugated points A and B intersects $k(O,r)$ at C and D, then $\mathscr{H}(A,B;C,D)$. Conversely if $\mathscr{H}(A,B;C,D)$, where $C,D \in k$ then A and B are conjugated with respect to k.*

Proof. Let C_1 and D_1 be the intersection points of OA with k. Since the inversion preserves the cross-ratio and $\mathscr{R}(C_1,D_1;A,A^*) = \mathscr{R}(C_1,D_1;A^*,A)$ we have

$$\mathscr{H}(C_1,D_1;A,A^*). \tag{7}$$

Let p be the line that contains A and intersects k at C and D. Let $E = CC_1 \cap DD_1$, $F = CD_1 \cap DC_1$. Since C_1D_1 is the diameter of k we have $C_1F \perp D_1E$ and $D_1F \perp C_1E$, hence F is the orthocenter of the triangle C_1D_1E. Let $B = EF \cap CD$ and $\bar{A}^* = EF \cap C_1D_1$. Since

$$C_1D_1A\bar{A}^* \overset{E}{\bar{\wedge}} CDAB \overset{F}{\bar{\wedge}} D_1C_1A\bar{A}^*$$

have $\mathscr{H}(C_1,D_1;A,\bar{A}^*)$ and $\mathscr{H}(C,D;A,B)$. (7) now implies two facts:

1° From $\mathscr{H}(C_1,D_1;A,\bar{A}^*)$ and $\mathscr{H}(C_1,D_1;A,A^*)$ we get $A^* = \bar{A}^*$, hence $A^* \in EF$. However, since $EF \perp C_1D_1$, the line $EF = a$ is the polar of A.

2° For the point B which belongs to the polar of A we have $\mathscr{H}(C,D;A,B)$. This completes the proof. □

Theorem 8 (Brianchon's theorem). *Assume that the hexagon $A_1A_2A_3A_4A_5A_6$ is circumscribed about the circle k. The lines A_1A_4, A_2A_5, and A_3A_6 intersect at a point.*

Proof. We will use the convention in which the points will be denoted by capital latin letters, and their repsective polars with the corresponding lowercase letters.

Denote by M_i, $i = 1,2,\ldots,6$, the points of tangency of A_iA_{i+1} with k. Since $m_i = A_iA_{i+1}$, we have $M_i \in a_i$, $M_i \in a_{i+1}$, hence $a_i = M_{i-1}M_i$.

Let $b_j = A_jA_{j+3}$, $j = 1,2,3$. Then $B_j = a_j \cap a_{j+3} = M_{j-1}M_j \cap M_{j+3}M_{j+4}$. We have to prove that there exists a point P such that $P \in b_1, b_2, b_3$, or analogously, that there is a line p such that $B_1, B_2, B_3 \in p$. In other words we have to prove that the points B_1, B_2, B_3 are colinear. However this immediately follows from the Pascal's theorem applied to $M_1M_3M_5M_4M_6M_2$. □

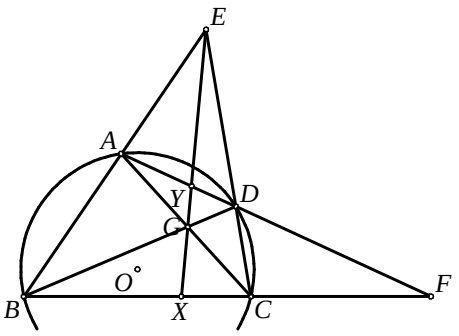
From the previous proof we see that the Brianchon's theorem is obtained from the Pascal's by replacing all the points with their polars and all lines by theirs poles.

Theorem 9 (Brokard). *The quadrilateral $ABCD$ is inscribed in the circle k with center O. Let $E = AB \cap CD$, $F = AD \cap BC$, $G = AC \cap BD$. Then O is the orthocenter of the triangle EFG.*

Proof. We will prove that EG is a polar of F. Let $X = EG \cap BC$ and $Y = EG \cap AD$. Then we also have

$$ADYF \overset{E}{\wedge} BCXF \overset{G}{\wedge} DAYF,$$

which implies the relations $\mathscr{H}(A,D;Y,F)$ and $\mathscr{H}(B,C;X,F)$. According to the properties of polar we have that the points X and Y lie on a polar of the point F, hence EG is a polar of the point F.



Since EG is a polar of F, we have $EG \perp OF$. Analogously we have $FG \perp OE$, thus O is the orthocenter of $\triangle EFG$. □

5 Problems

1. Given a quadrilateral $ABCD$, let $P = AB \cap CD$, $Q = AD \cap BC$, $R = AC \cap PQ$, $S = BD \cap PQ$. Prove that $\mathscr{H}(P,Q;R,S)$.

2. Given a triangle ABC and a point M on BC, let N be the point of the line BC such that $\angle MAN = 90^\circ$. Prove that $\mathscr{H}(B,C;M,N)$ if and only if AM is the bisector of the angle $\angle BAC$.

3. Let A and B be two points and let C be the point of the line AB. Using just a ruler find a point D on the line AB such that $\mathscr{H}(A,B;C,D)$.

4. Let A, B, C be the diagonal points of the quadrilateral $PQRS$, or equivalently $A = PQ \cap RS$, $B = QR \cap SP$, $C = PR \cap QS$. If only the points A, B, C, S, are given using just a ruler construct the points P, Q, R.

5. Assume that the incircle of $\triangle ABC$ touches the sides BC, AC, and AB at D, E, and F. Let M be the point such that the circle k_1 incscibed in $\triangle BCM$ touches BC at D, and the sides BM and CM at P and Q. Prove that the lines EF, PQ, BC are concurrent.

6. Given a triangle ABC, let D and E be the points on BC such that $BD = DE = EC$. The line p intersects AB, AD, AE, AC at K, L, M, N, respectively. Prove that $KN \geq 3LM$.

7. The point M_1 belongs to the side AB of the quadrilateral $ABCD$. Let M_2 be the projection of M_1 to the line BC from D, M_3 projection of M_2 to CD from A, M_4 projection of M_3 to DA from B, M_5 projection of M_4 to AB from C, etc. Prove that $M_{13} = M_1$.

8. (butterfly theorem) Points M and N belong to the circle k. Let P be the midpoint of the chord MN, and let AB and CD (A and C are on the same side of MN) be arbitrary chords of k passing through P. Prove that lines AD and BC intersect MN at points that are equidistant from P.

9. Given a triangle ABC, let D and E be the points of the sides AB and AC respectively such that $DE \| BC$. Let P be an interior point of the triangle ADE. Assume that the lines BP and CP intersect DE at F and G respectively. The circumcircles of $\triangle PDG$ and $\triangle PFE$ intersect at P and Q. Prove that the points A, P, and Q are colinear.

10. (IMO 1997 shortlist) Let $A_1A_2A_3$ be a non-isosceles triangle with the incenter I. Let C_i, $i = 1$, 2, 3, be the smaller circle through I tangent to both A_iA_{i+1} and A_iA_{i+2} (summation of indeces is done modulus 3). Let B_i, $i = 1$, 2, 3, be the other intersection point of C_{i+1} and C_{i+2}. Prove that the circumcenters of the triangles A_1B_1I, A_2B_2I, A_3B_3I are colinear.

11. Given a triangle ABC and a point T, let P and Q be the feet of perpendiculars from T to the lines AB and AC, respectively. Let R and S be the feet of perpendiculars from A to TC and TB, respectively. Prove that the intersection of PR and QS belongs to BC.

12. Given a triangle ABC and a point M, a line passing through M intersects AB, BC, and CA at C_1, A_1, and B_1, respectively. The lines AM, BM, and CM intersect the circumcircle of $\triangle ABC$ repsectively at A_2, B_2, and C_2. Prove that the lines A_1A_2, B_1B_2, and C_1C_2 intersect in a point that belongs to the circumcircle of $\triangle ABC$.

13. Let P and Q isogonaly conjugated points and assume that $\triangle P_1P_2P_3$ and $\triangle Q_1Q_2Q_3$ are their pedal triangles, respectively. Let $X_1 = P_2Q_3 \cap P_3Q_2$, $X_2 = P_1Q_3 \cap P_3Q_1$, $X_3 = P_1Q_2 \cap P_2Q_1$. Prove that the points X_1, X_2, X_3 belong to the line PQ.

14. If the points A and M are conjugated with respect to k, then the circle with diameter AM is orthogonal to k.

15. From a point A in the exterior of a circle k two tangents AM and AN are drawn. Assume that K and L are two points of k such that A, K, L are colinear. Prove that MN bisects the segment PQ.

16. The point isogonaly conjugated to the centroid is called the *Lemuan* point. The lines connected the vertices with the Lemuan point are called *symmedians*. Assume that the tangents from B and C to the circumcircle Γ of $\triangle ABC$ intersect at the point P. Prove that AP is a symmedian of $\triangle ABC$.

17. Given a triangle ABC, assume that the incircle touches the sides BC, CA, AB at the points M, N, P, respectively. Prove that AM, BN, and CP intersect in a point.

18. Let $ABCD$ be a quadrilateral circumscribed about a circle. Let M, N, P, and Q be the points of tangency of the incircle with the sides AB, BC, CD, and DA respectively. Prove that the lines AC, BD, MP, and NQ intersect in a point.

19. Let $ABCD$ be a cyclic quadrilateral whose diagonals AC and BD intersect at O; extensions of the sides AB and CD at E; the tangents to the circumcircle from A and D at K; and the tangents to the circumcircle at B and C at L. Prove that the points E, K, O, and L lie on a line.

20. Let $ABCD$ be a cyclic quadrilateral. The lines AB and CD intersect at the point E, and the diagonals AC and BD at the point F. The circumcircle of the triangles $\triangle AFD$ and $\triangle BFC$ intersect again at H. Prove that $\angle EHF = 90^\circ$.

6 Solutions

1. Let $T = AC \cap BD$. Consider the sequence of the perspectivities

$$PQRS \overset{A}{\barwedge} BDTS \overset{C}{\barwedge} QPRS.$$

Since the perspectivity preserves the cross-ratio $\mathscr{R}(P,Q;R,S) = \mathscr{R}(Q,P;R,S)$ we obtain that $\mathscr{H}(P,Q;R,S)$.

2. Let $\alpha = \angle BAC$, $\beta = \angle CBA$, $\gamma = \angle ACB$ and $\varphi = \angle BAM$. Using the sine theorem on $\triangle ABM$ and $\triangle ACM$ we get

$$\frac{BM}{MC} = \frac{BM}{AM}\frac{AM}{CM} = \frac{\sin\varphi}{\sin\beta}\frac{\sin\gamma}{\sin(\alpha-\varphi)}.$$

Similarly using the sine theorem on $\triangle ABN$ and $\triangle ACN$ we get

$$\frac{BN}{NC} = \frac{BN}{AN}\frac{AN}{CN} = \frac{\sin(90^\circ-\varphi)}{\sin(180^\circ-\beta)}\frac{\sin\gamma}{\sin(90^\circ+\alpha-\varphi)}.$$

Combining the previous two equations we get

$$\frac{BM}{MC} : \frac{BN}{NC} = \frac{\tan\varphi}{\tan(\alpha-\varphi)}.$$

Hence, $|\mathscr{R}(B,C;M,N)| = 1$ is equivalent to $\tan\varphi = \tan(\alpha-\varphi)$, i.e. to $\varphi = \alpha/2$. Since $B \neq C$ and $M \neq N$, the relation $|\mathscr{R}(B,C;M,N)| = 1$ is equivalent to $\mathscr{R}(B,C;M,N) = -1$, and the statement is now shown.

3. The motivation is the problem 1. Choose a point K outside AB and point L on AK different from A and K. Let $M = BL \cap CK$ and $N = BK \cap AM$. Now let us construct a point D as $D = AB \cap LN$. From the problem 1 we indeed have $\mathscr{H}(A,B;C,D)$.

4. Let us denote $D = AS \cap BC$. According to the problem 1 we have $\mathscr{H}(R,S;A,D)$. Now we construct the point $D = AS \cap BC$. We have the points A, D, and S, hence according to the previous problem we can construct a point R such that $\mathscr{H}(A,D;S,R)$. Now we construct $P = BS \cap CR$ and $Q = CS \cap BR$, which solves the problem.

5. It is well known (and is easy to prove using Ceva's theorem) that the lines AD, BE, and CF intersect at a point G (called a Gergonne point of $\triangle ABC$) Let $X = BC \cap EF$. As in the problem 1 we have $\mathscr{H}(B,C;D,X)$. If we denote $X' = BC \cap PQ$ we analogously have $\mathscr{H}(B,C;D,X')$, hence $X = X'$.

6. Let us denote $x = KL$, $y = LM$, $z = MN$. We have to prove that $x+y+z \geq 3y$, or equivalently $x+z \geq 2y$. Since $\mathscr{R}(K,N;L,M) = \mathscr{R}(B,C;D,E)$, we have

$$\frac{x}{y+z} : \frac{x+y}{z} = \frac{\overrightarrow{KL}}{\overrightarrow{LN}} : \frac{\overrightarrow{KM}}{\overrightarrow{MN}} = \frac{\overrightarrow{BD}}{\overrightarrow{DC}} : \frac{\overrightarrow{BE}}{\overrightarrow{EC}} = \frac{1}{2} : \frac{1}{2},$$

implying $4xz = (x+y)(y+z)$.

If it were $y > (x+z)/2$ we would have

$$x+y > \frac{3}{2}x + \frac{1}{2}z = 2\frac{1}{4}(x+x+x+z) \geq 2\sqrt[4]{xxxz},$$

and analogously $y+z > 2\sqrt[4]{xzzz}$ as well as $(x+y)(y+z) > 4xz$ which is a contradiction. Hence the assumption $y > (x+z)/2$ was false so we have $y \leq (x+z)/2$.

Let us analyze the case of equality. If $y = (x+z)/2$, then $4xz = (x+y)(x+z) = (3x+z)(x+3z)/4$, which is equivalent to $(x-z)^2 = 0$. Hence the equality holds if $x = y = z$. We leave to the reader to prove that $x = y = z$ is satisfied if and only if $p \parallel BC$.

7. Let $E = AB \cap CD$, $F = AD \cap BC$. Consider the sequence of perspectivities

$$ABEM_1 \overset{D}{\barwedge} FBCM_2 \overset{A}{\barwedge} DECM_3 \overset{B}{\barwedge} DAFM_4 \overset{C}{\barwedge} EABM_5. \tag{8}$$

According to the conditions given in the problem this sequence of perspectivites has two be applied three more times to arrive to the point M_{13}. Notice that the given sequence of perspectivities maps A to E, E to B, and B to A. Clearly if we apply (8) three times the points A, B, and E will be fixed while M_1 will be mapped to M_{13}. Thus $M_1 = M_{13}$.

8. Let X' be the point symmetric to Y with respect to P. Notice that

$$\begin{aligned}
\mathscr{R}(M,N;X,P) &= \mathscr{R}(M,N;P,Y) \quad (\text{from } MNXP \overset{D}{\barwedge} MNAC \overset{B}{\barwedge} MNPY) \\
&= \mathscr{R}(N,M;P,X') \quad (\text{ the reflection with the center } P \text{ preserves} \\
&\qquad\qquad \text{the ratio, hence it preserves the cross-ratio}) \\
&= \frac{1}{\mathscr{R}(N,M;X',P)} = \mathscr{R}(M,N;X',P),
\end{aligned}$$

where the last equality follows from the basic properties of the cross ratio. It follows that $X = X'$.

9. Let $J = DQ \cap BP$, $K = EQ \cap CP$. If we prove that $JK \parallel DE$ this would imply that the triangles BDJ and CEK are perspective with the respect to a center, hence with repsect to an axis as well (according to Desargue's theorem) which immediately implies that A, P, Q are colinear (we encourage the reader to verify this fact).

Now we will prove that $JK \| DE$. Let us denote $T = DE \cap PQ$. Applying the Menelaus theorem on the triangle DTQ and the line PF we get

$$\frac{\overrightarrow{DJ}}{\overrightarrow{JQ}} \frac{\overrightarrow{QP}}{\overrightarrow{PT}} \frac{\overrightarrow{TF}}{\overrightarrow{FD}} = -1.$$

Similarly from the triangle ETQ and the line PG:

$$\frac{\overrightarrow{EK}}{\overrightarrow{KQ}} \frac{\overrightarrow{QP}}{\overrightarrow{PT}} \frac{\overrightarrow{TG}}{\overrightarrow{GE}} = -1.$$

Dividing the last two equalities and using $DT \cdot TG = FT \cdot TE$ (T is on the radical axis of the circumcircles of $\triangle DPG$ and $\triangle FPE$), we get

$$\frac{\overrightarrow{DJ}}{\overrightarrow{JQ}} = \frac{\overrightarrow{EK}}{\overrightarrow{KQ}}.$$

Thus $JK \parallel DE$, q.e.d.

10. Apply the inversion with the respect to I. We leave to the reader to draw the inverse picture. Notice that the condition that I is the incentar now reads that the circumcircles $A_i^* A_{i+1}^* I$ are of the same radii. Indeed if R is the radius of the circle of inversion and r the distance between I and XY then the radius of the circumcircle of $\triangle IX^*Y^*$ is equal to R^2/r. Now we use the following statement that is very easy to prove: "Let k_1, k_2, k_3 be three circles such that all pass through the same point I, but no two of them are mutually tangent. Then the centers of these circles are colinear if and only if there exists another common point $J \neq I$ of these three circles."

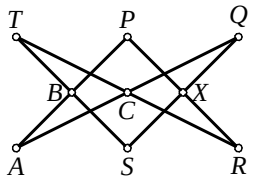
In the inverse picture this transforms into proving that the lines $A_1^* B_1^*$, $A_2^* B_2^*$, and $A_3^* B_3^*$ intersect at a point.

In order to prove this it is enough to show that the corresponding sides of the triangles $A_1^*A_2^*A_3^*$ and $B_1^*B_2^*B_3^*$ are parallel (then these triangles would be perspective with respect to the infinitely far line). Afterwards the Desargue's theorem would imply that the triangles are perspective with respect to a center. Let P_i^* be the incenter of $A_{i+1}^*A_{i+2}^*I$, and let Q_i^* be the foot of the perpendicular from I to $P_{i+1}^*P_{i+2}^*$. It is easy to prove that

$$\overrightarrow{A_1^*A_2^*} = 2\overrightarrow{Q_1^*Q_2^*} = -\overrightarrow{P_1^*P_2^*}.$$

Also since the circles $A_i^*A_{i+1}^*I$ are of the same radii, we have $P_1^*P_2^* \parallel B_1^*B_2^*$, hence $A_1^*A_2^* \parallel B_1^*B_2^*$.

11. We will prove that the intersection X of PR and QS lies on the line BC. Notice that the points P, Q, R, S belong to the circle with center AT. Consider the six points A, S, R, T, P, Q that lie on a circle. Using Pascal's theorem with respect to the diagram



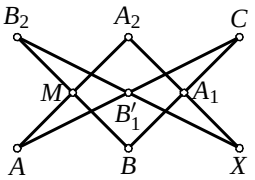
we get that the points $B, C,$ and $X = PR \cap QS$ are colinear.

12. *First solution, using projective mappings.* Let $A_3 = AM \cap BC$ and $B_3 = BM \cap AC$. Let X be the other intersection point of the line A_1A_2 with the circumcircle k of $\triangle ABC$. Let X' be the other intersection point of the line B_1B_2 with k. Consider the sequence of perspectivities

$$ABCX \overset{A_2}{\curlywedge} A_3BCA_1 \overset{M}{\curlywedge} AB_3CB_1 \overset{B_2}{\curlywedge} ABCX'$$

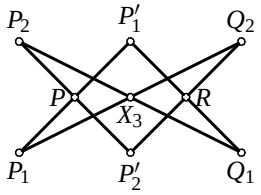
which has three fixed points $A, B, C,$ hence $X = X'$. Analogously the line C_1C_2 contains X and the problem is completely solved.

Second solution, using Pascal's theorem. Assume that the line A_1A_2 intersect the circumcircle of the trianlge ABC at A_2 and X. Let $XB_2 \cap AC = B_1'$. Let us apply the Pascal's theorem on the points A, B, C, A_2, B_2, X according the diagram:



It follows that the points A_1, B_1', and M are colinear. Hence $B_1' \in A_1M$. According to the definition of the point B_1' we have $B_1' \in AC$ hence $B_1' = A_1M \cap AC = B_1$. The conclusion is that the points X, B_1, B_2 are colinear. Analogously we prove that the points X, C_1, C_2 are colinear, hence the lines A_1A_2, B_1B_2, C_1C_2 intersect at X that belongs to the circumcircle of the triangle ABC.

13. It is well known (from the theory of pedal triangles) that pedal triangles corresponding to the isogonally conjugated points have the common circumcircle, so called *pedal circle* of the points P and Q. The center of that circle which is at the same time the midpoint of PQ will be denoted by R. Let $P_1' = PP_1 \cap Q_1R$ and $P_2' = PP_2 \cap Q_2R$ (the points P_1' and P_2' belong to the pedal circle of the point P, as point on the same diameters as Q_1 and Q_2 respectively). Using the Pascal's theorem on the points $Q_1, P_2, P_2', Q_2, P_1, P_1'$ in the order shown by the diagram



we get that the points P, R, X_1 are colinear or $X_1 \in PQ$. Analogously the points X_2, X_3 belong to the line PQ.

14. Let us recall the statement according to which the circle l is invariant under the inversion with respect to the circle k if and only if $l = k$ or $l \perp k$.

    Since the point M belongs to the polar of the point A with respect to k we have $\angle MA^*A = 90^\circ$ where $A^* = \psi_l(A)$. Therefore $A^* \in l$ where l is the circle with the radius AM. Analogously $M^* \in l$. However from $A \in l$ we get $A^* \in l^*$; $A^* \in l$ yields $A \in l^*$ (the inversion is inverse to itself) hence $\psi_l(A^*) = A$). Similarly we get $M \in l^*$ and $M^* \in l^*$. Notice that the circles l and l^* have the four common points A, A^*, M, M^*, which is exactly two too much. Hence $l = l^*$ and according to the statement mentioned at the beginning we conclude $l = k$ or $l \perp k$. The case $l = k$ can be easily eliminated, because the circle l has the diameter AM, and AM can't be the diameter of k because A and M are conjugated to each other.

    Thus $l \perp k$, q.e.d.

15. Let $J = KL \cap MN$, $R = l \cap MN$, $X_\infty = l \cap AM$. Since MN is the polar of A from $J \in MN$ we get $\mathscr{H}(K,L;J,A)$. From $KLJA \overset{M}{\barwedge} PQRX_\infty$ we also have $\mathscr{H}(P,Q;R,X_\infty)$. This implies that R is the midpoint of PQ.

16. Let Q be the intersection point of the lines AP and BC. Let Q' be the point of BC such that the ray AQ' is isogonal to the ray AQ in the triangle ABC. This exactly means that $\angle Q'AC = \angle BAQ$ i $\angle BAQ' = \angle QAC$.

    For an arbitrary point X of the segment BC, the sine theorem applied to triangles BAX and XAC yields

$$\frac{BX}{XC} = \frac{BX}{AX}\frac{AX}{XC} = \frac{\sin\angle BAX}{\sin\angle ABX}\frac{\sin\angle ACX}{\sin\angle XAC} = \frac{\sin\angle ACX}{\sin\angle ABX}\frac{\sin\angle BAX}{\sin\angle XAC} = \frac{AB}{AC}\frac{\sin\angle BAX}{\sin\angle XAC}.$$

    Applying this to $X = Q$ and $X = Q'$ and multiplying together afterwards we get

$$\frac{BQ}{QC}\frac{BQ'}{Q'C} = \frac{AB}{AC}\frac{\sin\angle BAQ}{\sin\angle QAC}\frac{AB}{AC}\frac{\sin\angle BAQ'}{\sin\angle Q'AC} = \frac{AB^2}{AC^2}. \tag{9}$$

    Hence if we prove $BQ/QC = AB^2/AC^2$ we would immediately have $BQ'/Q'C = 1$, making Q' the midpoint of BC. Then the line AQ is isogonaly conjugated to the median, implying the required statement.

    Since P belongs to the polars of B and C, then the points B and C belong to the polar of the point P, and we conclude that the polar of P is precisely BC. Consider the intersection D of the line BC with the tangent to the circumcircle at A. Since the point D belongs to the polars of A and P, AP has to be the polar of D. Hence $\mathscr{H}(B,C;D,Q)$. Let us now calculate the ratio BD/DC. Since the triangles ABD and CAD are similar we have $BD/AD = AD/CD = AB/AC$. This implies $BD/CD = (BD/AD)(AD/CD) = AB^2/AC^2$. The relation $\mathscr{H}(B,C;D,Q)$ implies $BQ/QC = BD/DC = AB^2/AC^2$, which proves the statement.

17. The statement follows from the Brianchon's theorem applied to $APBMCN$.

18. Applying the Brianchon's theorem to the hexagon $AMBCPD$ we get that the line MP contains the intersection of AB and CD. Analogously, applying the Brianchon's theorem to $ABNCDQ$ we get that NQ contains the same point.

19. The Brokard's theorem claims that the polar of $F = AD \cap BC$ is the line $f = EO$. Since the polar of the point on the circle is equal to the tangent at that point we know that $K = a \cap d$, where a and d are polars of the points A and D. Thus $k = AD$. Since $F \in AD = k$, we have $K \in f$ as well. Analogously we can prove that $L \in f$, hence the points E, O, K, L all belong to f.

20. Let $G = AD \cap BC$. Let k be the circumcircle of $ABCD$. Denote by k_1 and k_2 respectively the circumcircles of $\triangle ADF$ and $\triangle BCF$. Notice that AD is the radical axis of the circles k and k_1; BC the radical axis of k and k_2; and FH the radical axis of k_1 and k_2. According to the famous theorem these three radical axes intersect at one point G. In other words we have shown that the points F, G, H are colinear.

    Without loss of generality assume that F is between G and H (alternatively, we could use the oriented angles). Using the inscribed quadrilaterals $ADFH$ and $BCFH$, we get $\angle DHF = \angle DAF = \angle DAC$ and $\angle FHC = \angle FBC = \angle DBC$, hence $\angle DHC = \angle DHF + \angle FHC = \angle DAC + \angle DBC = 2\angle DAC = \angle DOC$. Thus the points D, C, H, and O lie on a circle. Similarly we prove that the points A, B, H, O lie on a circle.

    Denote by k_3 and k_4 respectively the circles circumscribed about the quadrilaterals $ABHO$ and $DCHO$. Notice that the line AB is the radical axis of the circles k and k_3. Simlarly CD and OH, respectively, are those of the pairs of circles (k, k_2) and (k_3, k_4). Thus these lines have to intersect at one point, and that has to be E. This proves that the points O, H, and E are colinear.

    According to the Brocard's theorem we have $FH \perp OE$, which according to $FH = GH$ and $OE = HE$ in turn implies that $GH \perp HE$, q.e.d.

terug naar echt bestand

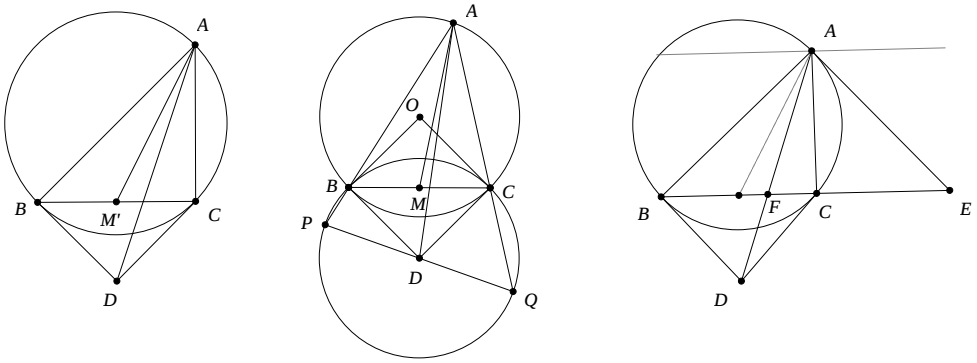
Lemmas in Euclidean Geometry[1]

Yufei Zhao

yufeiz@mit.edu

1. **Construction of the symmedian.**

Let ABC be a triangle and Γ its circumcircle. Let the tangent to Γ at B and C meet at D. Then AD coincides with a symmedian of $\triangle ABC$. (The *symmedian* is the reflection of the median across the angle bisector, all through the same vertex.)



We give three proofs. The first proof is a straightforward computation using Sine Law. The second proof uses similar triangles. The third proof uses projective geometry.

First proof. Let the reflection of AD across the angle bisector of $\angle BAC$ meet BC at M'. Then

$$\frac{BM'}{M'C} = \frac{AM' \frac{\sin \angle BAM'}{\sin \angle ABC}}{AM' \frac{\sin \angle CAM'}{\sin \angle ACB}} = \frac{\sin \angle BAM'}{\sin \angle ACD} \frac{\sin \angle ABD}{\sin \angle CAM'} = \frac{\sin \angle CAD}{\sin \angle ACD} \frac{\sin \angle ABD}{\sin \angle BAD} = \frac{CD}{AD} \frac{AD}{BD} = 1$$

Therefore, AM' is the median, and thus AD is the symmedian. □

Second proof. Let O be the circumcenter of ABC and let ω be the circle centered at D with radius DB. Let lines AB and AC meet ω at P and Q, respectively. Since $\angle PBQ = \angle DQC + \angle BAC = \frac{1}{2}(\angle BDC + \angle DOC) = 90^\circ$, we see that PQ is a diameter of ω and hence passes through D. Since $\angle ABC = \angle AQP$ and $\angle ACB = \angle APQ$, we see that triangles ABC and AQP are similar. If M is the midpoint of BC, noting that D is the midpoint of QP, the similarity implies that $\angle BAM = \angle QAD$, from which the result follows. □

Third proof. Let the tangent of Γ at A meet line BC at E. Then E is the pole of AD (since the polar of A is AE and the pole of D is BC). Let BC meet AD at F. Then point B, C, E, F are harmonic. This means that line AB, AC, AE, AF are harmonic. Consider the reflections of the four line across the angle bisector of $\angle BAC$. Their images must be harmonic too. It's easy to check that AE maps onto a line parallel to BC. Since BC must meet these four lines at harmonic points, it follows that the reflection of AF must pass through the midpoint of BC. Therefore, AF is a symmedian. □

[1]Updated July 26, 2008

Related problems:

(i) (Poland 2000) Let ABC be a triangle with $AC = BC$, and P a point inside the triangle such that $\angle PAB = \angle PBC$. If M is the midpoint of AB, then show that $\angle APM + \angle BPC = 180^\circ$.

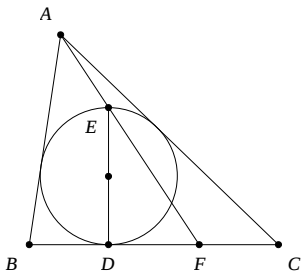
(ii) (IMO Shortlist 2003) Three distinct points A, B, C are fixed on a line in this order. Let Γ be a circle passing through A and C whose center does not lie on the line AC. Denote by P the intersection of the tangents to Γ at A and C. Suppose Γ meets the segment PB at Q. Prove that the intersection of the bisector of $\angle AQC$ and the line AC does not depend on the choice of Γ.

(iii) (Vietnam TST 2001) In the plane, two circles intersect at A and B, and a common tangent intersects the circles at P and Q. Let the tangents at P and Q to the circumcircle of triangle APQ intersect at S, and let H be the reflection of B across the line PQ. Prove that the points A, S, and H are collinear.

(iv) (USA TST 2007) Triangle ABC is inscribed in circle ω. The tangent lines to ω at B and C meet at T. Point S lies on ray BC such that $AS \perp AT$. Points B_1 and C_1 lies on ray ST (with C_1 in between B_1 and S) such that $B_1T = BT = C_1T$. Prove that triangles ABC and AB_1C_1 are similar to each other.

(v) (USA 2008) Let ABC be an acute, scalene triangle, and let M, N, and P be the midpoints of BC, CA, and AB, respectively. Let the perpendicular bisectors of AB and AC intersect ray AM in points D and E respectively, and let lines BD and CE intersect in point F, inside of triangle ABC. Prove that points A, N, F, and P all lie on one circle.

2. **Diameter of the incircle.**



Let the incircle of triangle ABC touch side BC at D, and let DE be a diameter of the circle. If line AE meets BC at F, then $BD = CF$.

Proof. Consider the dilation with center A that carries the incircle to an excircle. The diameter DE of the incircle must be mapped to the diameter of the excircle that is perpendicular to BC. It follows that E must get mapped to the point of tangency between the excircle and BC. Since the image of E must lie on the line AE, it must be F. That is, the excircle is tangent to BC at F. Then, it follows easily that $BD = CF$. □

Related problems:

(i) (IMO Shortlist 2005) In a triangle ABC satisfying $AB + BC = 3AC$ the incircle has centre I and touches the sides AB and BC at D and E, respectively. Let K and L be the symmetric points of D and E with respect to I. Prove that the quadrilateral $ACKL$ is cyclic.

(ii) (IMO 1992) In the plane let $\mathcal{C}$ be a circle, ℓ a line tangent to the circle $\mathcal{C}$, and M a point on ℓ. Find the locus of all points P with the following property: there exists two points Q, R on ℓ such that M is the midpoint of QR and $\mathcal{C}$ is the inscribed circle of triangle PQR.

(iii) (USAMO 1999) Let $ABCD$ be an isosceles trapezoid with $AB \parallel CD$. The inscribed circle ω of triangle BCD meets CD at E. Let F be a point on the (internal) angle bisector of $\angle DAC$ such that $EF \perp CD$. Let the circumscribed circle of triangle ACF meet line CD at C and G. Prove that the triangle AFG is isosceles.

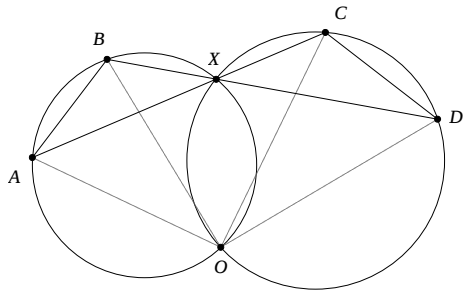
(iv) (USAMO 2001) Let ABC be a triangle and let ω be its incircle. Denote by D_1 and E_1 the points where ω is tangent to sides BC and AC, respectively. Denote by D_2 and E_2 the points on sides BC and AC, respectively, such that $CD_2 = BD_1$ and $CE_2 = AE_1$, and denote by P the point of intersection of segments AD_2 and BE_2. Circle ω intersects segment AD_2 at two points, the closer of which to the vertex A is denoted by Q. Prove that $AQ = D_2P$.

(v) (Tournament of Towns 2003 Fall) Triangle ABC has orthocenter H, incenter I and circumcenter O. Let K be the point where the incircle touches BC. If IO is parallel to BC, then prove that AO is parallel to HK.

(vi) (IMO 2008) Let $ABCD$ be a convex quadrilateral with $|BA| \neq |BC|$. Denote the incircles of triangles ABC and ADC by ω_1 and ω_2 respectively. Suppose that there exists a circle ω tangent to the ray BA beyond A and to the ray BC beyond C, which is also tangent to the lines AD and CD. Prove that the common external tangents of ω_1 and ω_2 intersect on ω.

3. **Dude, where's my spiral center?**

Let AB and CD be two segments, and let lines AC and BD meet at X. Let the circumcircles of ABX and CDX meet again at O. Then O is the center of the spiral similarity that carries AB to CD.



Proof. Since $ABOX$ and $CDXO$ are cyclic, we have $\angle OBD = \angle OAC$ and $\angle OCA = \angle ODB$. It follows that triangles AOC and BOD are similar. The result is immediate. □

Remember that spiral similarities always come in pairs: if there is a spiral similarity that carries AB to CD, then there is one that carries AC to BD.

Related problems:

(i) (IMO Shortlist 2006) Let $ABCDE$ be a convex pentagon such that

$$\angle BAC = \angle CAD = \angle DAE \qquad \text{and} \qquad \angle CBA = \angle DCA = \angle EDA.$$

Diagonals BD and CE meet at P. Prove that line AP bisects side CD.

(ii) (China 1992) Convex quadrilateral $ABCD$ is inscribed in circle ω with center O. Diagonals AC and BD meet at P. The circumcircles of triangles ABP and CDP meet at P and Q. Assume that points O, P, and Q are distinct. Prove that $\angle OQP = 90^\circ$.

(iii) Let $ABCD$ be a quadrilateral. Let diagonals AC and BD meet at P. Let O_1 and O_2 be the circumcenters of APD and BPC. Let M, N and O be the midpoints of AC, BD and O_1O_2. Show that O is the circumcenter of MPN.

(iv) (USAMO 2006) Let $ABCD$ be a quadrilateral, and let E and F be points on sides AD and BC, respectively, such that $AE/ED = BF/FC$. Ray FE meets rays BA and CD at S and T, respectively. Prove that the circumcircles of triangles SAE, SBF, TCF, and TDE pass through a common point.

(v) (IMO 2005) Let $ABCD$ be a given convex quadrilateral with sides BC and AD equal in length and not parallel. Let E and F be interior points of the sides BC and AD respectively such that $BE = DF$. The lines AC and BD meet at P, the lines BD and EF meet at Q, the lines EF and AC meet at R. Consider all the triangles PQR as E and F vary. Show that the circumcircles of these triangles have a common point other than P.

(vi) (IMO Shortlist 2002) Circles S_1 and S_2 intersect at points P and Q. Distinct points A_1 and B_1 (not at P or Q) are selected on S_1. The lines A_1P and B_1P meet S_2 again at A_2 and B_2 respectively, and the lines A_1B_1 and A_2B_2 meet at C. Prove that, as A_1 and B_1 vary, the circumcentres of triangles A_1A_2C all lie on one fixed circle.

(vii) (USA TST 2006) In acute triangle ABC, segments AD, BE, and CF are its altitudes, and H is its orthocenter. Circle ω, centered at O, passes through A and H and intersects sides AB and AC again at Q and P (other than A), respectively. The circumcircle of triangle OPQ is tangent to segment BC at R. Prove that $CR/BR = ED/FD$.

(viii) (IMO Shortlist 2006) Points A_1, B_1 and C_1 are chosen on sides BC, CA, and AB of a triangle ABC, respectively. The circumcircles of triangles AB_1C_1, BC_1A_1, and CA_1B_1 intersect the circumcircle of triangle ABC again at points A_2, B_2, and C_2, respectively ($A_2 \neq A, B_2 \neq B$, and $C_2 \neq C$). Points A_3, B_3, and C_3 are symmetric to A_1, B_1, C_1 with respect to the midpoints of sides BC, CA, and AB, respectively. Prove that triangles $A_2B_2C_2$ and $A_3B_3C_3$ are similar.

4. **Arc midpoints are equidistant to vertices and in/excenters**

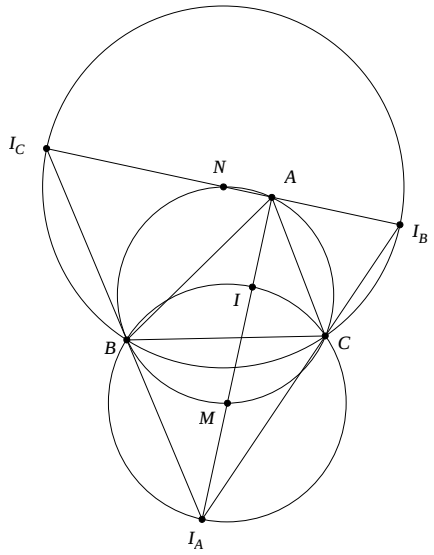
Let ABC be a triangle, I its incenter, and I_A, I_B, I_C its excenters. On the circumcircle of ABC, let M be the midpoint of the arc BC not containing A and let N be the midpoint of the arc BC containing A. Then $MB = MC = MI = MI_A$ and $NB = NC = NI_B = NI_C$.

Proof. Straightforward angle-chasing (do it yourself!). Another perspective is to consider the circumcircle of ABC as the nine-point-circle of $I_AI_BI_C$. □

Related problems:

(i) (APMO 2007) Let ABC be an acute angled triangle with $\angle BAC = 60^\circ$ and $AB > AC$. Let I be the incenter, and H the orthocenter of the triangle ABC. Prove that $2\angle AHI = 3\angle ABC$.

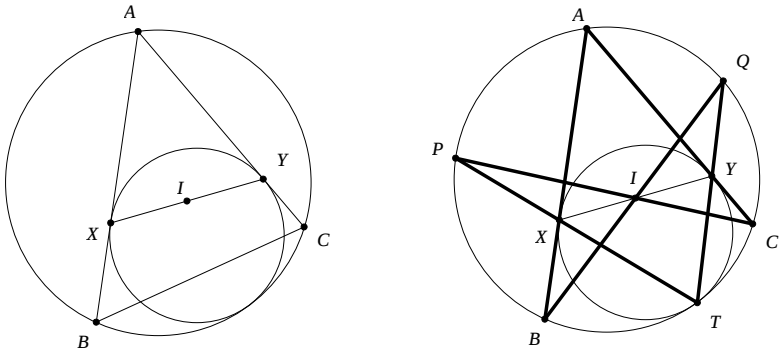
(ii) (IMO 2006) Let ABC be a triangle with incentre I. A point P in the interior of the triangle satisfies $\angle PBA + \angle PCA = \angle PBC + \angle PCB$. Show that $AP \geq AI$, and that equality holds if and only if $P = I$.



(iii) (Romanian TST 1996) Let $ABCD$ be a cyclic quadrilateral and let $\mathcal{M}$ be the set of incenters and excenters of the triangles BCD, CDA, DAB, ABC (16 points in total). Prove that there are two sets $\mathcal{K}$ and $\mathcal{L}$ of four parallel lines each, such that every line in $\mathcal{K} \cup \mathcal{L}$ contains exactly four points of $\mathcal{M}$.

5. **I is the midpoint of the touch-chord of the mixtilinear incircles**

Let ABC be a triangle and I its incenter. Let Γ be the circle tangent to sides AB, AC, as well as the circumcircle of ABC. Let Γ touch AB and AC at X and Y, respectively. Then I is the midpoint of XY.



Proof. Let the point of tangency between the two circles be T. Extend TX and TY to meet the circumcircle of ABC again at P and Q respectively. Note that P and Q are the midpoint of the arcs AB and AC. Apply Pascal's theorem to $BACPTQ$ and we see that X, I, Y are collinear. Since I lies on the angle bisector of $\angle XAY$ and $AX = AY$, I must be the midpoint of XY. □

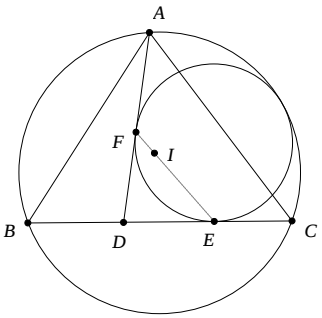
Related problems:

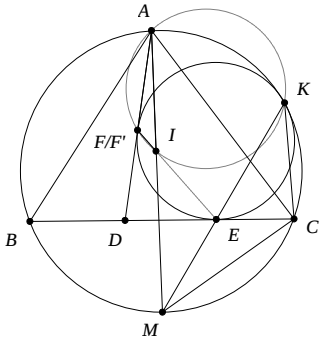
(i) (IMO 1978) In triangle ABC, $AB = AC$. A circle is tangent internally to the circumcircle of triangle ABC and also to sides AB, AC at P, Q, respectively. Prove that the midpoint of segment PQ is the center of the incircle of triangle ABC.

(ii) Let ABC be a triangle. Circle ω is tangent to AB and AC, and internally tangent to the circumcircle of triangle ABC. The circumcircle and ω are tangent at P. Let I be the incircle of triangle ABC. Line PI meets the circumcircle of ABC at P and Q. Prove that $BQ = CQ$.

6. **More curvilinear incircles.**

(A generalization of the previous lemma) Let ABC be a triangle, I its incenter and D a point on BC. Consider the circle that is tangent to the circumcircle of ABC but is also tangent to DC, DA at E, F respectively. Then E, F and I are collinear.





Proof. There is a "computational" proof using Casey's theorem[2] and transversal theorem[3]. You can try to work that out yourself. Here, we show a clever but difficult synthetic proof (communicated to me via Oleg Golberg).

Denote Ω the circumcircle of ABC and Γ the circle tangent tangent to the circumcircle of ABC and lines DC, DA. Let Ω and Γ touch at K. Let M be the midpoint of arc $\widehat{BC}$ on Ω not containing K. Then K, E, M are collinear (think: dilation with center K carrying Γ to Ω). Also, A, I, M are collinear, and $MI = MC$.

Let line EI meet Γ again at F'. It suffices to show that AF' is tangent to Γ.

Note that $\angle KF'E$ is subtended by $\widehat{KE}$ in Γ and $\angle KAM$ is subtended by $\widehat{KM}$ in Ω. Since $\widehat{KE}$ and $\widehat{KM}$ are homothetic with center K, we have $\angle KF'E = \angle KAM$, implying that A, K, I', F' are concyclic.

We have $\angle BCM = \angle CBM = \angle CKM$. So $\triangle MCE \sim \triangle MKC$. Hence $MC^2 = ME \cdot MK$. Since $MC = MI$, we have $MI^2 = ME \cdot MK$, implying that $\triangle MIE \sim \triangle MKI$. Therefore,

[2] **Casey's theorem**, also known as Generalized Ptolemy Theorem, states that if there are four circles $\Gamma_1, \Gamma_2, \Gamma_3, \Gamma_4$ (could be degenerated into a point) all touching a circle Γ such that their tangency points follow that order around the circle, then

$$t_{12}t_{34} + t_{23}t_{14} = t_{13}t_{24},$$

where t_{12} is the length of the common tangent between Γ_i and Γ_j (if Γ_i and Γ_j on the same side of Γ, then take their common external tangent, else take their common internal tangent.) I think the converse is also true—if both equations hold, then there is some circle tangent to all four circles.

[3] The **transversal theorem** is a criterion for collinearity. It states that if A, B, C are three collinear points, and P is a point not on the line ABC, and A', B', C' are arbitrary points on lines PA, PB, PC respectively, then A', B', C' are collinear if and only if

$$BC \cdot \frac{AP}{A'P} + CA \cdot \frac{BP}{B'P} + AB \cdot \frac{CP}{C'P} = 0,$$

where the lengths are directed. In my opinion, it's much easier to remember the proof than to memorize this huge formula. The simplest derivation is based on relationships between the areas of $[PAB], [PA'B']$, etc.

$\angle KEI = \angle AIK = \angle AF'K$ (since A, K, I, F' are concyclic). Therefore, AF' is tangent to Ω and the proof is complete. □

Related problems:

(i) (Bulgaria 2005) Consider two circles k_1, k_2 touching externally at point T. A line touches k_2 at point X and intersects k_1 at points A and B. Let S be the second intersection point of k_1 with the line XT . On the arc $\widehat{TS}$ not containing A and B is chosen a point C . Let CY be the tangent line to k_2 with $Y \in k_2$, such that the segment CY does not intersect the segment ST . If $I = XY \cap SC$. Prove that:

(a) the points C, T, Y, I are concyclic.

(b) I is the excenter of triangle ABC with respect to the side BC.

(ii) (Sawayama-Thébault[4]) Let ABC be a triangle with incenter I. Let D a point on side BC. Let P be the center of the circle that touches segments AD, DC, and the circumcircle of ABC, and let Q be the center of the circle that touches segments AD, BD, and the circumcircle of ABC. Show that P, Q, I are collinear.

(iii) Let P be a quadrilateral inscribed in a circle Ω, and let $\mathcal{Q}$ be the quadrilateral formed by the centers of the fourcircles internally touching O and each of the two diagonals of P. Show that the incenters of the four triangles having for sides the sides and diagonals of P form a rectangle R inscribed in Q.

(iv) (Romania 1997) Let ABC be a triangle with circumcircle Ω, and D a point on the side BC. Show that the circle tangent to Ω, AD and BD, and the circle tangent to Ω, AD and DC, are tangent to each other if and only if $\angle BAD = \angle CAD$.

(v) (Romania TST 2006) Let ABC be an acute triangle with $AB \neq AC$. Let D be the foot of the altitude from A and ω the circumcircle of the triangle. Let ω_1 be the circle tangent to AD, BD and ω. Let ω_2 be the circle tangent to AD, CD and ω. Let ℓ be the interior common tangent to both ω_1 and ω_2, different from CD. Prove that ℓ passes through the midpoint of BC if and only if $2BC = AB + AC$.

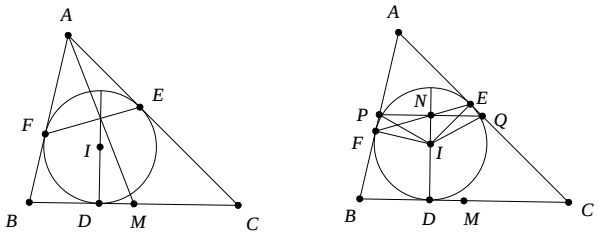
(vi) (AMM 10368) For each point O on diameter AB of a circle, perform the following construction. Let the perpendicular to AB at O meet the circle at point P. Inscribe circles in the figures bounded by the circle and the lines AB and OP. Let R and S be the points at which the two incircles to the curvilinear triangles AOP and BOP are tangent to the diameter AB. Show that $\angle RPS$ is independent of the position of O.

7. **Concurrent lines from the incircle.**

Let the incircle of ABC touch sides BC, CA, AB at D, E, F respectively. Let I be the incenter of ABC and M be the midpoint of BC. Then the lines EF, DI and AM are concurrent.

Proof. Let lines DI and EF meet at N. Construct a line through N parallel to BC, and let it meet sides AB and AC at P and Q, respectively. We need to show that A, N, M are collinear, so it suffices to show that N is the midpoint of PQ. We present two ways to finish this off, one using Simson's line, and the other using spiral similarities.

[4] A bit of history: this problem was posed by French geometer Victor Thébault (1882–1960) in the *American Mathematical Monthly* in 1938 (Problem 2887, 45 (1938) 482–483) and it remained unsolved until 1973. However, in 2003, Jean-Louis Ayme discovered that this problem was independently proposed and solved by instructor Y. Sawayama of the Central Military School of Tokyo in 1905! For more discussion, see Ayme's paper at `http://forumgeom.fau.edu/FG2003volume3/FG200325.pdf`



Simson line method: Consider the triangle APQ. The projections of the point I onto the three sides of APQ are D, N, F, which are collinear, I must lie on the circumcircle of APQ by Simson's theorem. But since AI is an angle bisector, $PI = QI$, thus $PN = QN$.

Spiral similarity method: Note that P, N, I, F are concyclic, so $\angle EFI = \angle QPI$. Similarly, $\angle PQI = \angle FEI$. So triangles PIQ and FIE are similar. Since $FI = EI$, we have $PI = QI$, and thus $PN = QN$. (c.f. Lemma 3) □

Related problems:

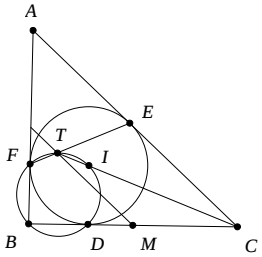
(i) (China 1999) In triangle ABC, $AB \neq AC$. Let D be the midpoint of side BC, and let E be a point on median AD. Let F be the foot of perpendicular from E to side BC, and let P be a point on segment EF. Let M and N be the feet of perpendiculars from P to sides AB and AC, respectively. Prove that M, E, and N are collinear if and only if $\angle BAP = \angle PAC$.

(ii) (IMO Shortlist 2005) The median AM of a triangle ABC intersects its incircle ω at K and L. The lines through K and L parallel to BC intersect ω again at X and Y. The lines AX and AY intersect BC at P and Q. Prove that $BP = CQ$.

8. **More circles around the incircle.**

Let I be the incenter of triangle ABC, and let its incircle touch sides BC, AC, AB at D, E and F, respectively. Let line CI meet EF at T. Then T, I, D, B, F are concyclic. Consequent results include: $\angle BTC = 90^\circ$, and T lies on the line connecting the midpoints of AB and BC.

An easier way to remember the third part of the lemma is: for a triangle ABC, draw a midline, an angle bisector, and a touch-chord, each generated from different vertex, then the three lines are concurrent.



Proof. Showing that I, T, E, B are concyclic is simply angle chasing (e.g. show that $\angle BIC = \angle BFE$). The second part follows from $\angle BTC = \angle BTI = \angle BFI = 90^\circ$. For the third part, note that if M is the midpoint of BC, then M is the midpoint of an hypotenuse of the right triangle BTC. So $MT = MC$. Then $\angle MTC = \angle MCT = \angle ACT$, so MT is parallel to AC, and so MT is a midline of the triangle. □

Related problems:

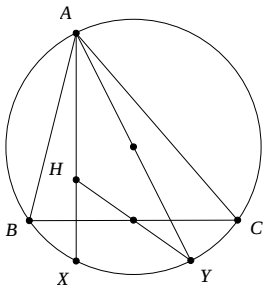
(i) Let ABC be an acute triangle whose incircle touches sides AC and AB at E and F, respectively. Let the angle bisectors of $\angle ABC$ and $\angle ACB$ meet EF at X and Y, respectively, and let the midpoint of BC be Z. Show that XYZ is equilateral if and only if $\angle A = 60^\circ$.

(ii) (IMO Shortlist 2004) For a given triangle ABC, let X be a variable point on the line BC such that C lies between B and X and the incircles of the triangles ABX and ACX intersect at two distinct points P and Q. Prove that the line PQ passes through a point independent of X.

(iii) Let points A and B lie on the circle Γ, and let C be a point inside the circle. Suppose that ω is a circle tangent to segments AC, BC and Γ. Let ω touch AC and Γ at P and Q. Show that the circumcircle of APQ passes through the incenter of ABC.

9. **Reflections of the orthocenter lie on the circumcircle.**

Let H be the orthocenter of triangle ABC. Let the reflection of H across the BC be X and the reflection of H across the midpoint of BC be Y. Then X and Y both lie on the circumcircle of ABC. Moreover, AY is a diameter of the circumcircle.



Proof. Trivial. Angle chasing. □

Related problems:

(i) Prove the existence of the nine-point circle. (Given a triangle, the nine-point circle is the circle that passes through the three midpoints of sides, the three feet of altitudes, and the three midpoints between the orthocenter and the vertices).

(ii) Let ABC be a triangle, and P a point on its circumcircle. Show that the reflections of P across the three sides of ABC lie on a lie that passes through the orthocenter of ABC.

(iii) (IMO Shortlist 2005) Let ABC be an acute-angled triangle with $AB \neq AC$, let H be its orthocentre and M the midpoint of BC. Points D on AB and E on AC are such that $AE = AD$ and D, H, E are collinear. Prove that HM is orthogonal to the common chord of the circumcircles of triangles ABC and ADE.

(iv) (USA TST 2005) Let $A_1A_2A_3$ be an acute triangle, and let O and H be its circumcenter and orthocenter, respectively. For $1 \leq i \leq 3$, points P_i and Q_i lie on lines OA_i and $A_{i+1}A_{i+2}$ (where $A_{i+3} = A_i$), respectively, such that OP_iHQ_i is a parallelogram. Prove that

$$\frac{OQ_1}{OP_1} + \frac{OQ_2}{OP_2} + \frac{OQ_3}{OP_3} \geq 3.$$

(v) (China TST quizzes 2006) Let ω be the circumcircle of triangle ABC, and let P be a point inside the triangle. Rays AP, BP, CP meet ω at A_1, B_1, C_1, respectively. Let A_2, B_2, C_2 be the images of A_1, B_1, C_1 under reflection about the midpoints of BC, CA, AB, respectively. Show that the orthocenter of ABC lies on the circumcircle of $A_2B_2C_2$.

10. **O and H are isogonal conjugates.**

Let ABC be a triangle, with circumcenter O, orthocenter H, and incenter I. Then AI is the angle bisector of $\angle HAO$.

Proof. Trivial. □

Related problems:

(i) (Crux) Points O and H are the circumcenter and orthocenter of acute triangle ABC, respectively. The perpendicular bisector of segment AH meets sides AB and AC at D and E, respectively. Prove that $\angle DOA = \angle EOA$.

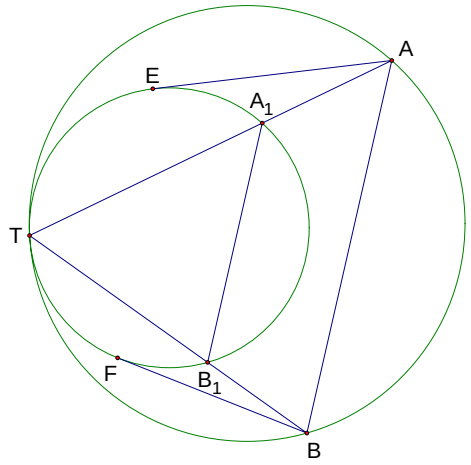
(ii) Show that $IH = IO$ if and only if one of $\angle A$, $\angle B$, $\angle C$ is 60°.

terug naar echt bestand

A Metric Relation and its Applications

Son Hong Ta

Lemma. Let γ be a circle and let A and B be two arbitrary points on it. A circle ρ touches γ internally at T. Denote by AE and BF the tangent lines to ρ at E and F, respectively. Then $\frac{TA}{TB} = \frac{AE}{BF}$.



Proof. Denote by A_1 and B_1 the second intersections of TA and TB with ρ, respectively. We know that A_1B_1 is parallel to AB. Therefore,

$$\left(\frac{AE}{TA_1}\right)^2 = \frac{AA_1 \cdot AT}{A_1T \cdot A_1T} = \frac{BB_1}{B_1T} \cdot \frac{BT}{B_1T} = \left(\frac{BF}{TB_1}\right)^2.$$

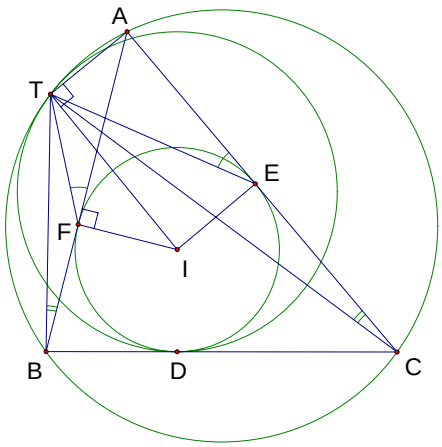
Hence,

$$\frac{AE}{TA_1} = \frac{BF}{TB_1} \implies \frac{AE}{BF} = \frac{TA_1}{TB_1} = \frac{TA}{TB},$$

which completes the proof. □

To illustrate how this lemma works, let us consider some examples. The following problem was proposed by Nguyen Minh Ha, in the Vietnamese Mathematics Magazine, in 2007.

Problem 1. Let Ω be the circumcircle of the triangle ABC and let D be the tangency point of its incircle $\rho(I)$ with the side BC. Let ω be the circle internally tangent to Ω at T, and to BC at D. Prove that $\measuredangle ATI = 90^\circ$.



Solution. Let E and F be the tangency points of $\rho(I)$ with sides CA and AB, respectively. According to the lemma,

$$\frac{TB}{TC} = \frac{BD}{CD} = \frac{BF}{CE}.$$

Therefore triangles TBF and TCE are similar. It follows that $\measuredangle TFA = \measuredangle TEA$, hence the points A, I, E, F, T lie on the same circle. It follows that $\measuredangle ATI = \measuredangle AFI = 90^\circ$ which completes our proof. □

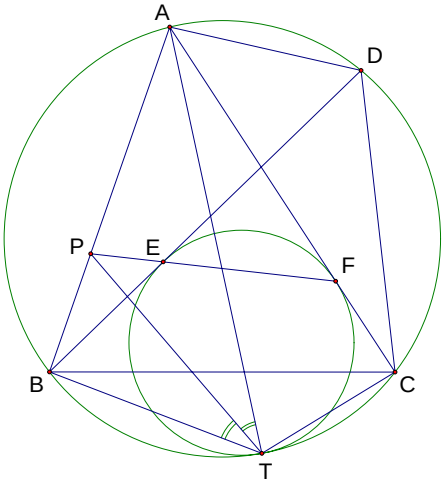
Problem 2. Let $ABCD$ be a quadrilateral inscribed in a circle Ω. Let ω be a circle internally tangent to Ω at T, and to DB and AC at E and F, respectively. Let P be the intersection of EF and AB. Prove that TP is the internal angle bisector of the angle $\measuredangle ATB$.

Solution. From our lemma, applied to circles Ω, ω and points A, B, we conclude that $\frac{AT}{BT} = \frac{AF}{BE}$, thus it suffices to prove that

$$\frac{AF}{BE} = \frac{AP}{PB}.$$

Indeed, notice that $\measuredangle PEB = \measuredangle AFP$, and from the Law of Sines, applied to triangles APF, BPE, we have

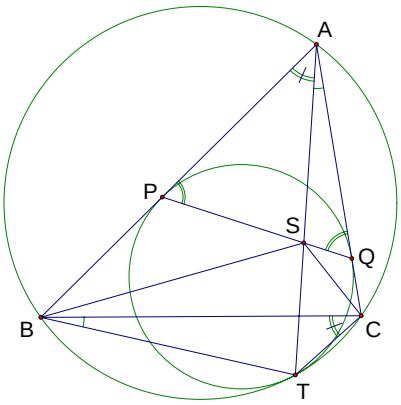
$$\frac{AP}{AF} = \frac{\sin \measuredangle AFP}{\sin \measuredangle APF} = \frac{\sin \measuredangle BEP}{\sin \measuredangle BPE} = \frac{BP}{BE}.$$



Therefore $\frac{AF}{BE} = \frac{AP}{PB}$, which completes our solution. □

The third problem comes from the Moldovan Team Selection Test in 2007, which can be found in [2] and [3].

Problem 3. Let ABC be a triangle and let Ω be its circumcircle. Circles ω is internally tangent to Ω at T, and to sides AB and AC at P and Q, respectively. Let S be the intersection of AT and PQ. Prove that $\measuredangle SBA = \measuredangle SCA$.



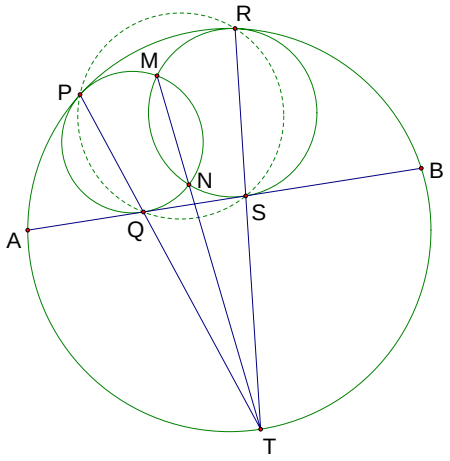
Solution. Using our lemma, we have

$$\frac{BP}{CQ} = \frac{BT}{CT} = \frac{\sin \measuredangle BCT}{\sin \measuredangle CBT} = \frac{\sin \measuredangle BAT}{\sin \measuredangle CAT} = \frac{PS}{QS}.$$

This fact implies that BPS and CQS are similar triangles which in turn implies that $\measuredangle SBA = \measuredangle SCA$. $\square$

Problem 4. Consider a circle (O) and a chord AB. Let circles (O_1), (O_2) be internally tangent to (O) and AB and let M and N their intersection. Prove that MN passes through the midpoint of the arc AB which does not contain M and N.

Solution. Denote by P and Q the tangency points of the circle (O_1) with (O) and AB, respectively. Let R and S be the tangency points of circle (O_2) with (O) and AB, respectively. Let T be the middle point of the arc AB which does not contain M and N.



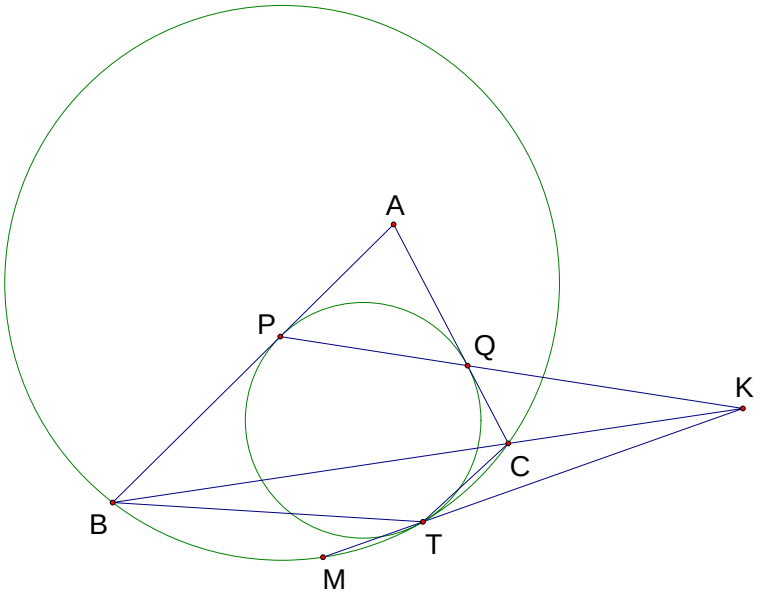
Applying the above lemma to circles (O), (O_1), and points A, B along with their tangent lines AQ, BQ to (O_1) we get $\frac{PA}{PB} = \frac{QA}{AB}$. This means that PQ passes through T. Similarly, RS passes through T. On the other hand, $\measuredangle PQA = \measuredangle QTA + \measuredangle QAT = \measuredangle PRA + \measuredangle ART = \measuredangle PRS$, therefore, points P, Q, R, S lie on a circle which we will denote by (O_3). We have that PQ is the radical axis of (O_1) and (O_3), RS is the radical axis of (O_2) and (O_3), and MN is the radical axis of (O_1) and (O_2). So, MN, PQ, and RS are concurrent at the radical center of the three circles. Hence, we deduce that MN passes through T, which is the midpoint of the arc AB that does not contain M and N. $\square$

We continue with a problem from the MOSP Tests 2007 [4].

Problem 5. Let ABC be a triangle. Circle ω passes through points B and C. Circle ω_1 is tangent internally to ω and also to the sides AB and AC at T, P, and Q, respectively. Let M be midpoint of arc BC (containing T) of ω. Prove that lines PQ, BC, and MT are concurrent.

Solution. Let $K = PQ \cap BC$ and let $K' = MT \cap BC$. Applying Menelaos' Theorem in triangle ABC we obtain

$$\frac{KB}{KB} \cdot \frac{QC}{QA} \cdot \frac{PA}{PB} = 1 \implies \frac{KB}{KC} = \frac{BP}{CQ}.$$



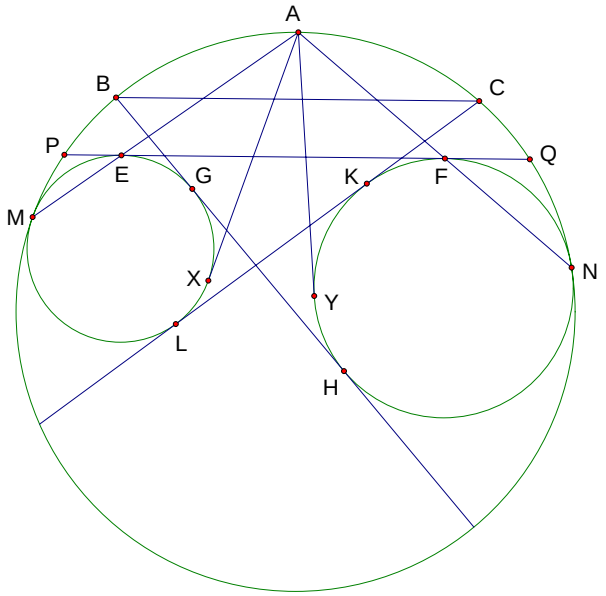
On the other hand, M is the midpoint of arc BC (containing T) of ω so MT is the external bisector of angle $\measuredangle BTC$, therefore $\frac{K'B}{K'C} = \frac{TB}{TC}$. Thus, we are left to prove that $\frac{BP}{CQ} = \frac{TB}{TC}$, which is true according to our lemma and we are done. □

The last problem was given in [5] and is also discussed and proved in [6]. Now, we will present another solution for this nice problem.

Problem 6. Circles (O_1) and (O_2) are internally tangent to a given circle (O) at M and N, respectively. Their internal common tangents intersect (O) at four points. Let B and C be two of them such that B and C lie on the same side with respect to O_1O_2. Prove that BC is parallel to an external common tangent of (O_1) and (O_2).

Solution. Draw the internal common tangents GH, KL of (O_1), (O_2) such that G and L lie on (O_1) and K and H lie on (O_2). Let EF be the external common tangent of (O_1), (O_2) such that E and B lie on the same side with respect to O_1O_2. Denote by P and Q the intersections of EF with (O). We will prove that BC is parallel to PQ. Denote by A be the midpoint of the arc PQ which does not contain M and N. Let AX and AY be the tangents at X and Y to the circles (O_1)

and (O_2). In the solution to Problem 4 we have proved that A, E, and M are collinear; A, F, and N are collinear, and the quadrilateral $MEFN$ is cyclic. Therefore, $AX^2 = AE \cdot AM = AF \cdot AN = AY^2$, i.e. $AX = AY$ (1).



Based on the lemma, $\frac{MA}{AX} = \frac{MB}{BG} = \frac{MC}{CL}$. On the other hand, by the Ptolemy's Theorem, $MA \cdot BC = MB \cdot AC = MC \cdot AB$, therefore

$$AX \cdot BC = BG \cdot AC = CL \cdot AB.$$

Similarly,

$$AY \cdot BC = BH \cdot AC + CK \cdot AB.$$

Thus $AC \cdot (BH - BG) = AB \cdot (CL - CK)$, i.e. $AC \cdot GH = AB \cdot KL$, which implies $AC = AB$. Hence, A is the midpoint of the arc BC of the circle (O). This means that BC is parallel to PQ and our solution is complete. □

References

[1] Mathlinks, *Nice geometry*, http://www.mathlinks.ro/viewtopic.php?t=170192

[2] Mathlinks, *A circle tangent to the circumcircle and two sides*, http://www.mathlinks.ro/viewtopic.php?t=140464

[3] Mathlinks, *Equal angle*,
http://www.mathlinks.ro/Forum/viewtopic.php?t=98968

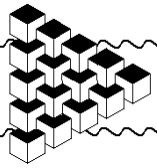
[4] 2007 Mathematical Olympiad Summer Program Tests, available at http://www.unl.edu/amc/a-activities/a6-mosp/archivemosp.shtml

[5] Shay Gueron, *Two Applications of the Generalized Ptolemy Theorem*, American Mathematical Monthly 2002.

[6] Mathlinks, *Parallel tangent*,
http://www.mathlinks.ro/viewtopic.php?t=15945

Son Hong Ta, High School for Gifted Students, Hanoi University of Education, Hanoi, Vietnam.
E-mail address: dam_xoan90@yahoo.com

terug naar echt bestand



©2007 The Author(s) and The IMO Compendium Group

Inversion

Dušan Djukić

Contents

1 General Properties

Inversion Ψ is a map of a plane or space without a fixed point O onto itself, determined by a circle k with center O and radius r, which takes point $A \neq O$ to the point $A' = \Psi(A)$ on the ray OA such that $OA \cdot OA' = r^2$. From now on, unless noted otherwise, X' always denotes the image of object X under a considered inversion.

Clearly, map Ψ is continuous and inverse to itself, and maps the interior and exterior of k to each other, which is why it is called "inversion". The next thing we observe is that $\triangle P'OQ' \sim \triangle QOP$ for all point $P, Q \neq O$ (for $\angle P'OQ' = \angle QOP$ and $OP'/OQ' = (r^2/OP)/(r^2/OQ) = OQ/OP$), with the ratio of similitude $\frac{r^2}{OP \cdot OQ}$. As a consequence, we have

$$\angle OQ'P' = \angle OPQ \quad \text{and} \quad P'Q' = \frac{r^2}{OP \cdot OQ} PQ.$$

What makes inversion attractive is the fact that it maps lines and circles into lines and circles. A line through O (O excluded) obviously maps to itself. What if a line p does not contain O? Let P be the projection of O on p and $Q \in p$ an arbitrary point of p. Angle $\angle OPQ = \angle OQ'P'$ is right, so Q' lies on circle k with diameter OP'. Therefore $\Psi(p) = k$ and consequently $\Psi(k) = p$. Finally, what is the image of a circle k not passing through O? We claim that it is also a circle; to show this, we shall prove that inversion takes any four concyclic points A, B, C, D to four concyclic points A', B', C', D'. The following angles are regarded as oriented. Let us show that $\angle A'C'B' = \angle A'D'B'$. We have $\angle A'C'B' = \angle OC'B' - \angle OC'A' = \angle OBC - \angle OAC$ and analogously $\angle A'D'B' = \angle OBD - \angle OAD$, which implies $\angle A'D'B' - \angle A'C'B' = \angle CBD - \angle CAD = 0$, as we claimed. To sum up:

- A line through O maps to itself.
- A circle through O maps to a line not containing O and vice-versa.
- A circle not passing through O maps to a circle not passing through O (not necessarily the same).

Remark. Based on what we have seen, it can be noted that inversion preserves angles between curves, in particular circles or lines. Maps having this property are called *conformal*.

When should inversion be used? As always, the answer comes with experience and cannot be put on a paper. Roughly speaking, inversion is useful in destroying "inconvenient" circles and angles on a picture. Thus, some pictures "cry" to be inverted:

- There are many circles and lines through the same point A. Invert through A.

Problem 1 (IMO 2003, shortlist). *Let $\Gamma_1,\Gamma_2,\Gamma_3,\Gamma_4$ be distinct circles such that Γ_1,Γ_3 are externally tangent at P, and Γ_2,Γ_4 are externally tangent at the same point P. Suppose that Γ_1 and Γ_2; Γ_2 and Γ_3; Γ_3 and Γ_4; Γ_4 and Γ_1 meet at A,B,C,D, respectively, and that all these points are different from P. Prove that*

$$\frac{AB\cdot BC}{AD\cdot DC}=\frac{PB^2}{PD^2}.$$

Solution. Apply the inversion with center at P and radius r; let $\widehat{X}$ denote the image of X. The circles $\Gamma_1,\Gamma_2,\Gamma_3,\Gamma_4$ are transformed into lines $\widehat{\Gamma_1},\widehat{\Gamma_2},\widehat{\Gamma_3},\widehat{\Gamma_4}$, where $\widehat{\Gamma_1}\parallel\widehat{\Gamma_3}$ and $\widehat{\Gamma_2}\parallel\widehat{\Gamma_4}$, and therefore $\widehat{A}\widehat{B}\widehat{C}\widehat{D}$ is a parallelogram. Further, we have $AB=\dfrac{r^2}{P\widehat{A}\cdot P\widehat{B}}\widehat{A}\widehat{B}$, $PB=\dfrac{r^2}{P\widehat{B}}$, etc. The equality to be proven becomes

$$\frac{P\widehat{D}^2}{P\widehat{B}^2}\cdot\frac{\widehat{A}\widehat{B}\cdot\widehat{B}\widehat{C}}{\widehat{A}\widehat{D}\cdot\widehat{D}\widehat{C}}=\frac{P\widehat{D}^2}{P\widehat{B}^2},$$

which holds because $\widehat{A}\widehat{B}=\widehat{C}\widehat{D}$ and $\widehat{B}\widehat{C}=\widehat{D}\widehat{A}$. $\triangle$

- There are many angles $\angle AXB$ with fixed A,B. Invert through A or B.

Problem 2 (IMO 1996, problem 2). *Let P be a point inside $\triangle ABC$ such that $\angle APB-\angle C=\angle APC-\angle B$. Let D,E be the incenters of $\triangle APB,\triangle APC$ respectively. Show that AP, BD, and CE meet in a point.*

Solution. Apply an inversion with center at A and radius r. Then the given condition becomes $\angle B'C'P'=\angle C'B'P'$, i.e., $B'P'=P'C'$. But $P'B'=\frac{r^2}{AP\cdot AB}PB$, so $AC/AB=PC/PB$. $\triangle$

Caution: Inversion may also bring new inconvenient circles and angles. Of course, keep in mind that not all circles and angles are inconvenient.

2 Problems

1. Circles k_1,k_2,k_3,k_4 are such that k_2 and k_4 each touch k_1 and k_3. Show that the tangency points are collinear or concyclic.

2. Prove that for any points A,B,C,D, $AB\cdot CD+BC\cdot DA\geq AC\cdot BD$, and that equality holds if and only if A,B,C,D are on a circle or a line in this order. (*Ptolemy's inequality*)

3. Let ω be the semicircle with diameter PQ. A circle k is tangent internally to ω and to segment PQ at C. Let AB be the tangent to k perpendicular to PQ, with A on ω and B on segment CQ. Show that AC bisects the angle $\angle PAB$.

4. Points A,B,C are given on a line in this order. Semicircles ω,ω_1,ω_2 are drawn on AC,AB,BC respectively as diameters on the same side of the line. A sequence of circles (k_n) is constructed as follows: k_0 is the circle determined by ω_2 and k_n is tangent to ω,ω_1,k_{n-1} for $n\geq 1$. Prove that the distance from the center of k_n to AB is $2n$ times the radius of k_n.

5. A circle with center O passes through points A and C and intersects the sides AB and BC of the triangle ABC at points K and N, respectively. The circumscribed circles of the triangles ABC and KBN intersect at two distinct points B and M. Prove that $\measuredangle OMB=90^\circ$. (*IMO 1985-5.*)

6. Let p be the semiperimeter of a triangle ABC. Points E and F are taken on line AB such that $CE=CF=p$. Prove that the circumcircle of $\triangle EFC$ is tangent to the excircle of $\triangle ABC$ corresponding to AB.

7. Prove that the nine-point circle of triangle ABC is tangent to the incircle and all three excircles. (*Feuerbach's theorem)*

8. The incircle of a triangle ABC is tangent to BC, CA, AB at M, N and P, respectively. Show that the circumcenter and incenter of $\triangle ABC$ and the orthocenter of $\triangle MNP$ are collinear.

9. Points A, B, C are given in this order on a line. Semicircles k and l are drawn on diameters AB and BC respectively, on the same side of the line. A circle t is tangent to k, to l at point $T \neq C$, and to the perpendicular n to AB through C. Prove that AT is tangent to l.

10. Let $A_1A_2A_3$ be a nonisosceles triangle with incenter I. Let C_i, $i = 1, 2, 3$, be the smaller circle through I tangent to A_iA_{i+1} and A_iA_{i+2} (the addition of indices being mod 3). Let B_i, $i = 1, 2, 3$, be the second point of intersection of C_{i+1} and C_{i+2}. Prove that the circumcenters of the triangles $A_1B_1I, A_2B_2I, A_3B_3I$ are collinear. (*IMO 1997 Shortlist*)

11. If seven vertices of a hexahedron lie on a sphere, then so does the eighth vertex.

12. A sphere with center on the plane of the face ABC of a tetrahedron $SABC$ passes through A, B and C, and meets the edges SA, SB, SC again at A_1, B_1, C_1, respectively. The planes through A_1, B_1, C_1 tangent to the sphere meet at a point O. Prove that O is the circumcenter of the tetrahedron $SA_1B_1C_1$.

13. Let KL and KN be the tangents from a point K to a circle k. Point M is arbitrarily taken on the extension of KN past N, and P is the second intersection point of k with the circumcircle of triangle KLM. The point Q is the foot of the perpendicular from N to ML. Prove that $\angle MPQ = 2\angle KML$.

14. The incircle Ω of the acute-angled triangle ABC is tangent to BC at K. Let AD be an altitude of triangle ABC and let M be the midpoint of AD. If N is the other common point of Ω and KM, prove that Ω and the circumcircle of triangle BCN are tangent at N. (*IMO 2002 Shortlist*)

3 Solutions

1. Let k_1 and k_2, k_2 and k_3, k_3 and k_4, k_4 and k_1 touch at A, B, C, D, respectively. An inversion with center A maps k_1 and k_2 to parallel lines k_1' and k_2', and k_3 and k_4 to circles k_3' and k_4' tangent to each other at C' and tangent to k_2' at B' and to k_4' at D'. It is easy to see that B', C', D' are collinear. Therefore B, C, D lie on a circle through A.

2. Applying the inversion with center A and radius r gives $AB = \frac{r^2}{AB'}$, $CD = \frac{r^2}{AC' \cdot AD'} C'D'$, etc. The required inequality reduces to $C'D' + B'C' \geq B'D'$.

3. Invert through C. Semicircle ω maps to the semicircle ω' with diameter $P'Q'$, circle k to the tangent to ω' parallel to $P'Q'$, and line AB to a circle l centered on $P'Q'$ which touches k (so it is congruent to the circle determined by ω'). Circle l intersects ω' and $P'Q'$ in A' and B' respectively. Hence $P'A'B'$ is an isosceles triangle with $\angle PAC = \angle A'P'C = \angle A'B'C = \angle BAC$.

4. Under the inversion with center A and squared radius $AB \cdot AC$ points B and C exchange positions, ω and ω_1 are transformed to the lines perpendicular to BC at C and B, and the sequence (k_n) to the sequence of circles (k_n') inscribed in the region between the two lines. Obviously, the distance from the center of k_n' to AB is $2n$ times its radius. Since circle k_n is homothetic to k_n' with respect to A, the statement immediately follows.

5. Invert through B. Points A', C', M' are collinear and so are K', N', M', whereas A', C', N', K' are on a circle. What does the center O of circle $ACNK$ map to? *Inversion does not preserve centers.* Let B_1 and B_2 be the feet of the tangents from B to circle $ACNK$. Their images B_1' and B_2' are the feet of the tangents from B to circle $A'C'N'K'$, and since O lies on the circle BB_1B_2,

its image O' lies on the line $B_1'B_2'$ - more precisely, it is at the midpoint of $B_1'B_2'$. We observe that M' is on the polar of point B with respect to circle $A'C'N'K'$, which is nothing but the line B_1B_2. It follows that $\angle OBM = \angle BO'M' = \angle BO'B_1' = 90^\circ$.

6. The inversion with center C and radius p maps points E and F and the excircle to themselves, and the circumcircle of $\triangle CEF$ to line AB which is tangent to the excircle. The statement follows from the fact that inversion preserves tangency.

7. We shall show that the nine-point circle ε touches the incircle k and the excircle k_a across A. Let A_1, B_1, C_1 be the midpoints of BC, CA, AB, and P, Q the points of tangency of k and k_a with BC, respectively. Recall that $A_1P = A_1Q$; this implies that the inversion with center A_1 and radius A_1P takes k and k_a to themselves. This inversion also takes ε to a line. It is not difficult to prove that this line is symmetric to BC with respect to the angle bisector of $\angle BAC$, so it also touches k and k_a.

8. The incenter of $\triangle ABC$ and the orthocenter of $\triangle MNP$ lie on the Euler line of the triangle ABC. The inversion with respect to the incircle of ABC maps points A, B, C to the midpoints of NP, PM, MN, so the circumcircle of ABC maps to the nine-point circle of $\triangle MNP$ which is also centered on the Euler line of MNP. It follows that the center of circle ABC lies on the same line.

9. An inversion with center T maps circles t and l to parallel lines t' and l', circle k and line n to circles k' and n' tangent to t' and l' (where $T \in n'$), and line AB to circle a' perpendicular to l' (because an inversion preserves angles) and passes through $B', C' \in l'$; thus a' is the circle with diameter $B'C'$. Circles k' and n' are congruent and tangent to l' at B' and C', and intersect a' at A' and T respectively. It follows that A' and T are symmetric with respect to the perpendicular bisector of $B'C'$ and hence $A'T \parallel l'$, so AT is tangent to l.

10. The centers of three circles passing through the same point I and not touching each other are collinear if and only if they have another common point. Hence it is enough to show that the circles A_iB_iI have a common point other than I. Now apply inversion at center I and with an arbitrary power. We shall denote by X' the image of X under this inversion. In our case, the image of the circle C_i is the line $B_{i+1}'B_{i+2}'$ while the image of the line $A_{i+1}A_{i+2}$ is the circle $IA_{i+1}'A_{i+2}'$ that is tangent to $B_i'B_{i+2}'$, and $B_i'B_{i+2}'$. These three circles have equal radii, so their centers P_1, P_2, P_3 form a triangle also homothetic to $\triangle B_1'B_2'B_3'$. Consequently, points A_1', A_2', A_3', that are the reflections of I across the sides of $P_1P_2P_3$, are vertices of a triangle also homothetic to $B_1'B_2'B_3'$. It follows that $A_1'B_1', A_2'B_2', A_3'B_3'$ are concurrent at some point J', i.e., that the circles A_iB_iI all pass through J.

11. Let $AYBZ, AZCX, AXDY, WCXD, WDYB, WBZC$ be the faces of the hexahedron, where A is the "eighth" vertex. Apply an inversion with center W. Points B', C', D', X', Y', Z' lie on some plane π, and moreover, C', X', D'; D', Y', B'; and B', Z', C' are collinear in these orders. Since A is the intersection of the planes YBZ, ZCX, XDY, point A' is the second intersection point of the spheres $WY'B'Z', WZ'C'X', WX'D'Y'$. Since the circles $Y'B'Z', Z'C'X', X'D'Y'$ themselves meet at a point on plane π, this point must coincide with A'. Thus $A' \in \pi$ and the statement follows.

12. Apply the inversion with center S and squared radius $SA \cdot SA_1 = SB \cdot SB_1 = SC \cdot SC_1$. Points A and A_1, B and B_1, and C and C_1 map to each other, the sphere through A, B, C, A_1, B_1, C_1 maps to itself, and the tangent planes at A_1, B_1, C_1 go to the spheres through S and A, S and B, S and C which touch the sphere $ABCA_1B_1C_1$. These three spheres are perpendicular to the plane ABC, so their centers lie on the plane ABC; hence they all pass through the point $\overline{S}$ symmetric to S with respect to plane ABC. Therefore $\overline{S}$ is the image of O. Now since $\angle SA_1O = \angle S\overline{S}A = \angle \overline{S}SA = \angle OSA_1$, we have $OS = OA_1$ and analogously $OS = OB_1 = OC_1$.

13. Apply the inversion with center M. Line MN' is tangent to circle k' with center O', and a circle through M is tangent to k' at L' and meets MN' again at K'. The line $K'L'$ intersects k' at P', and $N'O'$ intersects ML' at Q'. The task is to show that $\angle MQ'P' = \angle L'Q'P' = 2\angle K'ML'$.

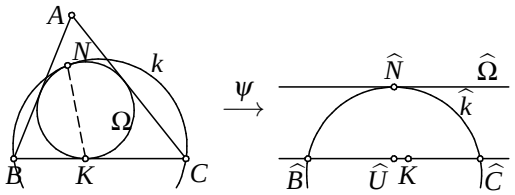
    Let the common tangent at L' intersect MN' at Y'. Since the peripheral angles on the chords $K'L'$ and $L'P'$ are equal (to $\angle K'L'Y'$), we have $\angle L'O'P' = 2\angle L'N'P' = 2\angle K'ML'$. It only remains to show that L', P', O', Q' are on a circle. This follows from the equality $\angle O'Q'L' = 90^\circ - \angle L'MK' = 90^\circ - \angle L'N'P' = \angle O'P'L'$ (the angles are regarded as oriented).

14. Let k be the circle through B, C that is tangent to the circle Ω at point N'. We must prove that K, M, N' are collinear. Since the statement is trivial for $AB = AC$, we may assume that $AC > AB$. As usual, $R, r, \alpha, \beta, \gamma$ denote the circumradius and the inradius and the angles of $\triangle ABC$, respectively.

    We have $\tan\angle BKM = DM/DK$. Straightforward calculation gives $DM = \frac{1}{2}AD = R\sin\beta\sin\gamma$ and $DK = \frac{DC-DB}{2} - \frac{KC-KB}{2} = R\sin(\beta-\gamma) - R(\sin\beta - \sin\gamma) = 4R\sin\frac{\beta-\gamma}{2}\sin\frac{\beta}{2}\sin\frac{\gamma}{2}$, so we obtain

$$\tan\angle BKM = \frac{\sin\beta\sin\gamma}{4\sin\frac{\beta-\gamma}{2}\sin\frac{\beta}{2}\sin\frac{\gamma}{2}} = \frac{\cos\frac{\beta}{2}\cos\frac{\gamma}{2}}{\sin\frac{\beta-\gamma}{2}}.$$

    To calculate the angle BKN', we apply the inversion ψ with center at K and power $BK\cdot CK$. For each object X, we denote by $\widehat{X}$ its image under ψ. The incircle Ω maps to a



    line $\widehat{\Omega}$ parallel to $\widehat{B}\widehat{C}$, at distance $\frac{BK\cdot CK}{2r}$ from $\widehat{B}\widehat{C}$. Thus the point $\widehat{N}'$ is the projection of the midpoint $\widehat{U}$ of $\widehat{B}\widehat{C}$ onto $\widehat{\Omega}$. Hence

$$\tan\angle BKN' = \tan\angle\widehat{B}K\widehat{N}' = \frac{\widehat{U}\widehat{N}'}{\widehat{U}K} = \frac{BK\cdot CK}{r(CK-BK)}.$$

    Again, one easily checks that $KB\cdot KC = bc\sin^2\frac{\alpha}{2}$ and $r = 4R\sin\frac{\alpha}{2}\cdot\sin\frac{\beta}{2}\cdot\sin\frac{\gamma}{2}$, which implies

$$\begin{aligned}\tan\angle BKN' &= \frac{bc\sin^2\frac{\alpha}{2}}{r(b-c)} \\ &= \frac{4R^2\sin\beta\sin\gamma\sin^2\frac{\alpha}{2}}{4R\sin\frac{\alpha}{2}\sin\frac{\beta}{2}\sin\frac{\gamma}{2}\cdot 2R(\sin\beta-\sin\gamma)} = \frac{\cos\frac{\beta}{2}\cos\frac{\gamma}{2}}{\sin\frac{\beta-\gamma}{2}}.\end{aligned}$$

    Hence $\angle BKM = \angle BKN'$, which implies that K, M, N' are indeed collinear; thus $N' \equiv N$.

terug naar echt bestand

Casey's Theorem and its Applications

Luis González
Maracaibo. Venezuela

July 2011

Abstract. We present a proof of the generalized Ptolemy's theorem, also known as Casey's theorem and its applications in the resolution of difficult geometry problems.

1 Casey's Theorem.

Theorem 1. Two circles $\Gamma_1(r_1)$ and $\Gamma_2(r_2)$ are internally/externally tangent to a circle $\Gamma(R)$ through A, B, respetively. The length δ_{12} of the common external tangent of Γ_1, Γ_2 is given by:

$$\delta_{12} = \frac{AB}{R}\sqrt{(R \pm r_1)(R \pm r_2)}$$

Proof. Without loss of generality assume that $r_1 \geq r_2$ and we suppose that Γ_1 and Γ_2 are internally tangent to Γ. The remaining case will be treated analogously. A common external tangent between Γ_1 and Γ_2 touches Γ_1, Γ_2 at A_1, B_1 and A_2 is the orthogonal projection of O_2 onto O_1A_1. (See Figure 1). By Pythagorean theorem for $\triangle O_1O_2A_2$, we obtain

$$\delta_{12}{}^2 = (A_1B_1)^2 = (O_1O_2)^2 - (r_1 - r_2)^2$$

Let $\angle O_1OO_2 = \lambda$. By cosine law for $\triangle OO_1O_2$, we get

$$(O_1O_2)^2 = (R - r_1)^2 + (R - r_2)^2 - 2(R - r_1)(R - r_2)\cos\lambda$$

By cosine law for the isosceles triangle $\triangle OAB$, we get

$$AB^2 = 2R^2(1 - \cos\lambda)$$

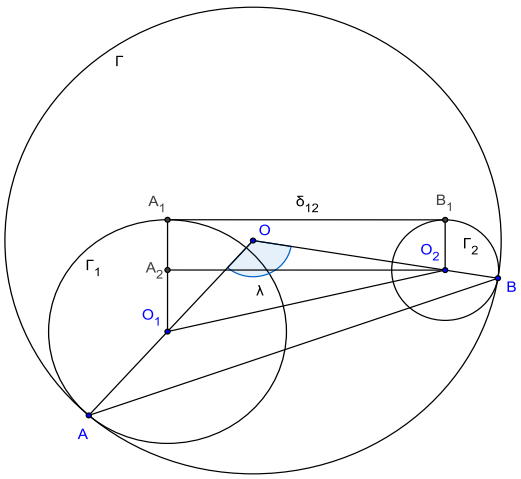


Figure 1: Theorem 1

Eliminating $\cos\lambda$ and O_1O_2 from the three previous expressions yields

$$\delta_{12}{}^2 = (R-r_1)^2 + (R-r_2)^2 - (r_1-r_2)^2 - 2(R-r_1)(R-r_2)\left(1 - \frac{AB^2}{2R^2}\right)$$

Subsequent simplifications give

$$\delta_{12} = \frac{AB}{R}\sqrt{(R-r_1)(R-r_2)} \quad (1)$$

Analogously, if Γ_1, Γ_2 are externally tangent to Γ, then we will get

$$\delta_{12} = \frac{AB}{R}\sqrt{(R+r_1)(R+r_2)} \quad (2)$$

If Γ_1 is externally tangent to Γ and Γ_2 is internally tangent to Γ, then a similar reasoning gives that the length of the common internal tangent between Γ_1 and Γ_2 is given by

$$\delta_{12} = \frac{AB}{R}\sqrt{(R+r_1)(R-r_2)} \quad (3)$$

Theorem 2 (Casey). Given four circles $\Gamma_i, i = 1, 2, 3, 4$, let δ_{ij} denote the length of a common tangent (either internal or external) between Γ_i and Γ_j. The four circles are tangent to a fith circle Γ (or line) if and only if for appropriate choice of signs,

$$\delta_{12} \cdot \delta_{34} \pm \delta_{13} \cdot \delta_{42} \pm \delta_{14} \cdot \delta_{23} = 0$$

The proof of the direct theorem is straightforward using Ptolemy's theorem for the quadrilateral $ABCD$ whose vertices are the tangency points of $\Gamma_1(r_1), \Gamma_2(r_2), \Gamma_3(r_3), \Gamma_4(r_4)$ with $\Gamma(R)$. We susbtitute the lengths of its sides and digonals in terms of the lenghts of the tangents δ_{ij}, by using the formulas (1), (2) and (3). For instance, assuming that all tangencies are external, then using (1), we get

$$\delta_{12} \cdot \delta_{34} + \delta_{14} \cdot \delta_{23} = \left(\tfrac{AB \cdot CD + AD \cdot BC}{R^2}\right) \sqrt{(R - r_1)(R - r_2)(R - r_3)(R - r_4)}$$

$$\delta_{12} \cdot \delta_{34} + \delta_{14} \cdot \delta_{23} = \left(\tfrac{AC \cdot BD}{R^2}\right) \sqrt{(R - r_1)(R - r_3)} \cdot \sqrt{(R - r_2)(R - r_4)}$$

$$\delta_{12} \cdot \delta_{34} + \delta_{14} \cdot \delta_{23} = \delta_{13} \cdot \delta_{42}.$$

Casey established that this latter relation is sufficient condition for the existence of a fith circle $\Gamma(R)$ tangent to $\Gamma_1(r_1), \Gamma_2(r_2), \Gamma_3(r_3), \Gamma_4(r_4)$. Interestingly, the proof of this converse is a much tougher exercise. For a proof you may see [1].

2 Some Applications.

I) $\triangle ABC$ **is isosceles with legs** $AB = AC = L$. **A circle** ω **is tangent to** $\overline{BC}$ **and the arc** BC **of the circumcircle of** $\triangle ABC$. **A tangent line from** A **to** ω **touches** ω **at** P. **Describe the locus of** P **as** ω **varies.**

Solution. We use Casey's theorem for the circles $(A), (B), (C)$ (with zero radii) and ω, all internally tangent to the circumcircle of $\triangle ABC$. Thus, if ω touches $\overline{BC}$ at Q, we have:

$$L \cdot CQ + L \cdot BQ = AP \cdot BC \Longrightarrow \ AP = \frac{L(BQ + CQ)}{BC} = L$$

The length AP is constant, i.e. Locus of P is the circle with center A and radius $AB = AC = L$.

II) (O) **is a circle with diameter** $\overline{AB}$ **and** P, Q **are two points on** (O) **lying on different sides of** $\overline{AB}$. T **is the orthogonal projection of** Q **onto** $\overline{AB}$. **Let** $(O_1), (O_2)$ **be the circles with diameters** TA, TB **and** PC, PD **are the tangent segments from** P **to** $(O_1), (O_2)$, **respectively. Show that** $PC + PD = PQ$. [2].

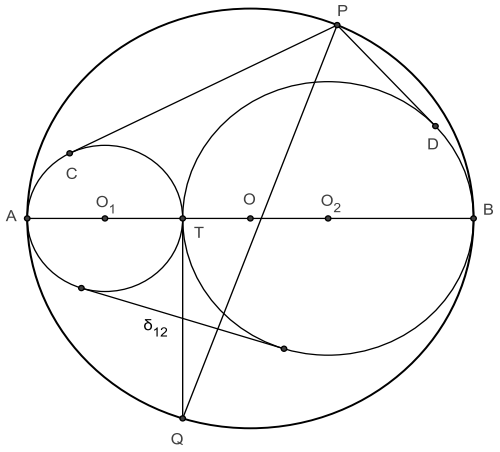


Figure 2: Application II

Solution. Let δ_{12} denote the length of the common external tangent of $(O_1), (O_2)$. We use Casey's theorem for the circles $(O_1), (O_2), (P), (Q)$, all internally tangent to (O).

$$PC \cdot QT + PD \cdot QT = PQ \cdot \delta_{12} \implies PC + PD = PQ \cdot \frac{\delta_{12}}{QT} = PQ \cdot \frac{\sqrt{TA \cdot TB}}{TQ} = PQ.$$

III) In $\triangle ABC$, let $\omega_A, \omega_B, \omega_C$ be the circles tangent to BC, CA, AB through their midpoints and the arcs BC, CA, AB of its circumcircle (not containing A, B, C). If $\delta_{BC}, \delta_{CA}, \delta_{AB}$ denote the lengths of the common external tangents between $(\omega_B, \omega_C), (\omega_C, \omega_A)$ and (ω_A, ω_B), respectively, then prove that

$$\delta_{BC} = \delta_{CA} = \delta_{AB} = \frac{a+b+c}{4}$$

Solution. Let $\delta_A, \delta_B, \delta_C$ denote the lengths of the tangents from A, B, C to $\omega_A, \omega_B, \omega_C$, respectively. By Casey's theorem for the circles $(A), (B), (C), \omega_B$, all tangent to the circumcircle of $\triangle ABC$, we get

$$\delta_B \cdot b = a \cdot AE + c \cdot CE \implies \delta_B = \tfrac{1}{2}(a+c)$$

Similarly, by Casey's theorem for $(A), (B), (C), \omega_C$ we'll get $\delta_C = \frac{1}{2}(a+b)$

Now, by Casey's theorem for $(B),(C),\omega_B,\omega_C$, we get $\delta_B\cdot\delta_C=\delta_{BC}\cdot a+BF\cdot BE\Longrightarrow$

$$\delta_{BC}=\frac{\delta_B\cdot\delta_C-BF\cdot BE}{a}=\frac{(a+c)(a+b)-bc}{4a}=\frac{a+b+c}{4}$$

By similar reasoning, we'll have $\delta_{CA}=\delta_{AB}=\frac{1}{4}(a+b+c)$.

IV) A circle $\mathcal{K}$ passes through the vertices B,C of $\triangle ABC$ and another circle ω touches $AB,AC,\mathcal{K}$ at P,Q,T, respectively. If M is the midpoint of the arc BTC of $\mathcal{K}$, show that BC,PQ,MT concur. [3]

Solution. Let R,ϱ be the radii of $\mathcal{K}$ and ω, respectively. Using formula (1) of Theorem 1 for $\omega,(B)$ and $\omega,(C)$. Both $(B),(C)$ with zero radii and tangent to $\mathcal{K}$ through B,C, we obtain:

$$TC^2=\frac{CQ^2\cdot R^2}{(R-\varrho)(R-0)}=\frac{CQ^2\cdot R}{R-\varrho}\ ,\ TB^2=\frac{BP^2\cdot R^2}{(R-\varrho)(R-0)}=\frac{BP^2\cdot R}{R-\varrho}\Longrightarrow\ \frac{TB}{TC}=\frac{BP}{CQ}$$

Let PQ cut BC at U. By Menelaus' theorem for $\triangle ABC$ cut by $\overline{UPQ}$ we have

$$\frac{UB}{UC}=\frac{BP}{AP}\cdot\frac{AQ}{CQ}=\frac{BP}{CQ}=\frac{TB}{TC}$$

Thus, by angle bisector theorem, U is the foot of the T-external bisector TM of $\triangle BTC$.

V) If D,E,F denote the midpoints of the sides BC,CA,AB of $\triangle ABC$. Show that the incircle (I) of $\triangle ABC$ is tangent to $\odot(DEF)$. (Feuerbach theorem).

Solution. We consider the circles $(D),(E),(F)$ with zero radii and (I). The notation δ_{XY} stands for the length of the external tangent between the circles $(X),(Y)$, then

$$\delta_{DE}=\frac{c}{2}\ ,\ \delta_{EF}=\frac{a}{2}\ ,\ \delta_{FD}=\frac{b}{2}\ ,\ \delta_{DI}=\left|\frac{b-c}{2}\right|\ ,\ \delta_{EI}=\left|\frac{a-c}{2}\right|\ ,\ \delta_{FI}=\left|\frac{b-a}{2}\right|$$

For the sake of applying the converse of Casey's theorem, we shall verify if, for some combination of signs $+$ and $-$, we get $\pm c(b-a)\pm a(b-c)\pm b(a-c)=0$, which is trivial. Therefore, there exists a circle tangent to $(D),(E),(F)$ and (I), i.e. (I) is internally tangent to $\odot(DEF)$. We use the same reasoning to show that $\odot(DEF)$ is tangent to the three excircles of $\triangle ABC$.

VI) $\triangle ABC$ is scalene and D,E,F are the midpoints of BC,CA,AB. The incircle (I) and 9 point circle $\odot(DEF)$ of $\triangle ABC$ are internally tangent through the Feuerbach point F_e. Show that one of the segments $\overline{F_eD},\overline{F_eE},\overline{F_eF}$ equals the sum of the other two. [4]

Solution. WLOG assume that $b \geq a \geq c$. Incircle (I, r) touches BC at M. Using formula (1) of Theorem 1 for (I) and (D) (with zero radius) tangent to the 9-point circle $(N, \frac{R}{2})$, we have:

$$F_eD^2 = \frac{DM^2 \cdot (\frac{R}{2})^2}{(\frac{R}{2} - r)(\frac{R}{2} - 0)} \Longrightarrow \ F_eD = \sqrt{\frac{R}{R-2r}} \cdot \frac{(b-c)}{2}$$

By similar reasoning, we have the expressions

$$F_eE = \sqrt{\frac{R}{R-2r}} \cdot \frac{(a-c)}{2} \ , \ F_eF = \sqrt{\frac{R}{R-2r}} \cdot \frac{(b-a)}{2}$$

Therefore, the addition of the latter expressions gives

$$F_eE + F_eF = \sqrt{\frac{R}{R-2r}} \cdot \frac{b-c}{2} = F_eD$$

VII) $\triangle ABC$ **is a triangle with** $AC > AB$. **A circle** ω_A **is internally tangent to its circumcircle** ω **and** AB, AC. S **is the midpoint of the arc** BC **of** ω, **which does not contain** A **and** ST **is the tangent segment from** S **to** ω_A. **Prove that**

$$\frac{ST}{SA} = \frac{AC - AB}{AC + AB} \quad [5]$$

Solution. Let M, N be the tangency points of ω_A with AC, AB. By Casey's theorem for $\omega_A, (B), (C), (S)$, all tangent to the circumcircle ω, we get

$$ST \cdot BC + CS \cdot BN = CM \cdot BS \Longrightarrow \ ST \cdot BC = CS(CM - BN)$$

If U is the reflection of B across AS, then $CM - BN = UC = AC - AB$. Hence

$$ST \cdot BC = CS(AC - AB) \ (\star)$$

By Ptolemy's theorem for $ABSC$, we get $SA \cdot BC = CS(AB + AC)$. Together with $(\star)$, we obtain

$$\frac{ST}{SA} = \frac{AC - AB}{AC + AB}$$

VIII) Two congruent circles $(S_1), (S_2)$ meet at two points. A line ℓ cuts (S_2) at A, C and (S_1) at B, D (A, B, C, D are collinear in this order). Two distinct circles ω_1, ω_2 touch the line ℓ and the circles $(S_1), (S_2)$ externally and internally respectively. If ω_1, ω_2 are externally tangent, show that $AB = CD$. [6]

Solution. Let $P \equiv \omega_1 \cap \omega_2$ and M, N be the tangency points of ω_1 and ω_2 with an external tangent. Inversion with center P and power $PB \cdot PD$ takes (S_1) and the line ℓ into themselves. The circles ω_1 and ω_2 go to two parallel lines k_1 and k_2 tangent to (S_1) and the circle (S_2) goes to another circle (S_2') tangent to k_1, k_2. Hence, (S_2) is congruent to its inverse (S_2'). Further, $(S_2), (S_2')$ are symmetrical about $P \implies PC \cdot PA = PB \cdot PD$.

By Casey's theorem for $\omega_1, \omega_2, (D), (B), (S_1)$ and $\omega_1, \omega_2, (A), (C), (S_2)$ we get:

$$DB = \frac{2PB \cdot PD}{MN} \ , \ AC = \frac{2PA \cdot PC}{MN}$$

Since $PC \cdot PA = PB \cdot PD \implies AC = BD \implies AB = CD$.

IX) $\triangle ABC$ is equilateral with side length L. Let (O, r) and (O, R) be the incircle and circumcircle of $\triangle ABC$. P is a point on (O, r) and P_1, P_2, P_3 are the projections of P onto BC, CA, AB. Circles $\mathcal{T}_1, \mathcal{T}_2$ and $\mathcal{T}_3$ touch BC, CA, AB through P_1, P_2, P_2 and (O, R) (internally), their centers lie on different sides of BC, CA, AB with respect to A, B, C. Prove that the sum of the lengths of the common external tangents of $\mathcal{T}_1, \mathcal{T}_2$ and $\mathcal{T}_3$ is a constant value.

Solution. Let δ_1 denote the tangent segment from A to $\mathcal{T}_1$. By Casey's theorem for $(A), (B), (C), \mathcal{T}_1$, all tangent to (O, R), we have $L \cdot BP_1 + L \cdot CP_1 = \delta_1 \cdot L \implies \delta_1 = L$. Similarly, we have $\delta_2 = \delta_3 = L$. By Euler's theorem for the pedal triangle $\triangle P_1P_2P_3$ of P, we get:

$$[P_1P_2P_3] = \frac{p(P, (O))}{4R^2}[ABC] = \frac{R^2 - r^2}{4R^2}[ABC] = \frac{3}{16}[ABC]$$

Therefore, we obtain

$$AP_2 \cdot AP_3 + BP_3 \cdot BP_1 + CP_1 \cdot CP_2 = \frac{2}{\sin 60^\circ}\left([ABC] - [P_1P_2P_3]\right) = \tfrac{13}{16}L^2. \ (\star)$$

By Casey's theorem for $(B), (C), \mathcal{T}_2, \mathcal{T}_3$, all tangent to (O, R), we get

$$\delta_2 \cdot \delta_3 = L^2 = BC \cdot \delta_{23} + CP_2 \cdot BP_3 = L \cdot \delta_{23} + (L - AP_1)(L - AP_2)$$

By cyclic exchange, we have the expressions:

$$L^2 = L \cdot \delta_{31} + (L - BP_3)(L - BP_1) \ , \ L^2 = L \cdot \delta_{12} + (L - CP_1)(L - CP_2)$$

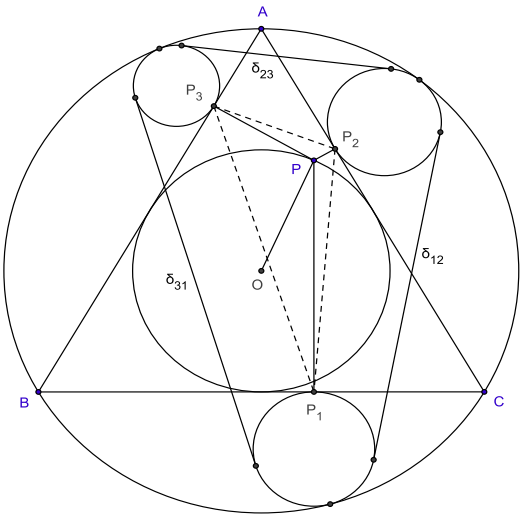


Figure 3: Application VII

Adding the three latter equations yields

$$3L^2 = L(\delta_{23} + \delta_{31} + \delta_{12}) + 3L^2 - 3L^2 + AP_3 \cdot AP_2 + BP_3 \cdot BP_1 + CP_1 \cdot CP_2$$

Hence, combining with $(\star)$ gives

$$\delta_{23} + \delta_{31} + \delta_{12} = 3L - \frac{13}{16}L = \frac{35}{16}L$$

3 Proposed Problems.

1) *Purser's theorem*: $\triangle ABC$ is a triangle with circumcircle (O) and ω is a circle in its plane. AX, BY, CZ are the tangent segments from A, B, C to ω. Show that ω is tangent to (O), if and only if

$$\pm AX \cdot BC \pm BY \cdot CA \pm CZ \cdot AB = 0$$

2) Circle ω touches the sides AB, AC of $\triangle ABC$ at P, Q and its circumcircle (O). Show that the midpoint of $\overline{PQ}$ is either the incenter of $\triangle ABC$ or the A-excenter of $\triangle ABC$, according to whether $(O), \omega$ are internally tangent or externally tangent.

3) $\triangle ABC$ is A-right with circumcircle (O). Circle Ω_B is tangent to the segments $\overline{OB}, \overline{OA}$ and the arc AB of (O). Circle Ω_C is tangent to the segments $\overline{OC}, \overline{OA}$ and the arc AC of (O). Ω_B, Ω_C touch $\overline{OA}$ at P, Q, respectively. Show that:

$$\frac{AB}{AC} = \frac{AP}{AQ}$$

4) *Gumma, 1874.* We are given a cirle (O, r) in the interior of a square $ABCD$ with side length L. Let (O_i, r_i) $i = 1, 2, 3, 4$ be the circles tangent to two sides of the square and (O, r) (externally). Find L as a fuction of r_1, r_2, r_3, r_4.

5) Two parallel lines τ_1, τ_2 touch a circle $\Gamma(R)$. Circle $k_1(r_1)$ touches Γ, τ_1 and a third circle $k_2(r_2)$ touches Γ, τ_2, k_1. We assume that all tangencies are external. Prove that $R = 2\sqrt{r_1 \cdot r_2}$.

6) *Victor Thébault. 1938.* $\triangle ABC$ has incircle (I, r) and circumcircle (O). D is a point on $\overline{AB}$. Circle $\Gamma_1(r_1)$ touches the segments $\overline{DA}, \overline{DC}$ and the arc CA of (O). Circle $\Gamma_2(r_2)$ touches the segments $\overline{DB}, \overline{DC}$ and the arc CB of (O). If $\angle ADC = \varphi$, show that:

$$r_1 \cdot \cos^2 \frac{\varphi}{2} + r_2 \cdot \sin^2 \frac{\varphi}{2} = r$$

References

[1] I. Shariguin, Problemas de Geometrié (Planimetrié), Ed. Mir, Moscu, 1989.

[2] Vittasko, Sum of two tangents, equal to the distance of two points, AoPS, 2011. http://www.artofproblemsolving.com/Forum/viewtopic.php?f=47&t=404640.

[3] My_name_is_math, Tangent circles concurrent lines, AoPs, 2011. http://www.artofproblemsolving.com/Forum/viewtopic.php?f=46&t=399496.

[4] Mathquark, Point [Feuerbach point of a triangle; FY + FZ = FX], AoPS, 2005. http://www.artofproblemsolving.com/Forum/viewtopic.php?f=49&t=24959.

[5] Virgil Nicula, ABC and circle tangent to AB, AC and circumcircle , AoPS, 2011. http://www.artofproblemsolving.com/Forum/viewtopic.php?f=47&t=357957.

[6] Shoki, Iran(3rd round)2009, AoPS, 2009. http://www.artofproblemsolving.com/Forum/viewtopic.php?f=46&t=300809.

terug naar echt bestand

GEOMETRY

for the

OLYMPIAD ENTHUSIAST

Bruce Merry

The South African Mathematical Society (SAMS) has the responsibility for selecting teams to represent South Africa at the Pan African Mathematics Olympiad (PAMO) and the International Mathematical Olympiad (IMO).

Team selection begins with the Old Mutual Mathematical Talent Search, a self-paced correspondence course in problem-solving that starts afresh in January each year. The best performers in the Talent Search are invited to attend mathematical camps in which they learn specialised problem-solving skills and write challenging Olympiad-level papers. Since the Pan African Maths Olympiad is not quite as daunting as the International version, the tradition has evolved that South African PAMO teams consist of students who have not previously been selected for an IMO team. The Inter-Provincial Mathematical Olympiad and the South African Mathematics Olympiad are closely linked with the PAMO and IMO selection programme.

To provide background reading for the Talent Search, the South African Mathematical Society has published a series of Mathematical Olympiad Training Notes that focus on mathematical topics and problem-solving skills needed in mathematical competitions and Olympiads. Six booklets have appeared to date:

- *The Pigeon-hole Principle*, by Valentine Goranko
- *Topics in Number Theory*, by Valentin Goranko
- *Inequalities for the Olympiad Enthusiast*, by Graeme West
- *Graph Theory for the Olympiad Enthusiast*, by Graeme West
- *Functional Equations for the Olympiad Enthusiast*, by Graeme West
- *Mathematical Induction for the Olympiad Enthusiast*, by David Jacobs

Though their primary target is the development of high-level problem-solving skills, these booklets can be read by anybody interested in the mathematics just beyond the high school curriculum. They are therefore particularly useful to teachers looking for enrichment material, and students who plan to study mathematics at university level and would like more of a challenge than the school curriculum provides.

For more information, write to

Old Mutual Mathematical Talent Search
Department of Mathematics and Applied Mathematics
University of Cape Town
7701 RONDEBOSCH

South Africa's participation in the Pan African and International Mathematical Olympiads is supported by Old Mutual and the Department of Science and Technology.

John Webb
January 2003

Geometry for the Olympiad Enthusiast

Bruce Merry

A booklet in this series was last published in 1996, and the series has been somewhat dormant since. Geometry has long been a gap in this series, and eventually I decided to address this gap. I started writing this booklet in December 2000. It was then put aside for three years, while I focused on my studies. In December 2003 I finally returned to finish the rather delayed project.

This booklet is primarily about classical, or Euclidean, geometry. Trigonometry is used as a tool, but is not explored in great depth, and coordinate geometry barely puts in an appearance. While tackling the exercises and geometry problems in general, one should remember that trigonometry and coordinate geometry are powerful tools. I simply did not have much to say about them.

The booklet assumes a knowledge of high-school geometry. If you have not completed the high-school syllabus, it would be a good idea to first find a textbook and work through both the theory and the exercises. The proofs included here are somewhat terse and you may need to fill in a few details yourself.

Some important results are left as problems, so you should at the very least read the problems (although you really should attempt to solve them, as well). The positioning of problems in the book is a good indicator of how you are expected to tackle them, although of course there are usually other solutions. There are two types of problems: exercises that deal very specifically with the topic in hand, and real olympiad problems. The olympiad problems are labelled with a star ($\star$). The exercises are generally easier than the olympiad problems, but some of them are quite challenging.

I would like to thank Dirk Laurie for writing his Geomplex diagram drawing package. This book would not have been possible without it. I would also like to thank Mark Berman, whose flair for geometry has always inspired me to find elegant solutions.

Contents

1 Techniques

Geometry is unlike many of the other areas of olympiad mathematics, requiring more intuition and less algebra. Nevertheless, it is important to do the basic groundwork as otherwise your intuition has nothing with which to work.

Here are some suggestions on ways to approach a geometry problem.

- Draw a quick diagram so that you can visualise the problem.
- Draw a neat and accurate diagram — this will often reveal additional facts which you could then try to prove.
- Draw a deliberately incorrect diagram (this could be your initial diagram), so that you don't accidentally assume the result because you referred to your accurate diagram (this is particularly important if you are proving concurrency or collinearity).
- It is very important to do as much investigation as you can. Try to relate as many angles and line segments as you can, even if you have several variables. Then look for similar or congruent triangles, parallel lines and so on. This on its own can be enough to solve some easier problems without even having to think.
- There are many approaches to attack geometry problems e.g. Euclidean geometry, coordinate geometry, complex numbers, vectors and trigonometry. Think about applying all the ones that you know to the problem and deciding which ones are most likely to work. Be guided by what you are asked to prove: for example, if you are asked to prove that two lines are parallel then coordinate geometry might work well, but if the problem involves lots of related angles then trigonometry may be a better approach.
- Don't be afraid to get your hands dirty with trigonometry, coordinate geometry or algebra. While such solutions might not be as "cool" as solutions that require an inspired construction, they are often easier to find and score the same number of points. However, doing as much as possible with Euclidean geometry first can make the equations simpler.
- Look for constructions that will give you similar triangles, special angles or allow you to restate the problem in a simpler way. For example, if you are asked to prove something about the sum of two lengths, try making a construction that places the two lengths end to end so that you only have to prove something about the length of a single line.

- Assume that the result is true, and see what follows from this. This may lead you to intermediate results which you can then try to prove.
- Always check that you haven't omitted any cases such as obtuse angles or constructions that are impossible in certain cases (for example, you can't take the intersection point of two lines if they are parallel). This booklet does a terrible job of this, because the special cases are almost always trivial. I'm lazy, the duplication costs of this booklet are high, the rainforests are dying, and this is not a competition. In a competition, you can expect to lose marks if your proof does not work in all cases.

2 Terminology and notation

There is some basic terminology for things that share some property. Concurrent lines pass through a common point, and collinear points lie on a common line. Concyclic points lie on a common circle; note that "A, B, C and D are concyclic" does not have the same meaning as "$ABCD$ is a cyclic quadrilateral", since the latter implies that the points lie in a particular order around the circle. Concentric circles have a common centre.

The humble triangle has possibly the richest terminology and notation. There are numerous "centres", generally the point of concurrency of certain lines, and a few have corresponding circles.

incentre The centre of the incircle (inscribed circle); the point of concurrency of the internal angle bisectors

circumcentre The centre of the circumcircle (circumscribed circle); the point of concurrency of the perpendicular bisectors

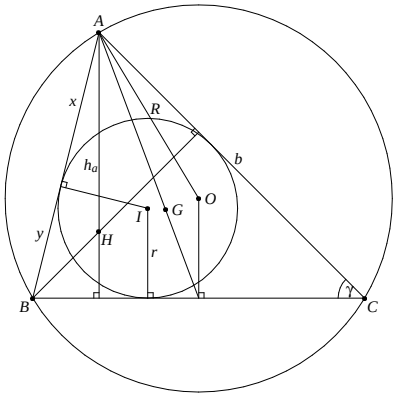
excircle The centre of an excircle (escribed circle); the point of concurrency of two external and one internal angle bisector

orthocentre The point of concurrency of the altitudes

centroid The point of concurrency of the medians (lines from a vertex to the midpoint of the opposite side)

Most of these terms should be familiar from high-school geometry. An unfamiliar term is a *cevian*: this is any line joining a vertex to the opposite side.

For this booklet (particularly section 6), we also introduce a lot of notation for triangles. Some of this is standard or mostly standard while some is not; you are advised to define any of these quantities in proofs, particularly K, x, y and z.



I the incentre

I_A the excentre opposite A

O the circumcentre

G the centroid

H the orthocentre

a the side opposite vertex A (similarly for B and C)

s the semiperimeter, $\frac{a+b+c}{2}$

x the tangent from A to the incircle, $\frac{-a+b+c}{2} = s-a$ (similarly for y and z)

R the radius of the circumcircle (circumradius)

r the radius of the incircle (inradius)

r_a the radius of the excircle opposite A

h_a the height of the altitude from A to BC

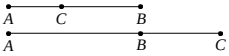
α the angle at A (similarly for β and γ)

K the area of the triangle

We also use the notation $|\triangle ABC|$ (or just $|ABC|$) to indicate the area of $\triangle ABC$.

3 Directed angles, line segments and area

In classical geometry, most quantities are undirected. That means that if you measure them in the opposite direction, they have the same value ($AB = BA$, $\angle ABC = \angle CBA$, and $|\triangle ABC| = |\triangle CBA|$). Most of the time this is a reasonable way of doing things. However, it occasionally has disadvantages. For example, if you know that A, B and C are collinear, and $AB = 5$, $BC = 3$, then what is AC? It could be either 2 or 8, depending on which way round they are on the line. The same problem arises when adding angles or areas.



Normally these situations are not important, because it is clear from a diagram which is correct. However, sometimes there are many different ways to draw the diagram, leading to a proof with many different cases. Another way to solve the problem is to treat the quantities as having a sign, indicating the direction. So now if you are told that $AB = 5$, $BC = 3$ then you can be sure that $AC = AB + BC = 8$. This is because both have the same sign, and hence are in the same direction. If C lay between A and B, then $AB = 5$, $BC = -3$ and so $AC = AB + BC = 2$. It could also be that $AB = -5$, $BC = 3$; the positive direction is generally arbitrary but must be consistent. What is important is that no matter in what order A, B and C lie, the equation $AC = AB + BC$ holds.

Directed line segments have somewhat limited use, because it only makes sense to compare lines that are parallel. Generally they are used when dealing with ratios or products of collinear line segments (see Menelaus' Theorem (6.3), for example). Directed angles and directed area are more often used.

A directed angle $\angle ABC$ is really a measure of the angle between the two lines AB and BC. Conventionally, it is the amount by which AB must be rotated anti-clockwise

to line up with BC. One effect of this is that while normal angles have a range of 360°, directed angles only have a range of 180°! This is because rotating a line by 180° leaves it back where it started, so 180° is equivalent to 0°. To indicate this, equivalent angles are sometimes written $\angle ABC \equiv \angle DEF$ rather than $\angle ABC = \angle DEF$. This limitation occasionally has disadvantages, and in particular it is not generally possible to combine trigonometry with directed angles (since the sin and cos functions only repeat every 360°). This is made up for by the special properties that directed angles do have:

1. $\angle AMC \equiv \angle AMB + \angle BMC$;
2. $\angle AXY \equiv \angle AXZ$ iff X, Y, Z are collinear
3. $\angle XYZ \equiv 0°$ iff X, Y, Z are collinear
4. $\angle ABC + \angle BCA + \angle CAB \equiv 0$;
5. $\angle PQS \equiv \angle PRS$ iff P, Q, R and S are concyclic.

Property 1 is simply the basis of directedness: the relative positions don't matter. Property 2 is trivial if Y and Z lie on the same side of X, and the fact that adjacent angles add up to 180° if not. Property 3 just restates the fact that rotating a line onto itself leads to no rotation. Property 4 is the result that angles in a triangle add up to 180°, but also brings in the fact that the three angles are either all clockwise or all anti-clockwise. Property 5 is the really interesting one: it is *simultaneously* the same segment theorem and the alternate segment theorem, depending on the ordering of the points on the circle. The problem below illustrates why having a single theorem can be so important.

Directed areas are used even less often than directed angles and line segments, but are sometimes useful when adding areas to compute the area of a more complex shape. Conventionally, a triangle ABC has positive area if A, B and C are arranged in anti-clockwise order, and negative if they are arranged in clockwise order.

Exercise 3.1. *Three circles, Γ_1, Γ_2 and Γ_3 intersect at a common point O. Γ_1 and Γ_2 intersect again at X, Γ_2 and Γ_3 intersect again at Y, and Γ_3 and Γ_1 intersect against at Z. A is a point on Γ_1 which does not lie on Γ_2 or Γ_3. AX intersects Γ_2 again at B, and BY intersects Γ_3 again at C. Prove that A, Z and C are collinear.*

Exercise 3.2 (Simpson Line). *Perpendiculars are dropped from a point P to the sides of $\triangle ABC$ to meet BC, CA, AB at D, E, F respectively. Show that D, E and F are collinear if and only if P lies on the circumcircle of $\triangle ABC$.*

You will find that directed angles in particular play a large role in the theorems in this book, and they are introduced early on for this purpose. Do not be led to believe that directed angles are so wonderful that they should be used for all problems: theorems try to make very general statements and use directed angles for generality, but most problems are constrained so that normal angles are adequate (e.g. points inside triangles or acute angles). Normal angles are easier to work with simply because one does not need to think about whether to write $\angle ABC$ or $\angle CBA$.

4 Trigonometry

Trigonometry is seldom required to solve a problem. After all, trigonometry is really just a way of reasoning about similar triangles. However, it is a very powerful reasoning tool, and if applied correctly can replace a page full of unlikely and ungainly constructions with a few lines of algebra. If applied incorrectly, however, it can have the opposite effect.

The first thing to do before applying any trigonometry is to reduce the number of variables to the minimum. Then choose the variables that you want to keep very carefully. The compound angle formulae below make it easy to expand out many trig expressions, but if you have chosen the wrong variables to start with the task is almost impossible.

The following angle formulae are invaluable in manipulating trigonometric expressions. In the formulae below, a $\mp$ indicates a sign that is opposite to the sign chosen in a $\pm$.

$$\sin(A \pm B) = \sin A \cos B \pm \cos A \sin B \tag{4.1}$$
$$\cos(A \pm B) = \cos A \cos B \mp \sin A \sin B \tag{4.2}$$
$$\tan(A \pm B) = \frac{\tan A \pm \tan B}{1 \mp \tan A \tan B} \tag{4.3}$$
$$\cot(A \pm B) = \frac{\cot A \cot B \mp 1}{\cot A \pm \cot B} \tag{4.4}$$
$$\sin A \sin B = \left[\cos(A-B) - \cos(A+B)\right]/2 \tag{4.5}$$
$$\sin A \cos B = \left[\sin(A-B) + \sin(A+B)\right]/2 \tag{4.6}$$
$$\cos A \cos B = \left[\cos(A-B) + \cos(A+B)\right]/2 \tag{4.7}$$
$$\sin A \pm \sin B = 2 \sin\left(\frac{A \pm B}{2}\right) \cos\left(\frac{A \mp B}{2}\right) \tag{4.8}$$
$$\cos A + \cos B = 2 \cos\left(\frac{A+B}{2}\right) \cos\left(\frac{A-B}{2}\right) \tag{4.9}$$

$$\cos A - \cos B = 2\sin\left(\frac{B+A}{2}\right)\sin\left(\frac{B-A}{2}\right) \tag{4.10}$$

You don't need to memorise any of these other than the first three, because all the others can be obtained from these with simple substitutions. You should be aware that these transformations exist and know how to derive them, so that you can do so in an olympiad if necessary (see the exercises).

You can also use these to derive other formulae; for example, you can calculate $\sin n\theta$ and $\cos n\theta$ in terms of $\sin\theta$ and $\cos\theta$ fairly easily (for small, known values of n).

Exercise 4.1. *Prove equations* (4.4) *to* (4.10)*.*

Exercise 4.2. *In a $\triangle ABC$ (which is not right-angled), prove that*

$$\tan A + \tan B + \tan C = \tan A \tan B \tan C.$$

4.1 The extended sine rule

The standard Sine Rule says that

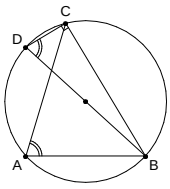
$$\frac{a}{\sin\alpha} = \frac{b}{\sin\beta} = \frac{c}{\sin\gamma}.$$

Theorem 4.1 (Extended Sine Rule). *In a triangle ABC,*

$$\frac{a}{\sin\alpha} = \frac{b}{\sin\beta} = \frac{c}{\sin\gamma} = 2R,$$

where R is the radius of the circumcircle.

Proof. Construct point D diametrically opposite B in the circumcircle of $\triangle ABC$. Then $\alpha = \angle CDB$ or $180° - \angle CDB$ and $\angle BCD = 90°$. It follows that $\frac{a}{\sin\alpha} = \frac{BC}{BC/BD} = 2R$, and similarly for $\frac{b}{\sin\beta}$ and $\frac{c}{\sin\gamma}$.



□

Exercise 4.3. *In a circle with centre O, AB and CD are diameters. From a point P on the circumference, perpendiculars PQ and PR are dropped onto AB and CD respectively. Prove that the length of QR is independent of the position of P.*

5 Circles

5.1 Cyclic quadrilaterals

A cyclic quadrilateral is a quadrilateral that can be inscribed in a circle. There are several results related to the angles of a cyclic quadrilateral that are covered in high school mathematics and which will not be repeated here. These results are still very important, and cyclic quadrilaterals appear in many unexpected places in olympiad problems.

Exercise 5.1 ($\star$). *Let $\triangle ABC$ have orthocentre H and let P be a point on its circumcircle. Let E be the foot of the altitude BH, let $PAQB$ and $PARC$ be parallelograms, and let AQ meet HR in X.*

(a) Show that H is the orthocentre of $\triangle AQR$.

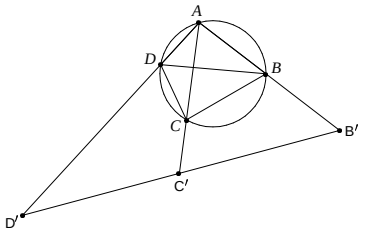
(b) Hence, or otherwise, show that EX is parallel to AP.

A result that is not normally taught in school is Ptolemy's Theorem. It is mainly useful if you have only one or two cyclic quadrilaterals, and lengths play a major role in the problem. It is also very useful when some more is known about the lengths. Equal lengths are particularly helpful as they can divide out of the equation.

Theorem 5.1 (Ptolemy's Theorem). *If $ABCD$ is a cyclic quadrilateral, then*

$$AB \cdot CD + BC \cdot AD = AC \cdot BD$$

Proof.



Choose an arbitrary constant K and construct B', C' and D' on AB, AC and AD respectively such that $AB \cdot AB' = AC \cdot AC' = AD \cdot AD' = K$.

Now consider $\triangle ABC$ and $\triangle AC'B'$. The angle at A is common and $\frac{AB}{AC'} = \frac{K/AB'}{K/AC} = \frac{AC}{AB'}$ and therefore the triangles are similar. It follows similarly that $\triangle ABD ||| \triangle AD'B'$ and $\triangle ACD ||| \triangle AD'C'$. Hence $\angle B'C'D' = \angle ABC + \angle ADC = 180°$ i.e. $B'C'D'$ is a straight line. From the similar triangles, we have $BC = B'C' \cdot \frac{AB}{AC'} = \frac{B'C' \cdot AB \cdot AC}{K}$, and similarly for CD and BD. Therefore

$$\begin{aligned} AC \cdot BD &= \tfrac{B'D'}{K}(AB \cdot AC \cdot AD) \\ &= (\tfrac{B'C'}{K} + \tfrac{C'D'}{K})(AB \cdot AC \cdot AD) \\ &= AB \cdot CD + AD \cdot BC \end{aligned}$$

This result relies on the fact that $B'C'D'$ is a straight line. If we had used a non-cyclic quadrilateral, this would not have been the case. This shows that the converse of Ptolemy's Theorem is also true. In fact the triangle inequality in $\triangle B'C'D'$ leads to *Ptolemy's Inequality*, which says that $AC \cdot BD \leq AB \cdot CD + AD \cdot BC$ for any quadrilateral $ABCD$, with equality precisely for cyclic quadrilaterals. □

Exercise 5.2. *Triangle ABC is equilateral. For any point P, show that $AP + BP \geq CP$ and determine when equality occurs.*

5.2 The Simpson line

The Simpson line was covered as exercise 3.2, but to emphasise its importance the statement is repeated here. A handy corollary is that the feet of perpendiculars from a point on the circumcircle cannot all meet the sides internally — which can limit the number of cases you need to consider.

Theorem 5.2 (The Simpson line). *Perpendiculars are dropped from a point P to the sides of $\triangle ABC$ to meet BC, CA, AB at D, E, F respectively. Show that D, E and F are collinear if and only if P lies on the circumcircle of $\triangle ABC$.*

This was exercise 3.2, so no proof is provided here.

Exercise 5.3. *From a point E on a median AD of $\triangle ABC$ the perpendicular EF is dropped to BC, and a point P is chosen on EF. Then perpendiculars PM and PN are drawn to the sides AB and AC.*

Now, it is most unlikely that M, E and N will lie in a straight line, but in the event that they do, prove that AP bisects $\angle A$.

5.3 Power of a point

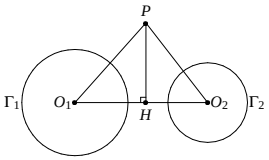
This section is based on the fact that if chords AB and CD of a circle intersect at a point P, then $PA \cdot PB = PC \cdot PD$ (even if P lies outside the circle). This is easily shown using similar triangles.

Consider fixing a point P and circle Γ and considering all possible chords AB that pass through P. Since $PA \cdot PB$ is equal for every pair of chords AB, it is equal for *all* such chords. This value is said to be the power of P with respect to Γ. The line segments are considered to be directed (see section 3), so P is negative inside the circle and positive outside of it. In fact by considering the chord that passes through O, the centre of Γ, it can be seen that the power of P is $d^2 - r^2$, where $d = OP$ and r is the radius of Γ. If P lies outside the circle then this also equals the square of the length of the tangent from P to Γ.

It is sometimes useful to know that the converse of the above result is true i.e. if $PA \cdot PB = PC \cdot PD$, where AB and CD pass through P, then A, B, C and D are concyclic (but only if using directed line segments).

5.3.1 The radical axis

Consider having two circles instead of one. What is the set of points which have the same power with respect to both circles? If the circles are concentric then no point will have the same power (since d will be the same and r different for every point), but the situation is less clear in general.



Consider two circles Γ_1 and Γ_2 with centres O_1 and O_2 with radii r_1 and r_2 respectively. Let P be a point which has equal powers with respect to Γ_1 and Γ_2, and let H be the foot of the perpendicular from P onto O_1O_2. Then

$$\begin{aligned}
& O_1P^2 - r_1^2 = O_2P^2 - r_2^2 && (5.1)\\
\iff \quad & O_1H^2 + HP^2 - r_1^2 = O_2H^2 + HP^2 - r_2^2 && (5.2)\\
\iff \quad & O_1H^2 - r_1^2 = O_2H^2 - r_2^2 && (5.3)\\
\iff \quad & O_1H^2 - r_1^2 = (O_2O_1 - HO_1)^2 - r_2^2 && (5.4)\\
\iff \quad & 2\cdot HO_1 \cdot O_2O_1 = {O_2O_1}^2 + r_1^2 - r_2^2 && (5.5)
\end{aligned}$$

We have eliminated P from the equation! In fact (5.3) shows that P has equal powers with respect to the circles iff H does. If $O_1O_2 \neq 0$ then we have a linear equation in HO_1 and so there is exactly one possibility for H (we are using directed line segments, so HO_1 uniquely determines H). Thus the locus of P is the line through H perpendicular to O_1O_2. This line is known as the *radical axis* of Γ_1 and Γ_2.

If the two circles intersect, the radical axis is easy to construct. The points of intersection both have zero power with respect to both circles, so both points lie on the radical axis. So the radical axis is simply the line through them.

Exercise 5.4. *Two circles are given. They do not intersect and neither lies inside the other. Show that the midpoints of the four common tangents are collinear.*

5.3.2 Radical centre

What happens when we consider three circles (say Γ_1, Γ_2 and Γ_3) instead of two? Firstly consider the case where the centres are not collinear. Then the radical axis of Γ_1 and Γ_2 will meet the radical axis of Γ_2 and Γ_3 at some point, say X (they will not be parallel because a radical axis is perpendicular to the line between the centres of the circles). Then from the definition of a radical axis, X has the same power with respect to all three circles and so it also lies on the radical axis of Γ_1 and Γ_2. The fact

that the three radical axes are concurrent at a point (known as the *radical centre*) can be used to solve concurrency problems.

If, however, the three centres are collinear, then all three radical axes are parallel. If they all coincide then all points on the common axis have equals powers with respect to the three circles; if not then no points do.

Exercise 5.5. *Show how to construct, using ruler and compass, the radical axis of two non-intersecting circles.*

Exercise 5.6 ($\star$). *Let A, B, C and D be four distinct points on a line, in that order. The circles with diameters AC and BD intersect at the points X and Y. The line XY meets BC at the point Z. Let P be a point on the line XY different from Z. The line CP intersects the circle with diameter AC at the points C and M, and the line BP intersects the circle with diameter BD at the points B and N. Prove that the lines AM, DN and XY are concurrent.*

6 Triangles

6.1 Introduction

A triangle would seem to be almost the simplest possible object in geometry, second only to the circle. It has only has two true degrees of freedom, since scaling a triangle up or down does not affect its properties. Yet the humble triangle contains an enormous amount of mathematics — in fact too much to fully explore here.

6.2 Tangents to the incircle

Let the lengths of the tangents to the incircle from A, B and C be x, y and z. Since $a = y+z$, $b = z+x$ and $c = x+y$, we can solve for x, y and z and get

$$x = \frac{-a+b+c}{2}, \quad y = \frac{a-b+c}{2}, \quad z = \frac{a+b-c}{2}.$$

This is the same notation that is introducted in section 2.

Exercise 6.1. *Determine the lengths of the tangents from B and C to the excircle opposite A.*

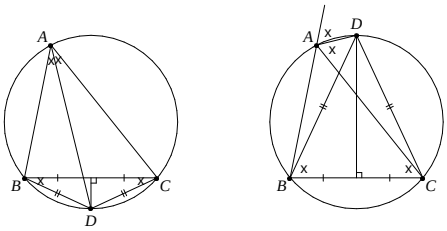
6.3 Triangles within triangles

There are specific names given to certain triangles formed from points of the original triangle:

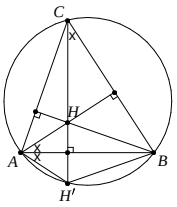
- The *medial* triangle has the midpoints of the original sides as its vertices.
- The *orthic* triangle has the feet of the altitudes as its vertices.
- A *pedal* triangle is the triangle formed by the feet of perpendiculars dropped from some point onto the three sides. If the point is the orthocentre, then this is the orthic triangle (and in fact some people use the term "pedal triangle" to refer to the orthic triangle).

6.4 Points on the circumcircle

Apart from the vertices, there are a few other points that are known to lie on the circumcircle. The first is the intersection point of a perpendicular bisector and the corresponding angle bisector. This is easily shown by taking the intersection of the perpendicular bisector and the circumcircle, which divides an arc (say BC) into two equal parts which subtend equal angles at A. This is also true (although less well known) in the case where the *external* angle bisector is used.



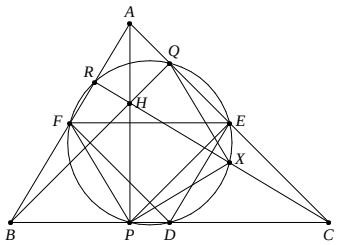
The second group of points that are known to lie on the circumcircle are the reflections of H (the orthocentre) in each of the three sides. This is an exercise in angle chasing, using the known results about the angles in cyclic quadrilaterals.



Exercise 6.2. *A rectangle $HOMF$ has $HO = 23$ and $OM = 7$. Triangle ABC has orthocentre H and circumcentre O. The midpoint of BC is M and F is the foot of the altitude from A. Determine the length of side BC.*

6.5 The nine-point circle

A rather interesting circle that arises in a triangle is the so-called nine-point circle. Let us examine the circumcircle of the triangle whose vertices are the midpoints of $\triangle ABC$ (the medial triangle). Firstly, what is its radius? The medial triangle is a half sized version of the original triangle (because of the midpoint theorem), so its circumradius will also be half that of the large triangle, i.e. it will be $\frac{R}{2}$.



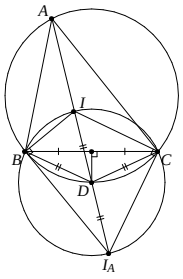
Now let us see what other points this circle passes through. From the diagram it appears that it passes through the feet of the altitudes, so let us prove this. Since F is the midpoint of the hypotenuse of $\triangle APB$, we have $\angle FPA = \angle FAP = 90° - \beta$. Similarly $\angle EPA = 90° - \gamma$ and so $\angle FPE = \alpha = \angle FDE$ (since $\triangle ABC \,|||\, \triangle DEF$). It follows that P lies on the circle. Similarly Q and R also lie on the circle.

Point X is the midpoint of HC, and it also appears to lie on the circle. HC is the diameter of the circle passing through H, Q, C and P, so X is the centre of this circle. It follows that $\angle PXQ = 2\angle PCQ = 2\gamma$. But $\angle PEQ = \angle PEF + \angle FEQ = \angle PDF + \angle FEQ = \gamma + \gamma$, so $\angle PEQ = \angle PXQ$ and so X lies on the circle. Similarly the midpoints of HA and HC lie on the circle.

Because there are nine well-defined points which lie on this circle, it is known as the nine-point circle.

6.6 Another circle

Consider that $\angle I_A BI = \angle I_A CI = 90°$; this shows that II_A is the diameter of a circle passing through I, I_A, B and C. Where is the centre of this circle? Well, any circle passing through B and C must have its centre on the perpendicular bisector of BC, and for II_A to be the diameter, the centre must also lie on the internal bisector of $\angle A$. Hence the centre is the intersection of these two lines. As shown above, the intersection also lies on the circumcircle of $\triangle ABC$.



Exercise 6.3 ($\star$). *In acute-angled triangle ABC the internal bisector of angle A meets the circumcircle of the triangle again at A_1. Points B_1 and C_1 are defined similarly. Let A_0 be the point of intersection of the line AA_1 with the external bisectors of angles B and C. Points B_0 and C_0 are defined similarly. Prove that*

(i) the area of the triangle $A_0B_0C_0$ is twice the area of the hexagon $AC_1BA_1CB_1$;

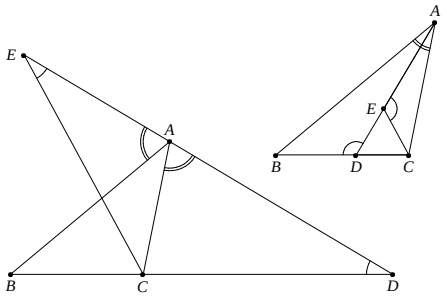
(ii) the area of the triangle $A_0B_0C_0$ is at least four times the area of the triangle ABC.

6.7 Theorems

Angle bisectors can be fairly tricky to deal with. The angle bisector theorem provides a way to compute the segments which the base is divided into.

Theorem 6.1 (Angle bisector theorem). *If D is the point of intersection of BC with an angle bisector of $\angle A$, then $\frac{DB}{DC} = \frac{AB}{AC}$.*

Proof. Construct E on AD such that $\angle AEC = \angle BDA$. Then $\triangle ABD \,|||\, \triangle ACE$ (two angles) and so $\frac{DB}{EC} = \frac{AB}{AC}$. But $\triangle ECD$ is isosceles, so $CE = CD$ and therefore $\frac{DB}{DC} = \frac{AB}{AC}$ as required.



□

Exercise 6.4. *In the right-hand diagram for the angle-bisector theorem, find a formula for the length BD in terms of the side lengths a, b and c.*

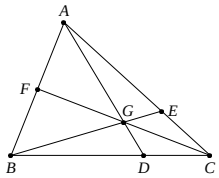
Exercise 6.5. *Given a line segment AB and a real number $r > 0$, find the locus of points P such that $\frac{AP}{BP} = r$.*

The theorems of Ceva and Menelaus are handy results when proving concurrency and collinearity respectively. They are particularly powerful because their converses are true, provided that the directions are taken into account. The converses are quite easy to prove by assuming them to be false, and then constructing two different points with the same uniquely defining properties.

Theorem 6.2 (Ceva's Theorem). *If AD, BE and CF are concurrent cevians of $\triangle ABC$ then*

$$\frac{BD}{DC} \cdot \frac{CE}{EA} \cdot \frac{AF}{FB} = 1$$

Proof.



Let G be the point of concurrency.

$$\frac{|\triangle ABD|}{|\triangle ACD|} = \frac{BD}{DC} \quad \text{(common height)}$$
$$\frac{|\triangle GBD|}{|\triangle GCD|} = \frac{BD}{DC} \quad \text{(common height)}$$
$$\therefore \quad \frac{|\triangle AGB|}{|\triangle CGA|} = \frac{BD}{DC}$$

We can show similar things for $\frac{CE}{EA}$ and $\frac{AF}{FC}$. Therefore

$$\frac{BD}{DC}\cdot\frac{CE}{EA}\cdot\frac{AF}{FB} = \frac{|\triangle AGB|}{|\triangle CGA|}\cdot\frac{|\triangle BGC|}{|\triangle AGB|}\cdot\frac{|\triangle CGA|}{|\triangle BGC|} = 1$$

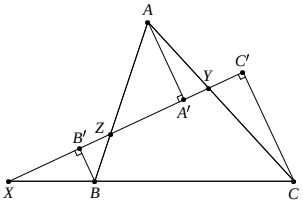
This proof has not explicitly invoked directed areas or line-segments, but if they are used it can be seen that the result will hold even if G lies outside of the triangle. □

Theorem 6.3 (Menelaus' Theorem). *If X, Y and Z and collinear and lie on sides BC, CA and AB (or their extensions) of $\triangle ABC$ respectively, then*

$$\frac{AZ}{ZB}\cdot\frac{BX}{XC}\cdot\frac{CY}{YA} = -1$$

(Note that the sign on the result is due to directed line segments, and indicates that the line cuts the sides themselves either twice or not at all.

Proof.



Drop perpendiculars from A, B and C to meet XYZ at A', B' and C'. From alternate angles, we have $\triangle AA'Z \,|||\, \triangle BB'Z$ and thus $\frac{AZ}{ZB} = \frac{AA'}{B'B}$. Similarly $\frac{BX}{XC} = \frac{BB'}{C'C}$ and $\frac{CY}{YA} = \frac{CC'}{A'A}$. Therefore

$$\frac{AZ}{ZB}\cdot\frac{BX}{XC}\cdot\frac{CY}{YA} = \frac{AA'}{B'B}\cdot\frac{BB'}{C'C}\cdot\frac{CC'}{A'A} = -1$$

□

Exercise 6.6. *Use Menelaus' Theorem to prove Ceva's Theorem.*

Exercise 6.7 ($\star$). *ABC is an isosceles triangle with $AB = AC$. Suppose that*

(i) M is the midpoint of BC and O is the point on the line AM such that $OB \perp AB$;

(ii) Q is an arbitrary point on the segment BC different from B and C;

(iii) E lies on the line AB and F lies on the line AC such that E, Q and F are distinct and collinear.

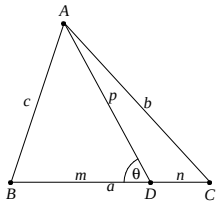
Prove that OQ is perpendicular to EF if and only if $QE = QF$.

Stewart's Theorem is a handy tool for dealing with the length of a cevian, which is otherwise difficult to work with.

Theorem 6.4 (Stewart's Theorem). *Suppose AD is a cevian in $\triangle ABC$. Let $p = AD$, $m = BD$ and $n = CD$. Then*

$$a(p^2 + mn) = b^2 m + c^2 n.$$

Proof.



Use the cosine rule in $\triangle ABD$:

$$c^2 = m^2 + p^2 - 2mp\cos\theta$$
$$\therefore \quad c^2 n = m^2 n + p^2 n - 2mnp\cos\theta \tag{6.1}$$

Do the same in $\triangle ACD$, noting that $\cos(180^\circ - \theta) = -\cos\theta$:

$$b^2 = n^2 + p^2 + 2np\cos\theta$$
$$\therefore \quad b^2 m = n^2 m + p^2 m + 2mnp\cos\theta \tag{6.2}$$

Now add (6.1) and (6.2):

$$b^2 m + c^2 n = m^2 n + n^2 m + p^2 n + p^2 m \tag{6.3}$$
$$= (m+n)(p^2 + mn) \tag{6.4}$$
$$= a(p^2 + mn) \tag{6.5}$$

□

In the special case that AD is a median, Stewart's Theorem reduces to $4p^2 + a^2 = 2(b^2 + c^2)$, which is known as Apollonius' Theorem.

Exercise 6.8. *In $\triangle ABC$, angle A is twice angle B. Prove that $a^2 = b(b+c)$.*

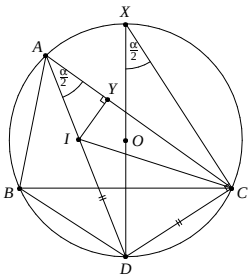
Theorem 6.5 (Euler's Formula).

$$OI^2 = R(R-2r)$$

As a corollary, we have Euler's Inequality:

$$R \geq 2r.$$

Proof. Extend the angle bisector from A to meet the circumcircle again at D. Also construct X diametrically opposite D on the circumcircle and construct Y as the foot of the perpendicular from I onto AC. We calculate the power of I with respect to the circumcircle (see section 5.3), which is equal to $OI^2 - R^2$ and also to $-AI \cdot ID$. From section 6.6, we have $ID = CD$.



Now we note that $\triangle DXC \,|||\, \triangle IAY$, and so $\frac{AI}{IY} = \frac{XD}{DC} \iff AI \cdot ID = 2rR$. Since $OI^2 - R^2 = -AI \cdot ID$, it follows that $OI^2 = R(R-2r)$ as required. □

Euler's Theorem provides a measure of the distance between the incentre and circumcentre. However it is most often invoked as Euler's Inequality.

Exercise 6.9 ($\star$). *Let r be the inradius and R the circumradius of ABC and let p be the inradius of the orthic triangle of triangle ABC. Prove that*

$$\frac{p}{R} \leq 1 - \frac{1}{3}\left(1 + \frac{r}{R}\right)^2.$$

6.8 Area

There are numerous formulae for the area of a triangle, and in many cases things can be discovered by equating them.

Theorem 6.6 (Heron's Formula).

$$K = \sqrt{sxyz}$$

Proof. This is probably the ugliest proof in this booklet. Here goes:

$$\begin{aligned}
16K^2 &= 4(ab\sin\gamma)^2 \\
&= 4a^2b^2(1-\cos^2\gamma) \\
&= 4a^2b^2\left[1-\left(\frac{a^2+b^2-c^2}{2ab}\right)^2\right] \\
&= 4a^2b^2-(a^2+b^2-c^2)^2 \\
&= (2ab-a^2-b^2+c^2)(2ab+a^2+b^2-c^2) \\
&= \left[c^2-(a-b)^2\right]\left[(a+b)^2-c^2\right] \\
&= (c-a+b)(c+a-b)(a+b+c)(a+b-c) \\
&= 16sxyz.
\end{aligned}$$

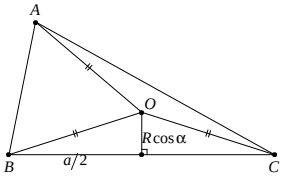
□

Theorem 6.7 (Triangle area formulae).

$$\begin{aligned}
K &= \tfrac{1}{2}ah_a = \tfrac{1}{2}bh_b = \tfrac{1}{2}ch_c && (6.6)\\
&= \tfrac{1}{2}ab\sin\gamma = \tfrac{1}{2}bc\sin\alpha = \tfrac{1}{2}ca\sin\beta && (6.7)\\
&= \frac{abc}{4R} && (6.8)\\
&= 2R^2\sin\alpha\sin\beta\sin\gamma && (6.9)\\
&= \tfrac{1}{2}R(a\cos\alpha+b\cos\beta+c\cos\gamma) && (6.10)\\
&= R(a\cos\beta\cos\gamma+b\cos\gamma\cos\alpha+c\cos\alpha\cos\beta) && (6.11)\\
&= rs && (6.12)\\
&= r_ax = r_by = r_cz && (6.13)\\
&= \sqrt{sxyz} \qquad \textit{(Heron's Formula)} && (6.14)
\end{aligned}$$

Proof. The first is the standard formula for the area of a triangle. The second is really the same formula, since $\sin\gamma = \frac{h_a}{b}$. The third is obtained using the extended sine rule ($\sin\gamma = \frac{c}{2R}$). The fourth is similarly obtained using the extended sine rule by converting all side lengths to sines.

Equation 6.9 is obtained by adding the areas of the isosceles triangles $\triangle BOC$, $\triangle COA$ and $\triangle AOB$. The base of $\triangle BOC$ is a and $\angle BOC = 2\angle BAC = 2\alpha$, so the height is $OC\cos\alpha = R\cos\alpha$. Adding up the areas gives the result.



The following equation is obtained from 6.9 by replacing a by $b\cos\gamma + c\cos\beta$ and similarly for b and c.

Equation 6.12 is obtained similarly to 6.9, but using I instead of O. The three triangles all have height r, so the area is $\frac{1}{2}(ra+rb+rc) = rs$. Equation 6.13 uses the excentre I_a instead; in this case one adds triangles ABI_a and ACI_a and subtracts triangle BCI_a.

Heron's Formula was covered earlier. □

Exercise 6.10. *An equilateral triangle has sides of length $4\sqrt{3}$. A point Q is located inside the triangle so that its perpendicular distances from two sides of the triangle are 1 and 2. What is the perpendicular distance to the third side?*

Exercise 6.11. *Prove that*

$$\frac{1}{r} = \frac{1}{r_a} + \frac{1}{r_b} + \frac{1}{r_c}.$$

There is one area more formula that is used with coordinate geometry.

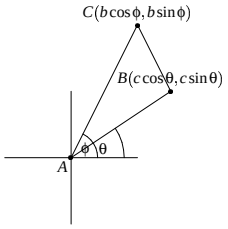
Theorem 6.8. *If one vertex of a triangle is at the origin and the other two are at (x_1, y_1) and (x_2, y_2), then*

$$K = \tfrac{1}{2}\left|x_1y_2 - x_2y_1\right|.$$

If the absolute value operator is removed, one gets a formula for directed area[1].

Proof. The proof below uses trigonometry. It is also possible to compute the area of the triangle by starting with a rectangle that bounds it, and subtracting right triangles. However, that approach requires several cases to be considered.

[1] The sign is used in computer graphics to determine whether three points are wound clockwise or anti-clockwise.



Assume without loss of generality that C makes a larger angle from the x-axis than B (swapping B and C simply negates the term inside the absolute value). Then $(x_1, y_1) = (c\cos\theta, c\sin\theta)$, $(x_2, y_2) = (b\cos\phi, b\sin\phi)$ and the area is

$$\begin{aligned}\tfrac{1}{2}bc\sin\alpha &= \tfrac{1}{2}bc\sin(\phi-\theta)\\ &= \tfrac{1}{2}bc(\sin\phi\cos\theta-\cos\theta\sin\phi\\ &= \tfrac{1}{2}(x_1y_2-x_2y_1).\end{aligned}$$

□

6.9 Inequalities

Inequalities in triangles are often best solved by first expressing all the quantities in terms of as few variables as possible (ideally, only two or three) and then using inequality techniques discussed in *Inequalities for the Olympiad Enthusiast* to finish the problem algebraically. Jensen's Inequality is particularly powerful when combined with trigonometric functions.

Theorem 6.9 (Jensen's Inequality). *A function f is said to be* convex *on an interval $[a,b]$ if $\frac{f(x)+f(y)}{2} \geq f\left(\frac{x+y}{2}\right)$ for all $x,y \in [a,b]$. If f is convex*[2] *on $[a,b]$ then for any $x_1, x_2, \ldots, x_n$ in $[a,b]$ we have*

$$f\left(\frac{x_1+\cdots+x_n}{n}\right) \leq \frac{f(x_1)+\cdots+f(x_n)}{n}.$$

The statement also holds if all inequality signs are reversed, in which case the function is termed concave.

[2] If you are familiar with calculus, a convex function is one that satisfies $f''(x) \geq 0$ for all $x \in [a,b]$.

Proof. Refer to page 18 of *Inequalities for the Olympiad Enthusiast*, by Graeme West. □

Exercise 6.12. *If α, β, γ are the angles of a triangle, then show that* $\sin\alpha + \sin\beta + \sin\gamma \leq \frac{3\sqrt{3}}{2}$.

One thing to keep in mind is the triangle inequality: if you reduce the problem to an inequality in a, b and c then it is possible (although not necessarily the case) that you will need to use the fact that the sum of any two is greater than the third. A technique that sometimes simplifies this to substituting $a = x+y$, $b = y+z$, $c = z+x$ in which case the triangle inequality is equivalent to $x,y,z > 0$. In some circles this has become known as the Ravi Substitution, after a Canadian IMO contestant (and later coach) Ravi Vakil. Although he did not invent the technique, he successfully applied it to an IMO problem.

There are a few other useful inequalities that are specific to triangles. The first is Euler's Inequality, mentioned above. The others are listed below.

Theorem 6.10. *In a triangle ABC,*

$$\frac{3\sqrt{3}}{2}R \geq s \qquad s^2 \geq 3\sqrt{3}K \qquad K \geq 3\sqrt{3}r^2.$$

In each case, equality occurs iff $\triangle ABC$ is equilateral.

Proof. We first prove that $\frac{3\sqrt{3}}{2}R \geq s$. From the extended sine rule, $\frac{a}{2R} = \sin\alpha$ and so

$$\begin{aligned}\frac{s}{R} &= \sin\alpha+\sin\beta+\sin\gamma\\ &\leq 3\sin\left(\frac{\alpha+\beta+\gamma}{3}\right) \qquad \text{(Jensen's Inequality)}\\ &= 3\sin 60^\circ\\ &= \frac{3\sqrt{3}}{2}.\end{aligned}$$

For the remaining inequalities, we express everything in terms of x, y and z. Thus

$$\begin{aligned}s^2 &= s^{3/2}\sqrt{s}\\ &= \sqrt{s(x+y+z)^3}\\ &\geq \sqrt{27sxyz} \qquad \text{(AM-GM)}\\ &= 3\sqrt{3}K \qquad \text{(Heron's Formula).}\end{aligned}$$

$$\begin{aligned} K &= \frac{r^2s^2}{K} \\ &\geq \frac{3\sqrt{3}r^2K}{K} \qquad \text{(from the previous step)} \\ &= 3\sqrt{3}r^2. \end{aligned}$$

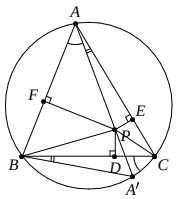
□

Theorem 6.11 (Erdős-Mordell). *Let P be a point inside triangle $\triangle ABC$, and let the feet of the perpendiculars from P to BC, CA, AB be D, E, F respectively. Then*

$$AP + BP + CP \geq 2(DP + EP + FP).$$

Proof. Extend AP to meet the circumcircle of $\triangle ABC$ at A'. Let $\angle BAP = \theta$ and $\angle CAP = \phi$. Note that $FP = AP\sin\theta$ and $EP = AP\sin\phi$, so $\frac{EP}{FP} = \frac{\sin\phi}{\sin\theta} = \frac{CA'}{BA'}$. Also note that $a \cdot AA' = b \cdot BA' + c \cdot CA'$ (from Ptolemy's Theorem in the cyclic quadrilateral $ACA'B$), so $AA' = \frac{b}{a} \cdot BA' + \frac{c}{a} \cdot CA'$. Now

$$\begin{aligned} AP &= \frac{FP}{\sin\theta} \\ &= \frac{FP \cdot 2R}{BA'} \qquad \text{(Extended Sine Rule)} \\ &\geq \frac{FP \cdot AA'}{BA'} \qquad (AA' \text{ is less than the diameter}) \\ &= \frac{FP(b \cdot BA' + c \cdot CA')}{a \cdot BA'} \\ &= \frac{b}{a} \cdot FP + \frac{c}{a} \cdot \frac{CA'}{BA'} \cdot FP \\ &= \frac{b}{a} \cdot FP + \frac{c}{a} \cdot EP. \end{aligned}$$



Now we can establish similar inequalities for BP and CP, and adding these gives

$$\begin{aligned} PA + PB + PC &\geq \left(\frac{b}{c} + \frac{c}{b}\right) PD + \left(\frac{c}{a} + \frac{a}{c}\right) PE + \left(\frac{a}{b} + \frac{b}{a}\right) PF \\ &\geq 2(PD + PE + PF). \qquad \text{(AM-GM)} \end{aligned}$$

□

Exercise 6.13. *Let ABC be a triangle and P be an interior point in ABC. Show that at least one of the angles PAB, PBC, PCA is less than or equal to 30 degrees.*

7 Transformations

A very powerful idea in geometry is that of a transformation. A transformation maps every point in space to some other point in space. Structures like lines or circles are transformed by applying the transformation to every point on them. They do not necessary maintain their shapes; in fact there is a transformation (inversion) which generally maps lines to circles! Each transformation will preserve certain properties of a diagram, and by translating the properties of the original into the transformed diagram one can obtain new information. Here a diagram is really just a set of points.

7.1 Affine transformations

The transformations we discuss here are all *affine*. That means that straight lines are mapped to straight lines, and lengths are scaled uniformly. The transformations presented here all preserve angles as well. These transformations can in fact be built up by combining reflections and scale changes, although this is not necessarily the best way to think about them.

7.2 Translations, rotations and reflections

The simplest transformation is a translation: every point simply moves a constant distance in a constant direction; this is like picking up a piece of paper and moving it, without rotating it. Rotations rotate all the points by some angle around a particular point; this is like sticking a pin in a piece of paper and then turning it. Reflections take all points and reflect them in a particular line; this is like picking up the piece of paper and putting it down upside-down (the paper would of course need to be thin enough for the diagram to be seen through the back).

While these are all quite straightforward, they can also be very powerful because they preserve so much. They are also closely related, as shown by the next problem.

Exercise 7.1. *In each of the following, show that the transformations exist using a concrete construction.*

(a) Show that any rotation or translation can be expressed as the combination of a pair of reflections, or vice versa.

(b) Show that two rotations, two translations or a translation and rotation can always be combined to produce a single translation or rotation.

(c) Show that any combination of translations, reflections and rotations yields either a rotation, a translation, or a translation followed by a reflection.

Exercise 7.2. *In acute-angled triangle ABC, a point P is given on side BC. Show how to find Q on CA and R on AB such that $\triangle PQR$ has the minimum perimeter.*

Exercise 7.3 ($\star$). *The point O is situated inside the parallelogram $ABCD$ so that $\angle AOB + \angle COD = 180^\circ$. Prove that $\angle OBC = \angle ODC$.*

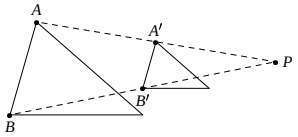
7.3 Homothetisms

So far we have discussed only *rigid* transforms, namely those that can be illustrated with a piece of paper. We now move on to scaling. Imagine drawing a diagram on a new T-shirt, and then letting the T-shirt shrink in the wash. Assume the ink doesn't run and that the T-shirt doesn't warp, you will have the same diagram, only smaller. All the angles and so on will be the same, although lengths will not.

A *homothetism* is a fancy name for scaling. One chooses a centre (sometimes called the "centre of similitude") and a scale factor. Every point is then kept in the same direction relative to the centre, but its distance from the centre is scaled by the scale factor. Like translations, homothetisms preserve orientation, angles, and ratios of lengths. However, lengths are scaled by the scale factor. The result below allows one to find the centre of a homothetism.

Theorem 7.1. *Let S and T be two similar figures which have the same orientation, but are not the same size. Then there is a homothetism that maps S to T.*

Proof. Pick a point A in S and its corresponding point A' in T. Now pick a second point B in S, not on AA', and its corresponding point in B[3]. Now if AA' and BB' are parallel then $AA'B'B$ would be a parallelogram, making $AB = A'B'$. But we assumed that S and T are of different sizes, which would give a contradiction. Hence AA' and BB' meet at a point, which we will call P. Now consider the homothetism with centre of similitude P and scale factor $\frac{A'P}{AP}$. It will clearly map A to A'; will it map B to B'? Yes, because $\triangle ABP \,|||\, \triangle A'B'P$ by parallel lines. If we can show that this homothetism maps the rest of S to T then we are done.



Let C be some arbitrary point in S. We aim to show that the homothetism maps C to its corresponding point C' in T. If C is A or B then we are done. If C lies on AB then C is uniquely defined by $\frac{AC}{BC}$ (with directed line segments). But homothetisms preserve ratios of lengths, and $\frac{A'C'}{B'C'} = \frac{AC}{BC}$ so C is mapped to C'. If C does not lie on AB then C is uniquely defined by the directed angles $\angle BAC$ and $\angle ABC$, and angles are preserved by homothetisms. □

The construction also suggests how the centre of similitude can be found in practice: take two pairs of corresponding points and find the intersection of the lines between them. For example, any two circles of different sizes satisfy the requirements, so a homothetism can be found between them. The points of tangency of the common tangent are corresponding points, since they have the same orientation relative to the centre. Hence the centre of similitude is the intersection of the common tangents.

What happens if we have non-overlapping circles, and use the *other* pair of common tangents? It turns out that this point is also a centre of similitude. However, this homothetism has a negative scale factor, which means that points are "sucked" through the centre and pushed out the other side. This also rotates the figure by 180°, although for a circle this isn't visible. The theorem above in fact applies to situations where the two figures have orientations that are out by 180°, in which case a negative scale factor is used. In this case the figures may even by the same size (since the scale factor is -1, not 1).

Exercise 7.4. *Let ABC be a triangle. Use a homothetism to show that*

(a) the medians of $\triangle ABC$ are concurrent;

[3] If no such B exists, then make some arbitrary construction in S and the corresponding construction in T to produce such a B.

(b) the point of concurrency (the centroid) divides the medians in a $2:1$ *ratio;*

(c) the orthocentre H*, the centroid* G *and the circumcentre* O *are collinear, with* $HG:GO=2:1$ *(this line is known as the Euler line). Assume that* H *and* O *exist (i.e. that the defining lines are concurrent).*

Exercise 7.5 ($\star$). *On a plane let* C *be a circle,* L *be a line tangent to the circle* C *and* M *be a point on* L*. Find the locus of all points* P *with the following property: there exist two points* Q,R *on* L *such that* M *is the midpoint of* QR *and* C *is the inscribed circle of triangle* PQR*.*

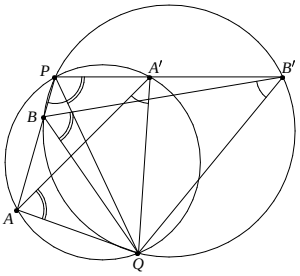
7.4 Spiral similarities

An even more general transformation than a homothetism is a spiral similarity. A spiral similarity combines the effects of a homothetism and a rotation: the plane is not only scaled around a centre P by some factor r, it is also rotated around P by an angle θ. A spiral similarity preserves pretty much the same things as homothetisms i.e. ratio of lengths and angles. However, corresponding lines are no longer parallel, but meet each other at an angle of θ. As for homothetisms, there is a result that makes it possible to find a spiral similarity given two similar figures.

Theorem 7.2. *Let* S *and* T *be two sets of points that are similar but have either different orientation or different size (or both). Then there is a spiral similarity that maps* S *to* T*.*

Proof. In the special case that S and T have the same orientation, there exists a homothetism, which is just a special case of a spiral similarity. So we assume that S and T have different orientations. We also include the case where S and T are oriented 180° apart in the special case, as this is a homothetism with negative scale factor.

Choose two arbitrary points A and B in S, and their corresponding points A' and B' in T. Let P be the intersection of AB and $A'B'$. Construct the circumcircles of $\triangle AA'P$ and $\triangle BB'P$, and let their second point of intersection be Q (Q exists because of the assumptions).



Now $\angle AQA' \equiv \angle APA' \equiv BPB' \equiv BQB'$, $\angle AA'Q \equiv \angle APQ \equiv \angle BPQ \equiv \angle BB'Q$ and similarly $\angle A'AQ \equiv B'BQ$. It follows that triangles $AA'Q$ and $BB'Q$ are directly similar[4]. Now consider the spiral similarity with centre Q, angle AQA' and scale factor $\frac{A'Q}{AQ}$. It will map A to A' by construction, and from the similar triangles it will map B to B'. We can now proceed to show that S is mapped to T, as was done in the corresponding theorem for homothetisms. □

Exercise 7.6. *Squares are constructed outwards on the sides of triangle* ABC*. Let* P*,* Q *and* R *be the centres of the squares opposite* A*,* B *and* C *respectively. Prove that* AP *and* QR *are equal and perpendicular.*

8 Miscellaneous problems

These problems all draw on the techniques in this book, but do not fit well into any particular section. They are mostly very challenging problems designed to give you practice.

Exercise 8.1 ($\star$). *$ABCD$ is a square. P is a point inside the square with $\angle ABP = \angle BAP = 15^\circ$. Show that $\triangle CDP$ is equilateral.*

Exercise 8.2 ($\star$). *A 6m tall statue stands on a pedestal, so that the foot of the statue is 2m above your head height. Determine how far from the statue you should stand so that it appears as large as possible in your vision.*[5]

[4] Two triangles are directly similar if they are similar and have the same clockwise/anti-clockwise orientation.

[5] In other words, maximise the angle formed by the foot of the statue, your head and the top of the statue.

Exercise 8.3 (⋆). *In an acute angled triangle ABC the interior bisector of $\angle A$ intersects BC at L and the circumcircle of $\triangle ABC$ again at N. From point L perpendiculars are drawn to AB and AC, the feet of these perpendiculars being K and M respectively. Prove that the quadrilateral $AKNM$ and the triangle ABC have equal areas.*

Exercise 8.4 (⋆). *ABC is a triangle. The internal bisector of the angle A meets the circumcircle again at P. Q and R are similarly defined relative to B and C. Prove that*

$$AP + BQ + CR > AB + BC + CA.$$

Exercise 8.5 (⋆). *A circle of radius r is inscribed in a triangle ABC with area K. The points of tangency with BC, CA and AC are X, Y and Z respectively. AX intersects the circle again in X'. Prove that $BC \cdot AX \cdot XX' = 4rK$.*

Exercise 8.6 (⋆). *A semicircle is drawn on one side of a straight line ℓ. C and D are points on the semicircle. The tangents at C and D meet ℓ again at B and A respectively, with the centre of the semicircle between them. Let E be the point of intersection of AC and BD, and F the point on ℓ such that EF is perpendicular to ℓ. Prove that EF bisects $\angle CFD$.*

Exercise 8.7 (⋆). *In $\triangle ABC$, let D and E be points on the side BC such that $\angle BAD = \angle CAE$. If M and N are, respectively, the points of tangency with BC of the incircles of $\triangle ABD$ and $\triangle ACE$, show that $\dfrac{1}{MB} + \dfrac{1}{MD} = \dfrac{1}{NC} + \dfrac{1}{NE}$.*

Exercise 8.8 (⋆). *Let P be a point inside $\triangle ABC$ such that*

$$\angle APB - \angle ACB = \angle APC - \angle ABC.$$

Let D, E be the incentres of $\triangle APB, \triangle APC$ respectively. Show that AP, BD and CE meet at a point.

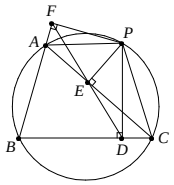
9 Solutions

3.1 Using classical geometry to solve this problem would result in an enormous number of different cases. However, directed angles hide all of that, and the result appears with a few lines of basic calculation:

$$\begin{aligned}
\angle AZC &\equiv \angle AZO + \angle OZC \\
&\equiv \angle AXO + \angle OYC \qquad \text{(concyclic points)} \\
&\equiv \angle BXO + \angle OYB \qquad \text{(collinear points)} \\
&\equiv \angle BXO + \angle OXB \qquad \text{(concyclic points)} \\
&\equiv \angle BXO - \angle BXO \qquad \text{(directed angles)} \\
&\equiv 0^\circ,
\end{aligned}$$

and hence A, Z and C are collinear.

3.2 Note that PC subtends right angles at D and E, and hence is the diameter of a circle passing through P, C, D and E. Similarly, P, A, F and E are concyclic.



$$\begin{aligned}
\angle DEF &\equiv \angle DEP + \angle PEF \\
&\equiv \angle DCP + \angle PAF \\
&\equiv \angle BCP - \angle BAP.
\end{aligned}$$

It follows that $\angle DEF \equiv 0^\circ \iff \angle BCP \equiv \angle BAP$. The first is a condition for D, E, F to be collinear and the second is a condition for P to lie on the circumcircle of $\triangle ABC$.

4.1 $\cot(A \pm B) = \frac{\cot A \cot B \mp 1}{\cot A \pm \cot B}$ can be shown by substituting $\tan\theta = \frac{1}{\cos\theta}$ into $\tan(A \pm B) = \frac{\tan A \pm \tan B}{1 \mp \tan A \tan B}$ and simplifying. The expressions for $\sin A \sin B$ and similar expressions can be proved simply by expanding the right hand side and cancelling terms. The final three equations are derived by making suitable substitutions into the previous three.

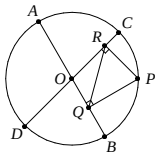
4.2 We first derive a general formula for $\tan(A+B+C)$.

$$\begin{aligned}
\tan(A+B+C) &= \tan[(A+B)+C] \\
&= \frac{\tan(A+B) + \tan C}{1 - \tan(A+B)\tan C} \\
&= \frac{\frac{\tan A + \tan B}{1 - \tan A \tan B} + \tan C}{1 - \frac{\tan A + \tan B}{1 - \tan A \tan B} \cdot \tan C}
\end{aligned}$$

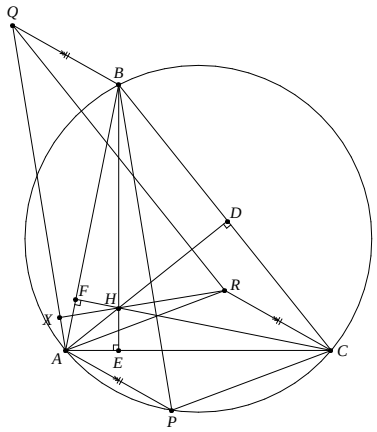
$$= \frac{\tan A + \tan B + (1 - \tan A \tan B)\tan C}{(1 - \tan A \tan B) - (\tan A + \tan B)\tan C}$$
$$= \frac{\tan A + \tan B + \tan C - \tan A \tan B \tan C}{1 - \tan A \tan B - \tan B \tan C - \tan C \tan A}.$$

However, we know that $\tan(A+B+C) = \tan 180^\circ = 0$, so the numerator must be 0. The result follows.

4.3 Suppose that P lies on the arc BC, as in the diagram. Then $OQPR$ is cyclic with diameter OP, so applying the extended sine rule in $\triangle OQR$ gives $QR = OP\sin\angle BOC$. Now $\angle BOC$ is fixed and OP is the radius of the circle, also fixed. So QR is fixed if P lies on the arc BC. But $\sin\angle BOC = \sin\angle COA = \sin\angle DOA = \sin\angle DOB$, so QR is constant wherever on the circle P may be.



5.1 This problem is fairly straight-forward as it consists almost entirely of angle chasing. The only difficulty is that P can lie anywhere on the circumcircle, which could give rise to multiple cases. We can get around this with directed angles. This diagram is thus only for reference. D and F are the feet of the altitudes from A and C in $\triangle ABC$.



(a) Firstly notice that since $PAQB$ and $PARC$ are parallelograms, BQ and CR are parallel and equal (and in the same direction), so $BCRQ$ is also a parallelogram. It follows that $RQ \parallel CB$ and hence $AH \perp RQ$. This shows that H lies on one altitude of $\triangle AQR$. If $RX \perp AQ$ then it would lie on another altitude we would be done.

Note that B, D, H and F are concyclic. Thus

$$\begin{aligned}
\angle AHC &\equiv \angle DHF && \text{(opposite angles)}\\
&\equiv \angle DBF && (D, H, F, B \text{ concyclic})\\
&\equiv \angle CBA &&\\
&\equiv \angle CPA && (A, B, C, P \text{ concyclic})\\
&\equiv \angle ARC && (AP \parallel RC, AR \parallel PC)
\end{aligned}$$

and therefore A, H, R and C are concyclic. Thus

$$\begin{aligned}
\angle AXR &\equiv \angle XAR + \angle ARX\\
&\equiv \angle XAB + \angle BAC + \angle CAR + \angle ARH\\
&\equiv \angle QAB + \angle FAC + \angle ACP + \angle ACH
\end{aligned}$$

$$\equiv \angle PBA + \angle ABP + \angle FAC + \angle ACF$$
$$\equiv \angle AFC$$
$$\equiv 90^\circ$$

and the result follows.

(b) This is just more angle chasing, using the fact that H, X, A and E are concyclic (because of the right angles).

$$\angle AEX \equiv \angle AHX$$
$$\equiv \angle AHR$$
$$\equiv \angle ACR$$
$$\equiv \angle PAC$$
$$\equiv \angle PAE$$

from which it follows that $XE \parallel AP$.

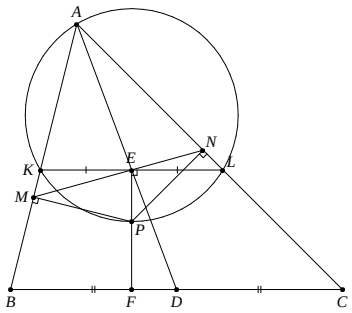
(Proposed for IMO 1996)

5.2 We use Ptolemy's Inequality:

$$AP \cdot BC + BP \cdot CA \geq CP \cdot AB$$
$$\iff \qquad AP + BP \geq CP \qquad \text{(since } AP = BP = CP\text{)}.$$

Equality occurs if and only if $ABPC$ is a cyclic quadrilateral.

5.3 Construct KL through E parallel to BC, with K and L on AB and AC respectively.



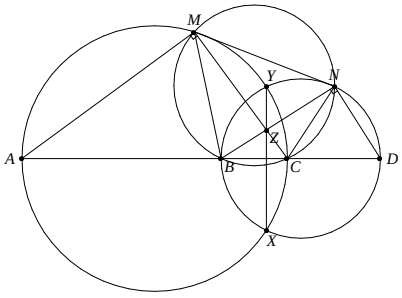
From similar triangles AKE and ABD, we have $KE = BD \cdot \frac{AE}{AD}$. Similarly, $EL = DC \cdot \frac{AE}{AD}$. But $BD = DC$, so $KE = EL$ and hence AE is a median of $\triangle AKL$. Also, $PE \perp KL$ (since $KL \parallel BC$), so M, E and N are the pedal points of P in triangle AKL. The Simpson Line theorem states that M, E and N are collinear if and only if P lies on the circumcircle of $\triangle AKL$. But the perpendicular bisector of KL and the angle bisector of $\angle A$ both meet the circumcircle at the middle of the arc KL, so P lies on the angle bisector of $\angle A$.

(*Crux Mathematicorum*, 1990, 293)

5.4 If P is one of the midpoints, then the lengths of the tangents from P to the two circles are equal. Since these lengths are the square roots of the power of P with respect to these two circles, P must lie on the radical axis. Since this is true for four midpoints, they are collinear because the radical axis is a straight line.

5.5 Call the given circles Γ_1 and Γ_2, and construct a third circle Γ_3 which intersects both Γ_1 and Γ_2. The position of Γ_3 is arbitrary, provided that the centres of the three circles are not collinear. The radical axes of (Γ_1,Γ_2) and (Γ_1,Γ_3) can be found by drawing lines through the intersection points. The intersection of these two lines is the radical centre of the three circles. The desired radical axis now passes through the radical centre and is perpendicular to the line of centres of Γ_1 and Γ_2, which can easily be constructed.

5.6 We use directed angles and line segments, since P may lie either inside or outside of the segment XY. It is also possible (but more tedious) to do the proof with two cases. The diagram below shows the one case.



Label the circle with diameter AC as Γ_1, and the circle with diameter BD as Γ_2. The point Z lies on the radical axis of the two circles, so it has equal power with respect to both. In particular, $ZM \cdot ZC = ZN \cdot ZB$, which prove that M, N, B and C are concyclic. Call this circle Γ_3. Now

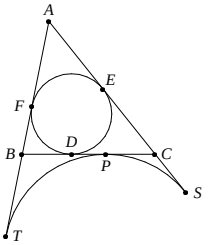
$$\begin{aligned}\angle MND &\equiv \angle MNB + \angle BND \\ &\equiv \angle MCB + 90^\circ \\ &\equiv \angle MCA + \angle AMC \\ &\equiv -\angle CAM \\ &\equiv \angle MAD.\end{aligned}$$

This proves that M, N, A and D are also concyclic; call this circle Γ_4. Finally, we note that AM, DN and XY are the three radical axes formed between the circles Γ_1, Γ_2 and Γ_4. These lines are not all parallel ($AM \parallel XY$ would require that $P = Z$), so they must coincide at the radical centre of the circles.

(IMO 1995, problem 1)

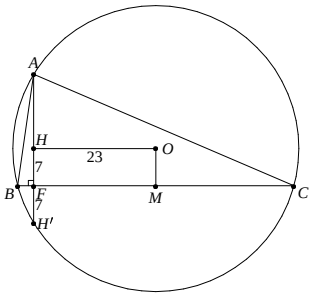
6.1 Let D, E and F be the points of tangency of the incircle with BC, CA, AB and let the excircle be tangent to the same sides at P, S and T respectively. Then from common tangents,

$$\begin{aligned}2ES = 2FT &= ES + FT \\ &= EC + CS + FB + BT \\ &= DC + CP + DB + BP \\ &= 2BC.\end{aligned}$$



Hence $ES = FT = BC = y + z$. Now $BP = BT = FT - BF = (y+z) - y = z$. Similarly, $CP = y$.

6.2 Since the altitude AF passes through H and $BC \perp AF$, it follows that BC and FM coincide. Let H' be the reflection of H in BC. H' is known to lie on the circumcircle of $\triangle ABC$, so $R = H'O = \sqrt{23^2 + 14^2}$. Hence $BM = \sqrt{H'O^2 - 7^2} = 26$ and $BC = 2BM = 52$.



6.3 Clearly, A_0, B_0 and C_0 are in fact I_A, I_B and I_C, and we will refer to them as such.

(i) We will show that $|\triangle II_AC| = 2|\triangle IA_1C|$ (refer to the diagram on page 15, where D is A_1). Results for five other pairs of triangles follow similarly, and adding them all up gives the desired result. Triangles II_AC and IA_1C have a common height, and bases II_A and IA_1. But these bases are the radius and diameter of the circle with diameter II_A, so the result follows.

(ii) It suffices to show that $|AC_1BA_1CB_1|$ is at least twice $|ABC|$, which is equivalent to showing that $|\triangle BCA_1| + |\triangle CAB_1| + |\triangle ABC_1| \geq |\triangle ABC|$. Let A_2, B_2 and C_2 be the reflections of H in BC, CA and AB. These points are known to lie on the circumcircle. When comparing the areas of triangles BCA_1 and BCA_2, we note that they share a common base but the height of $\triangle BCA_1$ is greater than or equal to that of $\triangle BCA_2$. Hence

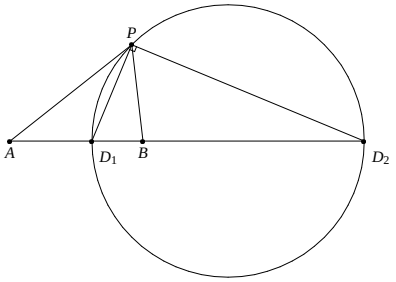
$$\begin{aligned}|\triangle BCA_1| + |\triangle CAB_1| + |\triangle ABC_1| &\geq |\triangle BCA_2| + |\triangle CAB_2| + |\triangle ABC_2| \\ &= |\triangle BCH| + |\triangle CAH| + |\triangle ABH| \\ &= |\triangle ABC|.\end{aligned}$$

(IMO 1989 Question 2)

6.4 Let $BD = m$ and $DC = n$. Then $m+n = a$ and $\frac{n}{m} = \frac{a-m}{m} = \frac{b}{c}$. Hence

$$BD = m = \frac{a}{1+\frac{b}{c}} = \frac{ac}{b+c}.$$

6.5 If $r = 1$, then $AP = BP$ and so the locus is simply the perpendicular bisector. Otherwise suppose $r > 1$ (the situation is symmetric if $r < 1$). Pick an arbitrary P not on AB which satisfies the condition. Let the internal and external angle bisectors of $\angle APB$ meet AB at D_1 and D_2 respectively. Then by the angle bisector theorem, $\frac{AD_1}{BD_1} = \frac{AD_2}{BD_2} = r$. D_1 and D_2 are the only two points on AB that satisfy this, so they are fixed independent of P. Also, $\angle D_1PD_2 = 90°$, so P must lie on the circle with diameter D_1D_2.

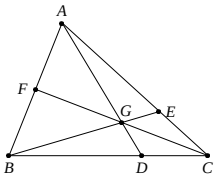


Conversely, suppose P lies on this circle. If P also lies on AB then $P = D_1$ or $P = D_2$, both of which satisfy the conditions. Otherwise let the internal and external bisectors of $\angle PAB$ meet AB at E_1 and E_2 respectively. If $\frac{AP}{BP} = \frac{AE_1}{BE_1} = \frac{AE_2}{BE_2} < r$ then E_1 lies closer to A than D_1 and E_2 lies further from A than D_2. But this means that $\angle E_1PE_2 > 90°$, which is a contradiction. Similarly, if $\frac{AP}{BP} > r$ then $\angle E_1PE_2 < 90°$, again a contradiction. Thus $\frac{AP}{BP} = r$, and this circle is precisely the locus of P.

This circle is known as an *Apollonius* circle.

6.6 Apply Menelaus to $\triangle ACD$ cut by line BGE:

$$\frac{AG}{GD} \cdot \frac{DB}{BC} \cdot \frac{CE}{EA} = -1. \tag{9.1}$$

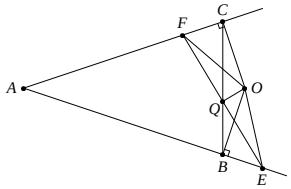


Similarly, one can apply it to $\triangle ABD$ cut by BGE:

$$\frac{AG}{GD} \cdot \frac{DC}{CB} \cdot \frac{BF}{FA} = -1. \tag{9.2}$$

Finally, dividing (9.1) by (9.2) and doing some re-arranging (while being careful with the sign conventions) gives Ceva's Theorem.

6.7 Without loss of generality, let $BQ \leq CQ$, giving the diagram below:



Suppose $OQ \perp EF$. Then $EBQO$ and $FCOQ$ are cyclic quadrilaterals, so $\angle BEO = 180° - \angle BQO = \angle CQO = \angle CFO$. But $BO = CO$, so $\triangle BEO \equiv \triangle CFO$. This gives $EO = FO$, making $\triangle EOF$ isosceles. But $OQ \perp EF$, so $EQ = QF$.

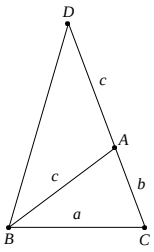
Now suppose that $QE = QF$. Apply Menelaus to triangle AEF, cut by line BQC:

$$1 = \frac{EQ}{QF} \cdot \frac{FC}{CA} \cdot \frac{AB}{BE} = \frac{FC}{BE}.$$

Hence $CF = BE$. Also, $BO = CO$, so $\triangle BEO \equiv \triangle CFO$ and hence $EO = FO$. Then $\triangle EOF$ is isosceles with $EQ = QF$, so $OQ \perp EF$.

(IMO 1994 question 2)

6.8 Construct D on the extension of AC such that $\angle ABD = \angle ABC$. Note that AB is then an angle bisector of $\triangle BDC$. Also, $\angle BDA = 2\angle ABC - \angle ABD = \angle ABD$, so triangle ABD is isosceles and $AD = c$. From the angle bisector theorem (or from $\triangle ABC \,|||\, \triangle BDC$), we find that $AD = \frac{ac}{b}$.



From Stewart's Theorem, we get

$$\begin{aligned} & (b+c)(c^2+bc) = \left(\frac{ac}{b}\right)^2 \cdot b + a^2 c \\ \Longrightarrow \quad & (b+c)^2 bc = a^2 c^2 + a^2 bc \\ \Longrightarrow \quad & b(b+c) = a^2, \end{aligned}$$

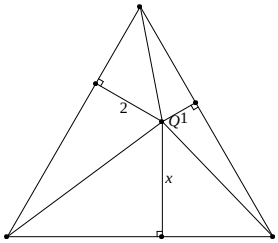
as required.

6.9 Let the orthic triangle be $A'B'C'$. We use Euler's Inequality twice, once on $\triangle ABC$ and once on $\triangle A'B'C'$. The vertices of the orthic triangle lie on the nine-point circle, so the circumradius of $\triangle A'B'C'$ is $R/2$. Thus

$$\begin{aligned} \frac{p}{R} &= \frac{1}{2} \cdot \frac{p}{R/2} \\ &\leq \frac{1}{4} \\ &= 1 - \frac{1}{3} \cdot \frac{3^2}{2} \\ &\leq 1 - \frac{1}{3}\left(1 + \frac{r}{R}\right)^2. \end{aligned}$$

(Proposed at IMO 1993)

6.10 The height of the triangle is 6, so the area is $12\sqrt{3}$. Let the required length be x, and consider the area as the sum of the areas of the triangles formed by Q and the vertices.



The total area is thus $2\sqrt{3}(1+2+x)$. Solving the equation $12\sqrt{3} = 2\sqrt{3}(1+2+x)$ gives $x = 3$.

6.11 We know that $s = x+y+z$. Divide through by K, recalling that $K = rs = r_a x = r_b y = r_c z$.

6.12 We first check that sin is concave on $[0°, 180°]$:

$$\frac{\sin x + \sin y}{2} = \sin\left(\frac{x+y}{2}\right) \cdot \cos\left(\frac{x-y}{2}\right) \leq \sin\left(\frac{x+y}{2}\right).$$

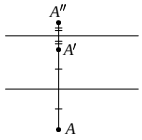
Thus

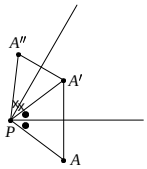
$$\sin\alpha + \sin\beta + \sin\gamma \leq 3\sin\left(\frac{\alpha+\beta+\gamma}{3}\right) = 3\sin 60° = \frac{3\sqrt{3}}{2}.$$

6.13 Suppose for a contradiction that these angles are all strictly greater than 30°. Drop perpendiculars from P onto BC, CA, AB to meet at D, E, F respectively. Then $2PF > PA$, $2PD > PB$ and $2PE > PC$. But then $PA + PB + PC < 2(PD + PE + PF)$, which contradicts the Erdős-Mordell Theorem.

(IMO 1991, question 5)

7.1 (a) When combining two reflections, there are two cases.





In the diagrams above, the first reflection maps A to A', and the second maps A' to A''.

(i) The lines of reflection are parallel, separated by a distance d. As can be seen from the diagram, the combination of the reflections is a translation by $2d$, perpendicular to the lines of reflection (the direction depends on the order in which the reflections are performed. Conversely, any translation can be expressed as the combination of two parallel reflections, suitably oriented, and with separation equal to half the distance of the translation.

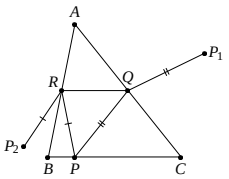
(ii) The lines of reflection are not parallel, and intersect at some point P with an angle of θ. From the diagram, it is now clear that any other point is rotated by an angle of 2θ around P, with the direction depending on the order of the rotations. Conversely, any rotation can be expressed as the combination of two reflections which pass through the centre of the rotation, and with an angle between them of half the rotation angle.

(b) Two translations trivially produce another translation, whose displacement is the vector sum of the original displacements. When one or both of the transformations is a rotation, express the transformations as pairs of reflections. We showed in part (a) that there is some freedom in the choice of reflections. We will have four reflections which are applied in order, say $b_2b_1a_2a_1$.[6] We can always choose the reflections such that a_2 and b_1 are the same. Identical reflections cancel out, so we are left with a_1b_2 which from (a) is equivalent to a rotation or translation.

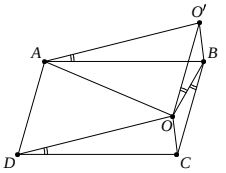
(c) We can transform all the rotations and translations into pairs of reflections, using part (a). We can then pair off these reflections and convert them back into translations and rotations, possibly leaving one reflection at the end. Now part (b) shows that we can reduce the sequence of translations and rotations to just one, which may be followed by a reflection. It remains to show that a rotation followed by a reflection is equivalent to a translation followed by a reflection. We do this by appending two identical (and hence cancelling) reflections to the sequence, at an angle we will choose in a moment. The sequence will now appear as $ccb(r_2r_1)$ where r_2r_1 is the rotation, and c is the newly added reflection. We choose c so that cb forms a rotation with angle exactly opposite to the angle of r_2r_1. Now $(cb)(r_2r_1)$ is the combination of two rotations that forms some translation, say T (it is a translation, not a rotation, because of the choice of angle). Thus the entire sequence is equivalent to cT i.e. a translation followed by a reflection.

[6] We write sequence of transformations from right to left. This is because they are functions, so applying ab to a point P actually means $a(b(P))$, with b being applied first.

7.2 Reflect P in CA to obtain P_1 and reflect P in AB to obtain P_2. Now $PQ+QR+RP = P_1Q+QR+RP_2$. This sum will clearly be smallest when P_1, Q, R and P_2 lie in a straight line. So choose Q and R to be the intersections of P_1P_2 with CA and AB.



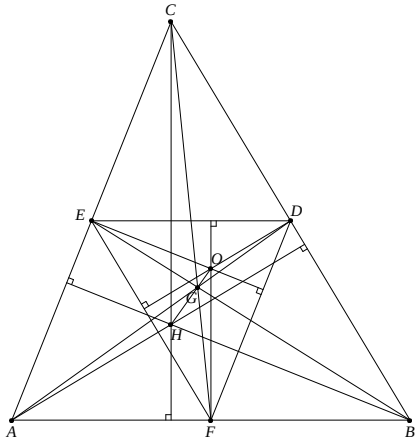
7.3 Having two supplementary angles vertically opposite each other is not very helpful. It would be more useful if we could get the angles to be either adjacent (to create a straight line) or opposite angles of a quadrilateral (to make it cyclic). One way to do this is to "pick up" triangle DOC and place DC on top of AB.



More formally, construct O' outside $ABCD$ such that $\triangle AO'B \equiv \triangle DOC$. Then $\angle AO'B + \angle AOB = 180°$, so $AO'BO$ is cyclic. Also, $OO'BC$ is a parallelogram because $O'B$ and OC are equal and parallel. Thus $\angle OBC = \angle BOO' = \angle BAO' = \angle ODC$.

(Canadian Mathematical Olympiad 1997)

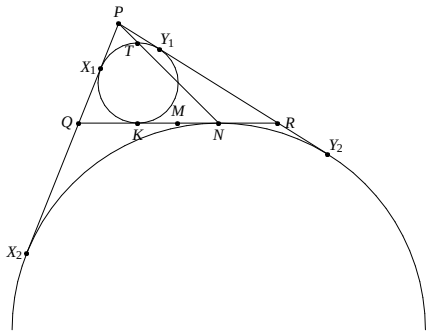
7.4 (a) Let D, E and F be the midpoints of BC, CA and AB respectively. From the Midpoint Theorem, $\triangle DEF \,|||\, \triangle ABC$ and is half the size. It is also oriented $180°$ relative to $\triangle ABC$. Thus there is a homothetism that maps $\triangle ABC$ to $\triangle DEF$, with scale factor $-\frac{1}{2}$. The centre of similitude must lie on AD, BE and CF, and hence these lines are concurrent.



(b) The homothetism maps AG to DG with scale factor $-\frac{1}{2}$, so $AG : GD = 2 : 1$. The result follows similarly for the other two medians.

(c) The line DO is perpendicular to BC, and hence also to EF. Similarly $EO \perp FD$ and $FO \perp DE$, so O is the orthocentre of $\triangle DEF$. Since the homothetism maps $\triangle ABC$ to $\triangle DEF$, it will also map H to O. This proves the collinearity, and the scale follows as in the previous section.

7.5 Start with an arbitrary pair (Q, R) for which P exists, and construct the excircle C_2 of $\triangle PQR$ opposite P (see diagram).



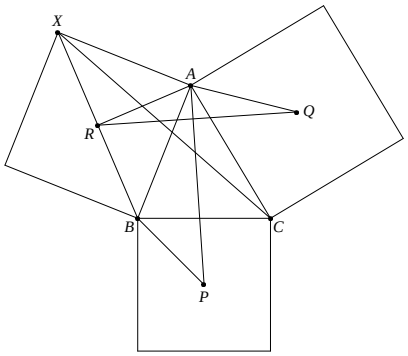
The incircle and excircle of $\triangle PQR$ must be homothetic, and P is the centre of the homothetism. Now let K be the point of tangency of C with L, and let T be the point diametrically opposite K. The corresponding point to T on C_2 must also be vertically above the centre in the diagram, i.e. it is N. But the line through corresponding points must pass through the centre of the homothetism, so P lies on NT.

From the solution to problem 6.1 (page 37), we have $QK = RN$, from which it follows that N and K are symmetrically placed about M. But K and M are fixed, so N must be fixed too.

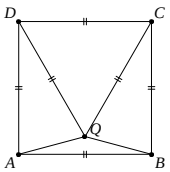
We have now established that any solution P must lie on NT. It is also clear that P must lie strictly beyond T. Conversely, suppose P' is some point on NT beyond T. Let L' be a line through P' and parallel to L, and consider moving a point P along L', finding Q and R on L such that C is the incircle of $\triangle PQR$. When P moves far to the left, the midpoint of QR will be far to the right, and vice versa. Since the midpoint shifts continuously, there is at least one point where it is M. We have shown above that this P must be the intersection of NT with L', namely P', and hence P' satisfies the desired properties. Therefore the locus is the portion of NT that lies strictly beyond T.

(IMO 1992, question 4)

7.6 Consider the spiral similarity with centre A, rotating clockwise (in the diagram) by 45° and scaling by $\sqrt{2}$. It will map Q to C and R to X. Now consider the spiral similarity with centre B that rotates anti-clockwise by 45° and scales by $\sqrt{2}$. It will map A to X and P to C. These two similarities thus map AP and QR to the same line. They both scale by the same amount ($\sqrt{2}$) and the difference of their angles is 90°, so AP and QR must be equal and perpendicular.



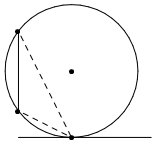
8.1 Construct Q inside the square with $\triangle CDQ$ equilateral. We aim to show that $P = Q$.



Now $\angle QDC = 60°$, so $\angle QDA = 30°$. But $QD = AD$, so $\triangle AQD$ is isosceles and thus $\angle DAQ = 75°$. This makes $\angle BAQ = 15°$, and similarly $\angle ABQ = 15°$. But then triangles ABP and ABQ have two common angles and a common side, so they are congruent. Both P and Q lie on the same side of AB (the inside

of the square), so P and Q must be the same. Triangle CDQ is equilateral by construction, so $\triangle CDP$ is equilateral.

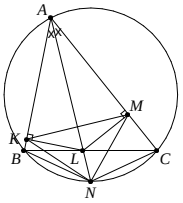
8.2 Construct a circle of radius 5m, with centre 5m above your head height and 4m from the statue. This circle will pass through the head and foot of the statue.



If your head lies on the circle you will have some constant viewing angle θ; with your head inside the circle the angle is larger, and with your head outside the circle it is smaller. But the circle is tangent to the line representing head-height, so the best angle is when your head is at this point of tangency. So you should stand 4m from the statue.

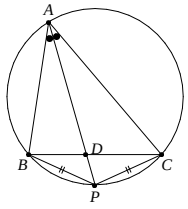
8.3 Firstly note that $\triangle ALK \equiv \triangle ALM$. Hence $AKLM$ is a kite and so $KM \perp AL$; thus $|AKNM| = \frac{1}{2} KM \cdot AN$. Since $ABNC$ is cyclic, $\triangle ABL \,|||\, \triangle ANC$ and hence $AN \cdot AL = AB \cdot AC$. Also, AL is the diameter of the circumcircle of $\triangle AKM$, so $\frac{KM}{AL} = \sin\alpha$. Substituting these into the above gives

$$\begin{aligned} |AKNM| &= \frac{1}{2} \cdot \frac{KM \cdot AB \cdot AC}{AL} \\ &= \frac{1}{2} \cdot AB \cdot AC \cdot \sin\alpha \\ &= |\triangle ABC| \end{aligned}$$



(IMO 1987 Question 2)

8.4 Let D be the point where the angle bisector from A cuts BC.

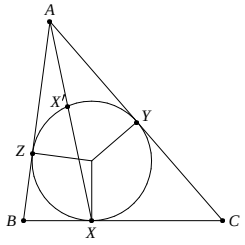


Since $\angle BAD = \angle PAC$ and $\angle DBA = \angle CPA$ we have $\triangle BAD ||| \triangle PAC$. Thus $\frac{c}{BD} = \frac{AP}{PC}$. From exercise 6.4 we have $BD = \frac{ac}{b+c}$. It follows that $AP = \frac{b+c}{a} \cdot PC$. But $PB = PC$ and so from the triangle inequality, $2PC > BC \iff PC > \frac{a}{2}$. Therefore $AP > \frac{b+c}{2}$.

Similarly $BQ > \frac{c+a}{2}$ and $CR > \frac{a+b}{2}$. Adding these inequalities gives the desired result.

(Australian Mathematics Olympiad 1985)

8.5 Firstly note that $AX \cdot AX'$ is the power of A with respect to the incircle, so it is equal to $AZ^2 = x^2$. Thus $a \cdot AX \cdot XX' = a \cdot AX^2 - ax^2$.



We can calculate $a \cdot AX^2$ using Stewart's Theorem:

$$BC(AX^2 + BX \cdot XC) = AC^2 \cdot BX + AB^2 \cdot CX$$

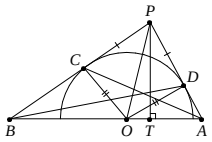
$$\begin{aligned} a(AX^2 + yz) &= b^2y + c^2z \\ a \cdot AX^2 &= (x+z)^2y + (x+y)^2z - (y+z)yz \\ &= x^2y + 2xyz + z^2y + x^2z + 2xyz + y^2z - y^2z - z^2y \\ &= x^2(y+z) + 4xyz \\ &= ax^2 + 4xyz. \end{aligned}$$

Now we can calculate $a \cdot AX^2 - ax^2$

$$\begin{aligned} a \cdot AX^2 - ax^2 &= 4xyz \\ &= \frac{4}{s} \cdot sxyz \\ &= \frac{4}{s} \cdot K^2 \\ &= \frac{4}{s} \cdot rsK \\ &= 4rK \text{ as desired.} \end{aligned}$$

(Arbelos May 1987)

8.6 This is a good example of a problem that becomes much easier with a good diagram (the diagram below is intentionally skewed). If AD and BC are extended to meet at P, then it appears that P, E and F are collinear. This would be a useful thing to know, so we attempt to prove it.



Let T be the foot of the perpendicular from P to AB and let O be the centre of the semicircle. $\triangle OCB ||| \triangle PTB$, so $\frac{CB}{TB} = \frac{BO}{BP}$. Similarly $\frac{DA}{TA} = \frac{AO}{AP}$. We want to prove that PT, AC and BD are concurrent, which by the converse of Ceva's Theorem would be true if

$$\frac{PC}{CB} \cdot \frac{BT}{TA} \cdot \frac{AD}{DP} = 1$$

Firstly, $PC = PD$ (equal tangents to the semicircle), and we can substitute the ratios found above to change this to $\frac{BP}{BO} \cdot \frac{AO}{AP} = 1$. However, this is true by the angle bisector theorem (PD is an angle bisector because $\triangle PCO \equiv \triangle PDO$). It follows that E lies on the altitude from A, and $F = T$.

Now notice that PO subtends right angles at C, D and F, so $PCFD$ is a cyclic quad. Thus $\angle DFP = \angle DCP$ and $\angle CFP = \angle CDP$, and since $PC = PD$ it follows that $\angle DFP = \angle CFP$. Therefore EF bisects $\angle CFD$.

(Proposed at IMO 1994)

8.7 The key to this problem is noticing that you can treat triangles ABD and ACE as completely separate, and ignore $\triangle ABC$. The only things these two triangles have in common is the angle at A and the height from A. Let these quantities be θ and h respectively. If we can express $\frac{1}{MB} + \frac{1}{MD}$ in terms of θ and h then we are done.

Let us rename D to C so that we are working with $\triangle ABC$ and can use the usual notation.

$$\begin{aligned}
\frac{1}{MB} + \frac{1}{MC} &= \frac{1}{y} + \frac{1}{z} \\
&= \frac{y+z}{yz} \\
&= \frac{a}{yz} \\
&= \frac{ahrsx}{hrsxyz} \\
&= \frac{ahrsx}{hrK^2} \quad \text{(Heron's Formula)} \\
&= \frac{2K^2x}{hrK^2} \\
&= \frac{x}{r} \cdot \frac{2}{h} \\
&= \frac{2}{h} \cot \frac{\theta}{2}.
\end{aligned}$$

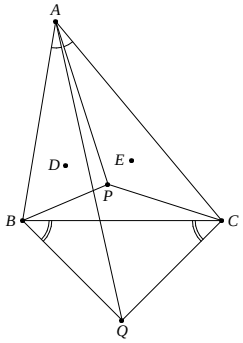
(Proposed at IMO 1993)

8.8 Construct Q so that $\angle BAQ = \angle PAC$ and $\angle ABQ = \angle APC$. Then by construction, $\triangle ABQ \,|||\, \triangle APC$. Now in $\triangle APB$ and $\triangle ACQ$:

- $\angle BAP = \angle BAC - \angle PAC = \angle QAC$
- $\frac{AC}{AQ} = \frac{AC}{AC \cdot AB/AP} = \frac{AP}{AB}$.

Hence $\triangle APB \,|||\, \triangle ACQ$. Now $\angle CBQ = \angle APC - \angle ABC = \angle APB - \angle ACB = \angle BCQ$, so $\triangle BCQ$ is isosceles. It follows that

$$\frac{AC}{PC} = \frac{AQ}{BQ} = \frac{AQ}{CQ} = \frac{AB}{BP}.$$



Now from the angle bisector theorem, BD will cut AP in the ratio $AB : BP$, and CE will cut AP in the ratio $AC : CP$. Since these ratios are the same, the three lines will be concurrent.

(IMO 1996 Question 2)

10 Recommended further reading

Geometric inequalities often require techniques from the world of standard inequalities. *Inequalities for the Olympiad Enthusiast*, by Graeme West (part of the same series as this booklet) provides some good material in this field.

This booklet is well under 100 pages, and as such cannot do proper justice to the rich field of classical geometry. A highly regarded and very readable reference is *Geometry Revisited*, by Coxeter and Greitzer.

A good source of problems are the yearbooks of the South African training program for the IMO (*South Africa and the* n*th IMO*, for $n \geq 35$). These contain problems

and solutions for all the problems used in the training problem, including many good geometry problems.

THE SOUTH AFRICAN COMMITTEE FOR THE PAN AFRICAN AND INTERNATIONAL MATHEMATICAL OLYMPIADS

Dr P Dankelmann (University of KwaZulu-Natal)
Dr S Hansraj (University of KwaZulu-Natal)
Mr D Hatton (University of Cape Town)
Professor N J H Heideman (University of Cape Town)
Professor D P Laurie (University of Stellenbosch)
Dr L R le Riche (University of Stellenbosch)
Professor P Maritz (University of Stellenbosch
Professor S Mabizela (Rhodes University)
Professor J Persens (University of the Western Cape)
Professor P Pillay (University of KwaZulu-Natal)
Professor J H Webb (University of Cape Town) - Convener

terug naar echt bestand

Driehoeken

Enkele speciale topics

Arne Smeets

Trainingsweekend
Februari 2008

Trilineaire en barycentrische coördinaten

Definitie van trilineaire coördinaten

Beschouw (in het vlak) een driehoek ΔABC en een punt P. Zijn $\tilde{x}$, $\tilde{y}$ en $\tilde{z}$ de afstanden van P tot BC, CA en AB, waar we $\tilde{x} > 0$ ($\tilde{x} < 0$) nemen indien P en A aan dezelfde kant (verschillende kanten) van BC liggen. Dan geldt duidelijk $a\tilde{x} + b\tilde{y} + c\tilde{z} = 2\,[\Delta ABC]$, waarbij $a = BC$, $b = CA$ en $c = AB$, en waarbij we oppervlaktes noteren met vierkante haakjes. Omwille van deze (ongewenste) "afhankelijkheid" tussen $\tilde{x}$, $\tilde{y}$ en $\tilde{z}$ is het opportuun om een *homogene versie* van $\tilde{x}$, $\tilde{y}$ en $\tilde{z}$ in te voeren. Dat doen we als volgt: we noemen (x, y, z) de *homogene driehoekscoördinaten* van P ten opzichte van $\triangle ABC$ indien

$$x : y : z = \tilde{x} : \tilde{y} : \tilde{z}.$$

Om met homogene driehoekscoördinaten te werken moeten we het vlak wel uitbreiden met een *rechte op oneindig*, de rechte met vergelijking $ax + by + cz = 0$. Het is niet moeilijk om na te gaan dat de homogene driehoekscoördinaten een punt dan op een eenduidige wijze bepalen.

Collineariteit met trilineaire coördinaten

Neem nu twee willekeurige punten $P_1 = (x_1, y_1, z_1)$ en $P_2 = (x_2, y_2, z_2)$ (ten opzichte van ΔABC). Met eenvoudig rekenwerk kunnen we aantonen dat $P = (x, y, z)$ collineair is met P_1 en P_2 als en slechts als

$$\begin{vmatrix} x & y & z \\ x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \end{vmatrix} = 0.$$

Dit geeft ons direct de vergelijking van de rechte door P_1 en P_2.
In de duale situatie kunnen we rechten ℓ_1 en ℓ_2 met vergelijkingen $\alpha_1 x + \beta_1 y + \gamma_1 z = 0$ en $\alpha_2 x + \beta_2 y + \gamma_2 z$ bekijken. De rechte ℓ met vergelijking $\alpha x + \beta y + \gamma z$ gaat door het snijpunt van ℓ_1 en ℓ_2 als en slechts als

$$\begin{vmatrix} \alpha & \beta & \gamma \\ \alpha_1 & \beta_1 & \gamma_1 \\ \alpha_2 & \beta_2 & \gamma_2 \end{vmatrix} = 0.$$

Op deze manier kunnen we dus concurrentie van drie rechten uitdrukken.

Definitie van barycentrische coördinaten

Gegeven P en ΔABC zoals in de definitie van trilineaire coördinaten, dan noemen we (X, Y, Z) de homogene barycentrische coördinaten van P als en slechts als voldaan is aan de voorwaarde

$$X : Y : Z = ax : by : cz.$$

Interpreteer barycentrische coördinaten in termen van oppervlakten. In sommige meetkundige situaties zijn trilineaire coördinaten erg handig, in andere situaties rekenen barycentrische coördinaten vlotter. Concurrentie en collineariteit worden hier op dezelfde manier uitgedrukt als in trilineaire coördinaten.

Trilineaire en barycentrische coördinaten van enkele speciale punten in driehoeken

Bekijk het volgende overzicht:

notatie	punt	trilineair	barycentrisch
I	middelpunt ingeschreven cirkel	$(1, 1, 1)$	(a, b, c)
G	zwaartepunt (barycentrum)	(bc, ca, ab)	$(1, 1, 1)$
O	middelpunt omgeschreven cirkel	$(\cos\alpha, \cos\beta, \cos\gamma)$	$(\sin 2\alpha, \sin 2\beta, \sin 2\gamma)$
H	hoogtepunt	$(\sec\alpha, \sec\beta, \sec\gamma)$	$(\tan\alpha, \tan\beta, \tan\gamma)$
K	punt van Lemoine	(a, b, c)	$\left(a^2, b^2, c^2\right)$

Over het punt van Lemoine volgt later meer.

Opgaven

(1) Leid de resultaten uit de bovenstaande tabel af.

(2) Het is bekend dat het hoogtepunt, het middelpunt van de omgeschreven cirkel en het zwaartepunt van een driehoek op een rechte liggen - de rechte van Euler. Verifieer dit met een berekening.

(3) Bepaal driehoeks- en barycentrische coördinaten voor enkele andere speciale punten in een driehoek die je nauw aan het hart liggen. Suggesties: de punten van Nagel, Gergonne, Feuerbach, ...

(4) In sommige formules in de tabel staan goniometrische uitdrukkingen. Hoe herschrijf je die zodat je uitdrukkingen overhoudt waarin enkel de lengtes van de zijden van de driehoek voorkomen?

(5) Geef een computationeel bewijs voor de stellingen van Ceva en Menelaos, met de juiste coördinaten.

(6) Beschouw in een driehoek de drie rechten die het midden van een zijde verbinden met het midden van de hoogtelijn op die zijde. Bewijs dat deze drie rechten concurrent zijn. Welk (bekend) punt hebben de drie rechten gemeenschappelijk? (Denk na: trilineaire of barycentrische coördinaten?)

(7) Zij ΔABC een driehoek, zij r de straal van de omgeschreven cirkel, zij O het middelpunt van de omgeschreven cirkel en zij H het hoogtepunt. Zijn D, E en F de spiegelbeelden van A, B en C in BC, CA en AB respectievelijk. Bewijs dat D, E en F collineair zijn als en slechts als $OH = 2r$.

(IMO Shortlist, 1998)

De synthetische oplossing voor deze opgave is bijzonder moeilijk. Voor iemand met ervaring met driehoeks- en barycentrische coördinaten en enkele leuke eigenschappen van driehoeken is deze opgave echter een peuleschil. Je zal hier de coördinaten van D, E en F moeten bepalen en collineariteit van D, E en F moeten vertalen in een determinant die gelijk is aan 0.

Daaraan heb je nog niet genoeg, je hebt twee identiteiten nodig die absoluut de moeite zijn om te onthouden:

- $OH^2 = 9r^2 - \left(a^2 + b^2 + c^2\right)$
- $\cos^2\alpha + \cos^2\beta + \cos^2\gamma + 2\cos\alpha\cos\beta\cos\gamma = 1$

Ik verwacht van jullie dat jullie beide identiteiten kunnen bewijzen (en ook onthouden). Voor de eerste identiteit kan dat heel vlot met vectorrekening: kies de oorsprong in het middelpunt van de omgeschreven cirkel. Kan je dan **OH** *schrijven in functie van* **OA**, **OB** *en* **OC** *door de ligging van H op de rechte van Euler te bekijken? Ik gebruik hier boldface om vectoren te noteren. Opfrissing van de rechte van Euler is te verkrijgen op aanvraag, en ik wil hier gerust over uitweiden als dit een gat in jullie meetkundige cultuur blijkt te zijn. De tweede identiteit is eenvoudiger te bewijzen.*

Ik besef dat dit een heel andere manier is om aan meetkunde te doen dan wat jullie normaal doen, maar ik vind het juist heel belangrijk dat jullie ook eens met deze technieken in aanraking komen.

(8) Zij ΔABC een scherphoekige driehoek. Zijn D en E de punten waar de hoogtelijnen door A en B de overstaande zijden snijden. Zijn P en Q de punten waar de bissectrices van $\angle A$ en $\angle B$ de overstaande zijden snijden. Zijn O en I de middelpunten van de om- en ingeschreven cirkels van de driehoek. Toon aan dat D, E en I collineair zijn als en slechts als P, Q en O collineair zijn.

(IMO Shortlist, 1997)

Isogonale conjugatie

Definitie

Zij ΔABC een driehoek en zij P een willekeurig punt. Door de rechten AP, BP en CP te spiegelen om de bissectrices van de hoeken $\angle A$, $\angle B$ en $\angle C$ krijgen we drie (meestal nieuwe) rechten. We kunnen op twee manieren inzien dat deze drie rechten concurrent zijn: met de goniometrische variant van de stelling van Ceva, maar ook met driehoekscoördinaten. *We werken de tweede methode uit als kleine oefening.* Het punt van concurrentie noemen we Q, en we noemen Q het *isogonaal geconjugeerde punt* van P. Als P op een van de zijden van de driehoek ligt krijgen we een "ontaarde" situatie: wat is dan het isogonaal geconjugeerde punt van P? Cruciale betrekkingen in alles wat volgt zijn de volgende:

- als $P = (x, y, z)$, dan is $Q = \left(x^{-1}, y^{-1}, z^{-1}\right) = (yz, zx, xy)$,
- als $P = (X, Y, Z)$, dan is $Q = \left(a^2X^{-1}, b^2Y^{-1}, c^2Z^{-1}\right) = (a^2YZ, b^2ZX, c^2XY)$.

We gebruiken (as usual) "lower case" voor trilineaire coördinaten, "upper case" voor barycentrische.

Algemene eigenschappen van isogonaal geconjugeerde punten

We formuleren eerst de *stelling van Steiner*. Zij ΔABC een driehoek. Zijn D en E punten op BC zodanig dat AD en AE elkaars spiegelbeeld zijn ten opzichte van de bissectrice van $\angle A$. Dan geldt dat

$$\frac{BD}{DC} \cdot \frac{BE}{EC} = \frac{AB^2}{CA^2}.$$

Deze gelijkheid kan bewezen worden door de sinusregel een honderdtal keer toe te passen. De link met isogonale conjugatie is, vermoed ik, duidelijk genoeg.

Interessant om even over na te denken is de volgende vraag. Zij P een punt en zijn U, V en W de snijpunten van AP, BP en CP met BC, CA en AB. Hoe krijg je de trilineaire (of barycentrische) coördinaten van U, V en W vanuit de coördinaten van P? Omgekeerd, hoe verkrijg je de coördinaten van P uit die van U, V en W? Het resultaat is krachtig in combinatie met de stelling van Steiner.
De constructie van de *voetpuntsdriehoek* van een punt ten opzichte van een gegeven driehoek is - hoop ik - bekend: indien nodig weid ik hier even over uit. Beschouw ΔABC en isogonaal geconjugeerde punten P en P' ten opzichte van ΔABC. Zijn $\triangle DEF$ en $\triangle D'E'F'$ de voetpuntsdriehoeken van P en P' ten opzichte van ΔABC, waarbij $D \in BC$, $E \in CA$ en $F \in AB$, en analoog voor $\triangle D'E'F'$. Dan geldt dat ΔDEF en $\Delta D'E'F'$ *dezelfde omgeschreven cirkel* hebben: gelijkvormige driehoeken geven

$$AE \cdot AE' = AF \cdot AF',\ BF \cdot BF' = BD \cdot BD',\ CE \cdot CE' = CD \cdot CD'.$$

Bijgevolg zijn $EE'FF'$, $FF'DD'$ en $EE'DD'$ koordenvierhoeken. Zijn γ_A, γ_B en γ_C de omgeschreven cirkels van deze vierhoeken. Bekijk bijvoorbeeld γ_A en γ_B. Als deze cirkels verschillend zouden zijn, dan zou de verzameling van punten die gelijke machten hebben ten opzichte van γ_A en γ_B een rechte zijn. We hebben echter drie niet-collineaire punten met gelijke machten ten opzichte van γ_A en γ_B, namelijk A, B en C. Bijgevolg moeten γ_A en γ_B wel samenvallen, en daarmee is ons resultaat bewezen.

Het punt van Lemoine

Het isogonaal geconjugeerde punt van het zwaartepunt van een driehoek noemen we het *Lemoine-punt* van de driehoek. De coördinaten van dit punt werden eerder gegeven: verifieer! Het punt van Lemoine heeft prachtige eigenschappen. De rechten die de hoekpunten van een driehoek verbinden met het punt van Lemoine noemen we de *symmedianen* van de driehoek. Verklaar de naamgeving! Symmedianen en het punt van Lemoine zijn erg nuttig: we bespreken hier enkel hun belangrijkste eigenschappen.

- In welke verhouding verdelen de symmedianen de zijden van een driehoek?
- Het punt van Lemoine is het zwaartepunt van zijn eigen voetpuntsdriehoek.
- Bekijk ΔABC en zij S het snijpunt van de raaklijnen in B en C aan de omgeschreven cirkel. Dan is AS een symmediaan in ΔABC. Dit geeft een eenvoudige constructie voor het punt van Lemoine.

  Inderdaad, zijn C_A en C_B de projecties van S op AC en BC. Dan hebben we dat $\angle SAC_A = \angle B$ en $\angle SBC_B = \angle A$. Omdat $AS = BC$ is $SC_A : SC_B = \sin\angle A : \sin\angle B = a : b$. Waarom volgt onze bewering over de rechte AS dan eigenlijk direct uit deze gelijkheid?
- Het Lemoine-punt van een gegeven driehoek is het punt waarvoor de som van de kwadraten van de afstanden tot de zijden van de driehoek een minimale waarde aanneemt. Verrassend!

  Noteer de afstanden van P tot de zijden van ΔABC met $\tilde{x}$, $\tilde{y}$ en $\tilde{z}$ zoals op de eerste pagina. Dan hebben we dat $\left(a^2+b^2+c^2\right)\left(\tilde{x}^2+\tilde{y}^2+\tilde{z}^2\right) \geq 4\left[\Delta ABC\right]^2$. We vinden zo een ondergrens voor $\tilde{x}^2+\tilde{y}^2+\tilde{z}^2$, en deze wordt enkel bereikt als $\tilde{x} : \tilde{y} : \tilde{z} = a : b : c$. Dit geeft het punt van Lemoine!

Opgaven

(1) Het middelpunt van de omgeschreven cirkel en het hoogtepunt van een driehoek zijn isogonaal geconjugeerde punten. Dit kan je heel elementair bewijzen, maar kan je ook afleiden uit de tabel (cfr. pagina 2).

(2) Zij ΔABC een scherphoekige driehoek en zijn D, E en F de voetpunten van de hoogtelijnen op BC, CA en AB. Zij t_A de rechte door A, loodrecht op EF, en definieer t_B en t_C op analoge wijze. Bewijs dat t_A, t_B en t_C concurrent zijn: wat is het punt van concurrentie?

(3) Zij ΔABC een driehoek met $AC = 2 \cdot AB$. Zij γ de omgeschreven cirkel en zij P het snijpunt van de raaklijnen aan γ in A en C. Bewijs: BP snijdt de middelloodlijn van BC op γ.

(4) Zijn A, B en C vaste punten op een rechte (in die volgorde). Zij γ een cirkel door A en C zodat AC geen diameter is van γ. Zij P het snijpunt van de raaklijnen aan γ in A en C. Zij Q het snijpunt van γ met het lijnstuk PB. Bewijs dat de bissectrice van $\angle AQC$ door een vast punt gaat, dat onafhankelijk is van de precieze ligging van de cirkel γ.

(IMO Shortlist, 2003)

(5) Zij ΔABC een scherphoekige driehoek en zijn P en Q isogonaal geconjugeerde punten binnen ΔABC. Zijn D, E en F zoals gewoonlijk de orthogonale projecties van P op BC, CA en AB. Bewijs dat $\angle DEF = 90^\circ$ als en slechts als Q het hoogtepunt is van $\triangle BDF$.

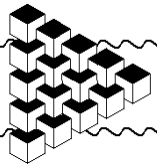
Als er nog tijd over is...

... dan zal ik met plezier wat dieper ingaan op de rechte van Euler en de negenpuntscirkel.

Aan de hand van de resultaten kunnen we dan de volgende uitspraak aantonen:

> Zij $\Delta A_1A_2A_3$ een driehoek die niet gelijkzijdig is. Zijn O en I de middelpunten van de om- en ingeschreven cirkels. Zijn T_1, T_2 en T_3 de punten waar de ingeschreven cirkel van $\Delta A_1A_2A_3$ raakt aan A_2A_3, A_3A_1 en A_1A_2. Dan ligt het hoogtepunt van $\Delta T_1T_2T_3$ op de rechte OI.

Dit zou een mooi besluit zijn van de sessie.



©2007 The Author(s) and The IMO Compendium Group

Complex Numbers in Geometry

Marko Radovanović
radmarko@yahoo.com

Contents

1 Introduction

When we are unable to solve some problem in plane geometry, it is recommended to try to do calculus. There are several techniques for doing calculations instead of geometry. The next text is devoted to one of them – the application of complex numbers.

The plane will be the complex plane and each point has its corresponding complex number. Because of that points will be often denoted by lowercase letters a, b, c, d, ..., as complex numbers.

The following formulas can be derived easily.

2 Formulas and Theorems

Theorem 1.

- *$ab \parallel cd$ if and only if* $\dfrac{a-b}{\overline{a}-\overline{b}} = \dfrac{c-d}{\overline{c}-\overline{d}}$.
- *a,b,c are colinear if and only if* $\dfrac{a-b}{\overline{a}-\overline{b}} = \dfrac{a-c}{\overline{a}-\overline{c}}$.
- *$ab \perp cd$ if and only if* $\dfrac{a-b}{\overline{a}-\overline{b}} = -\dfrac{c-d}{\overline{c}-\overline{d}}$.
- *$\varphi = \angle acb$ (from a to b in positive direction) if and only if* $\dfrac{c-b}{|c-b|} = e^{i\varphi}\dfrac{c-a}{|c-a|}$.

Theorem 2. *Properties of the unit circle:*

- *For a chord ab we have* $\dfrac{a-b}{\overline{a}-\overline{b}} = -ab$.
- *If c belongs to the chord ab then* $\overline{c} = \dfrac{a+b-c}{ab}$.
- *The intersection of the tangents from a and b is the point* $\dfrac{2ab}{a+b}$.
- *The foot of perpendicular from an arbitrary point c to the chord ab is the point* $p = \dfrac{1}{2}\left(a+b+c-ab\overline{c}\right)$.
- *The intersection of chords ab and cd is the point* $\dfrac{ab(c+d)-cd(a+b)}{ab-cd}$.

Theorem 3. *The points* a,b,c,d *belong to a circle if and only if*

$$\frac{a-c}{b-c} : \frac{a-d}{b-d} \in \boldsymbol{R}.$$

Theorem 4. *The triangles abc and pqr are similar and equally oriented if and only if*

$$\frac{a-c}{b-c} = \frac{p-r}{q-r}.$$

Theorem 5. *The area of the triangle abc is*

$$p = \frac{i}{4}\begin{vmatrix} a & \overline{a} & 1 \\ b & \overline{b} & 1 \\ c & \overline{c} & 1 \end{vmatrix} = \frac{i}{4}\left(a\overline{b}+b\overline{c}+c\overline{a}-\overline{a}b-\overline{b}c-\overline{c}a.\right)$$

Theorem 6.
- *The point c divides the segment ab in the ratio* $\lambda \neq -1$ *if and only if* $c = \dfrac{a+\lambda b}{1+\lambda}$.
- *The point t is the centroid of the triangle abc if and only if* $t = \dfrac{a+b+c}{3}$.
- *For the orthocenter h and the circumcenter o of the triangle abc we have* $h+2o = a+b+c$.

Theorem 7. *Suppose that the unit circle is inscribed in a triangle abc and that it touches the sides* bc, ca, ab, *respectively at* p,q,r.

- *It holds* $a = \dfrac{2qr}{q+r}, b = \dfrac{2rp}{r+p}$ *and* $c = \dfrac{2pq}{p+q}$;
- *For the orthocenter h of the triangle abc it holds*

$$h = \frac{2(p^2q^2+q^2r^2+r^2p^2+pqr(p+q+r))}{(p+q)(q+r)(r+p)}.$$

- *For the excenter o of the triangle abc it holds* $o = \dfrac{2pqr(p+q+r)}{(p+q)(q+r)(r+p)}$.

Theorem 8.
- *For each triangle abc inscribed in a unit circle there are numbers* u,v,w *such that* $a = u^2, b = v^2, c = w^2$, *and* $-uv, -vw, -wu$ *are the midpoints of the arcs* ab, bc, ca *(respectively) that don't contain* c,a,b.
- *For the above mentioned triangle and its incenter i we have* $i = -(uv+vw+wu)$.

Theorem 9. *Consider the triangle $\triangle$ whose one vertex is 0, and the remaining two are x and y.*

- *If h is the orthocenter of $\triangle$ then* $h = \dfrac{(\bar{x}y + x\bar{y})(x-y)}{x\bar{y} - \bar{x}y}$.
- *If o is the circumcenter of $\triangle$, then* $o = \dfrac{xy(\bar{x} - \bar{y})}{\bar{x}y - x\bar{y}}$.

3 Complex Numbers and Vectors. Rotation

This section contains the problems that use the main properties of the interpretation of complex numbers as vectors (Theorem 6) and consequences of the last part of theorem 1. Namely, if the point b is obtained by rotation of the point a around c for the angle φ (in the positive direction), then $b - c = e^{i\varphi}(a - c)$.

1. (Yug MO 1990, 3-4 grade) Let S be the circumcenter and H the orthocenter of $\triangle ABC$. Let Q be the point such that S bisects HQ and denote by T_1, T_2, and T_3, respectively, the centroids of $\triangle BCQ$, $\triangle CAQ$ and $\triangle ABQ$. Prove that

$$AT_1 = BT_2 = CT_3 = \frac{4}{3}R,$$

where R denotes the circumradius of $\triangle ABC$.

2. (BMO 1984) Let $ABCD$ be an inscribed quadrilateral and let H_A, H_B, H_C and H_D be the orthocenters of the triangles BCD, CDA, DAB, and ABC respectively. Prove that the quadrilaterals $ABCD$ and $H_AH_BH_CH_D$ are congruent.

3. (Yug TST 1992) The squares $BCDE$, $CAFG$, and $ABHI$ are constructed outside the triangle ABC. Let $GCDQ$ and $EBHP$ be parallelograms. Prove that $\triangle APQ$ is isosceles and rectangular.

4. (Yug MO 1993, 3-4 grade) The equilateral triangles BCB_1, CDC_1, and DAD_1 are constructed outside the triangle ABC. If P and Q are respectively the midpoints of B_1C_1 and C_1D_1 and if R is the midpoint of AB, prove that $\triangle PQR$ is isosceles.

5. In the plane of the triangle $A_1A_2A_3$ the point P_0 is given. Denote with $A_s = A_{s-3}$, for every natural number $s > 3$. The sequence of points P_0, P_1, P_2, ... is constructed in such a way that the point P_{k+1} is obtained by the rotation of the point P_k for an angle 120^o in the clockwise direction around the point A_{k+1}. Prove that if $P_{1986} = P_0$, then the triangle $A_1A_2A_3$ has to be isosceles.

6. (IMO Shortlist 1992) Let $ABCD$ be a convex quadrilateral for which $AC = BD$. Equilateral triangles are constructed on the sides of the quadrilateral. Let O_1, O_2, O_3, and O_4 be the centers of the triangles constructed on AB, BC, CD, and DA respectively. Prove that the lines O_1O_3 and O_2O_4 are perpendicular.

4 The Distance. Regular Polygons

In this section we will use the following basic relation for complex numbers: $|a|^2 = a\bar{a}$. Similarly, for calculating the sums of distances it is of great advantage if points are colinear or on mutually parallel lines. Hence it is often very useful to use rotations that will move some points in nice positions.

Now we will consider the regular polygons. It is well-known that the equation $x^n = 1$ has exactly n solutions in complex numbers and they are of the form $x_k = e^{i\frac{2k\pi}{n}}$, for $0 \le k \le n-1$. Now we have that $x_0 = 1$ and $x_k = \varepsilon^k$, for $1 \le k \le n-1$, where $x_1 = \varepsilon$.

Let's look at the following example for the illustration:

Problem 1. *Let $A_0A_1A_2A_3A_4A_5A_6$ be a regular 7-gon. Prove that*

$$\frac{1}{A_0A_1} = \frac{1}{A_0A_2} + \frac{1}{A_0A_3}.$$

Solution. As mentioned above let's take $a_k = \varepsilon^k$, for $0 \leq k \leq 6$, where $\varepsilon = e^{i\frac{2\pi}{7}}$. Further, by rotation around $a_0 = 1$ for the angle ε, i.e. $\omega = e^{i\frac{2\pi}{14}}$, the points a_1 and a_2 are mapped to a_1' and a_2' respectively. These two points are collinear with a_3. Now it is enough to prove that $\frac{1}{a_1'-1} = \frac{1}{a_2'-1} + \frac{1}{a_3-1}$. Since $\varepsilon = \omega^2$, $a_1' = \varepsilon(a_1-1)+1$, and $a_2' = \omega(a_2-1)+1$ it is enough to prove that

$$\frac{1}{\omega^2(\omega^2-1)} = \frac{1}{\omega(\omega^4-1)} + \frac{1}{\omega^6-1}.$$

After rearranging we get $\omega^6+\omega^4+\omega^2+1 = \omega^5+\omega^3+\omega$. From $\omega^5 = -\omega^{12}$, $\omega^3 = -\omega^{10}$, and $\omega = -\omega^8$ (which can be easily seen from the unit circle), the equality follows from $0 = \omega^{12}+\omega^{10}+\omega^8+\omega^6+\omega^4+\omega^2+1 = \varepsilon^6+\varepsilon^5+\varepsilon^4+\varepsilon^3+\varepsilon^2+\varepsilon+1 = \frac{\varepsilon^7-1}{\varepsilon-1} = 0$. $\triangle$

7. Let $A_0A_1\ldots A_{14}$ be a regular 15-gon. Prove that

$$\frac{1}{A_0A_1} = \frac{1}{A_0A_2} + \frac{1}{A_0A_4} + \frac{1}{A_0A_7}.$$

8. Let $A_0A_1\ldots A_{n-1}$ be a regular n-gon inscribed in a circle with radius r. Prove that for every point P of the circle and every natural number $m < n$ we have

$$\sum_{k=0}^{n-1} PA_k^{2m} = \binom{2m}{m} nr^{2m}.$$

9. (SMN TST 2003) Let M and N be two different points in the plane of the triangle ABC such that

$$AM : BM : CM = AN : BN : CN.$$

Prove that the line MN contains the circumcenter of $\triangle ABC$.

10. Let P be an arbitrary point on the shorter arc A_0A_{n-1} of the circle circumscribed about the regular polygon $A_0A_1\ldots A_{n-1}$. Let $h_1, h_2, \ldots, h_n$ be the distances of P from the lines that contain the edges A_0A_1, A_1A_2, …, $A_{n-1}A_0$ respectively. Prove that

$$\frac{1}{h_1} + \frac{1}{h_2} + \cdots + \frac{1}{h_{n-1}} = \frac{1}{h_n}.$$

5 Polygons Inscribed in Circle

In the problems where the polygon is inscribed in the circle, it is often useful to assume that the unit circle is the circumcircle of the polygon. In theorem 2 we can see lot of advantages of the unit circle (especially the first statement) and in practice we will see that lot of the problems can be solved using this method. In particular, we know that each triangle is inscribed in the circle and in many problems from the geometry of triangle we can make use of complex numbers. The only problem in this task is finding the circumcenter. For that you should take a look in the next two sections.

11. The quadrilateral $ABCD$ is inscribed in the circle with diameter AC. The lines AB and CD intersect at M and the tangets to the circle at B and C interset at N. Prove that $MN \perp AC$.

12. (IMO Shorlist 1996) Let H be the orthocenter of the triangle $\triangle ABC$ and P an arbitrary point of its circumcircle. Let E the foot of perpendicular BH and let $PAQB$ and $PARC$ be parallelograms. If AQ and HR intersect in X prove that $EX \| AP$.

13. Given a cyclic quadrilateral $ABCD$, denote by P and Q the points symmetric to C with respect to AB and AD respectively. Prove that the line PQ passes through the orthocenter of $\triangle ABD$.

14. (IMO Shortlist 1998) Let ABC be a triangle, H its orthocenter, O its incenter, and R the circumradius. Let D be the point symmetric to A with respect to BC, E the point symmetric to B with respect to CA, and F the point symmetric to C with respect to AB. Prove that the points D, E, and F are collinear if and only if $OH = 2R$.

15. (Rehearsal Competition in MG 2004) Given a triangle ABC, let the tangent at A to the circumscribed circle intersect the midsegment parallel to BC at the point A_1. Similarly we define the points B_1 and C_1. Prove that the points A_1, B_1, C_1 lie on a line which is parallel to the Euler line of $\triangle ABC$.

16. (MOP 1995) Let AA_1 and BB_1 be the altitudes of $\triangle ABC$ and let $AB \neq AC$. If M is the midpoint of BC, H the orthocenter of $\triangle ABC$, and D the intersection of BC and B_1C_1, prove that $DH \perp AM$.

17. (IMO Shortlist 1996) Let ABC be an acute-angled triangle such that $BC > CA$. Let O be the circumcircle, H the orthocenter, and F the foot of perpendicular CH. If the perpendicular from F to OF intersects CA at P, prove that $\angle FHP = \angle BAC$.

18. (Romania 2005) Let $A_0A_1A_2A_3A_4A_5$ be a convex hexagon inscribed in a circle. Let A_0', A_2', A_4' be the points on that circle such that

$$A_0A_0' \parallel A_2A_4, \quad A_2A_2' \parallel A_4A_0 \quad A_4A_4' \parallel A_2A_0.$$

Suppose that the lines $A_0'A_3$ and A_2A_4 intersect at A_3', the lines $A_2'A_5$ and A_0A_4 intersect at A_5', and the lines $A_4'A_1$ and A_0A_2 intersect at A_1'.
If the lines A_0A_3, A_1A_4, and A_2A_5 are concurrent, prove that the lines A_0A_3', A_4A_1' and A_2A_5' are concurrent as well.

19. (**Simson's line**) If A, B, C are points on a circle, then the feet of perpendiculars from an arbitrary point D of that circle to the sides of ABC are collinear.

20. Let A, B, C, D be four points on a circle. Prove that the intersection of the Simsons line corresponding to A with respect to the triangle BCD and the Simsons line corresponding to B w.r.t. $\triangle ACD$ belongs to the line passing through C and the orthocenter of $\triangle ABD$.

21. Denote by $l(S;PQR)$ the Simsons line corresponding to the point S with respect to the triangle PQR. If the points A, B, C, D belong to a circle, prove that the lines $l(A;BCD)$, $l(B;CDA)$, $l(C,DAB)$, and $l(D,ABC)$ are concurrent.

22. (Taiwan 2002) Let A, B, and C be fixed points in the plane, and D the mobile point of the circumcircle of $\triangle ABC$. Let I_A denote the Simsons line of the point A with respect to $\triangle BCD$. Similarly we define I_B, I_C, and I_D. Find the locus of the points of intersection of the lines I_A, I_B, I_C, and I_D when D moves along the circle.

23. (BMO 2003) Given a triangle ABC, assume that $AB \neq AC$. Let D be the intersection of the tangent to the circumcircle of $\triangle ABC$ at A with the line BC. Let E and F be the points on the bisectors of the segments AB and AC respectively such that BE and CF are perpendicular to BC. Prove that the points D, E, and F lie on a line.

24. (**Pascal's Theorem**) If the hexagon $ABCDEF$ can be inscribed in a circle, prove that the points $AB \cap DE$, $BC \cap EF$, and $CD \cap FA$ are colinear.

25. (**Brokard's Theorem**) Let $ABCD$ be an inscribed quadrilateral. The lines AB and CD intersect at E, the lines AD and BC intersect in F, and the lines AC and BD intersect in G. Prove that O is the orthocenter of the triangle EFG.

26. (Iran 2005) Let ABC be an equilateral triangle such that $AB = AC$. Let P be the point on the extention of the side BC and let X and Y be the points on AB and AC such that

$$PX \parallel AC, \quad PY \parallel AB.$$

Let T be the midpoint of the arc BC. Prove that $PT \perp XY$.

27. Let $ABCD$ be an inscribed quadrilateral and let K, L, M, and N be the midpoints of AB, BC, CA, and DA respectively. Prove that the orthocenters of $\triangle AKN$, $\triangle BKL$, $\triangle CLM$, $\triangle DMN$ form a parallelogram.

6 Polygons Circumscribed Around Circle

Similarly as in the previous chapter, here we will assume that the unit circle is the one inscribed in the given polygon. Again we will make a use of theorem 2 and especially its third part. In the case of triangle we use also the formulas from the theorem 7. Notice that in this case we know both the incenter and circumcenter which was not the case in the previous section. Also, notice that the formulas from the theorem 7 are quite complicated, so it is highly recommended to have the circumcircle for as the unit circle whenever possible.

28. The circle with the center O is inscribed in the triangle ABC and it touches the sides AB, BC, CA in M, K, E respectively. Denote by P the intersection of MK and AC. Prove that $OP \perp BE$.

29. The circle with center O is inscribed in a quadrilateral $ABCD$ and touches the sides AB, BC, CD, and DA respectively in K, L, M, and N. The lines KL and MN intersect at S. Prove that $OS \perp BD$.

30. (BMO 2005) Let ABC be an acute-angled triangle which incircle touches the sides AB and AC in D and E respectively. Let X and Y be the intersection points of the bisectors of the angles $\angle ACB$ and $\angle ABC$ with the line DE. Let Z be the midpoint of BC. Prove that the triangle XYZ is isosceles if and only if $\angle A = 60^\circ$.

31. (**Newtons Theorem**) Given an circumscribed quadrilateral $ABCD$, let M and N be the midpoints of the diagonals AC and BD. If S is the incenter, prove that M, N, and S are colinear.

32. Let $ABCD$ be a quadrilateral whose incircle touches the sides AB, BC, CD, and DA at the points M, N, P, and Q. Prove that the lines AC, BD, MP, and NQ are concurrent.

33. (Iran 1995) The incircle of $\triangle ABC$ touches the sides BC, CA, and AB respectively in D, E, and F. X, Y, and Z are the midpoints of EF, FD, and DE respectively. Prove that the incenter of $\triangle ABC$ belongs to the line connecting the circumcenters of $\triangle XYZ$ and $\triangle ABC$.

34. Assume that the circle with center I touches the sides BC, CA, and AB of $\triangle ABC$ in the points D, E, F, respectively. Assume that the lines AI and EF intersect at K, the lines ED and KC at L, and the lines DF and KB at M. Prove that LM is parallel to BC.

35. (25. Tournament of Towns) Given a triangle ABC, denote by H its orthocenter, I the incenter, O its circumcenter, and K the point of tangency of BC and the incircle. If the lines IO and BC are parallel, prove that AO and HK are parallel.

36. (IMO 2000) Let AH_1, BH_2, and CH_3 be the altitudes of the acute-angled triangle ABC. The incircle of ABC touches the sides BC, CA, AB respectively in T_1, T_2, and T_3. Let l_1, l_2, and l_3 be the lines symmetric to H_2H_3, H_3H_1, H_1H_2 with respect to T_2T_3, T_3T_1, and T_1T_2 respectively. Prove that the lines l_1, l_2, l_3 determine a triagnle whose vertices belong to the incircle of ABC.

7 The Midpoint of Arc

We often encounter problems in which some point is defined to be the midpoint of an arc. One of the difficulties in using complex numbers is distinguishing the arcs of the cirle. Namely, if we define the midpoint of an arc to be the intersection of the bisector of the corresponding chord with the circle, we are getting two solutions. Such problems can be relatively easy solved using the first part of the theorem 8. Moreover the second part of the theorem 8 gives an alternative way for solving the problems with incircles and circumcircles. Notice that the coordinates of the important points are given with the equations that are much simpler than those in the previous section. However we have a problem when calculating the points d, e, f of tangency of the incircle with the sides (calculate

them!), so in this case we use the methods of the previous section. In the case of the non-triangular polygon we also prefer the previous section.

37. (Kvant M769) Let L be the incenter of the triangle ABC and let the lines AL, BL, and CL intersect the circumcircle of $\triangle ABC$ at A_1, B_1, and C_1 respectively. Let R be the circumradius and r the inradius. Prove that:

$$\text{(a) } \frac{LA_1 \cdot LC_1}{LB} = R; \quad \text{(b) } \frac{LA \cdot LB}{LC_1} = 2r; \quad \text{(c) } \frac{S(ABC)}{S(A_1B_1C_1)} = \frac{2r}{R}.$$

38. (Kvant M860) Let O and R be respectively the center and radius of the circumcircle of the triangle ABC and let Z and r be respectively the incenter and inradius of $\triangle ABC$. Denote by K the centroid of the triangle formed by the points of tangency of the incircle and the sides. Prove that Z belongs to the segment OK and that $OZ : ZK = 3R/r$.

39. Let P be the intersection of the diagonals AC and BD of the convex quadrilateral $ABCD$ for which $AB = AC = BD$. Let O and I be the circumcenter and incenter of the triangle ABP. Prove that if $O \neq I$ then $OI \perp CD$.

40. Let I be the incenter of the triangle ABC for which $AB \neq AC$. Let O_1 be the point symmetric to the circumcenter of $\triangle ABC$ with respect to BC. Prove that the points A, I, O_1 are colinear if and only if $\angle A = 60^\circ$.

41. Given a triangle ABC, let A_1, B_1, and C_1 be the midpoints of BC, CA, and AB respecctively. Let P, Q, and R be the points of tangency of the incircle k with the sides BC, CA, and AB. Let P_1, Q_1, and R_1 be the midpoints of the arcs QR, RP, and PQ on which the points P, Q, and R divide the circle k, and let P_2, Q_2, and R_2 be the midpoints of arcs QPR, RQP, and PRQ respectively. Prove that the lines A_1P_1, B_1Q_1, and C_1R_1 are concurrent, as well as the lines A_1P_1, B_1Q_2, and C_1R_2.

8 Important Points. Quadrilaterals

In the last three sections the points that we've taken as initial, i.e. those with *known* coordinates have been "equally improtant" i.e. all of them had the same properties (they've been either the points of the same circle, or intersections of the tangents of the same circle, etc.). However, there are numerous problems where it is possible to distinguish one point from the others based on its influence to the other points. That point will be regarded as the origin. This is particularly useful in the case of quadrilaterals (that can't be inscribed or circumscribed around the circle) – in that case the intersection of the diagonals can be a good choice for the origin. We will make use of the formulas from the theorem 9.

42. The squares $ABB'B''$, $ACC'C''$, $BCXY$ are consctructed in the exterior of the triangle ABC. Let P be the center of the square $BCXY$. Prove that the lines CB'', BC'', AP intersect in a point.

43. Let O be the intersection of diagonals of the quadrilateral $ABCD$ and M, N the midpoints of the side AB and CD respectively. Prove that if $OM \perp CD$ and $ON \perp AB$ then the quadrilateral $ABCD$ is cyclic.

44. Let F be the point on the base AB of the trapezoid $ABCD$ such that $DF = CF$. Let E be the intersection of AC and BD and O_1 and O_2 the circumcenters of $\triangle ADF$ and $\triangle FBC$ respectively. Prove that $FE \perp O_1O_2$.

45. (IMO 2005) Let $ABCD$ be a convex quadrilateral whose sides BC and AD are of equal length but not parallel. Let E and F be interior points of the sides BC and AD respectively such that $BE = DF$. The lines AC and BD intersect at P, the lines BD and EF intersect at Q, and the lines EF and AC intersect at R. Consider all such triangles PQR as E and F vary. Show that the circumcircles of these triangles have a common point other than P.

46. Assume that the diagonals of $ABCD$ intersect in O. Let T_1 and T_2 be the centroids of the triangles AOD and BOC, and H_1 and H_2 orthocenters of $\triangle AOB$ and $\triangle COD$. Prove that $T_1T_2 \perp H_1H_2$.

9 Non-unique Intersections and Viete's formulas

The point of intersection of two lines can be determined from the system of two equations each of which corresponds to the condition that a point correspond to a line. However this method can lead us into some difficulties. As we mentioned before standard methods can lead to non-unique points. For example, if we want to determine the intersection of two circles we will get a quadratic equations. That is not surprising at all since the two circles have, in general, two intersection points. Also, in many of the problems we don't need both of these points, just the direction of the line determined by them. Similarly, we may already know one of the points. In both cases it is more convenient to use Vieta's formulas and get the sums and products of these points. Thus we can avoid "taking the square root of a complex number" which is very suspicious operation by itself, and usually requires some knowledge of complex analysis.

Let us make a remark: If we need explicitly coordinates of one of the intersection points of two circles, and we don't know the other, the only way to solve this problem using complex numbers is to set the given point to be one of the initial points.

47. Suppose that the tangents to the circle Γ at A and B intersect at C. The circle Γ_1 which passes through C and touches AB at B intersects the circle Γ at the point M. Prove that the line AM bisects the segment BC.

48. (Republic Competition 2004, 3rd grade) Given a circle k with the diameter AB, let P be an arbitrary point of the circle different from A and B. The projections of the point P to AB is Q. The circle with the center P and radius PQ intersects k at C and D. Let E be the intersection of CD and PQ. Let F be the midpoint of AQ, and G the foot of perpendicular from F to CD. Prove that $EP = EQ = EG$ and that A, G, and P are colinear.

49. (China 1996) Let H be the orthocenter of the triangle ABC. The tangents from A to the circle with the diameter BC intersect the circle at the points P and Q. Prove that the points P, Q, and H are colinear.

50. Let P be the point on the extension of the diagonal AC of the rectangle $ABCD$ over the point C such that $\angle BPD = \angle CBP$. Determine the ratio $PB : PC$.

51. (IMO 2004) In the convex quadrilateral $ABCD$ the diagonal BD is not the bisector of any of the angles ABC and CDA. Let P be the point in the interior of $ABCD$ such that

$$\angle PBC = \angle DBA \text{ and } \angle PDC = \angle BDA.$$

Prove that the quadrilateral $ABCD$ is cyclic if and only if $AP = CP$.

10 Different Problems – Different Methods

In this section you will find the problems that are not closely related to some of the previous chapters, as well as the problems that are related to more than one of the chapters. The useful advice is to carefully think of possible initial points, the origin, and the unit circle. As you will see, the main problem with solving these problems is the time. Thus if you are in competition and you want to use complex numbers it is very important for you to estimate the time you will spend. Having this in mind, it is very important to learn complex numbers as early as possible.

You will see several problems that use theorems 3, 4, and 5.

52. Given four circles k_1, k_2, k_3, k_4, assume that $k_1 \cap k_2 = \{A_1, B_1\}$, $k_2 \cap k_3 = \{A_2, B_2\}$, $k_3 \cap k_4 = \{A_3, B_3\}$, $k_4 \cap k_1 = \{A_4, B_4\}$. If the points A_1, A_2, A_3, A_4 lie on a circle or on a line, prove that the points B_1, B_2, B_3, B_4 lie on a circle or on a line.

53. Suppose that $ABCD$ is a parallelogram. The similar and equally oliented triangles CD and CB are constructed outside this parallelogram. Prove that the triangle FAE is similar and equally oriented with the first two.

54. Three triangles KPQ, QLP, and PQM are constructed on the same side of the segment PQ in such a way that $\angle QPM = \angle PQL = \alpha$, $\angle PQM = \angle QPK = \beta$, and $\angle PQK = \angle QPL = \gamma$. If $\alpha < \beta < \gamma$ and $\alpha + \beta + \gamma = 180^\circ$, prove that the triangle KLM is similar to the first three.

55. *(Iran, 2005) Let n be a prime number and H_1 a convex n-gon. The polygons $H_2, \ldots, H_n$ are defined recurrently: the vertices of the polygon H_{k+1} are obtained from the vertices of H_k by symmetry through k-th neighbour (in the positive direction). Prove that H_1 and H_n are similar.

56. Prove that the area of the triangles whose vertices are feet of perpendiculars from an arbitrary vertex of the cyclic pentagon to its edges doesn't depend on the choice of the vertex.

57. The points A_1, B_1, C_1 are chosen inside the triangle ABC to belong to the altitudes from A, B, C respectively. If

$$S(ABC_1) + S(BCA_1) + S(CAB_1) = S(ABC),$$

prove that the quadrilateral $A_1B_1C_1H$ is cyclic.

58. (IMO Shortlist 1997) The feet of perpendiculars from the vertices A, B, and C of the triangle ABC are D, E, end F respectively. The line through D parallel to EF intersects AC and AB respectively in Q and R. The line EF intersects BC in P. Prove that the circumcircle of the triangle PQR contains the midpoint of BC.

59. (BMO 2004) Let O be a point in the interior of the acute-angled triangle ABC. The circles through O whose centers are the midpoints of the edges of $\triangle ABC$ mutually intersect at K, L, and M, (different from O). Prove that O is the incenter of the triangle KLM if and only if O is the circumcenter of the triangle ABC.

60. Two circles k_1 and k_2 are given in the plane. Let A be their common point. Two mobile points, M_1 and M_2 move along the circles with the constant speeds. They pass through A always at the same time. Prove that there is a fixed point P that is always equidistant from the points M_1 and M_2.

61. (Yug TST 2004) Given the square $ABCD$, let γ be i circle with diameter AB. Let P be an arbitrary point on CD, and let M and N be intersections of the lines AP and BP with γ that are different from A and B. Let Q be the point of intersection of the lines DM and CN. Prove that $Q \in \gamma$ and $AQ : QB = DP : PC$.

62. (IMO Shortlist 1995) Given the triangle ABC, the circle passing through B and C intersect the sides AB and AC again in C' and B' respectively. Prove that the lines BB', CC', and HH' are concurrent, where H and H' orthocenters of the triangles ABC and $A'B'C'$ respectively.

63. (IMO Shortlist 1998) Let M and N be interior points of the triangle ABC such that $\angle MAB = \angle NAC$ and $\angle MBA = \angle NBC$. Prove that

$$\frac{AM \cdot AN}{AB \cdot AC} + \frac{BM \cdot BN}{BA \cdot BC} + \frac{CM \cdot CN}{CA \cdot CB} = 1.$$

64. (IMO Shortlist 1998) Let $ABCDEF$ be a convex hexagon such that $\angle B + \angle D + \angle F = 360^\circ$ and $AB \cdot CD \cdot EF = BC \cdot DE \cdot FA$. Prove that

$$BC \cdot AE \cdot FD = CA \cdot EF \cdot DB.$$

65. (IMO Shortlist 1998) Let ABC be a triangle such that $\angle A = 90^\circ$ and $\angle B < \angle C$. The tangent at A to its circumcircle ω intersect the line BC at D. Let E be the reflection of A with respect to BC, X the foot of the perpendicular from A to BE, and Y the midpoint of AX. If the line BY intersects ω in Z, prove that the line BD tangents the circumcircle of $\triangle ADZ$.

Hint: Use some inversion first...

66. (Rehearsal Competition in MG 1997, 3-4 grade) Given a triangle ABC, the points A_1, B_1 and C_1 are located on its edges BC, CA, and AB respectively. Suppose that $\triangle ABC \sim \triangle A_1B_1C_1$. If either

the orthocenters or the incenters of $\triangle ABC$ and $\triangle A_1B_1C_1$ coincide prove that the triangle ABC is equilateral.

67. (**Ptolomy's inequality**) Prove that for every convex quadrilateral $ABCD$ the following inequality holds

$$AB \cdot CD + BC \cdot AD \geq AC \cdot BD.$$

68. (China 1998) Find the locus of all points D such that

$$DA \cdot DB \cdot AB + DB \cdot DC \cdot BC + DC \cdot DA \cdot CA = AB \cdot BC \cdot CA.$$

11 Disadvantages of the Complex Number Method

The bigest difficulties in the use of the method of complex numbers can be encountered when dealing with the intersection of the lines (as we can see from the fifth part of the theorem 2, although it dealt with the chords of the circle). Also, the difficulties may arrise when we have more than one circle in the problem. Hence you should avoid using the comples numbers in problems when there are lot of lines in general position without some special circle, or when there are more then two circles. Also, the things can get very complicated if we have only two circles in general position, and only in the rare cases you are advised to use complex numbers in such situations. The problems when some of the conditions is the equlity with sums of distances between non-colinear points can be very difficult and pretty-much unsolvable with this method.

Of course, these are only the obvious situations when you can't count on help of complex numbers. There are numerous innocent-looking problems where the calculation can give us increadible difficulties.

12 Hints and Solutions

Before the solutions, here are some remarks:

- In all the problems it is assumed that the lower-case letters denote complex numbers corresponding to the points denoted by capital letters (sometimes there is an exception when the unit circle is the incircle of the triangle and its center is denoted by o).
- Some abbreviations are used for addressing the theorems. For example T1.3 denotes the third part of the theorem 1.
- The solutions are quite useless if you don't try to solve the problem by yourself.
- Obvious derivations and algebraic manipulations are skipped. All expressions that are somehow "equally" related to both a and b are probably divisible by $a-b$ or $a+b$.
- To make the things simpler, many conjugations are skipped. However, these are very straightforward, since most of the numbers are on the unit circle and they satisfy $\overline{a} = \dfrac{1}{a}$.
- If you still doesn't believe in the power of complex numbers, you are more than welcome to try these problems with other methods– but don't hope to solve all of them. For example, try the problem 41. Sometimes, complex numbers can give you shorter solution even when comparing to the elementar solution.
- The author has tried to make these solutions available in relatively short time, hence some mistakes are possible. For all mistakes you've noticed and for other solutions (with complex numbers), please write to me to the above e-mail address.

1. Assume that the circumcircle of the triangle abc is the unit circle, i.e. $s=0$ and $|a|=|b|=|c|=1$. According to T6.3 we have $h=a+b+c$, and according to T6.1 we conclude that $h+q=2s=0$, i.e. $q=-a-b-c$. Using T6.2 we get $t_1=\dfrac{b+c+q}{3}=-\dfrac{a}{3}$ and similarly $t_2=-\dfrac{b}{3}$ and $t_3=-\dfrac{c}{3}$. We now have $|a-t_1|=\left|a+\dfrac{a}{3}\right|=\left|\dfrac{4a}{3}\right|=\dfrac{4}{3}$ and similarly $|b-t_2|=|c-t_3|=\dfrac{4}{3}$. The proof is complete. We have assumed that $R=1$, but this is no loss of generality.

2. For the unit circle we will take the circumcircle of the quadrilateral $abcd$. According to T6.3 we have $h_a=b+c+d$, $h_b=c+d+a$, $h_c=d+a+b$, and $h_d=a+b+c$. In order to prove that $abcd$ and $h_ah_bh_ch_d$ are congruent it is enough to establish $|x-y|=|h_x-h_y|$, for all $x,y\in\{a,b,c,d\}$. This is easy to verify.

3. Notice that the point h ca be obtained by the rotation of the point a around b for the angle $\dfrac{\pi}{2}$ in the positive direction. Since $e^{i\frac{\pi}{2}}=i$, using T1.4 we get $(a-b)i=a-h$, i.e. $h=(1-i)a+ib$. Similarly we get $d=(1-i)b+ic$ and $g=(1-i)c+ia$. Since $BCDE$ is a square, it is a parallelogram as well, hence the midpoints of ce and bd coincide, hence by T6.1 we have $d+b=e+c$, or $e=(1+i)b-ic$. Similarly $g=(1+i)c-ia$. The quadrilaterals $beph$ and $cgqd$ are parallelograms implying that $p+b=e+h$ and $c+q=g+d$, or

$$p=ia+b-ic,\quad q=-ia+ib+c.$$

In order to finish the proof it is enough to show that q ca be obtained by the rotation of p around a by an angle $\dfrac{\pi}{2}$, which is by T1.4 equivalent to

$$(p-a)i=p-b.$$

The last identity is easy to verify.

4. The points b_1, c_1, d_1, are obtained by rotation of b, c, d around c, d, and a for the angle $\dfrac{\pi}{3}$ in the positive direction. If we denote $e^{i\pi/3}=\varepsilon$ using T1.4 we get

$$(b-c)\varepsilon=b_1-c,\quad (c-d)\varepsilon=c_1-d,\quad (d-a)\varepsilon=d_1-a.$$

Since p is the midpoint of b_1c_1 T6.1 gives

$$p=\frac{b_1+c_1}{2}=\frac{\varepsilon b+c+(1-\varepsilon)d}{2}.$$

Similarly we get $q=\dfrac{\varepsilon c+d+(1-\varepsilon)a}{2}$. Using T6.1 again we get $r=\dfrac{a+b}{2}$. It is enough to prove that q can be obtained by the rotation of p around r for the angle $\dfrac{\pi}{3}$, in the positive direction. The last is (by T1.4) equivalent to

$$(p-r)\varepsilon=q-r,$$

which follows from

$$p-r=\frac{-a+(\varepsilon-1)b+c+(1-\varepsilon)}{2},\quad q-r=\frac{-\varepsilon a-b+\varepsilon c+d}{2},$$

and $\varepsilon^2-\varepsilon+1=0$ (since $0=\varepsilon^3+1=(\varepsilon+1)(\varepsilon^2-\varepsilon+1)$).

5. Let $\varepsilon=e^{i\frac{2\pi}{3}}$. According to T1.4 we have $p_{k+1}-a_{k+1}=(p_k-a_{k+1})\varepsilon$. Hence

$$\begin{aligned}p_{k+1} &= \varepsilon p_k+(1-\varepsilon)a_{k+1}=\varepsilon(\varepsilon p_{k-1}+(1-\varepsilon)a_k)+(1-\varepsilon)a_{k+1}=\dots\\ &= \varepsilon^{k+1}p_0+(1-\varepsilon)\sum_{i=1}^{k+1}\varepsilon^{k+1-i}a_i.\end{aligned}$$

Now we have $p_{1996} = p_0 + 665(1-\varepsilon)(\varepsilon^2 a_1 + \varepsilon a_2 + a_3)$, since $\varepsilon^3 = 1$. That means $p_{1996} = p_0$ if and only if $\varepsilon^2 a_1 + \varepsilon a_2 + a_3 = 0$. Using that $a_1 = 0$ we conclude $a_3 = -\varepsilon a_2$, and it is clear that a_2 can be obtained by the rotation of a_3 around $0 = a_1$ for the angle $\dfrac{\pi}{3}$ in the positive direction.

6. Since the point a is obtained by the rotation of b around o_1 for the angle $\dfrac{2\pi}{3} = \varepsilon$ in the positive direction, T1.4 implies $(o_1 - b)\varepsilon = o_1 - a$, i.e. $o_1 = \dfrac{a - b\varepsilon}{1-\varepsilon}$. Analogously

$$o_2 = \frac{b - c\varepsilon}{1-\varepsilon}, \quad o_3 = \frac{c - d\varepsilon}{1-\varepsilon}, \quad o_4 = \frac{d - a\varepsilon}{1-\varepsilon}.$$

Since $o_1o_3 \perp o_2o_4$ is equivalent to $\dfrac{o_1 - o_3}{\overline{o_1} - \overline{o_3}} = -\dfrac{o_2 - o_4}{\overline{o_2} - \overline{o_4}}$, it is enouogh to prove that

$$\frac{a - c - (b-d)\varepsilon}{\overline{a-c} - \overline{(b-d)\varepsilon}} = -\frac{b - d - (c-a)\varepsilon}{\overline{b-d} - \overline{(c-a)\varepsilon}},$$

i.e. that $(a-c)\overline{b-d} - (b-d)\overline{b-d}\,\varepsilon + (a-c)\overline{a-c}\,\overline{\varepsilon} - (b-d)\overline{a-c}\,\varepsilon\overline{\varepsilon} = -\overline{a-c}(b-d) + (b-d)\overline{b-d}\,\overline{\varepsilon} - (a-c)\overline{a-c}\,\varepsilon + (a-c)\overline{b-d}\,\varepsilon\overline{\varepsilon}$. The last follows from $\overline{\varepsilon} = \dfrac{1}{\varepsilon}$ and $|a-c|^2 = (a-c)\overline{a-c} = |b-d|^2 = (b-d)\overline{b-d}$.

7. We can assume that $a_k = \varepsilon^k$ for $0 \le k \le 12$, where $\varepsilon = e^{i\frac{2\pi}{15}}$. By rotation of the points a_1, a_2, and a_4 around $a_0 = 1$ for the angles ω^6, ω^5, and ω^3 (here $\omega = e^{i\pi/15}$), we get the points a_1', a_2', and a_4', such that takve da su $a_0, a_7, a_1', a_2', a_4'$ kolinearne. Sada je dovoljno dokazati da je

$$\frac{1}{a_1' - 1} = \frac{1}{a_2' - 1} + \frac{1}{a_4' - 1} + \frac{1}{a_7 - 1}.$$

From T1.4 we have $a_1' - a_0 = (a_1 - a_0)\omega^6, a_2' - a_0 = (a_2 - a_0)\omega^5$ and $a_4' - a_0 = (a_4 - a_0)\omega^3$, as well as $\varepsilon = \omega^2$ and $\omega^{30} = 1$. We get

$$\frac{1}{\omega^6(\omega^2 - 1)} = \frac{1}{\omega^5(\omega^4 - 1)} + \frac{1}{\omega^3(\omega^8 - 1)} - \frac{\omega^{14}}{\omega^{16} - 1}.$$

Taking the common denominator and cancelling with $\omega^2 - 1$ we see that it is enough to prove that

$$\omega^8 + \omega^6 + \omega^4 + \omega^2 + 1 = \omega(\omega^{12} + \omega^8 + \omega^4 + 1) + \omega^3(\omega^8 + 1) - \omega^{20}.$$

Since $\omega^{15} = -1 = -\omega^{30}$, we have that $\omega^{15-k} = -\omega^{30-k}$. The required statement follows from $0 = \omega^{28} + \omega^{26} + \omega^{24} + \omega^{22} + \omega^{20} + \omega^{18} + \omega^{16} + \omega^{14} + \omega^{12} + \omega^{10} + \omega^8 + \omega^6 + \omega^4 + \omega^2 + 1 = \frac{\omega^{30} - 1}{\omega^2 - 1} = 0$.

8. [Obtained from Uroš Rajković] Take the complex plane in which the center of the polygon is the origin and let $z = e^{i\frac{\pi}{k}}$. Now the coordinate of A_k in the complex plane is z^{2k}. Let p ($|p| = 1$) be the coordinate of P. Denote the left-hand side of the equality by S. We need to prove that $S = \binom{2m}{m} \cdot n$. We have that

$$S = \sum_{k=0}^{n-1} PA_k^{2m} = \sum_{k=0}^{n-1} \left| z^{2k} - p \right|^{2m}$$

Notice that the arguments of the complex numbers $(z^{2k} - p) \cdot z^{-k}$ (where $k \in \{0, 1, 2, \ldots, n\}$) are equal to the argument of the complex number $(1 - p)$, hence

$$\frac{(z^{2k} - p) \cdot z^{-k}}{1 - p}$$

is a positive real number. Since $|z^{-k}| = 1$ we get:

$$S = \sum_{k=0}^{n-1} |z^{2k} - p|^{2m} = |1-p|^{2m} \cdot \sum_{k=0}^{n-1} \left(\frac{z^{2k}-p}{1-p} \right)^{2m} = |1-p|^{2m} \cdot \frac{\sum_{k=0}^{n-1} (z^{2k}-p)^{2m}}{(1-p)^{2m}}.$$

Since S is a positive real number we have:

$$S = \left| \sum_{k=0}^{n-1} (z^{2k} - p)^{2m} \right|.$$

Now from the binomial formula we have:

$$S = \left| \sum_{k=0}^{n-1} \left[\sum_{i=0}^{2m} \binom{2m}{i} \cdot z^{2ki} \cdot (-p)^{2m-i} \right] \cdot z^{-2mk} \right|.$$

After some algebra we get:

$$S = \left| \sum_{k=0}^{n-1} \sum_{i=0}^{2m} \binom{2m}{i} \cdot z^{2k(i-m)} \cdot (-p)^{2m-i} \right|,$$

or, equivalently

$$S = \left| \sum_{i=0}^{2m} \binom{2m}{i} \cdot (-p)^{2m-i} \cdot \sum_{k=0}^{n-1} z^{2k(i-m)} \right|.$$

Since for $i \neq m$ we have:

$$\sum_{k=0}^{n-1} z^{2k(i-m)} = \frac{z^{2n(i-m)} - 1}{z^{2(i-m)} - 1},$$

for $z^{2n(i-m)} - 1 = 0$ and $z^{2(i-m)} - 1 \neq 0$, we have

$$\sum_{k=0}^{n-1} z^{2k(i-m)} = 0.$$

For $i = m$ we have:

$$\sum_{k=0}^{n-1} z^{2k(i-m)} = \sum_{k=0}^{n-1} 1 = n.$$

From this we conclude:

$$S = \left| \binom{2m}{m} \cdot (-p)^m \cdot n \right| = \binom{2m}{m} \cdot n \cdot |(-p)^m|.$$

Using $|p| = 1$ we get

$$S = \binom{2m}{m} \cdot n$$

and that is what we wanted to prove.

9. Choose the circumcircle of the triangle abc to be the unit circle. Then $o = 0$ and $\overline{a} = \frac{1}{a}$. The first of the given relations can be written as

$$1 = \frac{|a-m||b-n|}{|a-n||b-m|} \Rightarrow 1 = \frac{|a-m|^2|b-n|^2}{|a-n|^2|b-m|^2} = \frac{(a-m)(\overline{a}-\overline{m})(a-n)(\overline{a}-\overline{n})}{(a-n)(\overline{a}-\overline{n})(b-m)(\overline{b}-\overline{m})}$$

After some simple algebra we get $(a-m)(\overline{a}-\overline{m})(b-n)(\overline{b}-\overline{n}) = (1-\frac{m}{a}-a\overline{m}+m\overline{m})(1-\frac{n}{b}-b\overline{n}+n\overline{n}) = 1-\frac{m}{a}-a\overline{m}+m\overline{m}-\frac{n}{b}+\frac{mn}{ab}+\frac{a\overline{m}n}{b}-\frac{m\overline{m}n}{b}-b\overline{n}+\frac{bm\overline{n}}{a}+ab\overline{m}\,\overline{n}-bm\overline{m}\,\overline{n}+n\overline{n}-\frac{mn\overline{n}}{a}-a\overline{m}n\overline{n}+m\overline{m}n\overline{n}$. The value of the expression $(a-n)(\overline{a}-\overline{n})(b-m)(\overline{b}-\overline{m})$ we can get from the prevoius one replacing every a with b and vice versa. The initial equality now becomes:

$$\begin{aligned} & 1-\frac{m}{a}-a\overline{m}+m\overline{m}-\frac{n}{b}+\frac{mn}{ab}+\frac{a\overline{m}n}{b}-\frac{m\overline{m}n}{b}-b\overline{n}+ \\ & \frac{bm\overline{n}}{a}+ab\overline{m}\,\overline{n}-bm\overline{m}\,\overline{n}+n\overline{n}-\frac{mn\overline{n}}{a}-a\overline{m}n\overline{n}+m\overline{m}n\overline{n} \\ = \quad & 1-\frac{m}{b}-b\overline{m}+m\overline{m}-\frac{n}{a}+\frac{mn}{ab}+\frac{b\overline{m}n}{a}-\frac{m\overline{m}n}{a}-a\overline{n}+\frac{am\overline{n}}{b}+ \\ & ab\overline{m}\,\overline{n}-am\overline{m}\,\overline{n}+n\overline{n}-\frac{mn\overline{n}}{b}-b\overline{m}n\overline{n}+m\overline{m}n\overline{n}. \end{aligned}$$

Subtracting and taking $a-b$ out gives

$$\frac{m}{ab}-\overline{m}-\frac{n}{ab}+\frac{(a+b)\overline{m}n}{ab}-\frac{m\overline{m}n}{ab}+\overline{n}-\frac{(a+b)m\overline{n}}{ab}+m\overline{m}\,\overline{n}+\frac{mn\overline{n}}{ab}-\overline{m}n\overline{n}=0.$$

Since $AM/CM = AN/CM$ holds as well we can get the expression analogous to the above when every b is exchanged with c. Subtracting this expression from the previous and taking $b-c$ out we get

$$-\frac{m}{abc}+\frac{n}{abc}-\frac{\overline{m}n}{bc}+\frac{m\overline{m}n}{abc}+\frac{m\overline{n}}{bc}-\frac{mn\overline{n}}{abc}=0.$$

Writing the same expression with ac instead of bc (this can be obtained from the initial conditions because of the symmetry), subtracting, and simplifying yields $m\overline{n}-n\overline{m}=0$. Now we have $\frac{m-o}{\overline{m}-\overline{o}} = \frac{n-o}{\overline{n}-\overline{o}}$, and by T1.2 the points m,n,o are colinear.

10. [Obtained from Uroš Rajković] First we will prove that for the points p, a, and b of the unit circle the distance from p to the line ab is equal to:

$$\frac{1}{2}|(a-p)(b-p)|.$$

Denote by q the foot of perpendicular from p to ab and use T2.4 to get:

$$q=\frac{1}{2}\Big(p+a+b-\frac{ab}{p}\Big).$$

Now the required distance is equal to:

$$|q-p|=\frac{1}{2}\left|-p+a+b-\frac{ab}{p}\right|.$$

Since $|p|=1$ we can multiply the expression on the right by $-p$ which gives us:

$$\left|\frac{1}{2}(p^2-(a+b)p+ab)\right|.$$

Now it is easy to see that the required distance is indeed equal to:

$$\frac{1}{2}|(a-p)(b-p)|.$$

If we denote $z = e^{i\frac{2\pi}{2n}}$, the coordinate of A_k is z^{2k}. Now we have:

$$2 \cdot h_k = |(z^{2k} - p)(z^{2k-2} - p)|.$$

The vector $(z^{2k} - p) \cdot z^{-k}$ is colinear with $1 - p$, nece

$$\frac{(z^{2k} - p) \cdot z^{-k}}{1 - p}$$

is a positive real number. Hence for $k \in \{1, 2, \cdots, n-1\}$ it holds:

$$h_k = \frac{(z^{2k} - p) \cdot (z^{2k-2} - p) \cdot z^{-(2k-1)}}{2 \cdot (1-p)^2} \cdot |1 - p|^2,$$

since $|z| = 1$. We also have:

$$h_n = \frac{(1 - p) \cdot (z^{2n-2} - p) \cdot z^{-(n-1)}}{2 \cdot (1-p)^2} \cdot |1 - p|^2.$$

We need to prove that:

$$\sum_{k=1}^{n-1} \frac{1}{\frac{(z^{2k} - p) \cdot (z^{2k-2} - p) \cdot z^{-(2k-1)}}{2 \cdot (1-p)^2} \cdot |1 - p|^2} = \frac{1}{\frac{(1 - p) \cdot (z^{2n-2} - p) \cdot z^{-(n-1)}}{2 \cdot (1-p)^2} \cdot |1 - p|^2}.$$

After cancelling and multiplying by z we get:

$$\sum_{k=1}^{n-1} \frac{z^{2k}}{(z^{2k} - p) \cdot (z^{2k-2} - p)} = \frac{-1}{(1 - p) \cdot (z^{2n-2} - p)},$$

since $z^n = -1$. Denote by S the left-hand side of the equality. We have:

$$S - \frac{1}{z^2} S = \sum_{k=1}^{n-1} \frac{(z^{2k} - p) - (z^{2k-2} - p)}{(z^{2k} - p) \cdot (z^{2k-2} - p)}.$$

This implies:

$$(1 - \frac{1}{z^2})S = \sum_{k=1}^{n-1} \left(\frac{1}{z^{2k-2} - p} - \frac{1}{z^{2k} - p} \right).$$

After simplifying we get:

$$(1 - \frac{1}{z^2})S = \frac{1}{1 - p} - \frac{1}{z^{2n-2} - p} = \frac{(z^{2n-2} - p) - (1 - p)}{(1 - p) \cdot (z^{2n-2} - p)}.$$

Since $z^{2n-2} = \frac{1}{z^2}$ (from $z^n = 1$) we get:

$$S = \frac{-1}{(1 - p) \cdot (z^{2n-2} - p)},$$

q.e.d.

11. Assume that the unit circle is the circumcircle of the quadrilateral $abcd$. Since ac is its diameter we have $c=-a$. Furthermore by T2.5 we have that

$$m=\frac{ab(c+d)-cd(a+b)}{ab-cd}=\frac{2bd+ad-ab}{d+b}.$$

According to T2.3 we have that $n=\dfrac{2bd}{b+d}$, hence $m-n=\dfrac{a(d-b)}{b+d}$ and $\overline{m}-\overline{n}=\dfrac{b-d}{a(b+d)}$. Now we have

$$\frac{m-n}{\overline{m}-\overline{n}}=-\frac{a-c}{\overline{a}-\overline{c}}=a^2,$$

hence according to T1.3 $mn\perp ac$, q.e.d.

12. Assume that the unit circle is the circumcircle of the triangle abc. Using T6.3 we have $h=a+b+c$, and using T2.4 we have $e=\dfrac{1}{2}\left(a+b+c-\dfrac{ac}{b}\right)$. Since $paqb$ is a parallelogram the midpoints of pq and ab coincide, and according to T6.1 $q=a+b-p$ and analogously $r=a+c-p$. Since the points x,a,q are colinear, we have (using T1.2)

$$\frac{x-a}{\overline{x}-\overline{a}}=\frac{a-q}{\overline{a}-\overline{q}}=\frac{p-b}{\overline{p}-\overline{b}}=-pb,$$

or, equivalently $\overline{x}=\dfrac{pb+a^2-ax}{abp}$. Since the points h,r,x are colinear as well, using the same theorem we get

$$\frac{x-h}{\overline{x}-\overline{h}}=\frac{h-r}{\overline{h}-\overline{r}}=\frac{b+p}{\overline{b}+\overline{p}}=bp,$$

i.e.

$$\overline{x}=\frac{x-a-b-c+p+\dfrac{bp}{a}+\dfrac{bp}{c}}{bp}.$$

Equating the expressions obtained for $\overline{x}$ we get

$$x=\frac{1}{2}\left(2a+b+c-p-\frac{bp}{c}\right).$$

By T1.1 it is sufficient to prove that

$$\frac{e-x}{\overline{e}-\overline{x}}=\frac{a-p}{\overline{a}-\overline{p}}=-ap.$$

The last follows from

$$e-x=\frac{1}{2}\left(p+\frac{bp}{c}-a-\frac{ac}{b}\right)=\frac{bcp+b^2p-abc-ac^2}{2bc}=\frac{(b+c)(bp-ac)}{2bc},$$

by conjugation.

13. We will assume that the circumcircle of the quadrilateral $abcd$ is the unit circle. Using T2.4 and T6.1 we get

$$p=a+b-\frac{ab}{c},\quad q=a+d+\frac{ad}{c}\quad (1).$$

Let H be the orthocenter of the triangle ABD. By T6.3 we have $h=a+b+d$, hence according to T1.2 it is enough to prove that

$$\frac{p-h}{\overline{p}-\overline{h}}=\frac{q-h}{\overline{q}-\overline{h}}.\quad (2)$$

Chaning for p from (1) we get

$$\frac{p-h}{\overline{p}-\overline{h}} = \frac{a+b-\dfrac{ab}{c}-a-b-d}{\dfrac{1}{a}+\dfrac{1}{b}-\dfrac{c}{ab}-\dfrac{1}{a}-\dfrac{1}{b}-\dfrac{1}{d}} = \frac{abd}{c},$$

and since this expression is symmetric with respect to b and d, (2) is clearly satisfied.

14. Assume that the unit circle is the circumcircle of the triangle abc and assume that a',b',c' are feet of perpendiculars from a,b,c respectively. From T2.4 we have

$$a' = \frac{1}{2}\Big(a+b+c-\frac{bc}{a}\Big), \quad b' = \frac{1}{2}\Big(a+b+c-\frac{ca}{b}\Big), \quad c' = \frac{1}{2}\Big(a+b+c-\frac{ab}{c}\Big).$$

Since a',b',c' are midpoints of ad,be,cf respectively according to T6.1 we have

$$d = b+c-\frac{bc}{a}, \quad e = a+c-\frac{ac}{b}, \quad f = a+b-\frac{ab}{c}.$$

By T1.2 the colinearity of the points d,e,f is equivalent to

$$\frac{d-e}{\overline{d}-\overline{e}} = \frac{f-e}{\overline{f}-\overline{e}}.$$

Since $d-e = b-a+\dfrac{ac}{b}-\dfrac{bc}{a} = (b-a)\dfrac{ab-c(a+b)}{ab}$ and similarly $f-e = (b-c)\dfrac{bc-a(b+c)}{bc}$, by conjugation and some algebra we get

$$\begin{aligned} 0 &= (a^2b+a^2c-abc)(c-a-b)-(c^2a+c^2b-abc)(a-b-c) \\ &= (c-a)(abc-a^2b-ab^2-a^2c-ac^2-b^2c-bc^2). \quad (1) \end{aligned}$$

Now we want to get the necessary and sufficient condition for $|h|=2$ (the radius of the circle is 1). After the squaring we get

$$\begin{aligned} 4 &= |h|^2 = h\overline{h} = (a+b+c)\Big(\frac{1}{a}+\frac{1}{b}+\frac{1}{c}\Big) \\ &= \frac{a^2b+ab^2+a^2c+ac^2+b^2c+bc^2+3abc}{abc}. \quad (2) \end{aligned}$$

Now (1) is equivalent to (2), which finishes the proof.

15. Assume that the unit circle is the circumcircle of the triangle abc. Let a',b',c' be the midpoints of bc,ca,ab. Since $aa_1 \perp ao$ and since a_1,b',c' are colinear, using T1.3 and T1.2, we get

$$\frac{a-a_1}{\overline{a}-\overline{a_1}} = -\frac{a-o}{\overline{a}-\overline{o}} = -a^2, \quad \frac{b'-c'}{\overline{b'}-\overline{c'}} = \frac{b'-a_1}{\overline{b'}-\overline{a_1}}.$$

From the first equality we have $\overline{a_1} = \frac{2a-a_1}{a^2}$, and since from T6.1 $b' = \dfrac{a+c}{2}$ and $c' = \dfrac{a+b}{2}$ we also have $\overline{a_1} = \dfrac{ab+bc+ca-aa_1}{2abc}$. By equating the above expressions we get $a_1 = \dfrac{a^2(a+b+c)-3abc}{a^2-2bc}$. Similarly $b_1 = \dfrac{b^2(a+b+c)-3abc}{2(b^2-ac)}$ and $c_1 = \dfrac{c^2(a+b+c)-3abc}{2(c^2-2ab)}$. Now we have

$$a_1-b_1 = \frac{a^2(a+b+c)-3abc}{2(a^2-bc)} - \frac{b^2(a+b+c)-3abc}{2(b^2-ac)} = -\frac{c(a-b)^3(a+b+c)}{2(a^2-bc)(b^2-ac)},$$

and it is easy to verify the condition for $a_1b_1 \perp ho$, which is according to T1.3:

$$\frac{a_1-b_1}{\overline{a_1}-\overline{b_1}} = -\frac{h-o}{\overline{h}-\overline{o}} = -\frac{(a+b+c)abc}{ab+bc+ca}.$$

Similarly $a_1c_1 \perp ho$, implying that the points a_1, a_2, and a_3 are colinear.

16. Assume that the unit circle is the circumcircle of the triangle abc. By T2.4 we have that $b_1 = \frac{1}{2}\left(a+b+c-\frac{ac}{b}\right)$ and $c_1 = \frac{1}{2}\left(a+b+c-\frac{ab}{c}\right)$, according to T6.1 $m = \frac{b+c}{2}$, and according to T6.3 $h = a+b+c$. Now we will determine the point d. Since d belongs to the chord bc according to T2.2 $\overline{d} = \frac{b+c-d}{bc}$. Furthermore, since the points b_1, c_1, and d are colinear, according to T1.2 we have

$$\frac{d-b_1}{\overline{d}-\overline{b_1}} = \frac{b_1-c_1}{\overline{b_1}-\overline{c_1}} = \frac{a\left(\frac{b}{c}-\frac{c}{b}\right)}{\frac{1}{a}\left(\frac{c}{b}-\frac{b}{c}\right)} = -a^2.$$

Now we have that $\overline{d} = \frac{a^2\overline{b_1}+b_1-d}{a^2}$, hence

$$d = \frac{a^2b+a^2c+ab^2+ac^2-b^2c-bc^2-2abc}{2(a^2-bc)}.$$

In order to prove that $dh \perp am$ (see T1.3) it is enough to prove that $\frac{d-h}{\overline{d}-\overline{h}} = -\frac{m-a}{\overline{m}-\overline{a}}$. This however follows from

$$\begin{aligned} d-h &= \frac{b^2c+bc^2+ab^2+ac^2-a^2b-a^2c-2a^3}{2(a^2-bc)} \\ &= \frac{(b+c-2a)(ab+bc+ca+a^2)}{2(a^2-bc)} \end{aligned}$$

and $m-a = \frac{b+c-2a}{2}$ by conjugation.

17. Assume that the unit circle is the circumcircle of the triangle abc. By T2.4 we have that $f = \frac{1}{2}\left(a+b+c-\frac{ab}{c}\right)$. Since a, c, p are colinear and ac is a chord of the unit circle, according to T2.2 we have $\overline{p} = \frac{a+c-p}{ac}$. Since $fo \perp pf$ using T1.3 we coclude

$$\frac{f-o}{\overline{f}-\overline{o}} = -\frac{p-f}{\overline{p}-\overline{f}}.$$

From the last two relations we have

$$p = f\frac{2ac\overline{f}-(a+c)}{ac\overline{f}-f} = \frac{\left(a+b+c-\frac{ab}{c}\right)c^2}{b^2+c^2}.$$

Let $\angle phf = \varphi$, then

$$\frac{f-h}{\overline{f}-\overline{h}} = e^{i2\varphi}\frac{p-h}{\overline{p}-\overline{h}}.$$

Since $p-h = -b\frac{ab+bc+ca+c^2}{b^2+c^2}$, and by conjugation

$$\overline{p}-\overline{h} = -\frac{c(ab+bc+ca+b^2)}{ab(b^2+c^2)},$$

$f-h=\dfrac{ab+bc+ca+c^2}{2c}$, $\overline{f}-\overline{h}=\dfrac{ab+bc+ca+c^2}{2abc}$, we see that $e^{i2\varphi}=\dfrac{c}{b}$. On the other hand we have $\dfrac{c-a}{\overline{c}-\overline{a}}=e^{i2\alpha}\dfrac{b-a}{\overline{b}-\overline{a}}$, and using T1.2 $e^{i2\alpha}=\dfrac{c}{b}$. We have proved that $\alpha=\pi+\varphi$ or $\alpha=\varphi$, and since the first is impossible, the proof is complete.

18. First we will prove the following useful lemma.

Lemma 1. *If a, b, c, a', b', c' are the points of the unit circle, then the lines aa',bb',cc' concurrent or colinear if and only if*

$$(a-b')(b-c')(c-a')=(a-c')(b-a')(c-b').$$

Proof. Let x be the intersection of aa' and bb', and let y be the intersection of the lines aa' and cc'. Using T2.5 we have

$$x=\frac{aa'(b+b')-bb'(a+a')}{aa'-bb'},\quad y=\frac{aa'(c+c')-cc'(a+a')}{aa'-cc'}.$$

Here we assumed that these points exist (i.e. that none of $aa'\parallel bb'$ and $aa'\parallel cc'$ holds). It is obvious that the lines aa', bb', cc' are concurrent if and only if $x=y$, i.e. if and only if

$$(aa'(b+b')-bb'(a+a'))(aa'-cc')=(aa'(c+c')-cc'(a+a'))(aa'-bb').$$

After simplifying we get $aa'b+aa'b'-abb'-a'b'b-bcc'-b'cc'=aa'c+aa'c'-bc'c-bb'c'-acc'-a'cc'$, and since this is equivalent to $(a-b')(b-c')(c-a')=(a-c')(b-a')(c-b')$, the lemma is proven. □

Now assume that the circumcircle of the hexagon is the unit circle. Using T1.1 we get

$$\frac{a_2-a_4}{\overline{a_2}-\overline{a_4}}=\frac{a_0-a_0'}{\overline{a_0}-\overline{a_0'}},\quad \frac{a_4-a_0}{\overline{a_4}-\overline{a_0}}=\frac{a_2-a_2'}{\overline{a_2}-\overline{a_2'}},\quad \frac{a_2-a_0}{\overline{a_2}-\overline{a_0}}=\frac{a_4-a_4'}{\overline{a_4}-\overline{a_4'}},$$

hence $a_0'=\dfrac{a_2a_4}{a_0}, a_2'=\dfrac{a_0a_4}{a_2}$ i $a_4'=\dfrac{a_0a_2}{a_4}$. Similarly, using T2.5 we get

$$a_3'=\frac{a_0'a_3(a_2+a_3)-a_2a_3(a_0'+a_3)}{a_0'a_3-a_2a_4}=\frac{a_4(a_3-a_2)+a_3(a_2-a_0)}{a_3-a_0}.$$

Analogously,

$$a_5'=\frac{a_0(a_5-a_4)+a_5(a_4-a_2)}{a_5-a_2},\quad a_1'=\frac{a_2(a_1-a_0)+a_1(a_0-a_4)}{a_1-a_4}.$$

Assume that the points a_3'',a_1'',a_5'' are the other intersection points of the unit circle with the lines a_0a_3', a_4a_1', a_2a_5' respectively. According to T1.2

$$\frac{a_3'-a_0}{\overline{a_3'}-\overline{a_0}}=\frac{a_3''-a_0}{\overline{a_3''}-\overline{a_0}}=-a_3''a_0,$$

and since $a_0-a_3'=\dfrac{a_3(2a_0-a_2-a_4)+a_2a_4-a_0^2}{a_3-a_0}$, we have

$$a_3''-a_4=\frac{(a_0-a_2)^2(a_3-a_4)}{a_0a_2(a_3-a_0)(\overline{a_0}-\overline{a_3'})},\quad a_3''-a_2=\frac{(a_0-a_4)^2(a_3-a_2)}{a_0a_4(a_3-a_0)(\overline{a_0}-\overline{a_3'})}.$$

Analogously we get

$$a_1''-a_0=a_3''-a_4=\frac{(a_2-a_4)^2(a_1-a_0)}{a_2a_4(a_1-a_4)(\overline{a_4}-\overline{a_1'})},$$

$$a_1''-a_2=a_3''-a_4=\frac{(a_4-a_0)^2(a_1-a_2)}{a_0a_4(a_1-a_4)(\overline{a_4}-\overline{a_1'})},$$
$$a_5''-a_0=a_3''-a_4=\frac{(a_2-a_4)^2(a_5-a_0)}{a_2a_4(a_5-a_0)(\overline{a_2}-\overline{a_5'})},$$
$$a_5''-a_4=a_3''-a_4=\frac{(a_0-a_2)^2(a_5-a_4)}{a_0a_2(a_5-a_4)(\overline{a_2}-\overline{a_5'})}.$$

Using the lemma and the concurrence of the lines a_0a_3, a_1a_4, and a_2a_5 (i.e. $(a_0-a_1)(a_2-a_3)(a_4-a_5)=(a_0-a_5)(a_2-a_1)(a_4-a_3)$) we get the concurrence of the lines a_0a_3'', a_4a_1'', and a_2a_5'', i.e. $(a_0-a_1'')(a_2-a_3'')(a_4-a_5'')=(a_0-a_5'')(a_2-a_1'')(a_4-a_3'')$, since they, obviously, intersect.

19. [Obtained from Uroš Rajković] Assume that the unit circle is the circumcircle of the triangle abc. If A_1, B_1, and C_1 denote the feet of the perpendiculars, we have from T2.4:

$$\begin{aligned} a_1 &= \frac{1}{2}\left(b+c+m-\frac{bc}{m}\right),\\ b_1 &= \frac{1}{2}\left(a+c+m-\frac{ac}{m}\right), \text{ and}\\ c_1 &= \frac{1}{2}\left(a+b+m-\frac{ab}{m}\right). \end{aligned}$$

We further get:

$$\frac{a_1-c_1}{b_1-c_1}=\frac{c-a+\dfrac{ab-bc}{m}}{c-b+\dfrac{ab-ac}{m}}=\frac{(c-a)(m-b)}{(c-b)(m-a)}=\frac{\overline{a}_1-\overline{c}_1}{\overline{b}_1-\overline{c}_1},$$

and, according to T1.2, the points A_1, B_1, and C_1 are colinear.

20. The quadrilateral $ABCD$ is cyclic, and we assume that it's circumcircle is the unti circle. Let a_1, a_2, and a_3 denote the feet of the perpendiculars from a to bc, cd, and db respectively. Denote by b_1, b_2, and b_3 the feet of the perpendiculars from b to ac, cd, and da respectively. According to T2.4 we have that

$$a_1=\frac{1}{2}\left(a+b+c-\frac{bc}{a}\right),\ a_2=\frac{1}{2}\left(a+b+d-\frac{bd}{a}\right),\ a_3=\frac{1}{2}\left(a+c+d-\frac{cd}{a}\right)$$
$$b_1=\frac{1}{2}\left(b+a+c-\frac{ac}{b}\right),\ b_2=\frac{1}{2}\left(b+c+d-\frac{cd}{b}\right),\ b_3=\frac{1}{2}\left(b+d+a-\frac{da}{b}\right)$$

The point x can be obtained from the condition for colinearity. First from the colinearity of x, a_1, a_2 and T1.2 we have that

$$\frac{x-a_1}{\overline{x}-\overline{a_1}}=\frac{a_1-a_2}{\overline{a_1}-\overline{a_2}}=\frac{\dfrac{1}{2}\left(c-d+\dfrac{bd}{a}-\dfrac{bc}{a}\right)}{\dfrac{1}{2}\left(\dfrac{1}{c}-\dfrac{1}{d}+\dfrac{a}{bd}-\dfrac{a}{bc}\right)}=\frac{bcd}{a},$$

and after simplifying

$$\overline{x}=\frac{x-\dfrac{1}{2}\left(a+b+c+d-\dfrac{abc+acd+abd+bcd}{a^2}\right)}{bcd}a.$$

Similarly from the colinearity of the points x, b_1, and b_2 we get

$$\overline{x}=\frac{x-\dfrac{1}{2}\left(a+b+c+d-\dfrac{abc+acd+abd+bcd}{b^2}\right)}{acd}b,$$

and from this we conclude

$$x=\frac{1}{2}\Big(a+b+c+d\Big).$$

Let $h=a+c+d$ (by T6) be the orthocenter of the triangle acd. In order to finish the proof, according to T1.2 it is enough to show that

$$\frac{x-c}{\overline{x}-\overline{c}}=\frac{h-c}{\overline{h}-\overline{c}}=\frac{a+b+d-c}{\overline{a}+\overline{b}+\overline{d}-\overline{c}}.$$

On the other hand $x-c=\frac{1}{2}\Big(a+b+d-c\Big)$, from which the equality is obvious.

21. Using the last problem we have that the intersection of the lines $l(a;bcd)$ and $l(b;cda)$ is the point $x=\frac{1}{2}\Big(a+b+c+d\Big)$, which is a symmetric expression, hence this point is the intersection of every two of the given lines.

22. Using the last two problems we get the locus of points is the set of all the points of the form $x=\frac{1}{2}\Big(a+b+c+d\Big)$, when d moves along the circle. That is in fact the circle with the radius $\frac{1}{2}$ and center $\frac{a+b+c}{2}$, which is the midpoint of the segment connecting the center of the given circle with the orthocenter of the triangle abc.

23. Assume that the unit circle is the circumcircle of the triangle abc. From T1.3 and the condition $ad\perp ao$ we have that

$$\frac{d-a}{\overline{d}-\overline{a}}=-\frac{a-o}{\overline{a}-\overline{o}}=-a^2,$$

and after simplifying $\overline{d}=\frac{2a-d}{a^2}$. Since the points b,c,d are colinear and bc is the chord of the unit circle, according to T2.2 $\overline{d}=\frac{b+c-d}{bc}$, and solving the given system we get $d=\frac{a^2(b+c)-2abc}{a^2-bc}$. Since e belongs to the perpendicular bisector of ab we have $oe\perp ab$. According to T1.3 and $\frac{e-o}{\overline{e}-\overline{o}}=-\frac{a-b}{\overline{a}-\overline{b}}=ab$, i.e. $\overline{e}=\frac{e}{ab}$. From $be\perp bc$, using T1.3 again we get $\frac{b-e}{\overline{b}-\overline{e}}=-\frac{b-c}{\overline{b}-\overline{c}}=bc$, or equivalently $\overline{e}=\frac{c-b+e}{bc}=\frac{e}{ab}$. Hence $e=\frac{a(c-b)}{c-a}$. Similarly we have $f=\frac{a(b-c)}{b-a}$. Using T1.2 we see that it is enough to prove that $\frac{d-f}{\overline{d}-\overline{f}}=\frac{f-e}{\overline{f}-\overline{e}}$. Notice that

$$\begin{aligned} d-f &= \frac{a^2(b+c)-2abc}{a^2-bc}-\frac{a(b-c)}{b-a}=\frac{a^2b^2+3a^2bc-ab^2c-2a^3b-abc^2}{(a^2-bc)(b-a)} \\ &= \frac{ab(a-c)(b+c-2a)}{(a^2-bc)(b-a)}, \end{aligned}$$

and similarly $d-e=\frac{ac(a-b)(b+c-2a)}{(a^2-bc)(c-a)}$. After conjugation we see that the required condition is easy to verify.

24. [Obtained from Uroš Rajković] Assume that the unit circle is the incircle of the hexagon $ABCDEF$. After conjugating and using T2.5 we get:

$$\overline{m}=\frac{a+b-(d+e)}{ab-de},\ \overline{n}=\frac{b+c-(e+f)}{bc-ef},\ \overline{p}=\frac{c+d-(f+a)}{cd-fa},$$

hence:

$$\overline{m}-\overline{n}=\frac{(b-e)(bc-cd+de-ef+fa-ab)}{(ab-de)(bc-ef)},$$

and analogously:

$$\overline{n}-\overline{p}=\frac{(c-f)(cd-de+ef-fa+ab-bc)}{(bc-ef)(cd-fa)}.$$

From here we get:

$$\frac{\overline{m}-\overline{n}}{\overline{n}-\overline{p}}=\frac{(b-e)(cd-fa)}{(f-c)(ab-de)}.$$

Since the numbers $\overline{a}$, $\overline{b}$, $\overline{c}$, $\overline{d}$, $\overline{e}$, and $\overline{f}$ are equal to $\frac{1}{a}$, $\frac{1}{b}$, $\frac{1}{c}$, $\frac{1}{d}$, $\frac{1}{e}$, and $\frac{1}{f}$, respectively, we see that it is easy to verify that the complex number on the left-hand side of the last equality equal to its complex conjugate, hence it is real. Now according to T1.2 the points M, N, and P are colinear, q.e.d.

25. Assume that the quadrilateral $abcd$ is inscribed in the unit circle. Using T2.5 we get

$$\begin{aligned} e &= \frac{ab(c+d)-cd(a+b)}{ab-cd}, \\ f &= \frac{ad(b+c)-bc(a+d)}{ad-bc}mboxand \\ g &= \frac{ac(b+d)-bd(a+c)}{ac-bd}. \quad (1) \end{aligned}$$

In order to prove that $o=0$ is the orthocenter of the triangle efg, it is enough to prove that $of \perp eg$ and $og \perp ef$. Because of the symmetry it is enough to prove one of these two relateions. Hence, by T1.3 it is enough to prove that

$$\frac{f-o}{\overline{f}-\overline{o}}=\frac{e-g}{\overline{e}-\overline{g}} \quad (2).$$

From (1) we have that

$$\frac{f-o}{\overline{f}-\overline{o}}=\frac{\frac{ad(b+c)-bc(a+d)}{ad-bc}}{\frac{(b+c)-(a+d)}{bc-ad}}=\frac{ad(b+c)-bc(a+d)}{a+d-(b+c)}, (3)$$

or equivalently

$$\begin{aligned} e-g &= \frac{(a-d)(ab^2d-ac^2d)+(b-c)(bcd^2-a^2bc)}{(ab-cd)(ac-bd)} \\ &= \frac{(a-d)(b-c)((b+c)ad-(a+d)bc)}{(ab-cd)(ac-bd)} \quad (4) \end{aligned}$$

and by conjugation

$$\overline{e}-\overline{g}=\frac{(a-d)(b-c)(b+c-(a+d))}{(ab-cd)(ac-bd)} \quad (5).$$

Comparing the expressions (3),(4), and (5) we derive the statement.

26. Assume that the unit circle is the circumcircle of the triangle abc and assume that $a=1$. Then $c=\overline{b}$ and $t=-1$. Since p belongs to the chord bc, using T2.2 we get that $\overline{p}=b+\frac{1}{b}-p$. Since x belongs to the chord ab, in the similar way we get $\overline{x}=\frac{1+b-x}{b}$. Since $px \parallel ac$ by T1.1 we have

$$\frac{p-x}{\overline{p}-\overline{x}}=\frac{a-c}{\overline{a}-\overline{c}}=-\frac{1}{b},$$

i.e. $\overline{x} = pb + \overline{p} - xb$. From this we get $x = \dfrac{b(p+1)}{b+1}$. Similarly we derive $y = \dfrac{p+1}{b+1}$. According to T1.3 it remains to prove that $\dfrac{x-y}{\overline{x}-\overline{y}} = -\dfrac{p-t}{\overline{p}-\overline{t}} = -\dfrac{p+1}{\overline{p}+1}$. This follows from $x-y = \dfrac{(p+1)(b-1)}{b+1}$ and by conjugation

$$\overline{x}-\overline{y} = \frac{(\overline{p}+1)\left(\frac{1}{b}-1\right)}{\frac{1}{b}+1} = -\frac{(\overline{p}+1)(b-1)}{b+1}.$$

27. Assume that the unit circle is the circumcircle of the quadrilateral $abcd$. Using T6.1 we have $k = \dfrac{a+b}{2}$, $l = \dfrac{b+c}{2}$, $m = \dfrac{c+a}{2}$ and $n = \dfrac{d+a}{2}$. We want to determine the coordinate of the orthocenter of the triangle akn. Let h_1 be that point and denote by h_2, h_3, and h_4 the orthocenters of bkl, clm, and dmn respectively. Then $kh_1 \perp an$ and $nh_1 \perp ak$. By T1.3 we get

$$\frac{k-h_1}{\overline{k}-\overline{h_1}} = -\frac{a-n}{\overline{a}-\overline{n}} \text{ and } \frac{n-h_1}{\overline{n}-\overline{h_1}} = -\frac{a-k}{\overline{a}-\overline{k}}. \quad (1)$$

Since

$$\frac{a-n}{\overline{a}-\overline{n}} = \frac{a-d}{\overline{a}-\overline{d}} = -ad,$$

we have that

$$\overline{h_1} = \frac{\overline{k}ad - k + h_1}{ad}.$$

Similarly from the second of the equations in (1) we get

$$\overline{h_1} = \frac{\overline{n}ab - n + h_1}{ab}.$$

Solving this system gives us that

$$h_1 = \frac{2a+b+d}{2}.$$

Symmetricaly

$$h_2 = \frac{2b+c+a}{2},\ h_3 = \frac{2c+d+b}{2},\ h_4 = \frac{2d+a+c}{2},$$

and since $h_1 + h_3 = h_2 + h_4$ using T6.1 the midpoints of the segments h_1h_3 and h_2h_4 coincide hence the quadrilateral $h_1h_2h_3h_4$ is a parallelogram.

28. Assume that the unit circle is the circumcircle of the triangle abc. By T2.3 we have that $a = \dfrac{2em}{e+m}$ i $b = \dfrac{2mk}{m+k}$. Let's find the point p. Since the points m, k, and p are colinear and mk is the chord of the unit circle, by T2.2 we have that $\overline{p} = \dfrac{m+k-p}{mk}$. Furthermore the points p, e, and c are colinear. However, in this problem it is more convenient to notice that $pe \perp oe$ and now using T1.3 we have

$$\frac{e-p}{\overline{e}-\overline{p}} = -\frac{e-o}{\overline{e}-\overline{o}} = -e^2$$

and after simplifying $\overline{p} = \dfrac{2e-p}{e^2}$. Equating the two expressions for $\overline{p}$ we get

$$p = e\frac{(m+k)e - 2mk}{e^2 - mk}.$$

In order to finish the proof using T1.3 it is enough to prove that $\frac{p-o}{\overline{p}-\overline{o}} = -\frac{e-b}{\overline{e}-\overline{b}}$. This will follow from

$$e-b = \frac{e(m+k)-2mk}{m+k},$$

and after conjugating $\overline{e}-\overline{b} = \frac{m+k-2e}{(m+k)e}$ and $\overline{p} = \frac{m+k-2e}{mk-e^2}$.

29. Assume that the circle inscribed in *abcd* is the unit one. From T2.3 we have that

$$a = \frac{2nk}{n+k}, \quad b = \frac{2kl}{k+l}, \quad c = \frac{2lm}{l+m}, \quad d = \frac{2mn}{m+n}. \quad (1)$$

Using T2.5 we get

$$s = \frac{kl(m+n)-mn(k+l)}{kl-mn}. \quad (2)$$

According to T1.1 it is enough to verify that

$$\frac{s-o}{\overline{s}-\overline{o}} = \frac{b-d}{\overline{b}-\overline{d}}.$$

From (1) we have that

$$b-d = 2\frac{kl(m+n)-mn(k+l)}{(k+l)(m+n)}, \quad (3)$$

and after conjugating

$$\overline{b}-\overline{d} = \frac{m+n-(k+l)}{(k+l)(m+n)}. \quad (4)$$

From (2) we have that

$$\frac{s}{\overline{s}} = \frac{kl(m+n)-mn(k+l)}{kl-mn}, \quad (5)$$

and comparing the expressions (3),(4), and (5) we finish the proof.

30. [Obtained from Uroš Rajković] Let P be the point of tangency of the incircle with the line BC. Assume that the incircle is the unit circle. By T2.3 the coordinates of A, B, and C are respectively

$$a = \frac{2qr}{q+r}, \quad b = \frac{2pr}{p+r} \text{ i } c = \frac{2pq}{p+q}.$$

Furthermore, using T6.1 we get $x = \frac{1}{2}(b+c) = \frac{pr}{p+r} + \frac{pq}{p+q}$, $y = \alpha b = \alpha\frac{2pr}{p+r}$, and $z = \beta c = \beta\frac{2pq}{p+q}$ $(\alpha,\beta \in R)$. The values of α and β are easy to compute from the conditions $y \in rq$ and $z \in rq$:

$$\alpha = \frac{(p+r)(q+r)}{2(p+q)r} \text{ i } \beta = \frac{(p+q)(r+q)}{2(p+r)q}.$$

From here we get the coordinates of y and z using p, q, and r:

$$y = \frac{p(q+r)}{(p+q)} \text{ and } z = \frac{p(r+q)}{(p+r)}.$$

We have to prove that:

$$\angle RAQ = 60^\circ \Longleftrightarrow XYZ \text{ is equilateral.}$$

The first condition is equivalent to $\angle QOR = 60^\circ$ i.e. with

$$r = q \cdot e^{i2\pi/3}.$$

The second condition is equivalent to $(z-x) = (y-x) \cdot e^{i\pi/3}$. Notice that:

$$y - x = \frac{p(q+r)}{(p+q)} - \left(\frac{pr}{p+r} + \frac{pq}{p+q}\right) = \frac{pr(r-q)}{(p+q)(p+r)} \text{ and}$$

$$z - x = \frac{p(p+q)}{(p+r)} - \left(\frac{pr}{p+r} + \frac{pq}{p+q}\right) = \frac{pq(q-r)}{(p+q)(p+r)}.$$

Now the second condition is equivalent to:

$$\frac{pq(q-r)}{(p+q)(p+r)} = \frac{pr(r-q)}{(p+q)(p+r)}\, e^{i\pi/3},$$

i.e. with $q = -r\, e^{i\pi/3}$. It remains to prove the equivalence:

$$r = q\, e^{i2\pi/3} \Longleftrightarrow q = -r\, e^{i\pi/3},$$

which obviously holds.

31. According to T1.1 it is enough to prove that

$$\frac{m-o}{\overline{m}-\overline{o}} = \frac{n-o}{\overline{n}-\overline{o}}.$$

If p,q,r,s are the points of tangency of the incircle with the sides ab, bc, cd, da respectively using T2.3 we get

$$m = \frac{a+c}{2} = \frac{ps}{p+s} + \frac{qr}{q+r} = \frac{pqs+prs+pqr+qrs}{(p+s)(q+r)},$$

and after conjugating $\overline{m} = \dfrac{p+q+r+s}{(p+s)(q+r)}$ and

$$\frac{m}{\overline{m}} = \frac{pqr+ps+prs+qrs}{p+q+r+s}.$$

Since the last expression is symmetric in p,q,r,s we conclude that $\dfrac{m}{\overline{m}} = \dfrac{n}{\overline{n}}$, as required.

32. Assume that the incircle of the quadrilateral $abcd$ is the unit circle. We will prove that the intersection of the lines mp and nq belongs to bd. Then we can conlude by symmetry that the point also belongs to ac, which will imply that the lines mp, nq, ac, and bd are concurrent. Using T2.3 we have that

$$b = \frac{2mn}{m+n}, \quad d = \frac{2pq}{p+q}.$$

If x is the intersection point of mp and nq, using T2.5 we get

$$x = \frac{mp(n+q) - nq(m+p)}{mp - nq}.$$

We have to prove that the points x, b, d are colinear, which is according to T1.2 equivalent to saying that

$$\frac{b-d}{\overline{b}-\overline{d}} = \frac{b-x}{\overline{b}-\overline{x}}.$$

This follows from $b-d=\dfrac{2mn}{m+n}-\dfrac{2pq}{p+q}=2\dfrac{mn(p+q)-pq(m+n)}{(m+n)(p+q)}$ and

$$\begin{aligned} b-x &= \frac{2mn}{m+n}-\frac{mp(n+q)-nq(m+p)}{mp-nq} \\ &= \frac{m^2np-mn^2q-m^2pq+n^2pq+m^2nq-mn^2p}{(mp-nq)(m+n)} \\ &= \frac{(m-n)(mn(p+q)-pq(m+n))}{(m+n)(mp-nq)}, \end{aligned}$$

by conjugation.

33. Assume that the unit circle is the incumcircle of the triangle abc. Using T7.3 we have that the circumecenter has the coordinate

$$o=\frac{2def(d+e+f)}{(d+e)(e+f)(f+d)}.$$

Let's calculate the coordinate of the circumcenter o_1 of the triangle xyz. First, according to T6.1 we have that $x=\dfrac{e+f}{2}$, $y=\dfrac{d+f}{2}$ and $z=\dfrac{d+e}{2}$. Moreover by T1.3 we have that $\dfrac{o_1-\dfrac{x+y}{2}}{\overline{o_1}-\overline{\dfrac{x+y}{2}}}=-\dfrac{x-y}{\overline{x}-\overline{y}}=\dfrac{(e-d)/2}{(\overline{e}-\overline{d})/2}=-ed$, and simplifying

$$\overline{o_1}=\frac{-\dfrac{f}{2}+\dfrac{ed}{2f}+o_1}{ed},$$

and similarly $\overline{o_1}=\dfrac{-\dfrac{d}{2}+\dfrac{ef}{2d}+o_1}{ef}$. By equating we get $o_1=\dfrac{e+f+d}{2}$. Now by T1.2 it is enough to prove that $\dfrac{o_1-i}{\overline{o_1}-\overline{i}}=\dfrac{o-i}{\overline{o}-\overline{i}}$, which can be easily obtained by conjugation of the previous expressions for o and o_1.

34. Assume that the incircle of the triangle abc is the unit circle. Using T7.1 we get $b=\dfrac{2fd}{f+d}$ and $c=\dfrac{2ed}{e+d}$. From some elemetary geometry we conclude that k is the midpoint of segment ef hence by T6.1 we have $k=\dfrac{e+f}{2}$. Let's calculate the coordinate of the point m. Since m belongs to the chord fd by T2.2 we have $\overline{m}=\dfrac{f+d-m}{fd}$. Similarly we have that the points b,m,k are colinear and by T1.2 we get $\dfrac{k-m}{\overline{k}-\overline{m}}=\dfrac{b-k}{\overline{b}-\overline{k}}$, i.e. $\overline{m}=m\dfrac{\overline{b}-\overline{k}}{b-k}+\dfrac{\overline{k}b-k\overline{b}}{b-k}$. Now equating the expressions for $\overline{m}$ one gets

$$m=\frac{(f+d)(b-k)+(k\overline{b}-\overline{k}b)fd}{(\overline{b}-\overline{k})fd+b-k}.$$

Since $b-k=\dfrac{3fd-de-f^2-ef}{2(f+d)}$ and $k\overline{b}-\overline{k}b=\dfrac{(e+f)(e-d)fd}{e(f+d)}$ we get

$$m=\frac{4ef^2d+efd^2-e^2d^2-e^2f^2-2f^2d^2-f^3e}{6efd-e^2d-ed^2-ef^2-e^2f-d^2f-df^2}$$

and symmetrically

$$n=\frac{4e^2fd+efd^2-f^2d^2-e^2f^2-2e^2d^2-e^3f}{6efd-e^2d-ed^2-ef^2-e^2f-d^2f-df^2}.$$

By T1.3 it is enough to prove that $\frac{m-n}{\overline{m}-\overline{n}}=-\frac{i-d}{\overline{i}-\overline{d}}=-d^2$. This however follows from

$$m-n=\frac{(e-f)(4efd-ed^2-fd^2-fe^2-f^2e)}{6efd-e^2d-ed^2-ef^2-e^2f-d^2f-df^2},$$

by conjugation.

35. Assume that the unit circle is the inrcumcircle of the triangle abc. Assume that k, l, and m are the points of tangency of the incircle with the sides bc, ca, and ab, respectively. By T7 we have that

$$o=\frac{2klm(k+l+m)}{(k+l)(l+m)(m+k)},\quad h=\frac{2(k^2l^2+l^2m^2+m^2k^2+klm(k+l+m))}{(k+l)(l+m)(m+k)}.$$

Since the segments io and bc are parallel we have that $io\perp ik$, which is by T1.3 equivalent to $\frac{o-i}{\overline{o}-\overline{i}}=-\frac{k-i}{\overline{k}-\overline{i}}=-k^2$. After conjugating the last expression for o becomes

$$klm(k+l+m)+k^2(kl+lm+mk)=0.\quad(*)$$

Let's prove that under this condition we have $ao\parallel hk$. According to T1.1 it is enough to prove that $\frac{a-o}{\overline{a}-\overline{o}}=\frac{h-k}{\overline{h}-\overline{k}}$. According to T7.1 we have that $a=\frac{2ml}{m+l}$, and

$$a-o=\frac{2ml}{m+l}-\frac{2klm(k+l+m)}{(k+l)(l+m)(m+k)}=\frac{2m^2l^2}{(k+l)(l+m)(m+k)}.$$

Now we get that it is enough to prove that

$$\frac{h-k}{\overline{h}-\overline{k}}=\frac{l^2m^2}{k^2}.$$

Notice that

$$\begin{aligned}
h-k &= \frac{2(k^2l^2+l^2m^2+m^2k^2+klm(k+l+m))}{(k+l)(l+m)(m+k)}-k\\
&= \frac{k^2l^2+k^2m^2+2l^2m^2+k^2lm+kl^2m+klm^2-k^2l-k^3m-k^2lm}{(k+l)(l+m)(m+k)}\\
&= \frac{klm(k+l+m)-k^2(k+l+m)+k^2l^2+2l^2m^2+m^2l^2}{(k+l)(l+m)(m+k)}\\
&= \Big(\text{according to (*)}\Big)=\frac{(kl+lm+mk)^2+l^2m^2}{(k+l)(l+m)(m+k)}\\
&= \Big(\text{according to (*)}\Big)=\frac{(kl+lm+mk)^2((k+l+m)^2+k^2)}{(k+l+m)^2(k+l)(l+m)(m+k)}.
\end{aligned}$$

After conjugating the last expression for $h-k$ we get

$$\overline{h}-\overline{k}=\frac{(k+l+m)^2+k^2}{(k+l)(l+m)(m+k)},$$

and using the last expression for $h-k$ we get

$$\frac{h-k}{\overline{h}-\overline{k}}=\frac{(kl+lm+mk)^2}{(k+l+m)^2}=\Big(\text{by (*)}\Big)=\frac{l^2m^2}{k^2},$$

which completes the proof.

36. Assume that the incircle of the triangle abc is the unit circle. Then using T7.1 we have $c=\dfrac{2t_1t_2}{t_1+t_2}$. Our goal is to first determine the point h_3. From $h_3t_3\perp it_3$ by T1.3 we have

$$\frac{h_3-t_3}{\overline{h_3}-\overline{t_3}}=-\frac{t_3-i}{\overline{t_3}-\overline{i}}=-t_3^2,$$

i.e. $\overline{h_3}=\dfrac{2t_3-h_3}{t_3^2}$. Furthermore from $ch_3\parallel it_3$ and T1.1 we have $\dfrac{h_3-c}{\overline{h_3}-\overline{c}}=\dfrac{t_3-i}{\overline{t_3}-\overline{i}}=t_3^2$. Writing the similar expression for $\overline{h_3}$ gives

$$h_3=\frac{1}{2}\Big(2t_3+c-\overline{c}t_3^2\Big)=t_3+\frac{t_1t_2-t_3^2}{t_1+t_2}.$$

Similarly we obtain $h_2=t_2+\dfrac{t_1t_3-t_2^2}{t_1+t_3}$. In order to determine the line symmetric to h_2h_3 with repsect to t_2t_3 it is enough to determine the points symmetric to h_2 and h_3 with respect to t_2t_3. Assume that p_2 and p_3 are these two points and let h_2' and h_3' be the feet of perpendiculars from h_2 and h_3 to the line t_2t_3 respectively. According to T2.4 we have $h_2'=\dfrac{1}{2}\Big(t_2+t_3-t_2t_3\overline{h_3}\Big)$ hence by T6.1

$$p_2=2h_2'-h_2=\frac{t_1(t_2^2+t_3^2)}{t_2(t_1+t_3)}$$

and symmetrically $p_3=\dfrac{t_1(t_2^2+t_3^2)}{t_3(t_1+t_2)}$. Furthermore

$$p_2-p_3=\frac{t_1^2(t_2^2+t_3^2)(t_3-t_2)}{t_1t_3(t_1+t_2)(t_1+t_3)},$$

and if the point x belongs to p_2p_3 by T1.2 the following must be satisfied:

$$\frac{x-p_2}{\overline{x}-\overline{p_2}}=\frac{p_2-p_3}{\overline{p_2}-\overline{p_3}}=-t_1^2.$$

Specifically if x belongs to the unit circle we also have $\overline{x}=\dfrac{1}{x}$, hence we get the quadratic equation

$$t_2t_3x^2-t_1(t_2^2+t_3^2)x+t_1^2t_2t_3=0.$$

Its solutions are $x_1=\dfrac{t_1t_2}{t_3}$ and $x_2=\dfrac{t_1t_3}{t_2}$ and these are the intersection points of the line p_2p_3 with the unit circle. Similarly we get $y_1=\dfrac{t_1t_2}{t_3}$, $y_2=\dfrac{t_2t_3}{t_1}$, and $z_1=\dfrac{t_3t_1}{t_2}$, $z_2=\dfrac{t_2t_3}{t_1}$, which finishes the proof.

37. Assume that the circumcircle of the triangle abc is the unit circle. Let u,v,w be the complex numbers described in T8. Using this theorem we get that $l=-(uv+vw+wu)$. By elementary geometry we know that the intersection of the line al and the circumcircle of the triangle abc is the midpoint of the arc bc which doesn't contain the point a. That means $a_1=-vw$ and similarly $b_1=-uw$ and $c_1=-uv$.

(a) The statement follows from the equality

$$1=\frac{|l-a_1|\cdot|l-c_1|}{|l-b|}=\frac{|u(v+w)|\cdot|w(u+v)|}{|uv+uw+vw+v^2|}=\frac{|v+w|\cdot|u+v|}{|(u+v)(v+w)|}=1.$$

(b) If x is the point of the tangency of the incircle with the side bc then x is the foot of the perpendicular from the point l l to the side bc and T2.4 implies $x=\frac{1}{2}\left(b+c+l-bc\bar{l}\right)$ and consequently $r=|l-x|=\frac{1}{2}\left|\frac{(u+v)(v+w)(w+u)}{u}\right|=\frac{1}{2}|(u+v)(v+w)(w+u)|$. Now the required equality follows from

$$\begin{aligned}\frac{|l-a|\cdot|l-b|}{|l-c_1|} &= \frac{|(u+v)(u+w)|\cdot|(u+v)(v+w)|}{|w(u+v)|}\\ &= |(u+v)(v+w)(w+u)|.\end{aligned}$$

(c) By T5 we have that

$$S(ABC)=\frac{i}{4}\begin{vmatrix}u^2 & 1/u^2 & 1\\ v^2 & 1/v^2 & 1\\ w^2 & 1/w^2 & 1\end{vmatrix}\ \text{i}\ S(A_1B_1C_1)=\frac{i}{4uvw}\begin{vmatrix}vw & u & 1\\ uw & v & 1\\ uv & w & 1\end{vmatrix},$$

hence

$$\begin{aligned}\frac{S(ABC)}{S(A_1B_1C_1)} &= \frac{u^4w^2+w^4v^2+v^4u^2-v^4w^2-u^4v^2-w^4u^2}{uvw(v^2w+uw^2+u^2v-uv^2-u^2w-vw^2)}\\ &= \frac{(u^2-v^2)(uw+vw-uv-w^2)(uw+vw+uv+w^2)}{uvw(u-v)(uv+w^2-uw-vw)}\\ &= -\frac{(u+w)(vw+uw+uv+w^2)}{uvw}\\ &= -\frac{(u+v)(v+w)(w+u)}{uvw}.\end{aligned}$$

Here we consider the oriented surface areas, and substracting the modulus from the last expression gives us the desired equality.

38. First solution. Assume that the circumcircle of the triangle abc is the unit circle and u,v,w are the complex numbers described in T8. Let d,e,f be the points of tangency of the incircle with the sides bc,ca,ab respectively. By T2.4 we have that $f=\frac{1}{2}\left(a+b+z-ab\bar{z}\right)=\frac{1}{2}\Big(u^2+v^2+w^2-uv-vw-wu+\frac{uv(u+v)}{2w}\Big)$. By symmetry we get the expressions for e and f and by T6.1 we get

$$k=\frac{1}{3}\Big(u^2+v^2+w^2-uv-vw-wu+\frac{uv(u+v)}{2w}+\frac{vw(v+w)}{2u}-\frac{wu(w+u)}{2v}\Big)=$$
$$=\frac{(uv+vw+wu)(u^2v+uv^2+uw^2+u^2w+v^2w+vw^2-4uvw)}{6uvw}.$$

Now it is easy to verify $\frac{z-o}{\bar{z}-\bar{o}}=\frac{k-o}{\bar{k}-\bar{o}}$, which is by T1.2 the condition for colinearity of the points z,k,o. Similarly we also have

$$\begin{aligned}\frac{|o-z|}{|z-k|} &= \frac{|uv+vw+wu|}{\left|\frac{(uv+vw+wu)(u^2v+uv^2+uw^2+u^2w+v^2w+vw^2+2uvw)}{6uvw}\right|}\\ &= \frac{6}{|(u+v)(v+w)(w+u)|}=\frac{6R}{2r}=\frac{3R}{r},\end{aligned}$$

which completes the proof.

Second solution. Assume that the incircle of the triangle abc is the unit circle and let d,e,f denote its points of tangency with the sides bc,ca,ab respectively. According to T7.3 we have that $o=\dfrac{2def(d+e+f)}{(d+e)(e+f)(f+d)}$ and according to T6.1 $k=\dfrac{d+e+f}{3}$. Now it is easy to verify that $\dfrac{o-z}{\overline{o}-\overline{z}}=\dfrac{k-z}{\overline{k}-\overline{z}}$ which is by T1.2 enough to establish the colinearity of the points o,z,k. We also have that

$$\frac{|o-z|}{|z-k|}=\frac{\left|\dfrac{d+e+f}{(d+e)(e+f)(f+d)}\right|}{\left|\dfrac{d+e+f}{3}\right|}=\frac{3}{|(d+e)(e+f)(f+d)|}=\frac{3R}{r}.$$

39. Assume that the circumcircle of the triangle abc is the unit circle and let u,v,w be the complex numbers described in T8 (here $p=w^2$). According to this theorem we have $i=-uv-vw-wu$. Since $|a-c|=|a-b|$ by T1.4 it holds

$$c-a=e^{i\angle cab}(b-a).$$

By the same theorem we have

$$\frac{-vw-u^2}{\overline{-vw}-\overline{u^2}}=e^{i2\frac{\angle pab}{2}}\frac{v^2-u^2}{\overline{v^2}-\overline{u^2}},$$

hence $e^{i\angle pab}=-\dfrac{w}{v}$. Now we have

$$c=\frac{u^2w+u^2v-v^2w}{v},$$

and symmetrically $d=\dfrac{v^2w+v^2u-u^2w}{u}$. By T1.3 it is enough to prove that

$$\frac{c-d}{\overline{c}-\overline{d}}=-\frac{o-i}{\overline{o}-\overline{i}}=-\frac{uv+vw+wu}{u+v+w}uvw.$$

This follows from $c-d=\dfrac{(u^2-v^2)(uv+vw+wu)}{uv}$ by conjugation.

40. Assume that the circumcircle of the triangle abc is the unit circle. By T8 there are numbers u,v,w such that $a=u^2,b=v^2,c=w^2$ and the incenter is $i=-(uv+vw+wu)$. If o' denotes the foot of the perpendicular from o to bc then by T2.4 we have $o'=\dfrac{1}{2}\Big(b+c\Big)$, and by T6.1 $o_1=2o'=b+c=v^2+w^2$. By T1.2 the points a,i,o_1 are colinear if and only if

$$\frac{o_1-a}{\overline{o_1}-\overline{a}}=\frac{a-i}{\overline{a}-\overline{i}}.$$

Since

$$\frac{o_1-a}{\overline{o_1}-\overline{a}}=\frac{o_1-a}{\overline{o_1}-\overline{a}}=\frac{v^2+w^2-u^2}{u^2(v^2+w^2)-v^2w^2}u^2v^2w^2 \text{ and}$$

$$\frac{a-i}{\overline{a}-\overline{i}}=\frac{u(u+v+w)+vw}{vw+uw+uv+u^2}u^2vw=u^2vw,$$

we get

$$v^3w+vw^3-u^2vw-(u^2v^2+u^2w^2-v^2w^2)=(vw-u^2)(v^2+w^2+vw)=0.$$

This means that either $vw = u^2$ or $v^2 + w^2 + vw = 0$. If $vw = u^2$ then by T6.1 the points u^2 and $-vw$ belong to the same radius hence abc is isosceles contrary to the assumption. This means that $v^2 + w^2 + vw = 0$. We now want to prove that the triangle with the vertices $o, -vw, w^2$ is equilateral. It is enough to prove that $1 = |w^2 + vw| = |v + w|$ which is equivalent to $1 = (v+w)(\overline{v} + \overline{w}) = \dfrac{(v+w)^2}{vw}$ and this to $v^2 + w^2 + vw = 0$. Since $\angle boc = 120^\circ$ we have $\alpha = 60^\circ$.

41. Assume that the incumcircle of the triangle abc is the unit circle. According to T8 there are complex numbers u, v, w such that $p = u^2, q = v^2, r = w^2$ and $p_1 = -vw, q_1 = -wu, r_1 = -uv$. Then $p_2 = vw, q_2 = wu, r_2 = uv$. By T7.1 we gave

$$a = \frac{2v^2w^2}{v^2+w^2},\ b = \frac{2w^2u^2}{w^2+u^2} \text{ i } c = \frac{2u^2w^2}{u^2+w^2},$$

hence by T6.1

$$a_1 = \frac{w^2u^2}{w^2+u^2} + \frac{u^2v^2}{u^2+v^2},\ b_1 = \frac{u^2v^2}{u^2+v^2} + \frac{v^2w^2}{v^2+w^2},\ c_2 = \frac{v^2w^2}{v^2+w^2} + \frac{w^2u^2}{w^2+u^2}.$$

If the point n is the intersection of the lines a_1p_1 and b_1q_1 then the triplets of points (n, a_1, p_1) and (n, b_1, q_1) are colinear and using T1.2 we get

$$\frac{n-a_1}{\overline{n}-\overline{a_1}} = \frac{a_1-p_1}{\overline{a_1}-\overline{p_1}}, \quad \frac{n-b_1}{\overline{n}-\overline{b_1}} = \frac{b_1-q_1}{\overline{b_1}-\overline{q_1}}.$$

Solving this system gives us

$$\begin{aligned} n \quad = \quad & \frac{u^4v^4+v^4w^4+w^4u^4}{(u^2+v^2)(v^2+w^2)(w^2+u^2)} + \\ & \frac{uvw(u^3v^2+u^2v^3+u^3w^2+u^2w^3+v^3w^2+v^2w^3)}{(u^2+v^2)(v^2+w^2)(w^2+u^2)} + \\ & \frac{3u^2v^2w^2(u^2+v^2+w^2)}{(u^2+v^2)(v^2+w^2)(w^2+u^2)} + \\ & \frac{2u^2v^2w^2(uv+vw+wu)}{(u^2+v^2)(v^2+w^2)(w^2+u^2)}. \end{aligned}$$

Since the above expression is symmetric this point belongs to c_1r_1. The second part of the problem can be solved similarly.

42. Assume that a is the origin. According to T1.4 we have $c'' - a = e^{i\pi/2}(c-a)$, i.e. $c'' = ic$. Similarly we get $b'' = -ib$. Using the same theorem we obtain $x - c = e^{i\pi/2}(b-c)$, i.e. $x = (1-i)c + ib$ hence by T6.1 $p = \dfrac{1+i}{2}b + \dfrac{1-i}{2}c$. Denote by q the intersection of the lines bc and ap. Then the points a, p, q are colinear as well as the points b, c'', q. Using T1.2 we get

$$\frac{a-p}{\overline{a}-\overline{p}} = \frac{a-q}{\overline{a}-\overline{q}}, \quad \frac{b-c''}{\overline{b}-\overline{c''}} = \frac{q-b}{\overline{q}-\overline{b}}.$$

From the first equation we conclude that $\overline{q} = q\dfrac{(1-i)\overline{b} + (1+i)\overline{c}}{(1+i)b + (1-i)c}$, and from the second we get the formula $\overline{q} = \dfrac{q(\overline{b} + i\overline{c}) - i(\overline{b}c + b\overline{c})}{b - ic}$. These two imply

$$q = \frac{i(\overline{b}c + b\overline{c})((1+i)b + (1-i)c)}{2(ib\overline{b} - 2b\overline{c} + 2\overline{b}c + 2ic\overline{c})} = \frac{(\overline{b}c + b\overline{c})((1+i)b + (1-i)c)}{(b-ic)(\overline{b} + i\overline{c})}.$$

Denote by q' the intersection of ap and cb''. Then the points a, p, q' are colinear as well as the points b'', c, q'. Hence by T1.2

$$\frac{a-p}{\overline{a}-\overline{p}} = \frac{a-q'}{\overline{a}-\overline{q'}}, \quad \frac{b''-c}{\overline{b''}-\overline{c}} = \frac{q-c}{\overline{q}-\overline{c}}.$$

The first equation gives $\overline{q'} = q' \dfrac{(1-i)\overline{b}+(1+i)\overline{c}}{(1+i)b+(1-i)c}$, and the second $\overline{q} = \dfrac{q(\overline{c}-i\overline{b})+i(\overline{b}c+b\overline{c})}{c+ib}$. By the equating we get

$$q' = \frac{(\overline{b}c+b\overline{c})((1+i)b+(1-i)c)}{(b-ic)(\overline{b}+i\overline{c})},$$

hence $q = q'$, q.e.d.

43. Assume that the origin is the intersection of the diagonals, i.e. $o = 0$. From the colinearity of a, o, c and b, o, d using T1.2 we get $a\overline{c} = \overline{a}c$ and $b\overline{d} = \overline{b}d$. By T6.1 we get $m = \dfrac{a+b}{2}$ and $n = \dfrac{c+d}{2}$. Since $om \perp cd$ and $on \perp ab$ by T1.3

$$\frac{\frac{c+d}{2}-o}{\frac{\overline{c+d}}{2}-\overline{o}} = -\frac{a-b}{\overline{a}-\overline{b}}, \quad \frac{\frac{a+b}{2}-o}{\frac{\overline{a+b}}{2}-\overline{o}} = -\frac{c-d}{\overline{c}-\overline{d}}.$$

From these two equations we get

$$c = \frac{da(\overline{a}b-2b\overline{b}+a\overline{b})}{b(\overline{a}b-2a\overline{a}+a\overline{b})} \text{ and } c = \frac{da(\overline{a}b+2b\overline{b}+a\overline{b})}{b(\overline{a}b+2a\overline{a}+a\overline{b})}.$$

The last two expressions give $(\overline{a}b+a\overline{b})(a\overline{a}-b\overline{b}) = 0$. We need to prove that the last condition is sufficient to guarantee that a, b, c, d belong to a circle. According to T3 the last is equivalent to

$$\frac{c-d}{\overline{c}-\overline{d}}\frac{b-a}{\overline{b}-\overline{a}} = \frac{b-d}{\overline{b}-\overline{d}}\frac{c-a}{\overline{c}-\overline{a}}.$$

Since the points b, d, o are colinear, by T1.2 $\dfrac{b-d}{\overline{b}-\overline{d}} = \dfrac{b-o}{\overline{b}-\overline{o}} = \dfrac{b}{\overline{b}}$ we get $\dfrac{a-c}{\overline{a}-\overline{c}} = \dfrac{a-o}{\overline{a}-\overline{o}} = \dfrac{a}{\overline{a}}$. If $a\overline{b}+\overline{a}b = 0$ then

$$c-d = d\frac{2ab(\overline{a}-\overline{b})}{b(\overline{a}b-2a\overline{a}+a\overline{b})},$$

and the last can be obtained by conjugation. If $a\overline{a} = b\overline{b}$, then

$$c-d = \frac{d(a-b)(\overline{a}b+a\overline{b})}{b(\overline{a}b-2a\overline{a}+a\overline{b})},$$

and in this case we can get the desired statement by conjugation.

44. Let f be the origin and let $d = \overline{c}$ (this is possible since $FC = FD$). According to T9.2 we have that

$$o_1 = \frac{ad(\overline{a}-\overline{d})}{\overline{a}d-a\overline{d}}, \quad o_2 = \frac{bc(\overline{b}-\overline{c})}{\overline{b}c-b\overline{c}}.$$

Since $cd \parallel af$ according to T1.1 $\dfrac{a-f}{\overline{a}-\overline{f}} = \dfrac{c-d}{\overline{c}-\overline{d}} = -1$, i.e. $\overline{a} = -a$ and similarly $\overline{b} = -b$. Now we have

$$o_1 = \frac{\overline{c}(a+c)}{c+\overline{c}}, \quad o_2 = \frac{c(b+\overline{c})}{c+\overline{c}}.$$

Let's denote the point e. From T1.2 using the colinearity of a,c,e and b,d,e we get the following two equations

$$\frac{a-c}{\overline{a}-\overline{c}}=\frac{e-a}{\overline{e}-\overline{a}},\quad \frac{b-d}{\overline{b}-\overline{d}}=\frac{e-b}{\overline{e}-\overline{b}}.$$

From these equations we get $\overline{e}=\dfrac{a(c+\overline{c})-e(a+\overline{c})}{a-c}$ and $\overline{e}=\dfrac{b(c+\overline{c})-e(b+c)}{b-\overline{c}}$. By equating these two we get

$$e=\frac{a\overline{c}-bc}{a+\overline{c}-b-c}.$$

Using T1.3 the condition $fe\perp o_1o_2$ is equivalent to $\dfrac{o_1-o_2}{\overline{o_1}-\overline{o_2}}=-\dfrac{f-e}{\overline{f}-\overline{e}}$, which trivially follows from $o_1-o_2=\dfrac{a\overline{c}-cb}{c+\overline{c}}$ by conjugation.

45. Assume that the point p is the origin. Let ac be the real axis and let $\angle cpd=\varphi$. Then $a=\alpha, b=\beta e^{i\varphi}, c=\gamma, d=\delta e^{i\varphi}$, where $\alpha,\beta,\gamma,\delta$ are some real numbers. Let $e^{i\varphi}=\Pi$. If $|a-f|=\varepsilon|a-d|$, then $|e-c|=\varepsilon|b-c|$ hence by T6.1 $a-f=\varepsilon(a-d)$ and $e-c=\varepsilon(b-c)$. Thus we have

$$f=\alpha(1-\varepsilon)+\varepsilon\delta\Pi,\ e=\gamma(1-\varepsilon)+\varepsilon\beta\Pi.$$

Since q belongs to pd we have that $q=\rho\Pi$ and since q also belongs to ef by T1.2 we have that $\dfrac{f-q}{\overline{f}-\overline{q}}=\dfrac{e-f}{\overline{e}-\overline{f}}$, hence

$$\frac{\alpha(1-\varepsilon)+(\varepsilon\delta-\rho)\Pi}{\alpha(1-\varepsilon)+(\varepsilon\delta-\rho)\dfrac{1}{\Pi}}=\frac{(1-\varepsilon)(\alpha-\gamma)+\varepsilon(\delta-\beta)\Pi}{(1-\varepsilon)(\alpha-\gamma)+\varepsilon(\delta-\beta)\dfrac{1}{\Pi}}.$$

After some algebra we get $(\Pi-\dfrac{1}{\Pi})(1-\varepsilon)\Big[(\alpha-\gamma)(\varepsilon\delta-\rho)-\varepsilon\alpha(\delta-\beta)\Big]=0$. Since $\Pi\neq\pm1$ (because $\angle CPD<180^\circ$) and $\varepsilon\neq1$ we get $\rho=\varepsilon\Big[\delta-\dfrac{\alpha(\delta-\beta)}{\alpha-\gamma}\Big]$. Similarly we get $\rho=(1-\varepsilon)\Big[\alpha-\dfrac{\delta(\alpha-\gamma)}{\delta-\beta}\Big]$, where ρ is the coordinate of the point r. By T9.2 we have

$$\begin{aligned}o_1 &= \frac{rq(\overline{r}-\overline{q})}{\overline{r}q-\overline{q}}=\frac{\rho\rho\Pi(\rho-\rho\frac{1}{\Pi})}{\rho\rho\Pi-\rho\rho\frac{1}{\Pi}}=\frac{\rho\Pi-\rho}{\Pi^2-1}\Pi\\ &= \frac{(1-\varepsilon)\Big[\alpha-\dfrac{\delta(\alpha-\gamma)}{\delta-\beta}\Big]\Pi-\varepsilon\Big[\delta-\dfrac{\alpha(\delta-\beta)}{\alpha-\gamma}\Big]}{\Pi^2-1}\Pi.\end{aligned}$$

For any other position of the point e on the line ad such that $ae=\varepsilon ad$ the corresponding center of the circle has the coordinate

$$o_2=\frac{(1-\varepsilon)\Big[\alpha-\dfrac{\delta(\alpha-\gamma)}{\delta-\beta}\Big]\Pi-\varepsilon\Big[\delta-\dfrac{\alpha(\delta-\beta)}{\alpha-\gamma}\Big]}{\Pi^2-1}\Pi.$$

Notice that the direction of the line o_1o_2 doesn't depend on ε and ε. Namely if we denote $A=\alpha-\dfrac{\delta(\alpha-\gamma)}{\delta-\beta}$ and $B=\delta-\dfrac{\alpha(\delta-\beta)}{\alpha-\gamma}$ we have

$$\frac{o_1-o_2}{\overline{o_1}-\overline{o_2}}=-\frac{A\Pi+B}{A+B\Pi}\Pi.$$

Thus for every three centers o_1, o_2, o_3 it holds $o_1o_2 \| o_2o_3$ hence all the centers are colinear. Since all the circles have a common point, the circles have another common point.

Remark. We have proved more than we've been asked. Namely two conditions $AD = BC$ and $BE = DF$ are substituted by one $BE/BC = DF/AD$.

Another advantage of this solutions is that we didn't have to guess what is the other intersection point.

46. Let o be the origin. According to the property T9.1 we have that $h_1 = \dfrac{(a-b)(\overline{a}b + a\overline{b})}{a\overline{b} - \overline{a}b}$, $h_2 = \dfrac{(c-d)(\overline{c}d + c\overline{d})}{c\overline{d} - \overline{c}d}$, and according to the theorem 6 $t_1 = \dfrac{a+c}{3}$, $t_2 = \frac{b+d}{3}$. Since the points a, c, and o are colinear as well as the points b, d, and o by T1.2 we have $\overline{c} = \dfrac{c\overline{a}}{a}$, $\overline{d} = \dfrac{d\overline{b}}{b}$, hence $h_2 = \dfrac{(c-d)(a\overline{b} + \overline{a}b)}{a\overline{b} - \overline{a}b}$. In order to prove that $t_1t_2 \perp h_1h_2$, by T1.3, it is enough to verify

$$\frac{t_1 - t_2}{\overline{t_1} - \overline{t_2}} = -\frac{h_1 - h_2}{\overline{h_1} - \overline{h_2}}.$$

This follows from

$$h_1 - h_2 = \frac{a\overline{b} + \overline{a}b}{a\overline{b} - \overline{a}b}\Big(a + c - b - d\Big),$$

by conjugation.

47. Let Γ be the unit circle. Using T2.3 we get $c = \dfrac{2ab}{a+b}$. Let o_1 be the center of Γ_1. Then $o_1b \perp ab$ (because ab is a tangent) hence by T1.3 $\dfrac{o_1 - b}{\overline{o_1} - \overline{b}} = -\dfrac{a-b}{\overline{a} - \overline{b}} = ab$. After simplifying $\overline{o_1} = \dfrac{o_1 + a - b}{ab}$. We have also $|o_1 - b| = |o_1 - c|$, and after squaring $(o_1 - b)(\overline{o_1} - \overline{b}) = (o_1 - c)(\overline{o_1} - \overline{c})$, i.e. $\overline{o_1} = \dfrac{o_1}{b^2} - \dfrac{a-b}{b(a+b)}$. Now we have

$$o_1 = \frac{ab}{a+b} + b.$$

Since the point m belongs to the unit circle it satisfies $\overline{m} = \dfrac{1}{m}$ and since it belongs to the circle with the center o_1 it satisfies $|o_1 - m| = |o_1 - b|$. Now we have

$$\overline{o_1}m^2 - \Big(\frac{o_1}{b} + \overline{o_1}b\Big)m + o_1 = 0.$$

This quadratic equation defines both m and b, and by Vieta's formulas we have $b + m = \dfrac{o_1}{\overline{o_1}b} + b$, i.e.

$$m = b\frac{2a+b}{a+2b}.$$

It remains to prove that the points a, m, and the midpoint of the segment bc colinear. The midpoint of bc is equal to $(b+c)/2$ by T6.1. According to T1.2 it is enough to prove that

$$\frac{a - \dfrac{b+c}{2}}{\overline{a} - \overline{\dfrac{b+c}{2}}} = \frac{a-m}{\overline{a} - \overline{m}} = -am,$$

which is easy to verify.

48. Assume that the circle k is unit and assume that $b = 1$. The $a = -1$ and since $p \in k$ we have $\overline{p} = \frac{1}{p}$. According to T2.4 we have that $q = \frac{1}{2}\left(p + \frac{1}{p}\right)$, and according to T6.1 we have that $f = \frac{\left(p+\frac{1}{p}\right)-1}{2} = \frac{(p-1)^2}{4p}$. Furthermore since c belongs to the circle with the center p and radius $|p-q|$ we have $|p-q| = |p-c|$ and after squaring

$$(p-q)(\overline{p}-\overline{q}) = (p-c)(\overline{p}-\overline{c}).$$

Since $c \in k$ we have $\overline{c} = \frac{1}{c}$. The relation $p - q = \frac{1}{2}\left(p - \frac{1}{p}\right)$ implies

$$4pc^2 - (p^4 + 6p^2 + 1)c + 4p^3 = 0.$$

Notice that what we obtained is the quadratic equation for c. Since d satisfies the same conditions we used for c, then the point d is the second solution of this quadratic equation. Now from Vieta's formulas we get

$$c + d = \frac{p^4 + 6p^2 + 1}{4p^3}, \quad cd = p^2.$$

Since the point g belongs to the chord cd by T2.2 we get

$$\overline{g} = \frac{c + d - g}{cd} = \frac{p^4 + 6p^2 + 1 - 4pg}{4p^3}.$$

From $gf \perp cd$ T1.3 gives $\frac{g-f}{\overline{g}-\overline{f}} = -\frac{c-d}{\overline{c}-\overline{d}} = cd = p^2$. Solving this system gives us

$$g = \frac{p^3 + 3p^2 - p + 1}{4p}.$$

The necessair and sufficient condition for colinearity of the points a, p, g is (according to T1.2) $\frac{a-g}{\overline{a}-\overline{g}} = \frac{a-p}{\overline{a}-\overline{p}} = p$. This easily follows from $a - g = \frac{p^3 + 3p^3 + 3p + 1}{4p}$ and by conjugating $\overline{a} - \overline{g} = \frac{1 + 3p + 3p^2 + p^3}{4p^2}$. Since e belongs to the chord cd we have by T2.2 $\overline{e} = \frac{c + d - g}{cd} = \frac{p^4 + 6p^2 + 1 - 4pe}{4p^3}$, and since $pe \perp ab$ T1.3 implies $\frac{e-p}{\overline{e}-\overline{p}} = -\frac{a-b}{\overline{a}-\overline{b}} = -1$, or equivalently $\overline{e} = p + \frac{1}{p} - e$. It follows that $e = \frac{3p^2+1}{4p}$. Since $p - q = \frac{p^2-1}{2p} = 2\frac{p^2-1}{4p} = 2\left(e - q\right)$, we get $|e-p| = |e-q|$. Furthermore since $g - e = \frac{p^2-1}{4}$ from $|p| = 1$, we also have $|e-q| = |g-e|$, which finishes the proof.

49. Assume that the circle with the diameter bc is unit and that $b = -1$. Now by T6.1 we have that $b + c = 0$, i.e. $c = 1$, and the origin is the midpoint of the segment bc. Since p belongs to the unit circle we have $\overline{p} = \frac{1}{p}$, and since $pa \perp p0$, we have according to T1.3 $\frac{a-p}{\overline{a}-\overline{p}} = -\frac{p-0}{\overline{p}-\overline{0}} = -p^2$. Simplification yields

$$\overline{a}\,p^2 - 2p + a = 0.$$

Since this quadratic equation defines both p and q, according to Vieta's formulas we have

$$p + q = \frac{2}{\overline{a}}, \quad pq = \frac{a}{\overline{a}}.$$

Let h' be the intersection of the perpendicular from a to bc with the line pq. Since $h' \in pq$ T2.2 gives $\overline{h'} = \dfrac{p+q-h'}{pq} = \dfrac{2-\overline{a}h}{a}$. Since $ah \perp bc$ according to T1.3 we have $\dfrac{a-h}{\overline{a}-\overline{h}} = -\dfrac{b-c}{\overline{b}-\overline{c}} = -1$, i.e. $\overline{h} = a+\overline{a}-h$. Now we get

$$h = \frac{a\overline{a}+a^2-2}{a-\overline{a}}.$$

It is enough to prove that $h' = h$, or $ch \perp ab$ which is by T1.3 equivalent to $\dfrac{h-c}{\overline{h}-\overline{c}} = -\dfrac{a-b}{\overline{a}-\overline{b}}$. The last easily follows from

$$h-1 = \frac{a\overline{a}+a^2-2-a+\overline{a}}{a-\overline{a}} = \frac{(a+1)(a+\overline{a}-2)}{a-\overline{a}}$$

and $a-b = a+1$ by conjugation.

50. Assume that the origin of our coordinate system is the intersection of the diagonals of the rectangle and that the line ab is parallel to the real axis. We have by T6.1 $c+a=0$, $d+b=0$, $c=\overline{b}$, and $d=\overline{a}$. Since the points $p,a,0$ are colinear T1.2 implies $\dfrac{p}{\overline{p}} = \dfrac{a}{\overline{a}}$, i.e. $\overline{p} = -\dfrac{b}{a}p$. Let $\varphi = \angle dpb = \angle pbc$. By T1.4 we have

$$\frac{c-p}{\overline{c}-\overline{p}} = e^{i2\varphi}\frac{b-p}{\overline{b}-\overline{p}}, \quad \frac{p-b}{\overline{p}-\overline{b}} = e^{i2\varphi}\frac{c-b}{\overline{c}-\overline{b}},$$

and after multiplying these equalities and expressing in terms of a and b

$$\frac{p+b}{bp+a^2} = \frac{a(p-b)^2}{(bp-a^2)^2}.$$

In the polynomial form this writes as

$$(b^2-ab)p^3+p^2(b^3-2a^2b-a^3+2ab^2)+p(a^4-2a^2b^2-ab^3+2a^3b)+a^4b-a^3b^2$$
$$=(b-a)(bp^3+(a^2+3ab+b^2)p^2-ap(a^2+3ab+b^2)-a^3b)=0.$$

Notice that a is one of those points p which satisfy the angle condition. Hence a is one of the zeroes of the polynomial. That means that p is the root of the polynomial which is obtained from the previous one after division by $p-a$ i.e. $bp^2+(a^2+3ab+b^2)p+a^2b=0$. Let's now determine the ratio $|p-b|:|p-c|$. From the previous equation we have $bp^2+a^2b = -(a^2+3ab+b^2)$, hence

$$\frac{PB^2}{PC^2} = \frac{(p-b)(\overline{p}-\overline{b})}{(p-c)(\overline{p}-\overline{c})} = \frac{bp^2-(a^2+b^2)p+a^2b}{bp^2+2abp+a^b} = \frac{-2(a^2+b^2+2ab)}{-(a^2+b^2+2ab)} = 2,$$

and the required ratio is $\sqrt{2}:1$.

51. Assume first that the quadrilateral $abcd$ is cyclic and that its cicrumcircle is the unit circle. If $\angle abd = \varphi$ and $\angle bda = \theta$ by T1.4 after squaring we have

$$\frac{d-b}{\overline{d}-\overline{b}} = e^{i2\varphi}\frac{a-b}{\overline{a}-\overline{b}}, \quad \frac{c-b}{\overline{c}-\overline{b}} = e^{i2\varphi}\frac{p-b}{\overline{p}-\overline{b}},$$
$$\frac{c-d}{\overline{c}-\overline{d}} = e^{i2\theta}\frac{p-d}{\overline{p}-\overline{d}}, \quad \frac{b-d}{\overline{b}-\overline{d}} = e^{i2\theta}\frac{a-d}{\overline{a}-\overline{d}}.$$

From the first of these equalities we get $e^{i2\varphi}\dfrac{a}{d}$, and from the fourth $e^{i2\theta} = \dfrac{b}{a}$. From the second equality we get $\overline{p} = \dfrac{ac+bd-pd}{abc}$, and from the third $\overline{p} = \dfrac{ac+bd-pb}{acd}$. Now it follows that

$$p = \frac{ac+bd}{b+d}.$$

We have to prove that $|a-p|^2=(a-p)(\overline{a}-\overline{p})=|c-p|^2=(c-p)(\overline{c}-\overline{p})$, which follows from

$$a-p=\frac{ab+ad-ac-bd}{b+d},\quad \overline{a}-\overline{p}=\frac{cd+bc-bd-ac}{ac(b+d)},$$

$$c-p=\frac{bc+cd-ac-bd}{b+d},\quad \overline{c}-\overline{p}=\frac{ad+ab-bd-ac}{ac(b+d)}.$$

Assume that $|a-p|=|c-p|$. Assume that the circumcircle of the triangle abc is unit. Squaring the last equality gives us that $a\overline{p}+\frac{p}{a}=c\overline{p}+\frac{p}{c}$, i.e. $(a-c)(\overline{p}-\frac{p}{ac})=0$. This means that $\overline{p}=\frac{p}{ac}$. Let d belong to the chord $d'c$. Then according to T2.2 $\overline{d}=\frac{c+d'-d}{cd'}$. By the condition of the problem we have $\angle dba=\angle cbp=\varphi$ and $\angle adb=\angle pdc=\theta$, and squaring in T1.4 yields

$$\frac{a-b}{\overline{a}-\overline{b}}=e^{i2\varphi}\frac{d-b}{\overline{d}-\overline{b}},\quad \frac{p-b}{\overline{p}-\overline{b}}=e^{i2\varphi}\frac{c-b}{\overline{c}-\overline{b}},$$

$$\frac{b-d}{\overline{b}-\overline{d}}=e^{i2\theta}\frac{a-d}{\overline{a}-\overline{d}},\quad \frac{c-d}{\overline{c}-\overline{d}}=e^{i2\theta}\frac{p-d}{\overline{p}-\overline{d}}.$$

Multiplying the first two equalities gives us

$$\frac{a-b}{\overline{a}-\overline{b}}\frac{c-b}{\overline{c}-\overline{b}}=ab^2c=\frac{p-b}{\overline{p}-\overline{b}}\frac{d-b}{\overline{d}-\overline{b}}.$$

After some algebra we conclude

$$p=\frac{ac+bd-b(ac\overline{d}+b)}{d-b^2\overline{d}}=\frac{bdd'+acd'-abd'-abc+abd-b^2d'}{cd'd-b^2d'+b^2d-b^2c}.$$

Since the pionts d,c,d' are colinear, according to T1.2 we get $\frac{d-c}{\overline{d}-\overline{c}}=\frac{c-d'}{\overline{c}-\overline{d'}}=-cd'$, and mylti-plying the third and fourth equality gives

$$(-cd')(d-a)(\overline{d}-\overline{b})(\overline{d}-\overline{p})-(\overline{d}-\overline{a})(d-b)(d-p)=0.$$

Substituting values for p gives us a polynomial f in d. It is of the most fourth degree and observing the coefficient next to d^4 of the left and right summand we get that the polynomial is of the degree at most 3. It is obvious that a and b are two of its roots. We will now prove that its third root is d' and that would imply $d=d'$. For $d=d'$ we get

$$p=\frac{bd'd+acd'-abc-b^2d'}{c(d'^2-b^2)}=\frac{ac+bd'}{b+d'},\quad d-p=\frac{d'^2-ac}{b+d'}$$

$$\overline{d}-\overline{p}=-bd'\frac{d'^2-ac}{ac(b+d')}\quad \frac{d-a}{\overline{d}-\overline{a}}=-d'a,\quad \frac{d-b}{\overline{d}-\overline{b}}=-d'b$$

and the statement is proved. Thus $d=d'$ hence the quadrilateral $abcd'$ is cyclic.

52. Since the rectangles $a_1b_2a_2b_1$, $a_2b_3a_3b_2$, $a_3b_4a_4b_3$, and a_4,b_1,a_1,b_4 are cyclic T3 implies that the numbers

$$\frac{a_1-a_2}{b_2-a_2}:\frac{a_1-b_1}{b_2-b_1},\quad \frac{a_2-a_3}{b_3-a_3}:\frac{a_2-b_2}{b_3-b_2},$$

$$\frac{a_3-a_4}{b_4-a_4}:\frac{a_3-b_3}{b_4-b_3},\quad \frac{a_4-a_1}{b_1-a_1}:\frac{a_4-b_4}{b_1-b_4},$$

are real. The product of the first and the third divided by the product of the second and the fourth is equal to

$$\frac{a_1-a_2}{a_2-a_3}\cdot\frac{a_3-a_4}{a_4-a_1}\cdot\frac{b_2-b_1}{b_3-b_2}\cdot\frac{b_4-b_3}{b_1-b_4},$$

and since the points a_1,a_2,a_3,a_4 lie on a circle according to the theorem 4 the number $\frac{a_1-a_2}{a_2-a_3}\cdot\frac{a_3-a_4}{a_4-a_1}$ is real, hence the number $\frac{b_2-b_1}{b_3-b_2}\cdot\frac{b_4-b_3}{b_1-b_4}$ is real as well. According to T3 the points b_1,b_2,b_3,b_4 are cyclic or colinear.

53. Assume that the origin is the intersection of the diagonals of the parallelogram. Then $c=-a$ and $d=-b$. Since the triangles cde and fbc are similar and equally orientged by T4

$$\frac{c-b}{b-f}=\frac{e-d}{d-c},$$

hence $f=\frac{be+c^2-bc-cd}{e-d}=\frac{be+a^2}{e+b}$. In order for triangles cde and fae to be similar and equally oriented (as well as for fbc and fae), according to T4 it is necessairy and sufficient that the following relation holds:

$$\frac{c-d}{d-e}=\frac{f-a}{a-e}.$$

The last equaliy follows from

$$f-a=\frac{be+a^2-ea-ab}{e+b}=\frac{(e-a)(b-a)}{e+b},$$

and $c-d=c+b$, $d-e=-(b+e)$, $c+b=b-a$.

54. Let $p=0$ and $q=1$. Since $\angle mpq=\alpha$, according to T1.4 we have that $\frac{q-p}{\overline{q}-\overline{p}}=e^{i2\alpha}\frac{m-p}{\overline{m}-\overline{p}}$, i.e. $\frac{m}{\overline{m}}=e^{i2\alpha}$. Since $\angle pqm=\beta$, the same theorem implies $\frac{m-q}{\overline{m}-\overline{q}}=e^{i2\beta}\frac{p-q}{\overline{p}-\overline{q}}$, i.e. $1=e^{i2\beta}\frac{m-1}{\overline{m}-1}$. Solving this system (with the aid of $e^{i2(\alpha+\beta+\gamma)}=1$) we get $m=\frac{e^{i2(\alpha+\gamma)}-1}{e^{i2\gamma}-1}$, and symmetrically $l=\frac{e^{i2(\beta+\gamma)}-1}{e^{i2\beta}-1}$, $k=\frac{e^{i2(\alpha+\beta)}-1}{e^{i2\alpha}-1}$. According to T4 in order to prove that the triangles klm and kpq are similar and equally oriented it is enough to prove that $\frac{k-l}{l-m}=\frac{k-p}{p-q}=-k$. The last follows from

$$\begin{aligned}\frac{k-l}{l-m} &= \frac{\dfrac{e^{i(2\alpha+4\beta)}-e^{i2\beta}-e^{i(2\alpha+2\beta)}+e^{i(2\beta+2\gamma)}+e^{i2\alpha}-1}{(e^{i2\alpha}-1)(e^{i2\beta}-1)}}{\dfrac{e^{i(2\beta+4\gamma)}-e^{i2\gamma}-e^{i(2\beta+2\gamma)}+e^{i(2\alpha+2\gamma)}+e^{i2\beta}-1}{(e^{i2\beta}-1)(e^{i2\gamma}-1)}}\\ &= \frac{e^{i2(\alpha+\beta)}(e^{i(2\beta+4\gamma)}-e^{i2\gamma}-e^{i(2\beta+2\gamma)}+e^{i(2\alpha+2\gamma)}+e^{i2\beta}-1)}{e^{i(2\beta+4\gamma)}-e^{i2\gamma}-e^{i(2\beta+2\gamma)}+e^{i(2\alpha+2\gamma)}+e^{i2\beta}-1}\cdot\\ &\quad\frac{e^{i2\gamma}-1}{e^{i2\alpha}-1}\\ &= \frac{1-e^{i2(\alpha+\beta)}}{e^{i2\alpha}-1}=-k.\end{aligned}$$

Since the triangles kpq,qlp,pqm are mutually similar and equally oriented the same holds for all four of the triangles.

55. Assume that the coordinates of the vertices of the i-th polygon are denoted by $a_1^{(i)}, a_2^{(i)}, \ldots, a_n^{(i)}$, respectively in positive direction. smeru. According to T6.1 and the given recurrent relation we have that for each i and k:

$$a_i^{(k+1)} = 2a_{i+k}^{(k)} - a_i^{(k)},$$

where the indices are modulo n. Our goal is to determine the value of $a_i^{(n)}$, using the values of $a_1^{(1)}, a_2^{(1)}, \ldots, a_n^{(1)}$. The following

$$\begin{aligned} a_i^{(k+1)} &= 2a_{i+k}^{(k)} - a_k^{(i)} = 4a_{i+k+k-1}^{(k-1)} - 2a_{i+k}^{(k-1)} - 2a_{i+k-1}^{(k-1)} + a_i^{(k-1)} \\ &= 4(2a_{i+k+k-1+k-2}^{(k-2)} - -a_{i+k+k-1}^{(k-2)}) - 2(2a_{i+k+k-2}^{(k-2)} - a_{i+k}^{(k-2)}) - \\ &\quad 2(2a_{i+k-1+k-2}^{(k-2)} - a_{i+k-1}^{(k-2)}) + 2a_{i+k-2}^{(k-2)} - a_i^{(k-2)} \\ &= 8a_{i+k+k-1+k-2}^{(k-2)} - 4(a_{i+k+k-1}^{(k-2)} + a_{i+k+k-2}^{(k-2)} + a_{i+k-1+k-2}^{(k-2)}) + \\ &\quad 2(a_{i+k}^{(k-2)} + a_{i+k-1}^{(k-2)} + a_{i+k-2}^{(k-2)}) - a_i^{(k-2)}, \end{aligned}$$

yields that

$$a_i^{(k)} = 2^{k-1}s_k^{(k)}(i) - 2^{k-2}s_{k-1}^{(k)}(i) + \ldots + (-1)^k s_0^{(k)}(i),$$

where $s_j^{(k)}(i)$ denotes the sum of all the numbers of the form $a_{i+s_k(j)}$ and $s_k(j)$ is one of the numbers obtained as the sum of exactly j different natural numbers not greater than n. Here we assume that $s_0^{(k)}(i) = a_i$. The last formula is easy to prove by induction. Particularly, the formula holds for $k = n$ hence

$$a_i^{(n)} = 2^{n-1}s_n^{(n)}(i) - 2^{n-2}s_{n-1}^{(n)}(i) + \ldots + (-1)^n s_0^{(n)}(i).$$

Now it is possible to prove that $s_l^{(n)}(i) = s_l^{(n)}(j)$, for each $1 \leq l \leq n-1$ which is not very difficult problem in the number theory. Since n is prime we have that $n+n-1+\ldots+1$ is divisible by n hence

$$\begin{aligned} a_i^{(n)} - a_j^{(n)} &= 2^{n-1}a_{i+n+n-1+\ldots+1}^{(1)} - 2^{n-1}a_{j+n+n-1+\ldots+1}^{(1)} + \\ &\quad (-1)^n a_i^{(1)} - (-1)^n a_j^{(1)} \\ &= (2^{n-1} + (-1)^n)(a_i^{(1)} - a_j^{(1)}), \end{aligned}$$

which by T4 finishes the proof.

56. Assume that the pentagon $abcde$ is inscribed in the unit circle and that x, y, and z are feet of perpendiculars from a to bc, cd, and de respectively. According to T2.4 we have that

$$x = \frac{1}{2}\Big(a+b+c-\frac{bc}{a}\Big), \quad y = \frac{1}{2}\Big(a+c+d-\frac{cd}{a}\Big), \quad z = \frac{1}{2}\Big(a+d+e-\frac{de}{a}\Big),$$

and according to T5 we have

$$S(xyz) = \frac{i}{4}\begin{vmatrix} x & \overline{x} & 1 \\ y & \overline{y} & 1 \\ z & \overline{z} & 1 \end{vmatrix} = \frac{i}{8}\begin{vmatrix} a+b+c-\frac{bc}{a} & \overline{a}+\overline{b}+\overline{c}-\frac{\overline{bc}}{\overline{a}} & 1 \\ a+c+d-\frac{cd}{a} & \overline{a}+\overline{c}+\overline{d}-\frac{\overline{cd}}{\overline{a}} & 1 \\ a+d+e-\frac{de}{a} & \overline{a}+\overline{d}+\overline{e}-\frac{\overline{de}}{\overline{a}} & 1 \end{vmatrix}.$$

Since the determinant is unchanged after substracting some columns from the others, we can substract the second column from the third, and the first from the second. After that we get

$$\begin{aligned} S(xyz) &= \frac{i}{8}\begin{vmatrix} a+b+c-\frac{bc}{a} & \overline{a}+\overline{b}+\overline{c}-\frac{\overline{b}\,\overline{c}}{\overline{a}} & 1 \\ \frac{(d-b)(a-c)}{a} & \frac{(d-b)(a-c)}{bcd} & 0 \\ \frac{(e-c)(a-d)}{a} & \frac{(e-c)(a-d)}{a} & 0 \end{vmatrix} \\ &= \frac{i(a-c)(d-b)(a-d)(e-c)}{8}\cdot \\ &\quad \begin{vmatrix} a+b+c-\frac{bc}{a} & \overline{a}+\overline{b}+\overline{c}-\frac{\overline{b}\,\overline{c}}{\overline{a}} & 1 \\ \frac{1}{a} & \frac{1}{bcd} & 0 \\ \frac{1}{a} & \frac{1}{a} & 0 \end{vmatrix}, \end{aligned}$$

and finally

$$\begin{aligned} S(xyz) &= \frac{i(a-c)(d-b)(a-d)(e-c)}{8}\Big(\frac{1}{acde}-\frac{1}{abcd}\Big) \\ &= \frac{i(a-c)(d-b)(a-d)(e-c)(b-e)}{8abcde}. \end{aligned}$$

Since the last expression is symmetric with respet to a,b,c,d, and e the given area doesn't depend on the choice of the vertex (in this case a).

57. Assume that the unit circle is the circumcircle of the triangle abc. Since $\dfrac{S(bca_1)}{S(abc)} = 1-\dfrac{|a-a_1|}{|a-a'|} = 1-\dfrac{a-a_1}{a-a'}$ (where a' is the foot of the perpendicular from a to bc), the given equality becomes

$$2 = \frac{a-a_1}{a-a'}+\frac{b-b_1}{b-b'}+\frac{c-c_1}{c-c'}.$$

According to T2.4 we have $a' = \dfrac{1}{2}\Big(a+b+c-\dfrac{bc}{a}\Big)$, hence

$$a-a' = \frac{1}{2}\Big(a+\frac{bc}{a}-b-c\Big) = \frac{(a-b)(a-c)}{2a}$$

and after writing the symmetric expressions we get

$$\begin{aligned} 2 &= \frac{2a(a-a_1)}{(a-b)(a-c)}+\frac{2b(b-b_1)}{(b-a)(b-c)}+\frac{2c(c-c_1)}{(c-a)(c-b)} \\ &= -2\frac{a(a-a_1)(b-c)+b(b-b_1)(c-a)+c(c-c_1)(a-b)}{(a-b)(b-c)(c-a)}, \end{aligned}$$

and after simplying

$$aa_1(b-c)+bb_1(c-a)+cc_1(a-b) = 0.$$

By T4 points a_1,b_1,c_1,h lie on a circle if and only if

$$\frac{a_1-c_1}{\overline{a_1}-\overline{c_1}}\frac{b_1-h}{\overline{b_1}-\overline{h}} = \frac{a_1-h}{\overline{a_1}-\overline{h}}\frac{b_1-c_1}{\overline{b_1}-\overline{c_1}}.$$

Since h is the orthocenter by T6.3 we have $h=a+b+c$, and since $aa_1 \perp bc$ T1.3 implies $\dfrac{a_1-a}{\overline{a_1}-\overline{a}} = -\dfrac{b-c}{\overline{b}-\overline{c}}$, i.e. $\overline{a_1} = \dfrac{bc+aa_1-a^2}{abc}$, and symmetrically $\overline{b_1} = \dfrac{ac+bb_1-b^2}{abc}$ and $\overline{c_1} = \dfrac{ab+cc_1-c^2}{abc}$. Similarly from $a_1h \perp bc$ and $b_1h \perp ac$

$$\frac{a_1-h}{\overline{a_1}-\overline{h}} = -\frac{b-c}{\overline{b}-\overline{c}} = bc, \quad \frac{b_1-h}{\overline{b_1}-\overline{h}} = -\frac{a-c}{\overline{a}-\overline{c}} = ac.$$

It is enough to prove that

$$\frac{a(a_1-c_1)}{aa_1-cc_1+(c-a)(a+b+c)} = \frac{b(b_1-c_1)}{bb_1-cc_1+(c-b)(a+b+c)}.$$

Notice that

$$a(b-c)a_1-a(b-c)c_1 = -b_1b(c-a)a-cc_1(a-b)a-a(b-c)c_1 = ab(c-a)(c_1-b_1),$$

and the result follows by the conjugation.

58. Assume that the unit circle is the circumcircle of the triangle abc. By T2.4 we have that $d = \frac{1}{2}\left(a+b+c-\frac{ab}{c}\right)$, $e = \frac{1}{2}\left(a+b+c-\frac{ac}{b}\right)$, and $f = \frac{1}{2}\left(a+b+c-\frac{bc}{a}\right)$. According to T6.1 we get $a_1 = \dfrac{b+c}{2}$ (where a_1 is the midpoint of the sidebc). Since q belongs to the chord ac T2.2 implies $\overline{q} = \dfrac{a+c-q}{ac}$, and since $qd \parallel ef$ T1.1 implies $\dfrac{q-d}{\overline{q}-\overline{d}} = \dfrac{e-f}{\overline{e}-\overline{f}} = -a^2$. Solving this system gives us

$$q = \frac{a^3+a^2b+abc-b^2c}{2ab}.$$

Symmetrically we get $r = \dfrac{a^3+a^2c+abc-bc^2}{2ac}$. Since p belongs to the chord bc T2.2 implies $\overline{p} = \dfrac{b+c-p}{bc}$, and from the colinearity of the points e, f, and p from T1.2 we conclude $\dfrac{p-e}{\overline{p}-\overline{e}} = \dfrac{e-f}{\overline{e}-\overline{f}} = -a^2$. After solving this system we get

$$p = \frac{a^2b+a^2c+ab^2+ac^2-b^2c-bc^2-2abc}{2(a^2-bc)} = \frac{b+c}{2} + \frac{a(b-c)^2}{2(a^2-bc)}.$$

By T4 it is sufficient to prove that

$$\frac{p-a_1}{p-r}\frac{q-r}{q-a_1} = \frac{\overline{p}-\overline{a_1}}{\overline{p}-\overline{r}}\frac{\overline{q}-\overline{r}}{\overline{q}-\overline{a_1}}.$$

Since

$$q-r = \frac{a(c-b)(a^2+bc)}{2abc}, \quad p-a_1 = \frac{a(b-c)^2}{2(a^2-bc)},$$

$$p-r = \frac{(a^2-c^2)(b^2c+abc-a^3-a^2c)}{2ac(a^2-bc)}, \quad q-a_1 = \frac{a^3+a^2b-b^2c-ab^2}{2ab}$$

the required statement follows by conjugation.

59. Let O be the circumcenter of the triangle abc. We will prove that O is the incenter as well. Assume that the circumcircle of the triangle abc is unit. According to T6.1 we have that $c_1 = \dfrac{a+b}{2}$,

$b_1=\dfrac{a+c}{2}$, and $a_1=\dfrac{b+c}{2}$. Assume that k_1,k_2,k_3 are the given circles with the centers a_1,b_1, and c_1. Let $k_1\cap k_2=\{k,o\}$, $k_2\cap k_3=\{m,o\}$, and $k_3\cap k_1=\{l,o\}$. Then we have $|a_1-k|=|a_1-o|$, $|b_1-k|=|b_1-o|$. After squaring $(a_1-k)(\overline{a_1}-\overline{k})=a_1\overline{a_1}$ and $(b_1-k)(\overline{b_1}-\overline{k})=b_1\overline{b_1}$. After solving this system we obtain

$$k=\frac{(a+c)(b+c)}{2c}.$$

Symmetrically we get $l=\dfrac{(b+c)(a+b)}{2b}$ and $m=\dfrac{(a+c)(a+b)}{2a}$. Let $\angle mko=\varphi$. According to T1.4 we have that $\dfrac{o-k}{\overline{o}-\overline{k}}=e^{i2\varphi}\dfrac{m-k}{\overline{m}-\overline{k}}$, and since $k-m=\dfrac{b(a^2-c^2)}{2ac}$, after conjugation $e^{i2\varphi}=-\dfrac{a}{b}$. If $\angle okl=\psi$, we have by T1.4 $\dfrac{o-k}{\overline{o}-\overline{k}}=e^{i2\psi}\dfrac{l-k}{\overline{l}-\overline{k}}$, hence $e^{i\psi}=-\dfrac{a}{b}$. Now we have $\varphi=\psi$ or $\varphi=\psi\pm\pi$, and since the second condition is impossible (why?), we have $\varphi=\psi$. Now it is clear that o is the incenter of the triangle klm.

For the second part of the problem assume that the circle is inscribed in the triangle klm is the unit circle and assume it touches the sides kl,km,lm at u,v,w respectively. According to T7.1 we have that

$$k=\frac{2uv}{u+v},\quad l=\frac{2uw}{w+u},\quad m=\frac{2vw}{v+w}.$$

Let a_1 be the circumcenter of the triangle kol. Then according to T9.2 we have

$$a_1=\frac{kl(\overline{k}-\overline{l})}{\overline{k}l-k\overline{l}}=\frac{2uvw}{k(u+v)(u+w)}$$

and symmetrically $b_1=\dfrac{2uvw}{(u+v)(v+w)}$ and $c_1=\dfrac{2uvw}{(w+u)(w+v)}$ (b_1 and c_1 are circumcenters of the triangles kom and mol respectively). Now T6.1 implies

$$a+b=2c_1,\quad b+c=2a_1,\quad a+c=2b_1,$$

and after solving this system we get $a=b_1+c_1-a_1$, $b=a_1+c_1-b_1$, and $c=a-1+b_1-c_1$. In order to finish the proof it is enough to establish $ab\perp oc_1$ (the other can be proved symmetrically), i.e. by T1.3 that $\dfrac{c_1-o}{\overline{c_1}-\overline{o}}=-\dfrac{a-b}{\overline{a}-\overline{b}}=-\dfrac{b_1-a_1}{\overline{b_1}-\overline{a_1}}$. The last easily follows from

$$b_1-a_1=\frac{2uvw(u-v)}{(u+v)(v+w)(w+u)},$$

by conjugation.

60. Let b and c be the centers of the circles k_1 and k_2 respectively and assume that bc is the real axis. If the points m_1 and m_2 move in the same direction using T1.4 we get that m_1 and m_2 satisfy

$$m_1-b=(a-b)e^{i\varphi},\quad m_2-c=(a-c)e^{i\varphi}.$$

If ω is the requested point, we must have $|\omega-m_1|=|\omega-m_2|$, and after squaring $(\omega-m_1)(\overline{\omega}-\overline{m_1})=(\omega-m_2)(\overline{\omega}-\overline{m_2})$. From the last equation we get

$$\overline{\omega}=\frac{m_1\overline{m_1}-m_2\overline{m_2}-\omega(\overline{m_1}-\overline{m_2})}{m_1-m_2}.$$

After simplification (with the usage of $\overline{b}=b$ and $\overline{c}=c$ where $e^{i\varphi}=z$)

$$\overline{w}(1-z)=2(b+c)-a-\overline{a}+az+\overline{a}\overline{z}-(b+c)(z+\overline{z})-(1-\overline{z})\omega.$$

Since $\overline{z} = \dfrac{1}{z}$, we have

$$(b+c-a-\overline{w})z^2 - (2(b+c)-a-\overline{a}-\omega-\overline{\omega})z + b + c - \overline{a} - \omega \equiv 0.$$

The last polynomial has to be identical to 0 hence each of its coefficients is 0, i.e. $\omega = b+c-\overline{a}$. From the previous relations we conclude that this point satisfies the conditions of the problem.
The problem is almost identical in the case of the oposite orientition.

61. Let γ be the unit circle and let $a=-1$. Then $b=1$, $c=1+2i$, and $d=-1+2i$. Since the points n,b,p are colinear we can use T1.2 to get

$$\frac{a-p}{\overline{a}-\overline{p}} = \frac{a-m}{\overline{a}-\overline{m}} = -am = m,$$

and after some algebra $\overline{p} = \dfrac{p+1-m}{m}$ (1). Since the points c,d,p are colinear using the same argument we get that

$$\frac{c-n}{\overline{c}-\overline{n}} = \frac{c-d}{\overline{c}-\overline{d}} = 1,$$

hence $\overline{p} = p-4i$. Comparing this with (1) one gets $p = 4i \cdot \dfrac{m}{m-1} - 1$. Furthermore, since the points b,n,p are colinear we have

$$\frac{p-1}{\overline{p}-\overline{1}} = \frac{1-n}{\overline{1}-\overline{n}} = n,$$

i.e.

$$n = \frac{m(1-2i)-1}{2i+1-m}.$$

Let q' be the intersection point of the circle γ and the line dm. If we show that the points q',n,c are colinear we would have $q=q'$ and $q \in \gamma$, which will finish the first part of the problem. Thus our goal is to find the coordinate of the point q'. Since q' belongs to the unit circle we have $q'\overline{q'} = 1$, and since d,m,q' are colinear, we have using T1.2 that

$$\frac{d-m}{\overline{d}-\overline{m}} = \frac{q'-m}{\overline{q'}-\overline{m}} = -q'm,$$

and after simplification

$$q' = -\frac{m+1-2i}{m(1+2i)+1}.$$

In order to prove that the points q',n,c are colinear it suffices to show that $\dfrac{q-c}{\overline{q}-\overline{c}} = \dfrac{n-q}{\overline{n}-\overline{q}} = -nq$, i.e. $n = \dfrac{q-1-2i}{(\overline{q}-1+2i)q}$, which is easy to verify. This proves the first part of the problem.

Now we are proving the second part. Notice that the required inequality is equivalent to $|q-a| \cdot |p-c| = |d-p| \cdot |b-q|$. From the previously computed values for p and q, we easily obtain

$$|q-a| = 2\left|\frac{m+1}{m(1+2i)+1}\right|,\ |p-c| = 2\left|\frac{m(1+i)+1-i}{m(1+2i)+1}\right|,$$

$$|d-p| = 2\left|\frac{m+1}{m+1}\right|,\ |b-q| = 2\left|\frac{m(i-1)+1+1}{m-1}\right|,$$

and since $-i((i-1)m+1+i) = m(1+i)+1-i$ the required equality obviously holds.

62. In this problem we have plenty of possibilities for choosing the unit circle. The most convenient choice is the circumcircle of $bcb'c'$ (try if you don't believe). According T2.5 we have that the intersection point x of bb' and cc' satisfy

$$x=\frac{bb'(c+c')-cc'(b+b')}{bb'-cc'}.$$

Since $bh\perp cb'$ and $ch\perp bc'$ T1.3 implies the following two equalities

$$\frac{b-h}{\overline{b}-\overline{h}}=-\frac{b'-c}{\overline{b'}-\overline{c}}=b'c,\quad \frac{c-h}{\overline{c}-\overline{h}}=-\frac{b-c'}{\overline{b}-\overline{c'}}=bc'.$$

From the first we get $\overline{h}=\dfrac{bh-b^2+b'c}{bb'c}$, and from the second $\overline{h}=\dfrac{ch-c^2+bc'}{bcc'}$. After equating the two relations we get

$$h=\frac{b'c'(b-c)+b^2c'-b'c^2}{bc'-b'c}.$$

Symmetrically we obtain $h'=\dfrac{bc(b'-c')+b'^2c-bc'^2}{b'c-bc'}$. It suffices to prove that the points h,h' and x are colinear, or after applying T1.2 we have to verify

$$\frac{h-h'}{\overline{h}-\overline{h'}}=\frac{h-x}{\overline{h}-\overline{x}}.$$

The last follwos from

$$\begin{aligned}
h-h' &= \frac{bc(b'-c')+b'c'(b-c)+bc'(b-c')+b'c(b'-c)}{bc'-b'c}\\
&= \frac{(b+b'-c-c')(bc'+b'c)}{bc'-b'c},\\
h-x &= \frac{b^2b'^2c'+b^3b'c'+b'c^2c'^2+b'c^3c'}{(bc'-b'c)(bb'-cc')}-\\
&\quad \frac{b^2b'cc'+b^2b'c'^2+bb'c^2c'+b'^2c^2c'}{(bc'-b'c)(bb'-cc')}\\
&= \frac{b'c'(b^2-c^2)(b'+b-c-c')}{(bc'-b'c)(bb'-cc')}
\end{aligned}$$

by conjugation.

63. From elementary geometry we know that $\angle nca=\angle mcb$ (such points m and n are called harmonic conjugates). Let $\angle mab=\alpha$, $\angle abm=\beta$, and $\angle mca=\gamma$. By T1.4 we have that

$$\frac{a-b}{|a-b|}=e^{i\alpha}\frac{a-m}{|a-m|},\ \frac{a-n}{|a-n|}=e^{i\alpha}\frac{a-c}{|a-c|},$$

$$\frac{b-c}{|b-c|}=e^{i\beta}\frac{b-n}{|b-n|},\ \frac{b-m}{|b-m|}=e^{i\beta}\frac{b-a}{|b-a|},$$

$$\frac{c-a}{|c-a|}=e^{i\gamma}\frac{c-n}{|c-n|},\quad \frac{c-m}{|c-m|}=e^{i\gamma}\frac{c-b}{|c-b|},$$

hence

$$\begin{aligned}
&\frac{AM\cdot AN}{AB\cdot AC}+\frac{BM\cdot BN}{BA\cdot BC}+\frac{CM\cdot CN}{CA\cdot CB}\\
=\ &\frac{(m-a)(n-a)}{(a-b)(a-c)}+\frac{(m-b)(n-b)}{(b-a)(b-c)}+\frac{(m-c)(n-c)}{(c-a)(c-b)}.
\end{aligned}$$

The last expression is always equal to 1 which finishes our proof.

64. Let $\angle A=\alpha$, $\angle B=\beta$, $\angle C=\gamma$, $\angle D=\delta$, $\angle E=\varepsilon$, and $\angle F=\varphi$. Applying T1.4 gives us

$$\frac{b-c}{|b-c|}=e^{i\beta}\frac{b-a}{|b-a|},\quad \frac{d-e}{|d-e|}=e^{i\delta}\frac{d-c}{|d-c|},\quad \frac{f-a}{|f-a|}=e^{i\varphi}\frac{f-e}{|f-e|}.$$

Multiplying these equalities and using the given conditions (from the conditions of the problem we read $e^{i(\beta+\delta+\varphi)}=1$) we get

$$(b-c)(d-e)(f-a)=(b-a)(d-c)(f-e).$$

From here we can immediately conclude that

$$(b-c)(a-e)(f-d)=(c-a)(e-f)(d-b),$$

and the result follows by placing the modulus in the last expression.

65. We first apply the inversion with repsect to the circle ω. The points a,b,c,e,z are fixed, and the point d is mapped to the intersection of the lines ae and bc. Denote that intersection by s. The circumcircle of the triangle azd is mapped to the circumcircle of the triangle azs, the line bd is mapped to the line bd, hence it is sufficient to prove that bd is the tangent to the circle circumscribed about azs. The last is equivalent to $az\perp sz$.

Let ω be the unit circle and let $b=1$. According to T6.1 we have $c=-1$ and $e=\overline{a}=\dfrac{1}{a}$. We also have $s=\dfrac{a+\overline{a}}{2}=\dfrac{a^2+1}{2a}$. Since $eb\perp ax$ using T1.3 we get

$$\frac{a-x}{\overline{a}-\overline{x}}=-\frac{e-b}{\overline{e}-\overline{b}}=-\frac{1}{a},$$

and since the point x belongs to the chord eb by T2.2 it satisfies $\overline{x}=\dfrac{1+\overline{a}-x}{\overline{a}}$. Solving this system gives sistema dobijamo $x=\dfrac{a^3+a^2+a-1}{2a^2}$. Since y is the midpoint of ax by T6.1

$$y=\frac{a+x}{2}=\frac{3a^3+a^2+a-1}{4a^2}.$$

Since the points b,y,z are colinear and z belongs to the unit circle according to T1.2 and T2.1 we get

$$\frac{b-y}{\overline{b}-\overline{y}}=\frac{b-z}{\overline{b}-\overline{z}}=-z.$$

After simplifying we get $z=\dfrac{1+3a^2}{(3+a^2)a}$. In order to prove that $az\perp zs$ by T1.3 it is sufficient to prove that

$$\frac{a-z}{\overline{a}-\overline{z}}=-\frac{s-z}{\overline{s}-\overline{z}}.$$

The last follows from

$$a-z=\frac{a^4-1}{a(3+a^2)},\quad s-z=\frac{a^4-2a^2+1}{2a(3+a^2)},$$

by conjugation.

66. Assume first that the orthocenters of the given triangles coincide. Assume that the circumcircle of abc is unit. According to T6.3 we have $h=a+b+c$. Consider the rotation with respect to h

for the angle ω in the negative direction. The point a_1 goes to the point a_1' such that a_1, a_1', and h are colinear. Assume that the same rotation maps b_1 to b_1' and c_1 to c_1'. Since the triangles abc and $a_1b_1c_1$ are similar and equally oriented we get that the points b, b_1', h are clinear as well as c, c_1', h. Moreover $a_1'b_1' \parallel ab$ (and similarly for $b_1'c_1'$ and $c_1'a_1'$). Now according to T1.4 $e^{i\omega}(a_1'-h) = (a_1-h)$ (since the rotation is in the negative direction), and since the points a, a_1', h are colinear, according to T1.2 we have $\dfrac{a_1'-h}{a-h} = \lambda \in \mathbf{R}$. This means that $a_1 = h + \lambda e^{i\omega}(a-h)$ and analogously

$$b_1 = h + \lambda e^{i\omega}(b-h), \quad c_1 = h + \lambda e^{i\omega}(c-h).$$

Since the point a_1 belongs to the chord bc of the unit circle, by T2.2 we get $\overline{a_1} = \dfrac{b+c-a_1}{bc}$. On the other hand by conjugation of the previous expression for a_1 we get $\overline{a_1} = \overline{h} + \lambda \dfrac{\overline{a}-\overline{h}}{e^{i\omega}}$. Solving for λ gives

$$\lambda = \frac{e^{i\omega}(a(a+b+c)+bc)}{a(b+c)(e^{ie\omega}+1)}. \quad (1)$$

Since λ has the same role in the formulas for b_1 also, we must also have

$$\lambda = \frac{e^{i\omega}(b(a+b+c)+ac)}{b(a+c)(e^{ie\omega}+1)}. \quad (2)$$

By equating (1) and (2) we get

$$\begin{aligned} & ab(a+c)(a+b+c) + b^2c(a+c) - ab(b+c)(a+b+c) - a^2c(b+c) \\ = \;& (a-b)(ab(a+b+c) - abc - ac^2 - bc^2) = (a^2-b^2)(ab-c^2). \end{aligned}$$

Since $a^2 \neq b^2$ we conclude $ab = c^2$. Now we will prove that this is necessair condition for triangle abc to be equilateral, i.e. $|a-b| = |a-c|$. After squaring the last expression we get that the triangle is equilateral if and only if $0 = \dfrac{(a-c)^2}{ac} - \dfrac{(a-b)^2}{ab} = \dfrac{(b-c)(a^2-bc)}{abc}$, and since $b \neq c$, this part of the problem is solved.

Assume now that the incenters of the given triangles coincide. Assume that the incircle of the triangle abc is unit and let d, e, f be the points of tangency of the incircle with the sides ab, bc, ca respectively. Similarly to the previous part of the problem we prove

$$a_1 = i + \lambda e^{i\omega}(a-i), \quad b_1 = i + \lambda e^{i\omega}(b-i), \quad c_1 = i + \lambda e^{i\omega}(c-i).$$

Together with the condition $i = 0$ T2.3 and conjugation imply $\overline{a_1} = \dfrac{2\lambda}{e^{i\omega}(e+f)}$. Also, since the points a_1, b, c are colinear we have $a_1d \perp di$ hence according to T1.3 $\dfrac{a_1-d}{\overline{a_1}-\overline{d}} = -\dfrac{d-i}{\overline{d}-\overline{i}} = -d^2$. Solving this system gives

$$\lambda = \frac{d(e+f)}{d^2 + efe^{i\omega}}.$$

Since λ has the same roles in the formulas for a_1 and b_1 we must have

$$\lambda = \frac{e(d+f)}{e^2 + dfe^{i\omega}},$$

and equating gives us

$$e^{i2\omega} = \frac{ed(e+d+f)}{f(de+ef+fd)}.$$

Symmetry implies $e^{i2\omega} = \dfrac{ef(e+d+f)}{d(de+ef+fd)}$ and since $f^2 \neq d^2$ we must have $e+d+f=0$. It is easy to prove that the triangle def is equilateral in this case as well as abc.

67. Since $(a-b)(c-d)+(b-c)(a-d)=(a-c)(b-d)$ the triangle inequality implies $|(a-b)(c-d)|+|(b-c)(a-d)| \geq |(a-c)(b-d)|$, which is exactly an expression of the required inequality. The equality holds if and only if the vectors $(a-b)(c-d)$, $(b-c)(a-d)$, and $(a-c)(c-d)$ are colinear. The first two of them are colinear if and only if

$$\frac{(a-b)(c-d)}{(b-c)(a-d)} \in \mathbf{R},$$

which is according to T3 precisely the condition that a,c,b,d belong to a circle. Similarly we prove that the other two vectors are colinear.

68. Since $(d-a)(d-b)(a-b)+(d-b)(d-c)(b-c)+(d-c)(d-a)(c-a)=(a-b)(b-c)(c-a)$, we have $|(d-a)(d-b)(a-b)|+|(d-b)(d-c)(b-c)|+|(d-c)(d-a)(c-a)| \geq |(a-b)(b-c)(c-a)|$ where the equality holds if and only if $(d-a)(d-b)(a-b), (d-b)(d-c)(b-c), (d-c)(d-a)(c-a)$ and $(a-b)(b-c)(c-a)$ are colinear. The condition for colinearity of the first two vectors can be expressed as

$$\frac{(d-a)(a-b)}{(d-c)(b-c)} = \frac{(\overline{d}-\overline{a})(\overline{a}-\overline{b})}{(\overline{d}-\overline{c})(\overline{b}-\overline{c})}.$$

Assume that the circumcircle of abc is unit. Now the given expression can be written as

$$d\overline{d}\,a - a^2\overline{d} - \frac{da}{c} + \frac{a^2}{c} = d\overline{d}\,c - c^2\overline{d} - \frac{dc}{a} + \frac{c^2}{a}$$

and after some algebra $d\overline{d}\,(a-c) = (a-c)\Big((a+c)\Big(\overline{d} + \dfrac{d}{ac} - \dfrac{a+c}{ac}\Big) + 1\Big)$ or

$$d\overline{d} = (a+c)\Big(\overline{d} + \frac{d}{ac} - \frac{a+c}{ac}\Big) + 1.$$

Similarly, from the colinearity of the first and the third vector we get $d\overline{d} = (b+c)\Big(\overline{d} + \dfrac{d}{bc} - \dfrac{b+c}{bc}\Big) + 1$. Substracting the last two expressions yields $(a-b)\Big(\overline{d} - \dfrac{d}{ab} + \dfrac{c^2-ab}{abc}\Big) = 0$, i.e.

$$\overline{d} - \frac{d}{ab} + \frac{c^2-ab}{abc} = 0.$$

Similarly $\overline{d} - \dfrac{d}{ac} + \dfrac{b^2-ac}{abc} = 0$ and after substracting and simplifying we get $d = a+b+c$. It is easy to verify that for $d = a+b+c$, i.e. the orthocenter of the triangle abc, all four of the above mentioned vectors colinear.

13 Problems for Indepent Study

For those who want more, here is the more. Many of the following problems are similar to the problems that are solved above. There are several quite difficult problems (towards the end of the list) which require more attention in choosing the known points, and more time. As in the case with solved problems, I tried to put lot of problems from math competitions from all over the world.

1. (Regional competition 2002, 2nd grade) In the acute-angled triangle ABC, B' and C' are feet of perpendiculars from the vertices B and C respectively. The circle with the diameter AB intersects the

line CC' at the points M and N, and the circle with the diameter AC intersects the line BB' at P and Q. Prove that the quadrilateral $MPNQ$ is cyclic.

2. (Yug TST 2002) Let $ABCD$ be a quadrilateral such that $\angle A = \angle B = \angle C$. Prove that the point D, the circumcenter, and the orthocenter of $\triangle ABC$ are colinear.

3. (Republic competition 2005, 4th grade) The haxagon $ABCDEF$ is inscribed in the circle k. If the lengths of the segments AB, CD, and EF are equal to the radius of the circle k prove that the midpoints of the remaining three edges form an equilateral trinagle.

4. (USA 1997) Three isosceles triangles BCD, CAE, and ABF with the bases BC, CA, and AB respectively are constructed in the exterior of the triangle ABC. Prove that the perpendiculars from A, B, and C to the lines EF, FD, and DE repsectively are concurrent.

5. Prove that the side length of the regular 9-gon is equal to the difference of the largest and the smallest diagonal.

6. If $h_1, h_2, \ldots, h_{2n}$ denote respectively the distances of an arbitrary point P of the circle k circumscribed about the polygon $A_1A_2\ldots A_{2n}$ from the lines that contain the edges A_1A_2, A_2A_3, ..., $A_{2n}A_1$, prove that $h_1h_3\cdots h_{2n-1} = h_2h_4\cdots h_{2n}$.

7. Let $d_1, d_2, \ldots, d_n$ denote the distances of the vertices $A_1, A_2, \ldots, A_n$ of the regular n-gon $A_1A_2\ldots A_n$ from an arbitrary point P of the smaller arc A_1A_n of the circumcircle. Prove that

$$\frac{1}{d_1d_2} + \frac{1}{d_2d_3} + \ldots + \frac{1}{d_{n-1}d_n} = \frac{1}{d_1d_n}.$$

8. Let $A_0A_1\ldots A_{2n}$ be a regular polygon, P a point of the smaller arc A_0A_{2n} of the circumcircle and m an integer such that $0 \leq m < n$. Prove that

$$\sum_{k=0}^{n} PA_{2k}^{2m+1} = \sum_{k=1}^{n} PA_{2k-1}^{2m+1}.$$

9. (USA 2000) Let $ABCD$ be a cyclic quadrilateral and let E and F be feet of perpendiculars from the intersection of the diagonals to the lines AB and CD respectively. Prove that EF if perpendicular to the line passing through the midpoints of AD and BC.

10. Prove that the midpoints of the altitudes of the traingle are colinear if and only if the triangle is rectangular.

11. (BMO 1990) The feet of preprendiculars of the acute angled triangle ABC are A_1, B_1, and C_1. If A_2, B_2, and C_2 denote the points of tangency of the incircle of $\triangle A_1B_1C_1$ prove that the Euler lines of the triangles ABC and $A_2B_2C_2$ coincide.

12. (USA 1993) Let $ABCD$ be a convex quadrilateral whose diagonals AC and BD are perpendicular. Assume that $AC \cup BD = E$. Prove that the points symmetric to E with respect to the lines AB, BC, CD, and DA form a cyclic quadrilateral.

13. (India 1998) Let AK, BL, CM be the altitudes of the triangle ABC, and let H be its orthocenter. Let P be the midpoint of the segment AH. If BH and MK intersect at the point S, and LP and AM in the point T, prove that TS is perpendicular to BC.

14. (Vietnam 1995) Let AD, BE, and CF be the altitudes of the triangle $\triangle ABC$. For each $k \in R$, $k \neq 0$, let A_1, B_1, and C_1 be such that $AA_1 = kAD$, $BB_1 = kBE$, and $CC_1 = kCF$. Find all k such that for every non-isosceles triangle ABC the triangles ABC and $A_1B_1C_1$ are similar.

15. (Iran 2005) Let ABC be a triangle and D,E,F the points on its edges BC, CA, AB respectively such that

$$\frac{BD}{DC}=\frac{CE}{EA}=\frac{AF}{FB}=\frac{1-\lambda}{\lambda}$$

where λ is a real number. Find the locus of circumcenters of the triangles DEF as $\lambda \in \mathbf{R}$.

16. Let H_1 and H_2 be feet of perpendiculars from the orthocenter H of the triangle ABC to the bisectors of external and internal angles at the vertex C. Prove that the line H_1H_2 contains the midpoint of the side AB.

17. Given an acute-angled triangle ABC and the point D in its interior, such that $\angle ADB = \angle ACB + 90^\circ$ and $AB \cdot CD = AD \cdot BC$. Find the ratio

$$\frac{AB \cdot CD}{AC \cdot BD}.$$

18. The lines AM and AN are tangent to the circle k, and an arbitrary line through A intersects k at K and L. Let l be an arbitrary line parallel to AM. Assume that KM and LM intersect the line l at P and Q, respectively. Prove that the line MN bisects the segment PQ.

19. The points D, E, and F are chosen on the edges BC, CA, and AB of the triangle ABC in such a way that $BD = CE = AF$. Prove that the triangles ABC and DEF have the common incenter if and only if ABC is equilateral.

20. Given a cyclic quadrilateral $ABCD$, prove that the incircles of the triangles ABC, BCD, CDA, DAB form an rectangle.

21. (India 1997) Let I be the incenter of the triangle ABC and let D and E be the midpoints of the segments AC and AB respectively. Assume that the lines AB and DI intersect at the point P, and the lines AC and EI at the point Q. Prove that $AP \cdot AQ = AB \cdot AC$ if and only if $\angle A = 60^\circ$.

22. Let M be an interior point of the square $ABCD$. Let A_1,B_1,C_1,D_1 be the intersection of the lines AM,BM,CM,DM with the circle circumscribed about the square $ABCD$ respectively. Prove that

$$A_1B_1 \cdot C_1D_1 = A_1D_1 \cdot B_1C_1.$$

23. Let $ABCD$ be a cyclic quadrilateral, $F = AC \cap BD$ and $E = AD \cap BC$. If M and N are the midpoints of the segments AB and CD prove that

$$\frac{MN}{EF}=\frac{1}{2}\cdot\left|\frac{AB}{CD}-\frac{CD}{AB}\right|.$$

24. (Vietnam 1994) The points A', B', and C' are symmetric to the points A, B, and C with respect to the lines BC, CA, and AB respectively. What are the conditions that $\triangle ABC$ has to satisfy in order for $\triangle A'B'C'$ to be equilateral?

25. Let O be the circumcenter of the triangle ABC and let R be its circumradius. The incircle of the triangle ABC touches the sides BC,CA,AB, at A_1,B_1,C_1 and its radius is r. Assume that the lines determined by the midpoints of AB_1 and AC_1, BA_1 and BC_1, CA_1 and CB_1 intersect at the points C_2, A_2, and B_2. Prove that the circumcenter of the triangle $A_2B_2C_2$ coincides with O, and that its circumradius is $R+\frac{r}{2}$.

26. (India 1994) Let $ABCD$ be a nonisosceles trapezoid such that $AB \parallel CD$ and $AB > CD$. Assume that $ABCD$ is circumscribed about the circle with the center I which tangets CD in E. Let M be the

midpoint of the segment AB and assume that MI and CD intersect at F. Prove that $DE = FC$ if and only if $AB = 2CD$.

27. (USA 1994) Assume that the hexagon $ABCDEF$ is inscribed in the circle, $AB = CD = EF$, and that the diagonals AD, BE, and CF are concurrent. If P is the intersection of the lines AD and CE, prove that $\frac{CP}{PE} = \left(\frac{AC}{CE}\right)^2$.

28. (Vietnam 1999) Let ABC be a troiangle. The points A', B', and C' are the midpoints of the arcs BC, CA, and AB, which don't contain A, B, and C, respectively. The lines $A'B'$, $B'C'$, and $C'A'$ partition the sides of the triangle into six parts. Prove that the "middle" parts are equal if and only if the triangle ABC is equilateral.

29. (IMO 1991 shortlist) Assume that in $\triangle ABC$ we have $\angle A = 60^\circ$ and that IF is parallel to AC, where I is the incenter and F belongs to the line AB. The point P of the segment BC is such that $3BP = BC$. Prove that $\angle BFP = \angle B/2$.

30. (IMO 1997 shortlist) The angle A is the smallest in the triangle ABC. The points B and C divide the circumcircle into two arcs. Let U be the interior point of the arc between B and C which doesn't contain A. The medians of the segments AB and AC intersect the line AU respectively at the points V and W. The lines BV and CW intersect at T. Prove that $AU = TB + TC$.

31. (Vietnam 1993) Let $ABCD$ be a convex quadrilateral such that AB is not parallel to CD and AD is not parallel to BC. The points P, Q, R, and S are chosen on the edges AB, BC, CD, and DA, respectively such that $PQRS$ is a parallelogram. Find the locus of centroids of all such quadrilaterals $PQRS$.

32. The incircle of the triangle ABC touches BC, CA, AB at E, F, G respectively. Let AA_1, BB_1, CC_1 the angular bisectors of the triangle ABC (A_1, B_1, C_1 belong to the corresponding edges). Let K_A, K_B, K_C respectively be the points of tangency of the other tangents to the incircle from A_1, B_1, C_1. Let P, Q, R be the midpoints of the segments BC, CA, AB. Prove that the lines PK_A, QK_B, RK_C intersect on the incircle of the triangle ABC.

33. Assume that I and I_a are the incenter and the excenter corresponding to the edge BC of the triangle ABC. Let II_a intersect the segment BC and the circumcircle of $\triangle ABC$ at A_1 and M respectively (M belongs to I_a and I) and let N be the midpoint of the arc MBA which contains C. Assume that S and T are intersections of the lines NI and NI_a with the circumcircle of $\triangle ABC$. Prove that the points S, T, and A_1 are colinear.

34. (Vietnam 1995) Let AD, BE, CF be the altitudes of the triangle ABC, and let A', B', C' be the points on the altitudes such that

$$\frac{AA'}{AD} = \frac{BB'}{BE} = \frac{CC'}{CF} = k.$$

Find all values for k such that $\triangle A'B'C' \sim \triangle ABC$.

35. Given the triangle ABC and the point T, let P and Q be the feet of perpendiculars from T to the lines AB and AC, respectively and let R and S be the feet of perpendiculars from A to the lines TC and TB, respectively. Prove that the intersection point of the lines PR and QS belongs to the line BC.

36. (APMO 1995) Let $PQRS$ be a cyclic quadrilateral such that the lines PQ and RS are not parallel. Consider the set of all the circles passing through P and Q and all the circles passing through R and S. Determine the set of all points A of tangency of the circles from these two sets.

37. (YugMO 2003, 3-4 grade) Given a circle k and the point P outside of it. The variable line s which contains point P intersects the circle k at the points A and B. Let M and N be the midpoints of

the arcs determined by the points A and B. If C is the point of the segment AB such that

$$PC^2 = PA \cdot PB,$$

prove that the measure of the angle $\angle MCN$ doesn't depend on the choice of s.

38. (YugMO 2002, 2nd grade) Let A_0,A_1,...,A_{2k}, respectevly be the points which divide the circle into $2k+1$ congruent arcs. The point A_0 is connected by the chords to all other points. Those $2k$ chords divide the circle into $2k+1$ parts. Those parts are colored alternatively in white and black in such a way that the number of white parts is by 1 bigger than the number of black parts. Prove that the surface area of teh black part is greater than the surface area of the white part.

39. (Vietnam 2003) The circles k_1 and k_2 touch each other at the point M. The radius of the circle k_1 is bigger than the radius of the circle k_2. Let A be an arbitrary point of k_2 which doesn't belong to the line connecting the centers of the circles. Let B and C be the points of k_1 such that AB and AC are its tangents. The lines BM and CM intersect k_2 again at E and F respectively. The point D is the intersection of the tangent at A with the line EF. Prove that the locus of points D (as A moves along the circle) is a line.

40. (Vietnam 2004) The circles k_1 and k_2 are given in the plane and they intersect at the points A and B. The tangents to k_1 at those points intersect at K. Let M be an arbitrary point of the circle k_1. Assume that $MA \cup k_2 = \{A,P\}$, $MK \cup k_1 = \{M,C\}$, and $CA \cup k_1 = \{A,Q\}$. Prove that the midpoint of the segment PQ belongs to the line MC and that PQ passes through a fixed point as M moves along k_1.

41. (IMO 2004 shortlist) Let $A_1A_2\ldots A_n$ be a regular n-gon. Assume that the points B_1, B_2, ..., B_{n-1} are determined in the following way:

- for $i = 1$ or $i = n-1$, B_i is the midpoint of the segment A_iA_{i+1};
- for $i \neq 1$, $i \neq n-1$, and S intersection of A_1A_{i+1} and A_nA_i, B_i is the intersection of the bisectors of the angle A_iS_{i+1} with A_iA_{i+1}.

Prove that $\angle A_1B_1A_n + \angle A_1B_2A_n + \ldots + \angle A_1B_{n-1}A_n = 180^\circ$.

69. (**Dezargue's Theorem**) The triangles are perspective with respect to a point if and only if they are perspective w.r.t to a line.

42. (IMO 1998 shortlist) Let ABC be a triangle such that $\angle ACB = 2\angle ABC$. Let D be the point of the segment BC such that $CD = 2BD$. The segment AD is extended over the point D to the point E for which $AD = DE$. Prove that

$$\angle ECB + 180^\circ = 2\angle EBC.$$

43. Given a triangle $A_1A_2A_3$ the line p passes through the point P and intersects the segments A_2A_3, A_3A_1, A_1A_2 at the points X_1, X_2, X_3, respectively. Let A_iP intersect the circumcircle of $A_1A_2A_3$ at R_i, for $i = 1,2,3$. Prove that X_1R_1, X_2R_2, X_3R_3 intersect at the point that belongs to the circumcircle of the triangle $A_1A_2A_3$.

44. The points O_1 and O_2 are the centers of the circles k_1 and k_2 that intersect. Let A be one of the intersection points of these circles. Two common tangents are constructed to these circles. BC are EF the chords of these circles with endpoints at the points of tangency of the common chords with the circles (C and F are further from A). If M and N are the midpoints of the segments BC and EF, prove that $\angle O_1AO_2 = \angle MAN = 2\angle CAF$.

45. (BMO 2002) Two circles of different radii intersect at points A and B. The common chords of these circles are MN and ST respectively. Prove that the orthocenters of $\triangle AMN$, $\triangle AST$, $\triangle BMN$, and $\triangle BST$ form a rectangle.

46. (IMO 2004 shortlist) Given a cyclic quadrilateral $ABCD$, the lines AD and BC intersect at E where C is between B and E. The diagonals AC and BD intersect at F. Let M be the midpoint of CD and let $N \neq M$ be the point of the circumcircle of the triangle ABM such that $AN/BN = AM/BM$. Prove that the points E, F, N are colinear.

47. (IMO 1994 shortlist) The diameter of the semicircle Γ belongs to the line l. Let C and D be the points on Γ. The tangents to Γ at C and D intersect the line l respectively at B and A such that the center of the semi-circle is between A and B. Let E be the intersection of the lines AC and BD, and F the foot of perpendicular from E to l. Prove that EF is the bisector of the angle $\angle CFD$.

terug naar echt bestand

Combinatorial Nullstellensatz

Noga Alon *

Abstract

We present a general algebraic technique and discuss some of its numerous applications in Combinatorial Number Theory, in Graph Theory and in Combinatorics. These applications include results in additive number theory and in the study of graph coloring problems. Many of these are known results, to which we present unified proofs, and some results are new.

1 Introduction

Hilbert's Nullstellensatz (see, e.g., [58]) is the fundamental theorem that asserts that if F is an algebraically closed field, and $f, g_1, \ldots, g_m$ are polynomials in the ring of polynomials $F[x_1, \ldots, x_n]$, where f vanishes over all common zeros of $g_1, \ldots, g_m$, then there is an integer k and polynomials $h_1, \ldots, h_m$ in $F[x_1, \ldots, x_n]$ so that

$$f^k = \sum_{i=1}^{n} h_i g_i.$$

In the special case $m = n$, where each g_i is a univariate polynomial of the form $\prod_{s \in S_i}(x_i - s)$, a stronger conclusion holds, as follows.

Theorem 1.1 *Let F be an arbitrary field, and let $f = f(x_1, \ldots, x_n)$ be a polynomial in $F[x_1, \ldots, x_n]$. Let $S_1, \ldots, S_n$ be nonempty subsets of F and define $g_i(x_i) = \prod_{s \in S_i}(x_i - s)$. If f vanishes over all the common zeros of $g_1, \ldots, g_n$ (that is; if $f(s_1, \ldots, s_n) = 0$ for all $s_i \in S_i$), then there are polynomials $h_1, \ldots, h_n \in F[x_1, \ldots, x_n]$ satisfying $deg(h_i) \leq deg(f) - deg(g_i)$ so that*

$$f = \sum_{i=1}^{n} h_i g_i.$$

Moreover, if $f, g_1, \ldots g_n$ lie in $R[x_1, \ldots, x_n]$ for some subring R of F then there are polynomials $h_i \in R[x_1, \ldots, x_n]$ as above.

*Department of Mathematics, Raymond and Beverly Sackler Faculty of Exact Sciences, Tel Aviv University, Tel Aviv, Israel and Institute for Advanced Study, Princeton, NJ 08540, USA. Research supported in part by a grant from the Israel Science Foundation, by a Sloan Foundation grant No. 96-6-2, by an NEC Research Institute grant and by the Hermann Minkowski Minerva Center for Geometry at Tel Aviv University.

As a consequence of the above one can prove the following,

Theorem 1.2 *Let F be an arbitrary field, and let $f = f(x_1, \ldots, x_n)$ be a polynomial in $F[x_1, \ldots, x_n]$. Suppose the degree $deg(f)$ of f is $\sum_{i=1}^{n} t_i$, where each t_i is a nonnegative integer, and suppose the coefficient of $\prod_{i=1}^{n} x_i^{t_i}$ in f is nonzero. Then, if $S_1, \ldots, S_n$ are subsets of F with $|S_i| > t_i$, there are $s_1 \in S_1, s_2 \in S_2, \ldots, s_n \in S_n$ so that*

$$f(s_1, \ldots, s_n) \neq 0.$$

In this paper we prove these two theorems, which may be called *Combinatorial Nullstellensatz*, and describe several combinatorial applications of them. After presenting the (simple) proofs of the above theorems in Section 2, we show, in Section 3 that the classical theorem of Chevalley and Warning on roots of systems of polynomials as well as the basic theorem of Cauchy and Davenport on the addition of residue classes follow as simple consequences. We proceed to describe additional applications in Additive Number Theory and in Graph Theory and Combinatorics in Sections 4,5,6,7 and 8. Many of these applications are known results, proved here in a unified way, and some are new. There are several known results that assert that a combinatorial structure satisfies certain combinatorial property if and only if an appropriate polynomial associated with it lies in a properly defined ideal. In Section 9 we apply our technique and obtain several new results of this form. The final Section 10 contains some concluding remarks and open problems.

2 The proofs of the two basic theorems

To prove Theorem 1.1 we need the following simple lemma proved, for example, in [13]. For the sake of completeness we include the short proof.

Lemma 2.1 *Let $P = P(x_1, x_2, \ldots, x_n)$ be a polynomial in n variables over an arbitrary field F. Suppose that the degree of P as a polynomial in x_i is at most t_i for $1 \leq i \leq n$, and let $S_i \subset F$ be a set of at least $t_i + 1$ distinct members of F. If $P(x_1, x_2, \ldots, x_n) = 0$ for all n-tuples $(x_1, \ldots, x_n) \in S_1 \times S_2 \times \ldots \times S_n$, then $P \equiv 0$.*

Proof. We apply induction on n. For $n = 1$, the lemma is simply the assertion that a non-zero polynomial of degree t_1 in one variable can have at most t_1 distinct zeros. Assuming that the lemma holds for $n - 1$, we prove it for n ($n \geq 2$). Given a polynomial $P = P(x_1, \ldots, x_n)$ and sets S_i satisfying the hypotheses of the lemma, let us write P as a polynomial in x_n- that is,

$$P = \sum_{i=0}^{t_n} P_i(x_1, \ldots, x_{n-1}) x_n^i,$$

where each P_i is a polynomial with x_j-degree bounded by t_j. For each fixed $(n-1)$-tuple

$$(x_1, \ldots, x_{n-1}) \in S_1 \times S_2 \times \ldots \times S_{n-1},$$

the polynomial in x_n obtained from P by substituting the values of $x_1, \ldots, x_{n-1}$ vanishes for all $x_n \in S_n$, and is thus identically 0. Thus $P_i(x_1, \ldots, x_{n-1}) = 0$ for all $(x_1, \ldots, x_{n-1}) \in S_1 \times \ldots \times S_{n-1}$. Hence, by the induction hypothesis, $P_i \equiv 0$ for all i, implying that $P \equiv 0$. This completes the induction and the proof of the lemma. □

Proof of Theorem 1.1. Define $t_i = |S_i| - 1$ for all i. By assumption,

$$f(x_1, \ldots, x_n) = 0 \quad \text{for every } n\text{-tuple} \quad (x_1, \ldots, x_n) \in S_1 \times S_2 \times \ldots \times S_n. \tag{1}$$

For each i, $1 \leq i \leq n$, let

$$g_i(x_i) = \prod_{s \in S_i} (x_i - s) = x_i^{t_i+1} - \sum_{j=0}^{t_i} g_{ij} x_i^j.$$

Observe that,

$$\text{if } x_i \in S_i \text{ then } g_i(x_i) = 0\text{- that is, } x_i^{t_i+1} = \textstyle\sum_{j=0}^{t_i} g_{ij} x_i^j. \tag{2}$$

Let $\overline{f}$ be the polynomial obtained by writing f as a linear combination of monomials and replacing, repeatedly, each occurrence of $x_i^{f_i}$ $(1 \leq i \leq n)$, where $f_i > t_i$, by a linear combination of smaller powers of x_i, using the relations (2). The resulting polynomial $\overline{f}$ is clearly of degree at most t_i in x_i, for each $1 \leq i \leq n$, and is obtained from f by subtracting from it products of the form $h_i g_i$, where the degree of each polynomial $h_i \in F[x_1, \ldots, x_n]$ does not exceed $deg(f) - deg(g_i)$ (and where the coefficients of each h_i are in the smallest ring containing all coefficients of f and $g_1, \ldots, g_n$.) Moreover, $\overline{f}(x_1, \ldots, x_n) = f(x_1, \ldots, x_n)$, for all $(x_1, \ldots, x_n) \in S_1 \times \ldots \times S_n$, since the relations (2) hold for these values of $x_1, \ldots, x_n$. Therefore, by (1), $\overline{f}(x_1, \ldots, x_n) = 0$ for every n-tuple $(x_1, \ldots, x_n) \in S_1 \times \ldots \times S_n$ and hence, by Lemma 2.1, $\overline{f} \equiv 0$. This implies that $f = \sum_{i=1}^n h_i g_i$, and completes the proof. □

Proof of Theorem 1.2. Clearly we may assume that $|S_i| = t_i + 1$ for all i. Suppose the result is false, and define $g_i(x_i) = \prod_{s \in S_i}(x_i - s)$. By Theorem 1.1 there are polynomials $h_1, \ldots, h_n \in F[x_1, \ldots, x_n]$ satisfying $deg(h_j) \leq \sum_{i=1}^n t_i - deg(g_j)$ so that

$$f = \sum_{i=1}^{n} h_i g_i.$$

By assumption, the coefficient of $\prod_{i=1}^n x_i^{t_i}$ in the left hand side is nonzero, and hence so is the coefficient of this monomial in the right hand side. However, the degree of $h_i g_i = h_i \prod_{s \in S_i}(x_i - s)$ is at most $deg(f)$, and if there are any monomials of degree $deg(f)$ in it they are divisible by $x_i^{t_i+1}$. It follows that the coefficient of $\prod_{i=1}^n x_i^{t_i}$ in the right hand side is zero, and this contradiction completes the proof. □

3 Two classical applications

The following theorem, conjectured by Artin in 1934, was proved by Chevalley in 1935 and extended by Warning in 1935. Here we present a very short proof using our Theorem 1.2 above. For simplicity, we restrict ourselves to the case of finite prime fields, though the proof easily extends to arbitrary finite fields.

Theorem 3.1 (cf., e.g., [52]) *Let p be a prime, and let*

$$P_1 = P_1(x_1, \ldots, x_n), P_2 = P_2(x_1, \ldots, x_n), \ldots, P_m = P_m(x_1, \ldots, x_n)$$

be m polynomials in the ring $Z_p[x_1, \ldots, x_n]$. If $n > \sum_{i=1}^{m} deg(P_i)$ and the polynomials P_i have a common zero $(c_1, \ldots, c_n)$, then they have another common zero.

Proof. Suppose this is false, and define

$$f = f(x_1, \ldots, x_n) = \prod_{i=1}^{m}(1 - P_i(x_1, \ldots, x_n)^{p-1}) - \delta \prod_{j=1}^{n} \prod_{c \in Z_p, c \neq c_j} (x_j - c),$$

where δ is chosen so that

$$f(c_1, \ldots, c_n) = 0. \tag{3}$$

Note that this determines the value of δ, and this value is nonzero. Note also that

$$f(s_1, \ldots, s_n) = 0 \tag{4}$$

for all $s_i \in Z_p$. Indeed, this is certainly true, by (3), if $(s_1, \ldots, s_n) = (c_1, \ldots, c_n)$. For other values of $(s_1, \ldots, s_n)$, there is, by assumption, a polynomial P_j that does not vanish on $(s_1, \ldots, s_n)$, implying that $1 - P_j(s_1, \ldots, s_n)^{p-1} = 0$. Similarly, since $s_i \neq c_i$ for some i, the product $\prod_{c \in Z_p, c \neq c_i}(s_i - c)$ is zero and hence so is the value of $f(s_1, \ldots, s_n)$.

Define $t_i = p - 1$ for all i and note that the coefficient of $\prod_{i=1}^{n} x_i^{t_i}$ in f is $-\delta \neq 0$, since the total degree of

$$\prod_{i=1}^{m}(1 - P_i(x_1, \ldots, x_n)^{p-1})$$

is $(p-1)\sum_{i=1}^{m} deg(P_i) < (p-1)n$. Therefore, by Theorem 1.2 with $S_i = Z_p$ for all i we conclude that there are $s_1, \ldots, s_n \in Z_p$ for which $f(s_1, \ldots, s_n) \neq 0$, contradicting (4) and completing the proof. □

The Cauchy-Davenport Theorem, which has numerous applications in Additive Number Theory, is the following.

Theorem 3.2 ([20]) *If p is a prime, and A, B are two nonempty subsets of Z_p, then*

$$|A + B| \geq min\{p, |A| + |B| - 1\}.$$

Cauchy proved this theorem in 1813, and applied it to give a new proof to a lemma of Lagrange in his well known 1770 paper that shows that any integer is a sum of four squares. Davenport formulated the theorem as a discrete analogue of a conjecture of Khintchine (which was proved a few years later by H. Mann) about the Schnirelman density of the sum of two sequences of integers. There are numerous extensions of this result, see, e.g., [45]. The proofs of Theorem 3.2 given by Cauchy and Davenport are based on the same combinatorial idea, and apply induction on $|B|$. A different, algebraic proof has recently been found by the authors of [10], [11], and its main advantage is that it extends easily and gives several related results. As shown below, this proof can be described as a simple application of Theorem 1.2.

Proof of Theorem 3.2. If $|A|+|B| > p$ the result is trivial, since in this case for every $g \in Z_p$ the two sets A and $g-B$ intersect, implying that $A+B = Z_p$. Assume, therefore, that $|A|+|B| \leq p$ and suppose the result is false and $|A+B| \leq |A|+|B|-2$. Let C be a subset of Z_p satisfying $A+B \subset C$ and $|C| = |A|+|B|-2$. Define $f = f(x,y) = \prod_{c \in C}(x+y-c)$ and observe that by the definition of C

$$f(a,b) = 0 \text{ for all } a \in A, b \in B. \tag{5}$$

Put $t_1 = |A|-1, t_2 = |B|-1$ and note that the coefficient of $x^{t_1}y^{t_2}$ in f is the binomial coefficient $\binom{|A|+|B|-2}{|A|-1}$ which is nonzero in Z_p, since $|A|+|B|-2 < p$. Therefore, by Theorem 1.2 (with $n = 2, S_1 = A, S_2 = B$), there is an $a \in A$ and a $b \in B$ so that $f(a,b) \neq 0$, contradicting (5) and completing the proof. $\Box$

4 Restricted sums

The first theorem in this section is a general result, first proved in [11]. Here we observe that it is a simple consequence of Theorem 1.2 above. We also describe some of its applications, proved in [11], which are extensions of the Cauchy Davenport Theorem.

Let p be a prime. For a polynomial $h = h(x_0, x_1, \ldots, x_k)$ over Z_p and for subsets $A_0, A_1, \ldots, A_k$ of Z_p, define

$$\oplus_h \sum_{i=0}^{k} A_i = \{a_0 + a_1 + \ldots + a_k : a_i \in A_i, \quad h(a_0, a_1, \ldots, a_k) \neq 0\}.$$

Theorem 4.1 ([11]) *Let p be a prime and let $h = h(x_0, \ldots, x_k)$ be a polynomial over Z_p. Let $A_0, A_1, \ldots, A_k$ be nonempty subsets of Z_p, where $|A_i| = c_i + 1$ and define $m = \sum_{i=0}^{k} c_i - deg(h)$. If the coefficient of $\prod_{i=0}^{k} x_i^{c_i}$ in*

$$(x_0 + x_1 + \cdots + x_k)^m h(x_0, x_1, \ldots, x_k)$$

is nonzero (in Z_p) then

$$|\oplus_h \sum_{i=0}^{k} A_i| \geq m+1$$

(and hence $m < p$).

Proof Suppose the assertion is false, and let E be a (multi-) set of m (not necessarily distinct) elements of Z_p that contains the set $\oplus_h \sum_{i=0}^{k} A_i$. Let $Q = Q(x_0, \ldots, x_k)$ be the polynomial defined as follows:

$$Q(x_0, \ldots, x_k) = h(x_0, x_1, \ldots x_k) \cdot \prod_{e \in E} (x_0 + \ldots + x_k - e).$$

Note that

$$Q(x_0, \ldots, x_k) = 0 \text{ for all } (x_0, \ldots, x_k) \in (A_0, \ldots, A_k). \tag{6}$$

This is because for each such $(x_0, \ldots, x_k)$ either $h(x_0, \ldots, x_k) = 0$ or $x_0 + \ldots + x_k \in \oplus_h \sum_{i=0}^{k} A_i \subset E$. Note also that $deg(Q) = m + deg(h) = \sum_{i=0}^{k} c_i$ and hence the coefficient of the monomial $x_0^{c_0} \cdots x_k^{c_k}$ in Q is the same as that of this monomial in the polynomial $(x_0 + \ldots + x_k)^m h(x_0, \ldots, x_k)$, which is nonzero, by assumption.

By Theorem 1.2 there are $x_0 \in A_0$, $x_1 \in A_1, \ldots, x_k \in A_k$ such that $Q(x_0, x_1, \ldots, x_k) \neq 0$, contradicting (6) and completing the proof. □

One of the applications of the last theorem is the following.

Proposition 4.2 *Let p be a prime, and let $A_0, A_1, \ldots, A_k$ be nonempty subsets of the cyclic group Z_p. If $|A_i| \neq |A_j|$ for all $0 \leq i < j \leq k$ and $\sum_{i=0}^{k} |A_i| \leq p + \binom{k+2}{2} - 1$ then*

$$|\{a_0 + a_1 + \ldots + a_k : a_i \in A_i, a_i \neq a_j \text{ for all } i \neq j\}| \geq \sum_{i=0}^{k} |A_i| - \binom{k+2}{2} + 1.$$

Note that the very special case of this proposition in which $k = 1$, $A_0 = A$ and $A_1 = A - \{a\}$ for an arbitrary element $a \in A$ implies that if $A \subset Z_p$ and $2|A| - 1 \leq p + 2$ then the number of sums $a_1 + a_2$ with $a_1, a_2 \in A$ and $a_1 \neq a_2$ is at least $2|A| - 3$. This easily implies the following theorem, conjectured by Erdős and Heilbronn in 1964 (cf., e.g., [25]). Special cases of this conjecture have been proved by various researchers ([49], [43], [50], [29]) and the full conjecture has recently been proved by Dias Da Silva and Hamidoune [21], using some tools from linear algebra and the representation theory of the symmetric group.

Theorem 4.3 ([21]) *If p is a prime, and A is a nonempty subset of Z_p, then*

$$|\{a + a' : a, a' \in A, a \neq a'\}| \geq min\{p, 2|A| - 3\}.$$

In order to deduce Proposition 4.2 from Theorem 4.1 we need the following Lemma which can be easily deduced from the known results about the Ballot problem (see, e.g., [44]), as well as from the known connection between this problem and the hook formula for the number of Young tableaux of a given shape. A simple, direct proof is given in [11].

Lemma 4.4 *Let $c_0, \ldots, c_k$ be nonnegative integers and suppose that $\sum_{i=0}^{k} c_i = m + \binom{k+1}{2}$, where m is a nonnegative integer. Then the coefficient of $\prod_{i=0}^{k} x_i^{c_i}$ in the polynomial*

$$(x_0 + x_1 + \ldots + x_k)^m \prod_{k \geq i > j \geq 0} (x_i - x_j)$$

is

$$\frac{m!}{c_0! c_1! \ldots c_k!} \prod_{k \geq i > j \geq 0} (c_i - c_j).$$

□

Let p be a prime, and let $A_0, A_1, \ldots, A_k$ be nonempty subsets of the cyclic group Z_p. Define

$$\oplus_{i=0}^{k} A_i = \{a_0 + a_1 + \ldots + a_k : a_i \in A_i, a_i \neq a_j \text{ for all } i \neq j\}.$$

In this notation, the assertion of Proposition 4.2 is that if $|A_i| \neq |A_j|$ for all $0 \leq i < j \leq k$ and $\sum_{i=0}^{k} |A_i| \leq p + \binom{k+2}{2} - 1$ then

$$|\oplus_{i=0}^{k} A_i| \geq \sum_{i=0}^{k} |A_i| - \binom{k+2}{2} + 1.$$

Proof of Proposition 4.2. Define

$$h(x_0, \ldots, x_k) = \prod_{k \geq i > j \geq 0} (x_i - x_j),$$

and note that for this h, the sum $\oplus_{i=0}^{k} A_i$ is precisely the sum $\oplus_h \sum_{i=0}^{k} A_i$. Suppose $|A_i| = c_i + 1$ and put

$$m = \sum_{i=0}^{k} c_i - \binom{k+1}{2} \quad (= \sum_{i=0}^{k} |A_i| - \binom{k+2}{2}).$$

By assumption $m < p$ and by Lemma 4.4 the coefficient of $\prod_{i=0}^{k} x_i^{c_i}$ in

$$h \cdot (x_0 + \ldots + x_k)^m$$

is

$$\frac{m!}{c_0! c_1! \ldots c_k!} \prod_{k \geq i > j \geq 0} (c_i - c_j),$$

which is nonzero modulo p, since $m < p$ and the numbers c_i are pairwise distinct. Since $m = \sum_{i=0}^{k} c_i + deg(h)$, the desired result follows from Theorem 4.1. □

An easy consequence of Proposition 4.2 is the following. See [11] for the detailed proof.

Theorem 4.5 *Let p be a prime, and let $A_0, \ldots, A_k$ be nonempty subsets of Z_p, where $|A_i| = b_i$, and suppose $b_0 \geq b_1 \ldots \geq b_k$. Define $b'_0, \ldots, b'_k$ by*

$$b'_0 = b_0 \quad \textit{and} \quad b'_i = min\{b'_{i-1} - 1, b_i\}, \ \textit{for } 1 \leq i \leq k. \tag{7}$$

If $b'_k > 0$ then

$$|\oplus_{i=0}^{k} A_i| \geq min\{p, \sum_{i=0}^{k} b'_i - \binom{k+2}{2} + 1\}.$$

Moreover, the above estimate is sharp for all possible values of $p \geq b_0 \geq \ldots \geq b_k$.

The following result of Dias da Silva and Hamidoune [21] is a simple consequence of (a special case of) the above theorem.

Theorem 4.6 ([21]) *Let p be a prime and let A be a nonempty subset of Z_p. Let $s^{\wedge}A$ denote the set of all sums of s distinct elements of A. Then $|s^{\wedge}A| \geq min\{p, s|A| - s^2 + 1\}$.*

Proof. If $|A| < s$ there is nothing to prove. Otherwise put $s = k + 1$ and apply Theorem 4.5 with $A_i = A$ for all i. Here $b'_i = |A| - i$ for all $0 \leq i \leq k$ and hence

$$|(k+1)^{\wedge}A| = |\oplus_{i=0}^{k} A_i| \geq min\{p, \sum_{i=0}^{k}(|A| - i) - \binom{k+2}{2} + 1\}$$

$$= min\{p, (k+1)|A| - \binom{k+1}{2} - \binom{k+2}{2} + 1\} = min\{p, (k+1)|A| - (k+1)^2 + 1\}.$$

□

Another easy application of Theorem 4.1 is the following result, proved in [10].

Proposition 4.7 *If p is a prime and A, B are two nonempty subsets of Z_p, then*

$$|\{a + b : a \in A, b \in B, ab \neq 1\}| \geq min\{p, |A| + |B| - 3\}.$$

The proof is by applying Theorem 4.1 with $k = 1$, $h = x_0x_1 - 1$, $A_0 = A$, $A_1 = B$, and $m = |A| + |B| - 4$. It is also shown in [10] that the above estimate is tight in all nontrivial cases. Additional extensions of the above proposition appear in [11].

5 Set addition in vector spaces over prime fields

A triple (r, s, n) of positive integers satisfies the *Hopf-Stiefel condition* if

$$\binom{n}{k} \quad \text{is even for every integer } k \text{ satisfying } \ n - r < k < s.$$

This condition arises in Topology. However, studying the combinatorial aspects of the well known Hurwitz problem, Yuzvinsky [59] showed that it has an interesting relation to a natural additive problem. he proved that in a vector space of infinite dimension over $GF(2)$, there exist two subsets $A, B \subset V$ satisfying $|A| = r$, $|B| = s$ and $|A + B| \leq n$ if and only if the triple (r, s, n) satisfies the Hopf-Stiefel condition.

Eliahou and Kervaire [23] have shown very recently that this can be proved using the algebraic technique of [10], [11], and generalized this result to an arbitrary prime p, thus obtaining a common generalization of Yuzvinsky's result and the Cauchy Davenport Theorem. Here is a description of their result, and a quick derivation of it from Theorem 1.2. It is worth noting that the same result also follows from the main result of Bollobás and Leader in [18], proved by a different, more combinatorial, approach.

Let us say that a triple (r, s, n) of positive integers satisfies the *Hopf-Stiefel condition with respect to a prime* p if

$$\binom{n}{k} \quad \text{is divisible by } p \text{ for every integer } k \text{ satisfying } \quad n - r < k < s. \tag{8}$$

Let $\beta_p(r, s)$ denote the smallest integer n for which the triple (r, s, n) satisfies (8). We note that it is not difficult to give a recursive formula for $\beta_p(r, s)$, which enables one to compute it quickly, given the representation of r and s in basis p.

Theorem 5.1 ([23], see also [18]) *If A and B are two finite nonempty subsets of a vector space V over $GF(p)$, and $|A| = r$, $|B| = s$, then $|A + B| \geq \beta_p(r, s)$.*

Proof. We may assume that V is finite, and identify it with the finite field F_q of the same cardinality over $GF(p)$. Viewing A and B as subsets of F_q, define $C = A + B$, and assume the assertion is false and $|C| = n < \beta_p(r, s)$. As in the previous section, define

$$Q(x, y) = \prod_{c \in C} (x + y - c),$$

where Q is a polynomial over F_q, and observe that $Q(a, b) = 0$ for all $a \in A, b \in B$. By the definition of $\beta_p(r, s)$ there is some k satisfying $n - r < k < s$ such that $\binom{n}{k}$ is not divisible by p. Therefore, the coefficient of $x^{n-k}y^k$ in the above polynomial is not zero, and since $|A| = r > n - k, |B| = s > k$ there are, by Theorem 1.2, $a \in A$ and $b \in B$ such that $Q(a, b) \neq 0$, contradiction. This completes the proof. □

The authors of [23] have also shown that the estimate in Theorem 5.1 is sharp for all possible r and s. In fact, if A is the set of r vectors whose coordinates correspond to the p-adic representation

of the integers $0, 1, \ldots, r-1$, and B is the set of s vectors whose coordinates correspond to the p-adic representation of the integers $0, 1, \ldots, s-1$, it is not too difficult to check that $A+B$ is the set of of all vectors whose coordinates correspond to the p-adic representation of the integers $0, 1, \ldots, \beta_p(r,s)-1$. For more details and several extensions, see [23].

6 Graphs, subgraphs and cubes

A well known conjecture of Berge and Sauer, proved by Taśkinov [53], asserts that any simple 4-regular graph contains a 3-regular subgraph. This assertion is easily seen to be false for graphs with multiple edges, but as shown in [6] one extra edge suffices to ensure a 3-regular subgraph in this more general case as well. This follows from the case $p = 3$ in the following result, which, as shown below, can be derived quickly from Theorem 1.2.

Theorem 6.1 ([6]) *For any prime p, any loopless graph $G = (V, E)$ with average degree bigger than $2p-2$ and maximum degree at most $2p-1$ contains a p-regular subgraph.*

Proof. Let $(a_{v,e})_{v \in V, e \in E}$ denote the incidence matrix of G defined by $a_{v,e} = 1$ if $v \in e$ and $a_{v,e} = 0$ otherwise. Associate each edge e of G with a variable x_e and consider the polynomial

$$F = \prod_{v \in V} [1 - (\sum_{e \in E} a_{v,e} x_e)^{p-1}] - \prod_{e \in E} (1 - x_e),$$

over $GF(p)$. Notice that the degree of F is $|E|$, since the degree of the first product is at most $(p-1)|V| < |E|$, by the assumption on the average degree of G. Moreover, the coefficient of $\prod_{e \in E} x_e$ in F is $(-1)^{|E|+1} \neq 0$. Therefore, by Theorem 1.2, there are values $x_e \in \{0, 1\}$ such that $F(x_e : e \in E) \neq 0$. By the definition of F, the above vector $(x_e : e \in E)$ is not the zero vector, since for this vector $F = 0$. In addition, for this vector, $\sum_{e \in E} a_{v,e} x_e$ is zero modulo p for every v, since otherwise F would vanish at this point. Therefore, in the subgraph consisting of all edges $e \in E$ for which $x_e = 1$ all degrees are divisible by p, and since the maximum degree is smaller than $2p$ all positive degrees are precisely p, as needed. □

The assertion of Theorem 6.1 is proved in [6] for prime powers p as well, but it is not known if it holds for every integer p. Combining this result with some additional combinatorial arguments, one can show that for every $k \geq 4r$, every loopless k-regular graph contains an r-regular subgraph. For more details and additional results, see [6].

Erdös and Sauer (c.f., e.g., [16], page 399) raised the problem of estimating the maximum number of edges in a simple graph on n vertices that contains no 3-regular subgraph. They conjectured that for every positive ϵ this number does not exceed $n^{1+\epsilon}$, provided n is sufficiently large as a function

of ϵ. This has been proved by Pyber [47], using Theorem 6.1. He proved that any simple graph on n vertices with at least $200n \log n$ edges contains a subgraph with maximum degree 5 and average degree more than 4. This subgraph contains, by Theorem 6.1, a 3-regular subgraph. On the other hand, Pyber, Rödl and Szemerédi [48] proved, by probabilistic arguments, that there are simple graphs on n vertices with at least $\Omega(n \log \log n)$ edges that contain no 3-regular subgraphs. Thus Pyber's estimate is not far from being best possible.

Here is another application of Theorem 1.2, which is not very natural, but demonstrates its versatility.

Proposition 6.2 *Let p be a prime, and let $G = (V, E)$ be a graph on a set of $|V| > d(p-1)$ vertices. Then there is a nonempty subset U of vertices of G such that the number of cliques of d vertices of G that intersect U is 0 modulo p.*

Proof. For each subset I of vertices of G, let $K(I)$ denote the number of copies of K_d in G that contain I. Associate each vertex $v \in V$ with a variable x_v, and consider the polynomial

$$F = \prod_{v \in V} (1 - x_v) - 1 + G,$$

where

$$G = [\sum_{\emptyset \neq I \subset V} (-1)^{|I|+1} K(I) \prod_{i \in I} x_i]^{p-1}$$

over $GF(p)$. Since $K(I)$ is obviously zero for all I of cardinality bigger than d, the degree of this polynomial is $|V|$, as the degree of G is at most $d(p-1) < |V|$. Moreover, the coefficient of $\prod_{v \in V} x_v$ in F is $(-1)^{|V|} \neq 0$. Therefore, by Theorem 1.2, there are $x_v \in \{0, 1\}$ for which $F(x_v : v \in V) \neq 0$. Since F vanishes on the all 0 vector, it follows that not all numbers x_v are zero, and hence that $G(x_v : v \in V) \neq 1$, implying, by Fermat's little Theorem that

$$\sum_{\emptyset \neq I \subset V} (-1)^{|I|+1} K(I) \prod_{i \in I} x_i \equiv 0 (\ mod\ p).$$

However, the left hand side of the last congruence is precisely the number of copies of K_d that intersect the set $U = \{v : \ x_v = 1\}$, by the Inclusion-Exclusion formula. Since U is nonempty, the desired result follows. □

The assertion of the last proposition can be proved for prime powers p as well. See also [8], [4] for some related results. Some versions of these results arise in the study of the minimum possible degree of a polynomial that represents the OR function of n variables in the sense discussed in [54] and its references.

We close this section with a simple geometric result, proved in [7] answering a question of Komjáth. As shown below, this result is also a simple consequence of Theorem 1.2.

Theorem 6.3 ([7]) *Let $H_1, H_2, \ldots, H_m$ be a family of hyperplanes in R^n that cover all vertices of the unit cube $\{0,1\}^n$ but one. Then $m \geq n$.*

Proof. Clearly we may assume that the uncovered vertex is the all zero vector. Let $(a_i, x) = b_i$ be the equation defining H_i, where $x = (x_1, x_2, \ldots, x_n)$, and (a, b) is the inner product between the two vectors a and b. Note that for every i, $b_i \neq 0$, since H_i does not cover the origin. Assume the assertion is false and $m < n$, and consider the polynomial

$$P(x) = (-1)^{n+m+1} \prod_{j=1}^{m} b_j \prod_{i=1}^{n} (x_i - 1) - \prod_{i=1}^{m} [(a_i, x) - b_i].$$

The degree of this polynomial is clearly n, and the coefficient of $\prod_{i=1}^{n} x_i$ in it is $(-1)^{n+m+1} \prod_{j=1}^{m} b_j \neq 0$. Therefore, by Theorem 1.2 there is a point $x \in \{0,1\}^n$ for which $P(x) \neq 0$. This point is not the all zero vector, as P vanishes on it, and therefore it is some other vertex of the cube. But in this case $(a_i, x) - b_i = 0$ for some i (as the vertex is covered by some H_i), implying that P does vanish on this point, a contradiction. □

The above result is clearly tight. Several extensions are proved in [7].

7 Graph Coloring

Graph coloring is arguably the most popular subject in graph theory. An interesting variant of the classical problem of coloring properly the vertices of a graph with the minimum possible number of colors arises when one imposes some restrictions on the colors available for every vertex. This variant received a considerable amount of attention that led to several fascinating conjectures and results, and its study combines interesting combinatorial techniques with powerful algebraic and probabilistic ideas. The subject, initiated independently by Vizing [57] and by Erdős, Rubin and Taylor [27], is usually known as the study of the *choosability* properties of a graph. Tarsi and the author developed in [13] an algebraic technique that has already been applied by various researchers to solve several problems in this area as well as problems dealing with traditional graph coloring. In this section we observe that the basic results of this technique can be derived from Theorem 1.2, and describe various applications. More details on some of these applications can be found in the survey [2].

We start with some notation and background. A *vertex coloring* of a graph G is an assignment of a color to each vertex of G. The coloring is *proper* if adjacent vertices receive distinct colors. The *chromatic number* $\chi(G)$ of G is the minimum number of colors used in a proper vertex coloring of G. An *edge coloring* of G is, similarly, an assignment of a color to each edge of G. It is *proper* if adjacent edges receive distinct colors. The minimum number of colors in a proper edge-coloring of

G is the *chromatic index* $\chi'(G)$ of G. This is clearly equal to the chromatic number of the line graph of G.

If $G = (V, E)$ is a (finite, directed or undirected) graph, and f is a function that assigns to each vertex v of G a positive integer $f(v)$, we say that G is *f-choosable* if, for every assignment of sets of integers $S(v) \subset Z$ to all the vertices $v \in V$, where $|S(v)| = f(v)$ for all v, there is a proper vertex coloring $c : V \mapsto Z$ so that $c(v) \in S(v)$ for all $v \in V$. The graph G is *k-choosable* if it is f-choosable for the constant function $f(v) \equiv k$. The *choice number* of G, denoted $ch(G)$, is the minimum integer k so that G is k-choosable. Obviously, this number is at least the classical chromatic number $\chi(G)$ of G. The choice number of the line graph of G, which we denote here by $ch'(G)$, is usually called the *list chromatic index* of G, and it is clearly at least the chromatic index $\chi'(G)$ of G.

As observed by various researchers, there are many graphs G for which the choice number $ch(G)$ is strictly larger than the chromatic number $\chi(G)$. A simple example demonstrating this fact is the complete bipartite graph $K_{3,3}$. If $\{u_1, u_2, u_3\}$ and $\{v_1, v_2, v_3\}$ are its two vertex-classes and $S(u_i) = S(v_i) = \{1, 2, 3\} \setminus \{i\}$, then there is no proper vertex coloring assigning to each vertex w a color from its class $S(w)$. Therefore, the choice number of this graph exceeds its chromatic number. In fact, it is not difficult to show that, for any $k \geq 2$, there are bipartite graphs whose choice number exceeds k. Moreover, in [2] it is proved, using probabilistic arguments, that for every k there is some finite $c(k)$ so that the choice number of every simple graph with minimum degree at least $c(k)$ exceeds k.

In view of this, the following conjecture, suggested independently by various researchers including Vizing, Albertson, Collins, Tucker and Gupta, which apparently appeared first in print in the paper of Bollobás and Harris ([17]), is somewhat surprising.

Conjecture 7.1 (The list coloring conjecture) *For every graph G, $ch'(G) = \chi'(G)$.*

This conjecture asserts that for *line graphs* there is no gap at all between the choice number and the chromatic number. Many of the most interesting results in the area are proofs of special cases of this conjecture, which is still wide open. An asymptotic version of it, however, has been proven by Kahn [38] using probabilistic arguments: for simple graphs of maximum degree d, $ch'(G) = (1 + o(1))d$, where the $o(1)$-term tends to zero as d tends to infinity. Since in this case $\chi'(G)$ is either d or $d + 1$, by Vizing's theorem [56], this shows that the list coloring conjecture is asymptotically nearly correct.

The *graph polynomial* $f_G = f_G(x_1, x_2, \ldots, x_n)$ of a directed or undirected graph $G = (V, E)$ on a set $V = \{v_1, \ldots, v_n\}$ of n vertices is defined by $f_G(x_1, x_2, \ldots, x_n) = \Pi\{(x_i - x_j) : i < j\ ,\ \{v_i, v_j\} \in E\}$. This polynomial has been studied by various researchers, starting already with Petersen [46] in 1891. See also, for example, [51], [40].

A subdigraph H of a directed graph D is called *Eulerian* if the indegree $d_H^-(v)$ of every vertex v of H is equal to its outdegree $d_H^+(v)$. Note that we do not assume that H is connected. H is *even* if it has an even number of edges, otherwise, it is *odd*. Let $EE(D)$ and $EO(D)$ denote the numbers of even and odd Eulerian subgraphs of D, respectively. (For convenience we agree that the empty subgraph is an even Eulerian subgraph.) The following result is proved in [13].

Theorem 7.2 *Let $D = (V, E)$ be an orientation of an undirected graph G, denote $V = \{1, 2, , \ldots, n\}$ and define $f : V \mapsto Z$ by $f(i) = d_i + 1$, where d_i is the outdegree of i in D. If $EE(D) \neq EO(D)$, then D is f-choosable.*

Proof (sketch): For $1 \leq i \leq n$, let $S_i \subset Z$ be a set of $d_i + 1$ distinct integers. The existence of a proper coloring of D assigning to each vertex i a color from its list S_i is equivalent to the existence of colors $c_i \in S_i$ such that $f_G(c_1, c_2, \ldots, c_n) \neq 0$.

Since the degree of f_G is $\sum_{i=1}^n d_i$, it suffices to show that the coefficient of $\prod_{i=1}^n x_i^{d_i}$ in f_G is nonzero in order to deduce the existence of such colors c_i from Theorem 1.2. This can be done by interpreting this coefficient combinatorially.

It is not too difficult to see that the coefficients of the monomials that appear in the standard representation of f_G as a linear combination of monomials can be expressed in terms of the orientations of G as follows. Call an orientation D of G *even* if the number of its directed edges (i, j) with $i > j$ is even, otherwise call it *odd*. For non-negative integers $d_1, d_2, \ldots, d_n$, let $DE(d_1, \ldots, d_n)$ and $DO(d_1, \ldots, d_n)$ denote, respectively, the sets of all even and odd orientations of G in which the outdegree of the vertex v_i is d_i, for $1 \leq i \leq n$. In this notation, one can check that

$$f_G(x_1, \ldots, x_n) = \sum_{d_1, \ldots, d_n \geq 0} (|DE(d_1, \ldots, d_n)| - |DO(d_1, \ldots, d_n)|) \Pi_{i=1}^n x_i^{d_i} \ .$$

Consider, now, the given orientation D which lies in $DE(d_1, \ldots, d_n) \cup DO(d_1, \ldots, d_n)$. For any orientation $D_2 \in DE(d_1, \ldots, d_n) \cup DO(d_1, \ldots, d_n)$, let $D \oplus D_2$ denote the set of all oriented edges of D whose orientation in D_2 is in the opposite direction. Since the outdegree of every vertex in D is equal to its outdegree in D_2, it follows that $D \oplus D_2$ is an Eulerian subgraph of D. Moreover, $D \oplus D_2$ is even as an Eulerian subgraph if and only if D and D_2 are both even or both odd. The mapping $D_2 \longrightarrow D \oplus D_2$ is clearly a bijection between $DE(d_1, \ldots, d_n) \cup DO(d_1, \ldots, d_n)$ and the set of all Eulerian subgraphs of D. In case D is even, it maps even orientations to even (Eulerian) subgraphs, and odd orientations to odd subgraphs. Otherwise, it maps even orientations to odd subgraphs, and odd orientations to even subgraphs. In any case,

$$\Big| |DE(d_1, \ldots, d_n)| - |DO(d_1, \ldots, d_n)| \Big| = |EE(D) - EO(D)|.$$

Therefore, the absolute value of the coefficient of the monomial $\Pi_{i=1}^{n} x_i^{d_i}$ in the standard representation of $f_G = f_G(x_1, \ldots, x_n)$ as a linear combination of monomials, is $|EE(D) - EO(D)|$. In particular, if $EE(D) \neq EO(D)$, then this coefficient is not zero and the desired result follows from Theorem 1.2. □

An interesting application of Theorem 7.2 has been obtained by Fleischner and Stiebitz in [28], solving a problem raised by Du, Hsu and Hwang in [22], as well as a strengthening of it suggested by Erdős.

Theorem 7.3 ([28]) *Let G be a graph on $3n$ vertices, whose set of edges is the disjoint union of a Hamilton cycle and n pairwise vertex-disjoint triangles. Then the choice number and the chromatic number of G are both 3.*

The proof is based on a subtle parity argument that shows that, if D is the digraph obtained from G by directing the Hamilton cycle as well as each of the triangles cyclically, then $EE(D) - EO(D) \equiv 2 (mod\ 4\ )$. The result thus follows from Theorem 7.2.

Another application of Theorem 7.2 together with some additional combinatorial arguments is the following result, that solves an open problem from [27].

Theorem 7.4 ([13]) *The choice number of every planar bipartite graph is at most 3.*

This is tight, since $ch(K_{2,4}) = 3$.

Recall that the list coloring conjecture (Conjecture 7.1) asserts that $ch'(G) = \chi'(G)$ for every graph G. In order to try to apply Theorem 7.2 for tackling this problem, it is useful to find a more convenient expression for the difference $EE(D) - EO(D)$, where D is the appropriate orientation of a given line graph. Such an expression is described in [2] for line graphs of d-regular graphs of chromatic index d. This expression is the sum, over all proper d-edge colorings of the graph, of an appropriately defined *sign* of the coloring. See [2] for more details, and [35] for a related discussion. Combining this with a known result of [55] (which asserts that for planar cubic graphs of chromatic index 3 all proper 3-edge colorings have the same sign), and with the Four Color Theorem, the following result, observed by F. Jaeger and M. Tarsi, follows immediately:

Corollary 7.5 *For every 2-connected cubic planar graph G, $ch'(G) = 3$.*

Note that the above result is a strengthening of the Four Color Theorem, which is well known to be equivalent to the fact that the chromatic index of any such graph is 3.

As shown in [24], it is possible to extend this proof to any d-regular planar multigraph with chromatic index d.

Another interesting application of the algebraic method described above appears in [33], where the authors apply it to show that the list coloring conjecture holds for complete graphs with an odd number of vertices, and to improve the error term in the asymptotic estimate of Kahn for the maximum possible list chromatic index of a simple graph with maximum degree d. Finally we mention that Galvin [30] proved recently that the list coloring conjecture holds for any bipartite multigraph, by an elementary, non-algebraic method.

8 The permanent lemma

The following lemma is a slight extension of a lemma proved in [12]. As shown below, it is an immediate corollary of Theorem 1.2 and has several interesting applications.

Lemma 8.1 (The permanent lemma) *Let $A = (a_{ij})$ be an n by n matrix over a field F, and suppose its permanent $Per(A)$ is nonzero (over F). Then for any vector $b = (b_1, b_2, \ldots, b_n) \in F^n$ and for any family of sets $S_1, S_2, \ldots, S_n$ of F, each of cardinality 2, there is a vector $x \in S_1 \times S_2 \times \ldots \times S_n$ such that for every i the i^{th} coordinate of Ax differs from b_i.*

Proof. The polynomial

$$P(x_1, x_2, \ldots, x_n) = \prod_{i=1}^{n} [\sum_{j=1}^{n} a_{ij} x_j - b_j]$$

is of degree n and the coefficient of $\prod_{i=1}^{n} x_i$ in it is $Per(A) \neq 0$. The result thus follows from Theorem 1.2. □

Note that in the special case $S_i = \{0, 1\}$ for every i the above lemma asserts that if the permanent of A is non-zero, then for any vector b, there is a subset of the column-vectors of A whose sum differs from b in all coordinates.

A conjecture of Jaeger asserts that for any field with more than 3 elements and for any nonsingular n by n matrix A over the field, there is a vector x so that both x and Ax have non-zero coordinates. Note that for the special case of fields of characteristic 2 this follows immediately from the Permanent Lemma. Simply take b to be the zero vector, let each S_i be an arbitrary subset of size 2 of the field that does not contain zero, and observe that in characteristic 2 the permanent and the determinant coincide, implying that $Per(A) \neq 0$. With slightly more work relying on some simple properties of the permanent function, the conjecture is proved in [12] for every non-prime field. It is still open for prime fields and, in particular, for $p = 5$.

Let $f(n, d)$ denote the minimum possible number f so that every set of f lattice points in the d-dimensional Euclidean space contains a subset of cardinality n whose centroid is also a lattice point.

The problem of determining or estimating $f(n,d)$ was suggested by Harborth [34], and studied by various authors.

It is convenient to reformulate the definition of $f(n,d)$ in terms of sequences of elements of the abelian group Z_n^d. In these terms, $f(n,d)$ is the minimum possible f so that every sequence of f members of Z_n^d contains a subsequence of size n the sum of whose elements (in the group) is 0.

By an old result of Erdős, Ginzburg and Ziv [26], $f(n,1) = 2n-1$ for all n. The main part in the proof of this statement is its proof for prime values of $n = p$, as the general case can then be easily proved by induction.

Proposition 8.2 ([26]) *For any prime p, any sequence of $2p-1$ members of Z_p contains a subsequence of cardinality p the sum of whose members is 0 (in Z_p).*

There are many proofs of this result. Here is one using the permanent lemma. Given $2p-1$ members of Z_p, renumber them $a_1, a_2, \ldots, a_{2p-1}$ such that $0 \leq a_1 \leq \ldots \leq a_{2p-1}$. If there is an $i \leq p-1$ such that $a_i = a_{i+p-1}$ then $a_i + a_{i+1} + \ldots + a_{i+p-1} = 0$, as needed. Otherwise, let A denote the $p-1$ by $p-1$ all 1 matrix, and define $S_i = \{a_i, a_{i+p-1}\}$ for all $1 \leq i \leq p-1$. Let $b_1, \ldots, b_{p-1}$ be the set of all elements of Z_p besides $-a_{2p-1}$. Since $Per(A) = (p-1)! \neq 0$, by Lemma 8.1, there are $s_i \in S_i$ such that the sum $\sum_{j=1}^{p-1} s_i$ differs from each b_j and is thus equal to $-a_{2p-1}$. Hence, in Z_p,

$$a_{2p-1} + \sum_{i=1}^{p-1} s_i = 0,$$

completing the proof. □

Kemnitz [39] conjectured that $f(n,2) = 4n-3$, observed that $f(n,2) \geq 4n-3$ for all n and proved his conjecture for $n = 2, 3, 5$ and 7. As in the one dimensional case, it suffices to prove this conjecture for prime values p. In [5] it is shown that $f(p,2) \leq 6p-5$ for every prime p. The details are somewhat complicated, but the main tool is again the Permanent Lemma mentioned above.

An *additive basis* in a vector space Z_p^n is a collection C of (not necessarily distinct) vectors, so that for every vector u in Z_p^n there is a subset of C the sum of whose elements is u. Motivated by the study of universal flows in graphs, Jaeger, Linial, Payan and Tarsi [36] conjectured that for every prime p there exists a constant $c(p)$, such that any union of $c(p)$ **linear** bases of Z_p^n contains an additive basis. This conjecture is still open, but in [9] it is shown that any union of $\lceil (p-1)\log_e n \rceil + p - 2$ linear bases of Z_p^n contains such an additive basis. Here, too, the permanent lemma plays a crucial role in the proof. The main idea is to observe how it can be applied to give equalities rather than inequalities (extending the very simple application described in the proof of Proposition 8.2 above.) Here is the basic approach. For a vector v of length n over Z_p, let v^* denote the tensor product of v with the all one vector of length $p-1$. Thus v^* is a vector of length $(p-1)n$ obtained by

concatenating $(p-1)$ copies of v. In this notation, the following result follows from the permanent lemma.

Lemma 8.3 *Let $S = (v_1, v_2, \ldots, v_{(p-1)n})$ be a sequence of $(p-1)n$ vectors of length n over Z_p, and let A be the $(p-1)n$ by $(p-1)n$ matrix whose columns are the vectors $v_1^*, v_2^*, \ldots, v_{(p-1)n}^*$. If $Per(A) \neq 0$ (over Z_p), then the sequence S is an additive basis of Z_p^n.*

Proof. For any vector $b = (b_1, b_2, \ldots, b_n)$, let u_b be the concatenation of the $(p-1)$ vectors $b+j, b+2j, \ldots, b+(p-1)j$, where j is the all one vector of length n. By the Permanent Lemma with all sets $S_i = \{0,1\}$, there is a subset $I \subset \{1,2,\ldots,(p-1)n\}$ such that the sum $\sum_{i\in I} v_i^*$ differs from u_b in all coordinates. This supplies $(p-1)$ forbidden values for every coordinate of the sum $\sum_{i\in I} v_i$, and hence implies that $\sum_{i\in I} v_i = b$. Since b was arbitrary, this completes the proof. □

In [9] it is shown that from any set consisting of all elements in the union of an appropriate number of linear bases of Z_p^n it is possible to choose $(p-1)n$ vectors satisfying the assumptions of the lemma. This is done by applying some properties of the permanent function. The details can be found in [9]. The following conjecture seems plausible, and would imply, if true, that the union of any set of p bases of Z_p^n is an additive basis.

Conjecture 8.4 *For any p nonsingular n by n matrices $A_1, A_2, \ldots, A_p$ over Z_p, there is an n by pn matrix C such that the pn by pn matrix*

$$M' = \begin{bmatrix} A_1 & A_2 & \ldots & A_{p-1} & A_p \\ A_1 & A_2 & \ldots & A_{p-1} & A_p \\ . & . & . & . & . \\ . & . & . & . & . \\ A_1 & A_2 & \ldots & A_{p-1} & A_p \\ & & C & & \end{bmatrix}$$

has a nonzero permanent over Z_p.

We close this section with a simple result about directed graphs. A *one-regular* subgraph of a digraph is a subgraph of it in which all outdegrees and all indegrees are precisely 1 (that is: a spanning subgraph which is a union of directed cycles.)

Proposition 8.5 *Let $D = (V, E)$ be a digraph containing a one-regular subgraph. Then, for any assignment of a set S_v of two reals for each vertex v of V, there is a choice $c(v) \in S_v$ for every v, so that for every vertex u the sum $\sum_{v:\ (u,v)\in E} c(v) \neq 0$.*

Proof. Let $A = (a_{u,v})$ be the adjacency matrix of D defined by $a_{u,v} = 1$ iff $(u,v) \in E$ and $a_{u,v} = 0$ otherwise. By the assumption, the permanent of A over the reals is strictly positive. The result thus follows from the permanent lemma. □

9 Ideals of polynomials and combinatorial properties

There are several known results that assert that a combinatorial structure satisfies a certain combinatorial property if and only if an appropriate polynomial associated with it lies in a properly defined ideal. Here are three known results of this type, all applying the graph polynomial defined in Section 7.

Theorem 9.1 (Li and Li, [40]) *A graph G does not contain an independent set of $k+1$ vertices if and only if the graph polynomial f_G lies in the ideal generated by all graph polynomials of unions of k pairwise vertex disjoint complete graphs that span its set of vertices.*

Theorem 9.2 (Kleitman and Lovász, [41], [42]) *A graph G is not k colorable if and only if the graph polynomial f_G lies in the ideal generated by all graph polynomials of complete graphs on $k+1$ vertices.*

Theorem 9.3 (Alon and Tarsi, [13]) *A graph G on the n vertices $\{1, 2, \ldots, n\}$ is not k colorable if and only if the graph polynomial f_G lies in the ideal generated by the polynomials $x_i^k - 1$, $(1 \leq i \leq n)$.*

Here is a quick proof of the last theorem, using Theorem 1.1.

Proof of Theorem 9.3. If f_G lies in the ideal generated by the polynomials $x_i^k - 1$ then it vanishes whenever each x_i attains a value which is a k^{th} root of unity. This means that in any coloring of the vertices of G by the k^{th} roots of unity, there is a pair of adjacent vertices that get the same color, implying that G is not k-colorable.
Conversely, suppose G is not k-colorable. Then f_G vanishes whenever each of the polynomials $g_i(x_i) = x_i^k - 1$ vanishes, and thus, by Theorem 1.1, f_G lies in the ideal generated by these polynomials. □

As described in Section 7, there are several interesting combinatorial consequences that can be derived from (some versions of) Theorem 9.3, but even without any consequences, such theorems are interesting in their own. One reason for this is that these theorems characterize $coNP$-complete properties, which, according to the common belief that the complexity classes NP and $coNP$ differ, cannot be checked by a polynomial time algorithm.

Using Theorem 1.1 it is not difficult to generate results of this type. We illustrate this with two examples, described below. Many other results can be formulated and proved in a similar manner. It would be nice to deduce any interesting combinatorial consequences of these results or their relatives.

The *bandwidth* of a graph $G = (V, E)$ on n vertices is the minimum integer k such that there is a bijection $f : V \mapsto \{1, 2, \ldots, n\}$ satisfying $|f(u) - f(v)| \leq k$ for every edge $uv \in E$. This invariant has been studied extensively by various researchers. See, e.g., [19] for a survey.

Proposition 9.4 *The bandwidth of a graph $G = (V, E)$ on a set $V = \{1, 2, \ldots, n\}$ of n vertices is at least $k + 1$ if and only if the polynomial*

$$Q_{G,k}(x_1, \ldots, x_n) = \prod_{1 \le i < j \le n} (x_i - x_j) \quad \prod_{ij \in E, i<j} \prod_{k<|l|<n} (x_i - x_j - l)$$

lies in the ideal generated by the polynomials

$$\{g_i(x_i) = \prod_{j=1}^{n} (x_i - j), \quad 1 \le i \le n\}.$$

Proof. If $Q_{G,k}$ lies in the above mentioned ideal, then it vanishes whenever we substitute a value in $\{1, 2, \ldots, n\}$ for each x_i. In particular, it vanishes when we substitute distinct values for these variables, implying that there is some edge $ij \in E$ for which $|x_i - x_j| > k$, and hence the bandwidth of G exceeds k.

Conversely, assume the bandwidth of G exceeds k. We claim that in this case $Q_{G,k}(x_1, \ldots, x_n)$ vanishes whenever each x_i attains a value in $\{1, 2, \ldots, n\}$. Indeed, if two of the variables attain the same value, the first product $(\prod_{1 \le i < j \le n}(x_i - x_j))$ in the definition of $Q_{G,k}$ vanishes. Else, the numbers x_i form a permutation of the members of $\{1, 2, \ldots, n\}$ and thus, by the assumption on the bandwidth, there is some edge $ij \in E$ for which $|x_i - x_j| > k$, implying that the polynomial vanishes in this case as well. Therefore, $Q_{G,k}$ vanishes whenever each x_i lies in $\{1, 2, \ldots, n\}$ and thus, by Theorem 1.1, it lies in the ideal generated by the polynomials $g_i(x_i)$, completing the proof. □

A *hypergraph* H is a pair (V, E), where V is a finite set, whose elements are called *vertices*, and E is a collection of subsets of V, called *edges*. It is *k-uniform* if each edge contains precisely k vertices. Thus, a 2-uniform hypergraph is simply a graph. H is 2-*colorable* if there is a vertex coloring of H with two colors so that no edge is monochromatic.

Proposition 9.5 *The 3-uniform hypergraph $H = (V, E)$ is not 2-colorable if and only if the polynomial*

$$\prod_{e \in E} [(\sum_{v \in e} x_v)^2 - 9]$$

lies in the ideal generated by the polynomials $\{x_v^2 - 1 : \ v \in V\}$.

Proof. The proof is similar to the previous one. If the polynomial lies in that ideal, then it vanishes whenever each x_v attains a value in $\{-1, 1\}$, implying that some edge is monochromatic in each vertex coloring by $\{-1, 1\}$, and hence implying that H is not 2-colorable. Conversely, if H is not 2-colorable, then in every vertex coloring by the numbers -1 and $+1$ some edge is monochromatic, implying that the polynomial vanishes in each such point, and thus showing, by Theorem 1.1, that it lies in the above ideal. □

Note that since the properties characterized in any of the theorems in this section are $coNP$-complete, it is possible to use the usual reductions and obtain, for each $coNP$-complete problem, a characterization in terms of some ideals of polynomials. In most cases, however, the known reductions are somewhat complicated, and would thus lead to cumbersome polynomials which are not likely to imply any interesting consequences. The results mentioned here are in terms of relatively simple polynomials, and are therefore more likely to be useful.

10 Concluding remarks

The discussion in Section 7 as well as that in Section 9 raises the hope that the polynomial approach might be helpful in the study of the Four Color Theorem. This certainly deserves more attention. Further results in the study of the List Coloring Conjecture (Conjecture 7.1) using the algebraic technique are also desirable.

Most proofs presented in this paper are based on the two basic theorems, proved in Section 2, whose proofs are algebraic, and hence non-constructive in the sense that they supply no efficient algorithm for solving the corresponding algorithmic problems.

In the classification of algorithmic problems according to their complexity, it is customary to try and identify the problems that can be solved efficiently, and those that *probably* cannot be solved efficiently. A class of problems that can be solved efficiently is the class P of all problems for which there are deterministic algorithms whose running time is polynomial in the length of the input. A class of problems that probably cannot be solved efficiently are all the NP-complete problems. An extensive list of such problems appears in [31]. It is well known that if any of them can be solved efficiently, then so can all of them, since this would imply that the two complexity classes P and NP are equal.

Is it possible to modify the algebraic proofs given here so that they yield efficient ways of solving the corresponding algorithmic problems? It seems likely that such algorithms do exists. This is related to questions regarding the complexity of search problems that have been studied by several researchers. See, e.g., [37].

In the study of complexity classes like P and NP one usually considers only decision problems, i.e., problems for which the only two possible answers are "yes" or "no." However, the definitions extend easily to the so called "search" problems, which are problems where a more elaborate output is sought. The search problems corresponding to the complexity classes P and NP are sometimes denoted by FP and FNP.

Consider, for example, the obvious algorithmic problem suggested by Theorem 6.1 (for $p = 3$,

say). Given a simple graph with average degree that exceeds 4 and maximum degree 5, it contains, by this theorem, a 3-regular subgraph. Can we find such a subgraph in polynomial time ?

It seems plausible that finding such a subgraph should not be a very difficult task. However, our proof provides no efficient algorithm for accomplishing this task. The situation is similar with many other algorithmic problems corresponding to the various results presented here. Can we, given an input graph satisfying the assumptions of Theorem 7.3 and given a list of three colors for each of its vertices, find, in polynomial time, a proper vertex coloring assigning each vertex a color from its class ? Similarly, can we color properly the edges of any given planar cubic 2-connected graph using given lists of three colors per edge, in polynomial time ?

These problems remain open. Note, however, that any efficient procedure that finds, for a given input polynomial that satisfies the assumptions of Theorem 1.2, a point $(s_1, s_2, \ldots, s_n)$ satisfying its conclusion, would provide efficient algorithms for most of these algorithmic problems. It would thus be interesting to find such an efficient procedure. See also [1] for a related discussion for other algorithmic problems.

Another computational aspect suggested by the results in Section 9 is the complexity of the representation of polynomials in the form that shows they lie in certain ideals. Thus, for example, by Proposition 9.5, a 3-uniform hypergraph is not 2-colorable iff the polynomial associated with it in that proposition is a linear combination with polynomial coefficients of the polynomials $x_v^2 - 1$. Since the problem of deciding whether such a given input hypergraph is not 2-colorable is $coNP$-complete, the existence of a representation like this that can be checked in polynomial time would imply that the complexity classes NP and $coNP$ coincide, and this is believed not to be the case by most researchers.

In this paper we developed and discussed a technique in which polynomials are applied for deriving combinatorial consequences. There are several other known proof-techniques in Combinatorics which are based on properties of polynomials. The most common and successful one is based on a dimension argument. This is the method of proving an upper bound for the size of a collection of combinatorial structures satisfying certain prescribed properties by associating each structure with a polynomial in some space of polynomials, showing that these polynomials are linearly independent, and then deducing the required bound from the dimension of the corresponding space. There are many interesting results proved in this manner; see, e.g., [32], [14], [15] and [3] for surveys of results of this type.

References

[1] N. Alon, *Non-constructive proofs in Combinatorics*, Proc. of the International Congress of Mathematicians, Kyoto 1990, Japan, Springer Verlag, Tokyo (1991), 1421-1429.

[2] N. Alon, *Restricted colorings of graphs*, in "Surveys in Combinatorics", Proc. 14^{th} British Combinatorial Conference, London Mathematical Society Lecture Notes Series 187, edited by K. Walker, Cambridge University Press, 1993, 1-33.

[3] N. Alon, *Tools from higher algebra*, in: *Handbook of Combinatorics*, (edited by R. Graham, M. Grötschel and L. Lovász), Elseveir and MIT Press (1995), 1749-1783.

[4] N. Alon and Y. Caro, *On three zero-sum Ramsey-type problems*, J. Graph Theory 17 (1993), 177-192.

[5] N. Alon and M. Dubiner, *Zero-sum sets of prescribed size*, in: "Combinatorics, Paul Erdös is Eighty", Bolyai Society, Mathematical Studies, Keszthely, Hungary, 1993, 33-50.

[6] N. Alon, S. Friedland and G. Kalai, *Regular subgraphs of almost regular graphs*, J. Combinatorial Theory Ser. B 37 (1984), 79-91. Also: N. Alon, S. Friedland and G. Kalai, *Every 4-regular graph plus an edge contains a 3-regular subgraph*, J. Combinatorial Theory Ser. B 37 (1984), 92-93.

[7] N. Alon and Z. Füredi, *Covering the cube by affine hyperplanes*, European J. Combinatorics 14 (1993), 79-83.

[8] N. Alon, D. Kleitman, R. Lipton, R. Meshulam, M. Rabin and J. Spencer, *Set systems with no union of cardinality* 0 *modulo* m, Graphs and Combinatorics 7 (1991), 97-99.

[9] N. Alon, N. Linial and R. Meshulam, *Additive bases of vector spaces over prime fields*, J. Combinatorial Theory Ser. A 57 (1991), 203-210.

[10] N. Alon, M. B. Nathanson, and I. Z. Ruzsa, *Adding distinct congruence classes modulo a prime*, Amer. Math. Monthly 102 (1995), 250-255.

[11] N. Alon, M. B. Nathanson, and I. Z. Ruzsa, *The polynomial method and restricted sums of congruence classes*, J. Number Theory 56 (1996), 404-417.

[12] N. Alon and M. Tarsi, *A nowhere-zero point in linear mappings*, Combinatorica 9 (1989), 393-395.

[13] N. Alon and M. Tarsi, *Colorings and orientations of graphs*, Combinatorica 12 (1992), 125-134.

[14] L. Babai and P. Frankl, **Linear Algebra Methods in Combinatorics**, to appear.

[15] A. Blokhuis, *Polynomials in Finite Geometries and Combinatorics*, in "Surveys in Combinatorics", Proc. 14^{th} British Combinatorial Conference, London Mathematical Society Lecture Notes Series 187, edited by K. Walker, Cambridge University Press, 1993, 35-52.

[16] B. Bollobás, **Extremal Graph Theory**, Academic Press, 1978.

[17] B. Bollobás and A. J. Harris, *List colorings of graphs*, Graphs and Combinatorics 1 (1985), 115-127.

[18] B. Bollobás and I. Leader, *Sums in the grid*, Discrete Math. 162 (1996), 31-48.

[19] F. R. K. Chung, *Labelings of graphs*, *Selected Topics in Graph Theory* 3, Academic Press (1988), 151-168.

[20] H. Davenport, *On the addition of residue classes*, J. London Math. Soc. 10 (1935), 30–32, 1935.

[21] J. A. Dias da Silva and Y. O. Hamidoune, *Cyclic spaces for Grassmann derivatives and additive theory*, Bull. London Math. Soc. 26 (1994), 140-146.

[22] D. Z. Du, D. F. Hsu and F. K. Hwang, *The Hamiltonian property of consecutive-d digraphs*, Mathematical and Computer Modelling 17 (1993), 61-63.

[23] S. Eliahou and M. Kervaire, *Sumsets in vector spaces over finite fields*, J. Number Theory 71 (1998), 12-39.

[24] M. N. Ellingham and L. Goddyn, *List edge colorings of some* 1*-factorable multigraphs*, Combinatorica 16 (1996), 343-352.

[25] P. Erdős and R. L. Graham, **Old and New Problems and Results in Combinatorial Number Theory**, L'Enseignement Mathématique, Geneva, 1980.

[26] P. Erdős, A. Ginzburg and A. Ziv, *Theorem in the additive number theory*, Bull. Research Council Israel 10F (1961), 41-43.

[27] P. Erdős, A. L. Rubin and H. Taylor, *Choosability in graphs*, Proc. West Coast Conf. on Combinatorics, Graph Theory and Computing, Congressus Numerantium XXVI, 1979, 125-157.

[28] H. Fleischner and M. Stiebitz, *A solution to a coloring problem of P. Erdős*, Discrete Math. 101 (1992), 39-48.

[29] G. A. Freiman, L. Low, and J. Pitman, *The proof of Paul Erdős' conjecture of the addition of different residue classes modulo a prime number*, In: *Structure Theory of Set Addition*, CIRM Marseille (1993), 99-108.

[30] F. Galvin, *The list chromatic index of a bipartite multigraph*, J. Combinatorial Theory Ser. B 63 (1995), 153-158.

[31] M. R. Garey and D. S. Johnson, **Computers and Intractability, A guide to the Theory of NP-Completeness**, W. H. Freeman and Company, New York, 1979.

[32] C. Godsil, *Tools from linear algebra*, in: *Handbook of Combinatorics*, (edited by R. Graham, M. Grötschel and L. Lovász), Elseveir and MIT Press (1995), 1705-1748.

[33] R. Häggkvist and J. Janssen, *New bounds on the list chromatic index of the complete graph and other simple graphs*, Combin., Prob. and Comput. 6 (1997), 295-313.

[34] H. Harborth, *Ein Extremalproblem für Gitterpunkte*, J. Reine Angew. Math. 262/263 (1973), 356-360.

[35] F. Jaeger, *On the Penrose number of cubic diagrams*, Discrete Math. 74 (1989), 85-97.

[36] F. Jaeger, N. Linial, C. Payan and M. Tarsi, *Group connectivity of graphs- a nonhomogeneous analogue of nowhere-zero flow*, J. Combinatorial Theory Ser. B 56 (1992), 165-182.

[37] D. S. Johnson, C. H. Papadimitriou and M. Yannakakis, *How easy is local search?*, JCSS 37 (1988), 79-100.

[38] J. Kahn, *Asymptotically good list colorings*, J. Combinatorial Theory Ser. A 73 (1996), 1–59.

[39] A. Kemnitz, *On a lattice point problem*, Ars Combinatoria 16b (1983), 151-160.

[40] S. Y. R. Li and W. C. W. Li, *Independence numbers of graphs and generators of ideals*, Combinatorica 1 (1981), 55-61.

[41] L. Lovász, *Bounding the independence number of a graph*, in: Bonn Workshop on Combinatorial Optimization, (A. Bachem, M. Grötschel and B. Korte, eds.), Mathematics Studies 66, Annals of Discrete Mathematics 16, North Holland, Amsterdam, 1982, 213-223.

[42] L. Lovász, *Stable sets and polynomials*, Discrete Math. 124 (1994), 137-153.

[43] R. Mansfield, *How many slopes in a polygon?* Israel J. Math. 39 (1981), 265–272.

[44] M. P. A. Macmahon, **Combinatory Analysis**, Chelsea Publishing Company, 1915, Chapter V.

[45] M. B. Nathanson, **Additive Number Theory: Inverse Theorems and the Geometry of Sumsets**, Springer-Verlag, New York, 1996.

[46] J. Petersen, *Die Theorie der regulären Graphs*, Acta Math. 15 (1891), 193-220.

[47] L. Pyber, *Regular subgraphs of dense graphs*, Combinatorica 5 (1985), 347-349.

[48] L. Pyber, V. Rödl and E. Szemerédi, *Dense Graphs without 3-regular Subgraphs* , J. Combinatorial Theory Ser. B 63 (1995), 41-54.

[49] U.-W. Rickert, *Über eine Vermutung in der additiven Zahlentheorie*, PhD thesis, Tech. Univ. Braunschweig, 1976.

[50] Ö. J. Rödseth, *Sums of distinct residues mod p*, Acta Arith. 65 (1994), 181-184.

[51] D. E. Scheim, *The number of edge* 3*-colorings of a planar cubic graph as a permanent*, Discrete Math. 8 (1974), 377-382.

[52] W. Schmidt, **Equations over Finite Fields, an Elementary Approach**, Lecture Notes in Mathematics, Vol. 536, Springer, Berlin, 1976.

[53] V. A. Taśkinov, *Regular subgraphs of regular graphs*, Soviet Math. Dokl. 26 (1982), 37-38.

[54] S. C. Tsai, *Lower bounds on representing Boolean functions as polynomials in* Z_m, SIAM J. Discrete Math. 9 (1996), 55-62.

[55] L. Vigneron, *Remarques sur les réseaux cubiques de classe* 3 *associés au probléme des quatre couleurs*, C. R. Acad. Sc. Paris, t. 223 (1946), 770-772.

[56] V. G. Vizing, *On an estimate on the chromatic class of a p-graph* (in Russian), Diskret. Analiz. 3 (1964), 25-30.

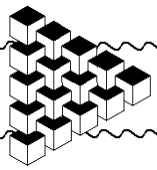
[57] V. G. Vizing, *Coloring the vertices of a graph in prescribed colors* (in Russian), Diskret. Analiz. No. 29, Metody Diskret. Anal. v. Teorii Kodov i Shem 101 (1976), 3-10.

[58] B. L. van der Waerden, **Modern Algebra**, Julius Springer, Berlin, 1931.

[59] S. Yuzvinsky, *Orthogonal pairings of Euclidean spaces*, Michigan Math. J. 28 (1981), 109-119.

terug naar echt bestand

The IMO Compendium Group
Olympiad Training Materials
www.imomath.com



©2007 The Author(s) and The IMO Compendium Group

Classical Inequalities

Ivan Matić

Contents

1 Introduction

This section will start with some basic facts and exercises. Frequent users of this discipline can just skim over the notation and take a look at formulas that talk about generalities in which the theorems will be shown.

The reason for starting with basic principles is the intention to show that the theory is simple enough to be completely derived on 20 pages without using any high-level mathematics. If you take a look at the first theorem and compare it with some scary inequality already mentioned in the table of contents, you will see how huge is the path that we will bridge in so few pages. And that will happen on a level accessible to a beginning high-school student. Well, maybe I exaggerated in the previous sentence, but the beginning high-school student should read the previous sentence again and forget about this one.

Theorem 1. *If x is a real number, then $x^2 \geq 0$. The equality holds if and only if $x = 0$.*

No proofs will be omitted in this text. Except for this one. We have to acknowledge that this is very important inequality, everything relies on it, ..., but the proof is so easy that it makes more sense wasting the space and time talking about its triviality than actually proving it. Do you know how to prove it? Hint: "A friend of my friend is my friend"; "An enemy of my enemy is my friend". It might be useful to notice that "An enemy of my friend is my enemy" and "A friend of my enemy is my enemy", but the last two facts are not that useful for proving theorem 1.

I should also write about the difference between "$\geq$" and "$>$"; that something weird happens when both sides of an inequality are multiplied by a negative number, but I can't imagine myself doing that. People would hate me for real.

Theorem 2. *If $a, b \in \mathbb{R}$ then:*

$$a^2 + b^2 \geq 2ab. \tag{1}$$

The equality holds if and only if $a = b$.

Proof. After subtracting $2ab$ from both sides the inequality becomes equivalent to $(a-b)^2 \geq 0$, which is true according to theorem 1. □

Problem 1. *Prove the inequality $a^2 + b^2 + c^2 \geq ab + bc + ca$, if a, b, c are real numbers.*

Solution. If we add the inequalities $a^2 + b^2 \geq 2ab$, $b^2 + c^2 \geq 2bc$, and $c^2 + a^2 \geq 2ca$ we get $2a^2 + 2b^2 + 2c^2 \geq 2ab + 2bc + 2ca$, which is equivalent to what we are asked to prove. △

Problem 2. *Find all real numbers a, b, c, and d such that*

$$a^2 + b^2 + c^2 + d^2 = a(b + c + d).$$

Solution. Recall that $x^2 + y^2 \geq 2xy$, where the equality holds if and only if $x = y$. Applying this inequality to the pairs of numbers $(a/2, b)$, $(a/2, c)$, and $(a/2, d)$ yields:

$$\frac{a^2}{4} + b^2 \geq ab, \;\; \frac{a^2}{4} + c^2 \geq ac, \;\; \frac{a^2}{4} + d^2 \geq ad.$$

Note also that $a^2/4 > 0$. Adding these four inequalities gives us $a^2 + b^2 + c^2 + d^2 \geq a(b + c + d)$. Equality can hold only if all the inequalities were equalities, i.e. $a^2 = 0$, $a/2 = b$, $a/2 = c$, $a/2 = d$. Hence $a = b = c = d = 0$ is the only solution of the given equation. △

Problem 3. *If a, b, c are positive real numbers that satisfy $a^2 + b^2 + c^2 = 1$, find the minimal value of*

$$S = \frac{a^2 b^2}{c^2} + \frac{b^2 c^2}{a^2} + \frac{c^2 a^2}{b^2}.$$

Solution. If we apply the inequality $x^2 + y^2 \geq 2xy$ to the numbers $x = \dfrac{ab}{c}$ and $y = \dfrac{bc}{a}$ we get

$$\frac{a^2 b^2}{c^2} + \frac{b^2 c^2}{a^2} \geq 2b^2. \tag{2}$$

Similarly we get

$$\frac{b^2 c^2}{a^2} + \frac{c^2 a^2}{b^2} \geq 2c^2, \text{ and} \tag{3}$$

$$\frac{c^2 a^2}{b^2} + \frac{a^2 b^2}{c^2} \geq 2a^2. \tag{4}$$

Summing up (2), (3), and (4) gives $2\left(\frac{a^2b^2}{c^2} + \frac{b^2c^2}{a^2} + \frac{c^2a^2}{b^2}\right) \geq 2(a^2 + b^2 + c^2) = 2$, hence $S \geq 1$. The equality holds if and only if $\dfrac{ab}{c} = \dfrac{bc}{a} = \dfrac{ca}{b}$, i.e. $a = b = c = \dfrac{1}{\sqrt{3}}$. △

Problem 4. *If x and y are two positive numbers less than* 1*, prove that*

$$\frac{1}{1-x^2}+\frac{1}{1-y^2}\geq\frac{2}{1-xy}.$$

Solution. Using the inequality $a+b\geq 2\sqrt{ab}$ we get $\frac{1}{1-x^2}+\frac{1}{1-y^2}\geq\frac{2}{\sqrt{(1-x^2)(1-y^2)}}$. Now we notice that $(1-x^2)(1-y^2)=1+x^2y^2-x^2-y^2\leq 1+x^2y^2-2xy=(1-xy)^2$ which implies $\frac{2}{\sqrt{(1-x^2)(1-y^2)}}\geq\frac{2}{1-xy}$ and this completes the proof. $\triangle$

Since the main focus of this text is to present some more advanced material, the remaining problems will be harder then the ones already solved. For those who want more of the introductory-type problems, there is a real hope that this website will soon get some text of that sort. However, nobody should give up from reading the rest, things are getting very interesting.

Let us return to the inequality (1) and study some of its generalizations. For $a,b\geq 0$, the consequence $\frac{a+b}{2}\geq\sqrt{ab}$ of (1) is called the Arithmetic-Geometric mean inequality. Its left-hand side is called the arithmetic mean of the numbers a and b, and its right-hand side is called the geometric mean of a and b. This inequality has its analogue:

$$\frac{a+b+c}{3}\geq\sqrt[3]{abc},\ a,b,c\geq 0.$$

More generally, for a sequence $x_1,\ldots,x_n$ of positive real numbers, the Arithmetic-Geometric mean inequality holds:

$$\frac{x_1+x_2+\cdots+x_n}{n}\geq\sqrt[n]{x_1\cdot x_2\cdots x_n}. \tag{5}$$

These two inequalities are highly non-trivial, and there are variety of proofs to them. We did (5) for $n=2$. If you try to prove it for $n=3$, you would see the real trouble. What a person tortured with the case $n=3$ would never suspect is that $n=4$ is much easier to handle. It has to do something with 4 being equal $2\cdot 2$ and $3\neq 2\cdot 2$. I believe you are not satisfied by the previous explanation but you have to accept that the case $n=3$ comes after the case $n=4$. The induction argument follows these lines, but (un)fortunately we won't do it here because that method doesn't allow generalizations that we need.

Besides (5) we have the inequality between quadratic and arithmetic mean, namely

$$\sqrt{\frac{x_1^2+x_2^2+\cdots+x_n^2}{n}}\geq\frac{x_1+x_2+\cdots+x_n}{n}. \tag{6}$$

The case of equality in (5) and (6) occurs if and only if all the numbers $x_1,\ldots,x_n$ are equal.

Arithmetic, geometric, and quadratic means are not the only means that we will consider. There are infinitely many of them, and there are infinitely many inequalities that generalize (5) and (6). The beautiful thing is that we will consider all of them at once. For appropriately defined means, a very general inequality will hold, and the above two inequalities will ended up just being consequences.

Definition 1. *Given a sequence $x_1,x_2,\ldots,x_n$ of positive real numbers, the mean of order r, denoted by $M_r(x)$ is defined as*

$$M_r(x)=\left(\frac{x_1^r+x_2^r+\cdots+x_n^r}{n}\right)^{\frac{1}{r}}. \tag{7}$$

Example 1. *$M_1(x_1,\ldots,x_n)$ is the arithmetic mean, while $M_2(x_1,\ldots,x_n)$ is the geometric mean of the numbers $x_1,\ldots,x_n$.*

M_0 can't be defined using the expression (7) but we will show later that as r approaches 0, M_r will approach the geometric mean. The famous mean inequality can be now stated as

$$M_r(x_1,\ldots,x_n) \leq M_s(x_1,\ldots,x_n), \quad \text{for } 0 \leq r \leq s.$$

However we will treat this in slightly greater generality.

Definition 2. *Let $m = (m_1,\ldots,m_n)$ be a fixed sequence of non-negative real numbers such that $m_1+m_2+\cdots+m_n = 1$. Then the weighted mean of order r of the sequence of positive reals $x = (x_1,\ldots,x_n)$ is defined as:*

$$M_r^m(x) = (x_1^r m_1 + x_2^r m_2 + \cdots + x_n^r m_n)^{\frac{1}{r}}. \tag{8}$$

Remark. Sequence m is sometimes called a sequence of masses, but more often it is called a measure, and $M_r^m(x)$ is the L^r norm with repsect to the Lebesgue integral defined by m. I didn't want to scare anybody. I just wanted to emphasize that this hard-core math and not something coming from physics.

We will prove later that as r tends to 0, the weighted mean $M_r^m(x)$ will tend to the weighted geometric mean of the sequence x defined by $G^m(x) = x_1^{m_1}\cdot x_2^{m_2}\cdots x_n^{m_n}$.

Example 2. *If $m_1 = m_2 = \cdots = \frac{1}{n}$ then $M_r^m(x) = M_r(x)$ where $M_r(x)$ is previously defined by the equation (7).*

Theorem 3 (General Mean Inequality)**.** *If $x = (x_1,\ldots,x_n)$ is a sequence of positive real numbers and $m = (m_1,\ldots,m_n)$ another sequence of positive real numbers satisfying $m_1+\cdots+m_n = 1$, then for $0 \leq r \leq s$ we have $M_r^m(x) \leq M_s^m(x)$.*

The proof will follow from the Hölders inequality.

2 Convex Funtions

To prove some of the fundamental results we will need to use convexity of certain functions. Proofs of the theorems of Young, Minkowski, and Hölder will require us to use very basic facts – you should be fine if you just read the definition 3 and example 3. However, the section on Karamata's inequality will require some deeper knowledge which you can find here.

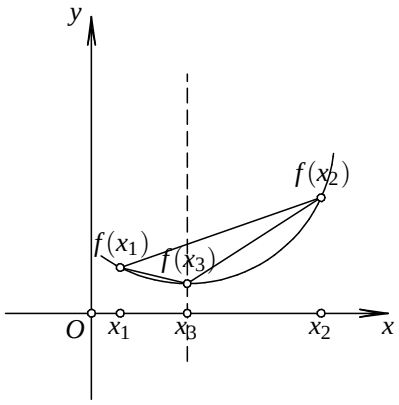
Definition 3. *The function $f : [a,b] \to \mathbb{R}$ is convex if for any $x_1, x_2 \in [a,b]$ and any $\lambda \in (0,1)$ the following inequality holds:*

$$f(\lambda x_1 + (1-\lambda)x_2) \leq \lambda f(x_1) + (1-\lambda) f(x_2). \tag{9}$$

Function is called concave if $-f$ is convex. If the inequality in (9) is strict then the function is called strictly convex.

Now we will give a geometrical interpretation of convexity. Take any $x_3 \in (x_1, x_2)$. There is $\lambda \in (0,1)$ such that $x_2 = \lambda x_1 + (1-\lambda)x_3$. Let's paint in green the line passing through x_3 and parallel to the y axis. Let's paint in red the chord connecting the points $(x_1, f(x_1))$ and $(x_2, f(x_2))$. Assume that the green line and the red chord intersect at the yellow point. The y coordinate (also called the height) of the yellow point is:

$$\lambda f(x_1) + (1-\lambda) f(x_2).$$



The inequality (9) means exactly that the the green line will intersect the graph of a function below the red chord. If f is strictly convex then the equality can hold in (9) if and only if $x_1 = x_2$.

Example 3. *The following functions are convex: e^x, x^p (for $p \geq 1$, $x > 0$), $\frac{1}{x}$ ($x \neq 0$), while the functions $\log x$ ($x > 0$), $\sin x$ ($0 \leq x \leq \pi$) , $\cos x$ ($-\pi/2 \leq x \leq \pi/2$) are concave.*

All functions mentioned in the previous example are elementary functions, and proving the convexity/concavity for them would require us to go to the very basics of their foundation, and we will not do that. In many of the examples and problems respective functions are slight modifications of elementary functions. Their convexity (or concavity) is something we don't have to verify. However, we will develop some criteria for verifying the convexity of more complex combinations of functions.

Let us take another look at our picture above and compare the slopes of the three drawn lines. The line connecting $(x_1, f(x_1))$ with $(x_3, f(x_3))$ has the smallest slope, while the line connecting $(x_3, f(x_3))$ with $(x_2, f(x_2))$ has the largest slope. In the following theorem we will state and prove that the convex function has always an "increasing slope".

Theorem 4. *Let $f : [a,b] \to \mathbb{R}$ be a convex function and $a \leq x_1 < x_3 < x_2 \leq b$. Then*

$$\frac{f(x_3) - f(x_1)}{x_3 - x_1} \leq \frac{f(x_2) - f(x_1)}{x_2 - x_1} \leq \frac{f(x_2) - f(x_3)}{x_2 - x_3}. \tag{10}$$

Proof. We can write $x_3 = \lambda x_1 + (1-\lambda)x_2$ for some $\lambda \in (0,1)$. More precisely $\lambda = \frac{x_2 - x_3}{x_2 - x_1}$, and $1 - \lambda = \frac{x_3 - x_1}{x_2 - x_1}$. From (9) we get

$$f(x_3) \leq \frac{x_2 - x_3}{x_2 - x_1} f(x_1) + \frac{x_3 - x_1}{x_2 - x_1} f(x_2).$$

Subtracting $f(x_1)$ from both sides of the last inequality yields $f(x_3) - f(x_1) = -\frac{x_3 - x_1}{x_2 - x_1} f(x_1) + \frac{x_3 - x_1}{x_2 - x_1} f(x_2)$ giving immediately the first inequality of (10). The second inequality of (10) is obtained in an analogous way. □

The rest of this chapter is using some of the properties of limits, continuity and differentiability. If you are not familiar with basic calculus, you may skip that part, and you will be able to understand most of what follows. The theorem 6 is the tool for verifying the convexity for differentiable functions that we mentioned before. The theorem 5 will be used it in the proof of Karamata's inequality.

Theorem 5. *If $f : (a,b) \to \mathbb{R}$ is a convex function, then f is continuous and at every point $x \in (a,b)$ it has both left and right derivative $f'_-(x)$ and $f'_+(x)$. Both f'_- and f'_+ are increasing functions on (a,b) and $f'_-(x) \leq f'_+(x)$.*

Solution. The theorem 10 implies that for fixed x the function $\varphi(t) = \frac{f(t)-f(x)}{t-x}$, $t \neq x$ is an increasing function bounded both by below and above. More precisely, if t_0 and t_1 are any two numbers from (a,b) such that $t_0 < x < t_1$ we have:

$$\frac{f(x)-f(t_0)}{x-t_0} \leq \varphi(t) \leq \frac{f(t_1)-f(x)}{t_1-x}.$$

This specially means that there are $\lim_{t\to x-} \varphi(t)$ and $\lim_{t\to x+} \varphi(t)$. The first one is precisely the left, and the second one – the right derivative of φ at x. Since the existence of both left and right derivatives implies the continuity, the statement is proved. □

Theorem 6. *If $f : (a,b) \to \mathbb{R}$ is a twice differentiable function. Then f is convex on (a,b) if and only if $f''(x) \geq 0$ for every $x \in (a,b)$. Moreover, if $f''(x) > 0$ then f is strictly convex.*

Proof. This theorem is the immediate consequence of the previous one. □

3 Inequalities of Minkowski and Hölder

Inequalities presented here are sometimes called weighted inequalities of Minkowski, Hölder, and Cauchy-Schwartz. The standard inequalities are easily obtained by placing $m_i = 1$ whenever some m appears in the text below. Assuming that the sum $m_1 + \cdots + m_n = 1$ one easily get the generalized (weighted) mean inequalities, and additional assumption $m_i = 1/n$ gives the standard mean inequalities.

Lemma 1. *If $x,y > 0$, $p > 1$ and $\alpha \in (0,1)$ are real numbers, then*

$$(x+y)^p \leq \alpha^{1-p}x^p + (1-\alpha)^{1-p}y^p. \tag{11}$$

The equality holds if and only if $\frac{x}{\alpha} = \frac{y}{1-\alpha}$.

Proof. For $p > 1$, the function $\varphi(x) = x^p$ is strictly convex hence $(\alpha a + (1-\alpha)b)^p \leq \alpha a^p + (1-\alpha)b^p$. The equality holds if and only if $a = b$. Setting $x = \alpha a$ and $y = (1-\alpha)b$ we get (11) immediately. □

Lemma 2. *If $x_1, x_2, \ldots, x_n, y_1, y_2, \ldots, y_n$ and $m_1, m_2, \ldots, m_n$ are three sequences of positive real numbers and $p > 1$, $\alpha \in (0,1)$, then*

$$\sum_{i=1}^{n}(x_i+y_i)^p m_i \leq \alpha^{1-p}\sum_{i=1}^{n}x_i^p m_i + (1-\alpha)^{1-p}\sum_{i=1}^{n}y_i^p m_i. \tag{12}$$

The equality holds if and only if $\frac{x_i}{y_i} = \frac{\alpha}{1-\alpha}$ for every i, $1 \leq i \leq n$.

Proof. From (11) we get $(x_i+y_i)^p \leq \alpha^{1-p}x_i^p + (1-\alpha)^{1-p}y_i^p$. Multiplying by m_i and adding as $1 \leq i \leq n$ we get (12). The equality holds if and only if $\frac{x_i}{y_i} = \frac{\alpha}{1-\alpha}$. □

Theorem 7 (Minkowski)**.** *If $x_1, x_2, \ldots, x_n, y_1, y_2, \ldots, y_n$, and $m_1, m_2, \ldots, m_n$ are three sequences of positive real numbers and $p > 1$, then*

$$\left(\sum_{i=1}^{n}(x_i+y_i)^p m_i\right)^{1/p} \leq \left(\sum_{i=1}^{n}x_i^p m_i\right)^{1/p} + \left(\sum_{i=1}^{n}y_i^p m_i\right)^{1/p}. \tag{13}$$

The equality holds if and only if the sequences (x_i) and (y_i) are proportional, i.e. if and only if there is a constant λ such that $x_i = \lambda y_i$ for $1 \leq i \leq n$.

Proof. For any $\alpha \in (0,1)$ we have inequality (12). Let us write

$$A = \left(\sum_{i=1}^{n} x_i^p m_i\right)^{1/p}, \quad B = \left(\sum_{i=1}^{n} y_i^p m_i\right)^{1/p}.$$

In new terminology (12) reads as

$$\sum_{i=1}^{n} (x_i + y_i)^p m_i \leq \alpha^{1-p} A^p + (1-\alpha)^{1-p} B^p. \tag{14}$$

If we choose α such that $\frac{A}{\alpha} = \frac{B}{1-\alpha}$, then (11) implies $\alpha^{1-p}A^p + (1-\alpha)^{1-p}B^p = (A+B)^p$ and (14) now becomes

$$\sum_{i=1}^{n} (x_i + y_i)^p m_i = \left[\left(\sum_{i=1}^{n} x_i^p m_i\right)^{1/p} + \left(\sum_{i=1}^{n} y_i^p m_i\right)^{1/p}\right]^p$$

which is equivalent to (13). □

Problem 5 (SL70)**.** *If $u_1, \ldots, u_n, v_1, \ldots, v_n$ are real numbers, prove that*

$$1 + \sum_{i=1}^{n} (u_i + v_i)^2 \leq \frac{4}{3}\left(1 + \sum_{i=1}^{n} u_i^2\right)\left(1 + \sum_{i=1}^{n} v_i^2\right).$$

When does equality hold?

Solution. Let us set $a = \sqrt{\sum_{i=1}^{n} u_i^2}$ and $b = \sqrt{\sum_{i=1}^{n} v_i^2}$. By Minkowski's inequality (for $p = 2$) we have $\sum_{i=1}^{n} (u_i + v_i)^2 \leq (a+b)^2$. Hence the LHS of the desired inequality is not greater than $1 + (a+b)^2$, while the RHS is equal to $4(1+a^2)(1+b^2)/3$. Now it is sufficient to prove that

$$3 + 3(a+b)^2 \leq 4(1+a^2)(1+b^2).$$

The last inequality can be reduced to the trivial $0 \leq (a-b)^2 + (2ab-1)^2$. The equality in the initial inequality holds if and only if $u_i/v_i = c$ for some $c \in \mathbb{R}$ and $a = b = 1/\sqrt{2}$. △

Theorem 8 (Young)**.** *If $a, b > 0$ and $p, q > 1$ satisfy $\frac{1}{p} + \frac{1}{q} = 1$, then*

$$ab \leq \frac{a^p}{p} + \frac{b^q}{q}. \tag{15}$$

Equality holds if and only if $a^p = b^q$.

Proof. Since $\varphi(x) = e^x$ is a convex function we have that $e^{\frac{1}{p}x + \frac{1}{q}y} \leq \frac{1}{p}e^x + \frac{1}{q}e^y$. The equality holds if and only if $x = y$, and the inequality (15) is immediately obtained by placing $a = e^{x/p}$ and $b = e^{y/q}$. The equality holds if and only if $a^p = b^q$. □

Lemma 3. *If $x_1, x_2, \ldots, x_n, y_1, y_2, \ldots, y_n, m_1, m_2, \ldots, m_n$ are three sequences of positive real numbers and $p, q > 1$ such that $\frac{1}{p} + \frac{1}{q} = 1$, and $\alpha > 0$, then*

$$\sum_{i=1}^{n} x_i y_i m_i \leq \frac{1}{p} \cdot \alpha^p \cdot \sum_{i=1}^{n} x_i^p m_i + \frac{1}{q} \cdot \frac{1}{\alpha^q} \cdot \sum_{i=1}^{n} y_i^q m_i. \tag{16}$$

The equality holds if and only if $\frac{\alpha^p x_i^p}{p} = \frac{y_i^q}{q\alpha^q}$ for $1 \leq i \leq n$.

Proof. From (15) we immediately get $x_iy_i = (\alpha x_i)\frac{y_i}{\alpha} \le \frac{1}{p}\cdot\alpha^p x_i^p + \frac{1}{q}\cdot\frac{1}{\alpha^q}y_i^q$. Multiplying by m_i and adding as $i = 1,2,\ldots,n$ we get (16). The inequality holds if and only if $\frac{\alpha^p x_i^p}{p} = \frac{y_i^q}{q\alpha^q}$ for $1 \le i \le n$. □

Theorem 9 (Hölder)**.** *If $x_1,x_2,\ldots,x_n,y_1,y_2,\ldots,y_n,m_1,m_2,\ldots,m_n$ are three sequences of positive real numbers and $p,q > 1$ such that $\frac{1}{p}+\frac{1}{q} = 1$, then*

$$\sum_{i=1}^{n} x_iy_im_i \le \left(\sum_{i=1}^{n} x_i^p m_i\right)^{1/p}\cdot\left(\sum_{i=1}^{n} y_i^q m_i\right)^{1/q}. \tag{17}$$

The equality holds if and only if the sequences (x_i^p) and (y_i^q) are proportional.

Proof. The idea is very similar to the one used in the proof of Minkowski's inequality. The inequality (16) holds for any positive constant α. Let

$$A = \left(\alpha^p\sum_{i=1}^{n} x_i^p m_i\right)^{1/p},\ B = \left(\frac{1}{\alpha^q}\sum_{i=1}^{n} y_i^q m_i\right)^{1/q}.$$

By Young's inequality we have that $\frac{1}{p}A^p + \frac{1}{q}B^q = AB$ if $A^p = B^q$. Equivalently $\alpha^p\sum_{i=1}^n x_i^p m_i = \frac{1}{\alpha^q}\sum_{i=1}^n y_i^q m_i$. Choosing such an α we get

$$\sum_{i=1}^{n} x_iy_im_i \le \frac{1}{p}A^p + \frac{1}{q}B^q = AB = \left(\sum_{i=1}^{n} x_i^p m_i\right)^{1/p}\cdot\left(\sum_{i=1}^{n} y_i^q m_i\right)^{1/q}. \ \square$$

Problem 6. *If $a_1,\ldots,a_n$ and $m_1,\ldots,m_n$ are two sequences of positive numbers such that $a_1m_1 + \cdots + a_nm_n = \alpha$ and $a_1^2m_1 + \cdots + a_n^2m_n = \beta^2$, prove that $\sqrt{a_1}m_1 + \cdots + \sqrt{a_n}m_n \ge \frac{\alpha^{3/2}}{\beta}$.*

Solution. We will apply Hölder's inequality on $x_i = a_i^{1/3}$, $y_i = a_i^{2/3}$, $p = \frac{3}{2}$, $q = 3$:

$$\alpha = \sum_{i=1}^{n} a_im_i \le \left(\sum_{i=1}^{n} a_i^{1/2}m_i\right)^{2/3}\cdot\left(\sum_{i=1}^{n} a_i^2m_i\right)^{1/3} = \left(\sum_{i=1}^{n}\sqrt{a_i}m_i\right)^{2/3}\cdot\beta^{2/3}.$$

Hence $\sum_{i=1}^n\sqrt{a_i}m_i \ge \frac{\alpha^{3/2}}{\beta}$. △

Proof of the theorem 3. $M_r^m = \left(\sum_{i=1}^n x_i^r\cdot m_i\right)^{1/r}$. We will use the Hölders inequality for $y_i = 1$, $p = \frac{s}{r}$, and $q = \frac{p}{1-p}$. Then we get

$$M_r^m \le \left(\sum_{i=1}^{n} x_i^{rp}\cdot m_i\right)^{\frac{1}{pr}}\cdot\left(\sum_{i=1}^{n} 1^q\cdot m_i\right)^{p/(1-p)} = M_s. \ \square$$

Problem 7. *(SL98) Let x, y, and z be positive real numbers such that $xyz = 1$. Prove that*

$$\frac{x^3}{(1+y)(1+z)} + \frac{y^3}{(1+z)(1+x)} + \frac{z^3}{(1+x)(1+y)} \ge \frac{3}{4}.$$

Solution. The given inequality is equivalent to

$$x^3(x+1)+y^3(y+1)+z^3(z+1) \geq \frac{3}{4}(1+x+y+z+xy+yz+zx+xyz).$$

The left-hand side can be written as $x^4+y^4+z^4+x^3+y^3+z^3 = 3M_4^4+3M_3^3$. Using $xy+yz+zx \leq x^2+y^2+z^2 = 3M_2^2$ we see that the right-hand side is less than or equal to $\frac{3}{4}(2+3M_1+3M_2^2)$. Since $M_1 \geq 3\sqrt[3]{xyz} = 1$, we can further say that the right-hand side of the required inequality is less than or equal to $\frac{3}{4}(5M_1+3M_2^2)$. Since $M_4 \geq M_3$, and $M_1 \leq M_2 \leq M_3$, the following inequality would imply the required statement:

$$3M_3^4+3M_3^3 \geq \frac{3}{4}(5M_3+3M_3^2).$$

However the last inequality is equivalent to $(M_3-1)(4M_3^2+8M_3+5) \geq 0$ which is true because $M_3 \geq 1$. The equality holds if and only if $x=y=z=1$. $\triangle$

Theorem 10 (Weighted Cauchy-Schwartz)**.** *If x_i, y_i are real numbers, and m_i positive real numbers, then*

$$\sum_{i=1}^{n} x_i y_i m_i \quad \leq \quad \sqrt{\sum_{i=1}^{n} x_i^2 m_i} \cdot \sqrt{\sum_{i=1}^{n} y_i^2 m_i}. \tag{18}$$

Proof. After noticing that $\sum_{i=1}^n x_i y_i m_i \leq \sum_{i=1}^n |x_i| \cdot |y_i| m_i$, the rest is just a special case ($p=q=2$) of the Hölder's inequality. □

Problem 8. *If a, b, and c are positive numbers, prove that*

$$\frac{a}{b}+\frac{b}{c}+\frac{c}{a} \geq \frac{(a+b+c)^2}{ab+bc+ca}.$$

Solution. We will apply the Cauchy-Schwartz inequality with $x_1 = \sqrt{\frac{a}{b}}$, $x_2 = \sqrt{\frac{b}{c}}$, $x_3 = \sqrt{\frac{c}{a}}$, $y_1 = \sqrt{ab}$, $y_2 = \sqrt{bc}$, and $y_3 = \sqrt{ca}$. Then

$$\begin{aligned} a+b+c &= x_1y_1+x_2y_2+x_3y_3 \leq \sqrt{x_1^2+x_2^2+x_3^2} \cdot \sqrt{y_1^2+y_2^2+y_3^2} \\ &= \sqrt{\frac{a}{b}+\frac{b}{c}+\frac{c}{a}} \cdot \sqrt{ab+bc+ca}. \end{aligned}$$

Theorem 11. *If $a_1, \ldots, a_n$ are positive real numbers, then*

$$\lim_{r \to 0} M_r(a_1, \ldots, a_n) = a_1^{m_1} \cdot a_2^{m_2} \cdots a_n^{m_n}.$$

Proof. This theorem is given here for completeness. It states that as $r \to 0$ the mean of order r approaches the geometric mean of the sequence. Its proof involves some elementary calculus, and the reader can omit the proof.

$$M_r(a_1, \ldots, a_n) = e^{\frac{1}{r}\log(a_1^r m_1 + \cdots + a_n^r m_n)}.$$

Using the L'Hospitale's theorem we get

$$\begin{aligned} \lim_{r \to 0} \frac{1}{r} \log(a_1^r m_1 + \cdots + a_n^r m_n) &= \lim_{r \to 0} \frac{m_1 a_1^r \log a_1 + \cdots + m_n a_n^r \log a_n}{a_1^r m_1 + \cdots + a_n^r m_n} \\ &= m_1 \log a_1 + \cdots + m_n \log a_n \\ &= \log\left(a_1^{m_1} \cdots a_n^{m_n}\right). \end{aligned}$$

The result immediately follows. □

4 Inequalities of Schur and Muirhead

Definition 4. *Let $\sum!F(a_1,\ldots,a_n)$ be the sum of $n!$ summands which are obtained from the function $F(a_1,\ldots,a_n)$ making all permutations of the array (a).*

We will consider the special cases of the functio F, i.e. when $F(a_1,\ldots,a_n)=a_1^{\alpha_1}\cdot\cdots\cdot a_n^{\alpha_n}$, $\alpha_i\geq 0$.

If (α) is an array of exponents and $F(a_1,\ldots,a_n)=a_1^{\alpha_1}\cdot\cdots\cdot a_n^{\alpha_n}$ we will use $T[\alpha_1,\ldots,\alpha_n]$ instead of $\sum!F(a_1,\ldots,a_n)$, if it is clear what is the sequence (a).

Example 4. *$T[1,0,\ldots,0]=(n-1)!\cdot(a_1+a_2+\cdots+a_n)$, and $T[\frac{1}{n},\frac{1}{n},\ldots,\frac{1}{n}]=n!\cdot\sqrt[n]{a_1\cdot\cdots\cdot a_n}$. The AM-GM inequality is now expressed as:*

$$T[1,0,\ldots,0]\geq T\left[\frac{1}{n},\ldots,\frac{1}{n}\right].$$

Theorem 12 (Schur)**.** *For $\alpha\in\mathbb{R}$ and $\beta>0$ the following inequality holds:*

$$T[\alpha+2\beta,0,0]+T[\alpha,\beta,\beta]\geq 2T[\alpha+\beta,\beta,0]. \tag{19}$$

Proof. Let (x,y,z) be the sequence of positive reals for which we are proving (19). Using some elementary algebra we get

$$\begin{aligned}&\frac{1}{2}T[\alpha+2\beta,0,0]+\frac{1}{2}T[\alpha,\beta,\beta]-T[\alpha+\beta,\beta,0]\\ =\;& x^\alpha(x^\beta-y^\beta)(x^\beta-z^\beta)+y^\alpha(y^\beta-x^\beta)(y^\beta-z^\beta)+z^\alpha(z^\beta-x^\beta)(z^\beta-y^\beta).\end{aligned}$$

Without loss of generality we may assume that $x\geq y\geq z$. Then in the last expression only the second summand may be negative. If $\alpha\geq 0$ then the sum of the first two summands is ≥ 0 because $x^\alpha(x^\beta-y^\beta)(x^\beta-z^\beta)\geq x^\alpha(x^\beta-y^\beta)(y^\beta-z^\beta)\geq y^\alpha(x^\beta-y^\beta)(y^\beta-z^\beta)=-y^\alpha(x^\beta-y^\beta)(y^\beta-z^\beta)$. Similarly for $\alpha<0$ the sum of the last two terms is ≥ 0. □

Example 5. *If we set $\alpha=\beta=1$, we get*

$$x^3+y^3+z^3+3xyz\geq x^2y+xy^2+y^2z+yz^2+z^2x+zx^2.$$

Definition 5. *We say that the array (α) majorizes array (α'), and we write that in the following way $(\alpha')\prec(\alpha)$, if we can arrange the elements of arrays (α) and (α') in such a way that the following three conditions are satisfied:*

1. *$\alpha_1'+\alpha_2'+\cdots+\alpha_n'=\alpha_1+\alpha_2+\cdots+\alpha_n$;*
2. *$\alpha_1'\geq\alpha_2'\geq\cdots\geq\alpha_n'$ i $\alpha_1\geq\alpha_2\geq\cdots\geq\alpha_n$.*
3. *$\alpha_1'+\alpha_2'+\cdots+\alpha_\nu'\leq\alpha_1+\alpha_2+\cdots+\alpha_\nu$, for all $1\leq\nu<n$.*

Clearly, $(\alpha)\prec(\alpha)$.

Theorem 13 (Muirhead)**.** *The necessairy and sufficient condition for comparability of $T[\alpha]$ and $T[\alpha']$, for all positive arrays (a), is that one of the arrays (α) and (α') majorizes the other. If $(\alpha')\prec(\alpha)$ then*

$$T[\alpha']\leq T[\alpha].$$

Equality holds if and only if (α) and (α') are identical, or when all a_is are equal.

Proof. First, we prove the necessity of the condition. Setting that all elements of the array a are equal to x, we get that

$$x^{\sum \alpha_i'} \leq x^{\sum \alpha_i}.$$

This can be satisfied for both large and small xs only if the condition 1 from the definition is satisfied. Now we put $a_1 = \cdots, a_\nu = x$ and $a_{\nu+1} = \cdots = a_n = 1$. Comparing the highest powers of x in expressions $T[\alpha]$ and $T[\alpha']$, knowing that for sufficiently large x we must have $T[\alpha'] \leq T[\alpha]$, we conclude that $\alpha_1' + \cdots + \alpha_\nu' \leq \alpha_1 + \cdots + \alpha_\nu$.

Now we will proof the sufficiency of the condition. The statement will follow from the following two lemmas. We will define one linear operation L on the set of the exponents (α). Suppose that α_k and α_l are two different exponents of (α) such that $\alpha_k > \alpha_l$. We can write

$$\alpha_k = \rho + \tau, \quad \alpha_l = \rho - \tau \quad (0 < \tau \leq \rho).$$

If $0 \leq \sigma < \tau \leq \rho$, define the array $(\alpha') = L(\alpha)$ in the following way:

$$\begin{cases} \alpha_k' = \rho + \sigma = \frac{\tau+\sigma}{2\tau}\alpha_k + \frac{\tau-\sigma}{2\tau}\alpha_l, \\ \alpha_l' = \rho - \sigma = \frac{\tau-\sigma}{2\tau}\alpha_k + \frac{\tau+\sigma}{2\tau}\alpha_l, \\ \alpha_\nu' = \alpha_\nu, \quad (\nu \neq k, \nu \neq l). \end{cases}$$

The definition of this mapping doesn't require that some of the arrays (α) and (α') is in non-decreasing order.

Lemma 4. *If $(\alpha') = L(\alpha)$, then $T[\alpha'] \leq T[\alpha]$, and equality holds if and only if all the elements of (a) are equal.*

Proof. We may rearrange the elements of the sequence such that $k = 1$ i $l = 2$. Then we have

$$\begin{aligned} & T[\alpha] - T[\alpha'] \\ = \; & \sum\!{}^{!} a_3^{\alpha_3} \cdots a_n^{\alpha_n} \cdot (a_1^{\rho+\tau} a_2^{\rho-\tau} + a_1^{\rho-\tau} a_2^{\rho+\tau} - a_1^{\rho+\sigma} a_2^{\rho-\sigma} - a_1^{\rho-\sigma} a_2^{\rho+\sigma}) \\ = \; & \sum\!{}^{!} (a_1 a_2)^{\rho-\tau} a_3^{\alpha_3} \cdots a_n^{\alpha_n} (a_1^{\tau+\sigma} - a_2^{\tau+\sigma})(a_1^{\tau-\sigma} - a_2^{\tau-\sigma}) \geq 0. \end{aligned}$$

Eaquality holds if and only if a_is are equal. □

Lemma 5. *If $(\alpha') \prec (\alpha)$, but (α') and (α) are different, then (α') can be obtained from (α) by succesive application of the transformation L.*

Proof. Denote by m the number of differences $\alpha_\nu - \alpha_\nu'$ that are $\neq 0$. m is a positive integer and we will prove that we can apply operation L in such a way that after each of applications, number m decreases (this would imply that the procedure will end up after finite number of steps). Since $\sum(\alpha_\nu - \alpha_\nu') = 0$, and not all of differences are 0, there are positive and negative differences, but the first one is positive. We can find such k and l for which:

$$\alpha_k' < \alpha_k, \quad \alpha_{k+1}' = \alpha_{k+1}, \ldots, \alpha_{l-1}' = \alpha_{l-1}, \quad \alpha_l' > \alpha_l.$$

($\alpha_l - \alpha_l'$ is the first negative difference, and $\alpha_k - \alpha_k'$ is the last positive difference before this negative one). Let $\alpha_k = \rho + \tau$ and $\alpha_l = \rho - \tau$, define σ by

$$\sigma = \max\{|\alpha_k' - \rho|, |\alpha_l' - \rho|\}.$$

At least one of the following two equalities is satisfied:

$$\alpha_l' - \rho = -\sigma, \quad \alpha_k' - \rho = \sigma,$$

because $\alpha'_k > \alpha'_l$. We also have $\sigma < \tau$, because $\alpha'_k < \alpha_k$ i $\alpha'_l > \alpha_l$. Let

$$\alpha''_k = \rho + \sigma, \;\; \alpha''_l = \rho - \sigma, \;\; \alpha''_\nu = \alpha_\nu \;\; (\nu \neq k, \nu \neq l).$$

Now instead of the sequence (α) we will consider the sequence (α''). Number m has decreased by at least 1. It is easy to prove that the sequence (α'') is increasing and majorizes (α'). Repeating this procedure, we will get the sequence (α') which completes the proof of the second lemma, and hence the Muirhead's theorem. □ □

Example 6. *AM-GM is now the consequence of the Muirhead's inequality.*

Problem 9. *Prove that for positive numbers a, b and c the following equality holds:*

$$\frac{1}{a^3+b^3+abc} + \frac{1}{b^3+c^3+abc} + \frac{1}{c^3+a^3+abc} \leq \frac{1}{abc}.$$

Solution. After multiplying both left and right-hand side of the required inequality with $abc(a^3+b^3+abc)(b^3+c^3+abc)(c^3+a^3+abc)$ we get that the original inequality is equivalent to

$$\begin{gathered}\tfrac{3}{2}T[4,4,1] + 2T[5,2,2] + \tfrac{1}{2}T[7,1,1] + \tfrac{1}{2}T[3,3,3] \leq \\ \leq \tfrac{1}{2}T[3,3,3] + T[6,3,0] + \tfrac{3}{2}T[4,4,1] + \tfrac{1}{2}T[7,1,1] + T[5,2,2]\end{gathered}$$

which is true because Muirhead's theorem imply that $T[5,2,2] \leq T[6,3,0]$. △

More problems with solutions using Muirhead's inequality can be found in the section "Problems".

5 Inequalities of Jensen and Karamata

Theorem 14 (Jensen's Inequality)**.** *If f is convex function and $\alpha_1, \ldots, \alpha_n$ sequence of real numbers such that $\alpha_1 + \cdots + \alpha_n = 1$, than for any sequence $x_1, \ldots, x_n$ of real numbers, the following inequality holds:*

$$f(\alpha_1 x_1 + \cdots + \alpha_n x_n) \leq \alpha_1 f(x_1) + \cdots + \alpha_n f(x_n).$$

Remark. If f is concave, then $f(\alpha_1 x_1 + \cdots + \alpha_n x_n) \geq \alpha_1 f(x_1) + \cdots + \alpha_n f(x_n)$.

Example 7. *Using Jensen's inequality prove the generalized mean inequality, i.e. that for every two sequences of positive real numbers $x_1, \ldots, x_n$ and $m_1, \ldots, m_n$ such that $m_1 + \cdots + m_n = 1$ the following inequality holds:*

$$m_1 x_1 + m_2 x_2 + \cdots + m_n x_n \geq x_1^{m_1} \cdot x_2^{m_2} \cdots x_n^{m_n}.$$

Theorem 15 (Karamata's inequalities)**.** *Let f be a convex function and $x_1, \ldots, x_n$, $y_1, y_2, \ldots, y_n$ two non-increasing sequences of real numbers. If one of the following two conditions is satisfied:*

(a) $(y) \prec (x)$*;*

(b) $x_1 \geq y_1$, $x_1 + x_2 \geq y_1 + y_2$, $x_1 + x_2 + x_3 \geq y_1 + y_2 + y_3$, …, $x_1 + \cdots + x_{n-1} \geq y_1 + \cdots + y_{n-1}$, $x_1 + \cdots + x_n \geq y_1 + \cdots + y_n$ *and f is increasing;*

then

$$\sum_{i=1}^{n} f(x_i) \geq \sum_{i=1}^{n} f(y_i). \tag{20}$$

Proof. Let $c_i = \frac{f(y_i)-f(x_i)}{y_i-x_i}$, for $y_i \neq x_i$, and $c_i = f'_+(x_i)$, for $x_i = y_i$. Since f is convex, and x_i, y_i are decreasing sequences, c_i is non-increasing (because is represents the "slope" of f on the interval between x_i and y_i). We now have

$$\begin{aligned}\sum_{i=1}^{n} f(x_i) - \sum_{i=1}^{n} f(y_i) &= \sum_{i=1}^{n} c_i(x_i - y_i) = \sum_{i=1}^{n} c_i x_i - \sum_{i=1}^{n} c_i y_i \\ &= \sum_{i=1}^{n} (c_i - c_{i+1})(x_1 + \cdots + x_i) \\ &\quad - \sum_{i=1}^{n} (c_i - c_{i+1})(y_1 + \cdots + y_i), \end{aligned} \tag{21}$$

here we define c_{n+1} to be 0. Now, denoting $A_i = x_1 + \cdots + x_i$ and $B_i = y_1 + \cdots + y_i$ (21) can be rearranged to

$$\sum_{i=1}^{n} f(x_i) - \sum_{i=1}^{n} f(y_i) = \sum_{i=1}^{n-1} (c_i - c_{i+1})(A_i - B_i) + c_n \cdot (A_n - B_n).$$

The sum on the right-hand side of the last inequality is non-negative because c_i is decreasing and $A_i \geq B_i$. The last term $c_n(A_n - B_n)$ is zero under the assumption (a). Under the assumption (b) we have that $c_n \geq 0$ (f is increasing) and $A_n \geq B_n$ and this implies (20). □

Problem 10. *If $a_1 \geq a_2 \geq \cdots \geq a_n$ and $b_1 \geq b_2 \geq \cdots \geq b_n$ are two sequences of positive real numbers which satisfy the following conditions:*

$$a_1 \geq b_2,\ a_1a_2 \geq b_1b_2,\ a_1a_2a_3 \geq b_1b_2b_3,\ \cdots \geq a_1a_2\cdots a_n \geq b_1b_2\cdots b_n,$$

prove that

$$a_1 + a_2 + \cdots + a_n \geq b_1 + b_2 + \cdots + b_n.$$

Solution. Let $a_i = e^{x_i}$ and $b_i = e^{y_i}$. We easily verify that the conditions (b) of the Karamata's theorem are satisfied. Thus $\sum_{i=1}^{n} e^{y_i} \geq \sum_{i=1}^{n} e^{x_i}$ and the result immediately follows. △

Problem 11. *If $x_1, \ldots, x_n \in [-\pi/6, \pi/6]$, prove that*

$$\cos(2x_1 - x_2) + \cos(2x_2 - x_3) + \cdots + \cos(2x_n - x_1) \leq \cos x_1 + \cdots + \cos x_n.$$

Solution. Rearrange $(2x_1 - x_2, 2x_2 - x_3, \ldots, 2x_n - x_1)$ and $(x_1, \ldots, x_n)$ in two non-increasing sequences $(2x_{m_1} - x_{m_1+1}, 2x_{m_2} - x_{m_2+1}, \ldots, 2x_{m_n} - x_{m_n+1})$ and $(x_{k_1}, x_{k_2}, \ldots, x_{k_n})$ (here we assume that $x_{n+1} = x_1$. We will verify that condition (a) of the Karamata's inequality is satisfied. This follows from

$$\begin{aligned} & (2x_{m_1} - x_{m_1+1} + \cdots + 2x_{m_l} - x_{m_l+1}) - (x_{k_1} + \cdots + x_{k_l}) \\ \geq\ & (2x_{k_1} - x_{k_1+1} + \cdots + 2x_{k_l} - x_{k_l+1}) - (x_{k_1} + \cdots + x_{k_l}) \\ =\ & (x_{k_1} + \cdots x_{k_l}) - (x_{k_1+1} + \cdots + x_{k_l+1}) \geq 0. \end{aligned}$$

The function $f(x) = -\cos x$ is convex on $[-\pi/2, \pi/2]$ hence Karamata's inequality holds and we get

$$-\cos(2x_1 - x_2) - \cdots - \cos(2x_n - x_1) \geq -\cos x_1 - \cdots - \cos x_n,$$

which is obviously equivalent to the required inequality. △

6 Chebyshev's inequalities

Theorem 16 (Chebyshev's inequalities)**.** *Let $a_1 \geq a_2 \geq \cdots \geq a_n$ and $b_1 \geq b_2 \geq \cdots \geq b_n$ be real numbers. Then*

$$n\sum_{i=1}^{n} a_i b_i \geq \left(\sum_{i=1}^{n} a_i\right)\left(\sum_{i=1}^{n} b_i\right) \geq n\sum_{i=1}^{n} a_i b_{n+1-i}. \tag{22}$$

The two inequalities become equalities at the same time when $a_1 = a_2 = \cdots = a_n$ or $b_1 = b_2 = \cdots = b_n$.

The Chebyshev's inequality will follow from the following generalization (placing $m_i = \frac{1}{n}$ for the left part, and the right inequality follows by applying the left on a_i and $c_i = -b_{n+1-i}$).

Theorem 17 (Generalized Chebyshev's Inequality)**.** *Let $a_1 \geq a_2 \geq \cdots \geq a_n$ and $b_1 \geq b_2 \geq \cdots \geq b_n$ be any real numbers, and $m_1, \ldots, m_n$ non-negative real numbers whose sum is 1. Then*

$$\sum_{i=1}^{n} a_i b_i m_i \geq \left(\sum_{i=1}^{n} a_i m_i\right)\left(\sum_{i=1}^{n} b_i m_i\right). \tag{23}$$

The inequality become an equality if and only if $a_1 = a_2 = \cdots = a_n$ or $b_1 = b_2 = \cdots = b_n$.

Proof. From $(a_i - a_j)(b_i - b_j) \geq 0$ we get:

$$\sum_{i,j} (a_i - a_j)(b_i - b_j) m_i m_j \geq 0. \tag{24}$$

Since $(\sum_{i=1}^{n} a_i m_i) \cdot (\sum_{i=1}^{n} b_i m_i) = \sum_{i,j} a_i b_j m_i m_j$, (24) implies that

$$\begin{aligned} 0 &\leq \sum_{i,j} a_i b_i m_i m_j - \sum_{i,j} a_i b_j m_i m_j - \sum_{i,j} a_j b_i m_j m_i + \sum_{i,j} a_j b_j m_i m_j \\ &= 2\left[\sum_{i} a_i b_i m_i - \left(\sum_{i} a_i m_i\right)\left(\sum_{i} b_i m_i\right)\right]. \square \end{aligned}$$

Problem 12. *Prove that the sum of distances of the orthocenter from the sides of an acute triangle is less than or equal to $3r$, where the r is the inradius.*

Solution. Denote $a = BC$, $b = CA$, $c = AB$ and let S_{ABC} denote the area of the triangle ABC. Let d_A, d_B, d_C be the distances from H to BC, CA, AB, and A', B', C' the feet of perpendiculars from A, B, C. Then we have $ad_a + bd_b + cd_c = 2(S_{BCH} + S_{ACH} + S_{ABH}) = 2P$. On the other hand if we assume that $a \geq b \geq c$, it is easy to prove that $d_A \geq d_B \geq d_C$. Indeed, $a \geq b$ implies $\angle A \geq \angle B$ hence $\angle HCB' \leq \angle HCA'$ and $HB' \leq HA'$. The Chebyshev's inequality implies

$$(a+b+c)r = 2P = ad_a + bd_b + cd_c \geqslant \frac{1}{3}(a+b+c)(d_a+d_b+d_c). \triangle$$

7 Problems

1. If $a, b, c, d > 0$, prove that

$$\frac{a}{b+c} + \frac{b}{c+d} + \frac{c}{d+a} + \frac{d}{a+b} \geq 2.$$

2. Prove that
$$\frac{a^3}{a^2+ab+b^2}+\frac{b^3}{b^2+bc+c^2}+\frac{c^3}{c^2+ca+a^2}\geq\frac{a+b+c}{3},$$
for $a,b,c>0$.

3. If $a,b,c,d,e,f>0$, prove that
$$\frac{ab}{a+b}+\frac{cd}{c+d}+\frac{ef}{e+f}\leq\frac{(a+c+e)(b+d+f)}{a+b+c+d+e+f}.$$

4. If $a,b,c\geq 1$, prove that
$$\sqrt{a-1}+\sqrt{b-1}+\sqrt{c-1}\leq\sqrt{c(ab+1)}.$$

5. Let $a_1,a_2,\ldots,a_n,b_1,b_2,\ldots,b_n$ be positive real numbers. Prove that
$$\left(\sum_{i\neq j}a_ib_j\right)^2\geq\left(\sum_{i\neq j}a_ia_j\right)\left(\sum_{i\neq j}b_ib_j\right).$$

6. If $\frac{1}{x}+\frac{1}{y}+\frac{1}{z}=1$ for $x,y,z>0$, prove that
$$(x-1)(y-1)(z-1)\geq 8.$$

7. Let $a,b,c>0$ satisfy $abc=1$. Prove that
$$\frac{1}{\sqrt{b+\frac{1}{a}+\frac{1}{2}}}+\frac{1}{\sqrt{c+\frac{1}{b}+\frac{1}{2}}}+\frac{1}{\sqrt{a+\frac{1}{c}+\frac{1}{2}}}\geq\sqrt{2}.$$

8. Given positive numbers a,b,c,x,y,z such that $a+x=b+y=c+z=S$, prove that $ay+bz+cx<S^2$.

9. Let a,b,c be positive real numbers. Prove the inequality
$$\frac{a^2}{b}+\frac{b^2}{c}+\frac{c^2}{a}\geq a+b+c+\frac{4(a-b)^2}{a+b+c}.$$

10. Determine the maximal real number a for which the inequality
$$x_1^2+x_2^2+x_3^2+x_4^2+x_5^2\geq a(x_1x_2+x_2x_3+x_3x_4+x_4x_5)$$
holds for any five real numbers x_1,x_2,x_3,x_4,x_5.

11. If $x,y,z\geq 0$ and $x+y+z=1$, prove that
$$0\leq xy+yz+zx-2xyz\leq\frac{7}{27}.$$

12. Let a,b and c be positive real numbers such that $abc=1$. Prove that
$$\frac{1}{a^3(b+c)}+\frac{1}{b^3(c+a)}+\frac{1}{c^3(a+b)}\geq\frac{3}{2}.$$

13. If a, b and c are positive real numbers, prove that:

$$\frac{a^3}{b^2-bc+c^2}+\frac{b^3}{c^2-ca+a^2}+\frac{c^3}{a^2-ab+b^2}\geq 3\cdot\frac{ab+bc+ca}{a+b+c}.$$

14. (IMO05) Let x, y and z be positive real numbers such that $xyz\geq 1$. Prove that

$$\frac{x^5-x^2}{x^5+y^2+z^2}+\frac{y^5-y^2}{y^5+z^2+x^2}+\frac{z^5-z^2}{z^5+x^2+y^2}\geq 0.$$

15. Let $a_1,\ldots,a_n$ be positive real numbers. Prove that

$$\frac{a_1^3}{a_2}+\frac{a_2^3}{a_3}+\cdots+\frac{a_n^3}{a_1}\geq a_1^2+a_2^2+\cdots+a_n^2.$$

16. Let $a_1,\ldots,a_n$ be positive real numbers. Prove that

$$(1+a_1)(1+a_2)\cdots(1+a_n)\leq\left(1+\frac{a_1^2}{a_2}\right)\cdot\left(1+\frac{a_2^2}{a_3}\right)\cdot\cdots\cdot\left(1+\frac{a_n^2}{a_1}\right).$$

17. If a, b, and c are the lengths of the sides of a triangle, s its semiperimeter, and $n\geq 1$ an integer, prove that

$$\frac{a^n}{b+c}+\frac{b^n}{c+a}+\frac{c^n}{a+b}\geq\left(\frac{2}{3}\right)^{n-2}\cdot s^{n-1}.$$

18. Let $0<x_1\leq x_2\leq\cdots\leq x_n$ $(n\geq 2)$ and

$$\frac{1}{1+x_1}+\frac{1}{1+x_2}+\cdots+\frac{1}{1+x_n}=1.$$

Prove that

$$\sqrt{x_1}+\sqrt{x_2}+\cdots+\sqrt{x_n}\geq(n-1)\left(\frac{1}{\sqrt{x_1}}+\frac{1}{\sqrt{x_2}}+\cdots+\frac{1}{\sqrt{x_n}}\right).$$

19. Suppose that any two members of certain society are either *friends* or *enemies*. Suppose that there is total of n members, that there is total of q pairs of friends, and that in any set of three persons there are two who are enemies to each other. Prove that there exists at least one member among whose enemies we can find at most $q\cdot\left(1-\frac{4q}{n^2}\right)$ pairs of friends.

20. Given a set of unit circles in the plane whose total area is S. Prove that among those circles there exist certain number of non-intersecting circles whose total area is $\geq\frac{2}{9}S$.

8 Solutions

1. Denote by L the left-hand side of the required inequality. If we add the first and the third summand of L we get

$$\frac{a}{b+c}+\frac{c}{d+a}=\frac{a^2+c^2+ad+bc}{(b+c)(a+d)}.$$

We will bound the denominator of the last fraction using the inequality $xy \le (x+y)^2/4$ for appropriate x and y. For $x = b+c$ and $y = a+d$ we get $(b+c)(a+d) \le (a+b+c+d)^2/4$. The equality holds if and only if $a+d = b+c$. Therefore

$$\frac{a}{b+c} + \frac{c}{d+a} \ge 4\frac{a^2+c^2+ad+bc}{(a+b+c+d)^2}.$$

Similarly $\frac{b}{c+d} + \frac{d}{a+b} \ge 4\frac{b^2+d^2+ab+cd}{(a+b+c+d)^2}$ (with the equality if and only if $a+b = c+d$) implying

$$\begin{aligned} & \frac{a}{b+c} + \frac{b}{c+d} + \frac{c}{d+a} + \frac{d}{a+b} \\ \ge\ & 4\frac{a^2+b^2+c^2+d^2+ad+bc+ab+cd}{(a+b+c+d)^2} \\ =\ & 4\frac{a^2+b^2+c^2+d^2+(a+c)(b+d)}{[(a+c)+(b+d)]^2}. \end{aligned}$$

In order to solve the problem it is now enough to prove that

$$2\frac{a^2+b^2+c^2+d^2+(a+c)(b+d)}{[(a+c)+(b+d)]^2} \ge 1. \tag{25}$$

After multiplying both sides of (25) by $[(a+c)+(b+d)]^2 = (a+c)^2+(b+d)^2$ it becomes equivalent to $2(a^2+b^2+c^2+d^2) \ge (a+c)^2+(b+d)^2 = a^2+b^2+c^2+d^2+2ac+2bd$. It is easy to see that the last inequality holds because many terms will cancel and the remaining inequality is the consequence of $a^2+c^2 \ge 2ac$ and $b^2+d^2 \ge 2bc$. The equality holds if and only if $a = c$ and $b = d$.

2. We first notice that

$$\frac{a^3-b^3}{a^2+ab+b^2} + \frac{b^3-c^3}{b^2+bc+c^2} + \frac{c^3-a^3}{c^2+ca+a^2} = 0.$$

Hence it is enough to prove that

$$\frac{a^3+b^3}{a^2+ab+b^2} + \frac{b^3+c^3}{b^2+bc+c^2} + \frac{c^3+a^3}{c^2+ca+a^2} \ge \frac{2(a+b+c)}{3}.$$

However since $3(a^2-ab+b^2) \ge a^2+ab+b^2$,

$$\frac{a^3+b^3}{a^2+ab+b^2} = (a+b)\frac{a^2-ab+b^2}{a^2+ab+b^2} \ge \frac{a+b}{3}.$$

The equality holds if and only if $a = b = c$.

Second solution. First we prove that

$$\frac{a^3}{a^2+ab+b^2} \ge \frac{2a-b}{3}. \tag{26}$$

Indeed after multiplying we get that the inequality is equivalent to $a^3+b^3 \ge ab(a+b)$, or $(a+b)(a-b)^2 \ge 0$ which is true. After adding (26) with two similar inequalities we get the result.

3. We will first prove that

$$\frac{ab}{a+b}+\frac{cd}{c+d}\leq\frac{(a+c)(b+d)}{a+b+c+d}. \tag{27}$$

As is the case with many similar inequalities, a first look at (27) suggests to multiply out both sides by $(a+b)(c+d)(a+b+c+d)$. That looks scary. But we will do that now. In fact you will do, I will not. I will just encourage you and give moral support (try to imagine me doing that). After you multiply out everything (do it twice, to make sure you don't make a mistake in calculation), the result will be rewarding. Many things cancel out and what remains is to verify the inequality $4abcd\leq a^2d^2+b^2c^2$ which is true because it is equivalent to $0\leq(ad-bc)^2$. The equality holds if and only if $ad=bc$, or $\frac{a}{b}=\frac{c}{d}$.

Applying (27) with the numbers $A=a+c$, $B=b+d$, $C=e$, and $D=f$ yields:

$$\frac{(a+c)(b+d)}{a+b+c+d}+\frac{ef}{e+f}\leq\frac{(A+C)(B+D)}{A+B+C+D}=\frac{(a+c+e)(b+d+f)}{a+b+c+d+e+f},$$

and the required inequality is proved because (27) can be applied to the first term of the left-hand side. The equality holds if and only if $\frac{a}{b}=\frac{c}{d}=\frac{e}{f}$.

4. To prove the required inequality we will use the similar approach as in the previous problem. First we prove that

$$\sqrt{a-1}+\sqrt{b-1}\leq\sqrt{ab}. \tag{28}$$

Squaring both sides gives us that the original inequality is equivalent to

$$\begin{aligned} & a+b-2+2\sqrt{(a-1)(b-1)}\leq ab \\ \Leftrightarrow\quad & 2\sqrt{(a-1)(b-1)}\leq ab-a-b+2=(a-1)(b-1)+1. \end{aligned} \tag{29}$$

The inequality (29) is true because it is of the form $x+1\geq 2\sqrt{x}$ for $x=(a-1)(b-1)$.

Now we will apply (28) on numbers $A=ab+1$ and $B=c$ to get

$$\sqrt{ab}+\sqrt{c-1}=\sqrt{A-1}+\sqrt{B-1}\leq\sqrt{AB}=\sqrt{(ab+1)c}.$$

The first term of the left-hand side is greater than or equal to $\sqrt{a-1}+\sqrt{b-1}$ which proves the statement. The equality holds if and only if $(a-1)(b-1)=1$ and $ab(c-1)=1$.

5. Let us denote $p=\sum_{i=1}^n a_i, q=\sum_{i=1}^n b_i, k=\sum_{i=1}^n a_i^2, l=\sum_{i=1}^n b_i^2$, and $m=\sum_{i=1}^n a_ib_i$. The following equalities are easy to verify:

$$\sum_{i\neq j}a_ib_j=pq-m,\ \sum_{i\neq j}a_ia_j=p^2-k,\text{ and }\sum_{i\neq j}b_ib_j=q^2-l,$$

so the required inequality is equivalent to

$$(pq-m)^2\geq(p^2-k)(q^2-l)\Leftrightarrow lp^2-2qm\cdot p+m^2+q^2k-kl\geq 0.$$

Consider the last expression as a quadratic equation in p, i.e. $\varphi(p)=lp^2-2qm\cdot p+q^2k-kl$. If we prove that its discriminant is less than or equal to 0, we are done. That condition can be written as:

$$q^2m^2-l(m^2+q^2k-kl)\leq 0\Leftrightarrow(lk-m^2)(q^2-l)\geq 0.$$

The last inequality is true because $q^2 - l = \sum_{i \neq j} b_i b_j > 0$ (b_i are positive), and $lk - m^2 \geq 0$ (Cauchy-Schwartz inequality). The equality holds if and only if $lk - m^2 = 0$, i.e. if the sequences (a) and (b) are proportional.

6. This is an example of a problem where we have some conditions on x, y, and z. Since there are many reciprocals in those conditions it is natural to divide both sides of the original inequality by xyz. Then it becomes

$$\left(1 - \frac{1}{x}\right) \cdot \left(1 - \frac{1}{y}\right) \cdot \left(1 - \frac{1}{z}\right) \geq \frac{8}{xyz}. \tag{30}$$

However $1 - \frac{1}{x} = \frac{1}{y} + \frac{1}{z}$ and similar relations hold for the other two terms of the left-hand side of (30). Hence the original inequality is now equivalent to

$$\left(\frac{1}{y} + \frac{1}{z}\right) \cdot \left(\frac{1}{z} + \frac{1}{x}\right) \cdot \left(\frac{1}{x} + \frac{1}{y}\right) \geq \frac{8}{xyz},$$

and this follows from $\frac{1}{x} + \frac{1}{y} \geq 2\frac{1}{\sqrt{xy}}$, $\frac{1}{y} + \frac{1}{z} \geq 2\frac{1}{\sqrt{yz}}$, and $\frac{1}{z} + \frac{1}{x} \geq 2\frac{1}{\sqrt{zx}}$. The equality holds if and only if $x = y = z = 3$.

7. Notice that

$$\frac{1}{2} + b + \frac{1}{a} + \frac{1}{2} > 2\sqrt{\frac{1}{2} \cdot \left(b + \frac{1}{a} + \frac{1}{2}\right)}.$$

This inequality is strict for any two positive numbers a and b. Using the similar inequalities for the other two denominators on the left-hand side of the required inequality we get:

$$\begin{aligned} & \frac{1}{\sqrt{b + \frac{1}{a} + \frac{1}{2}}} + \frac{1}{\sqrt{c + \frac{1}{b} + \frac{1}{2}}} + \frac{1}{\sqrt{a + \frac{1}{c} + \frac{1}{2}}} \\ > \quad & \sqrt{2}\left(\frac{1}{1 + \frac{1}{a} + b} + \frac{1}{1 + \frac{1}{b} + c} + \frac{1}{1 + \frac{1}{c} + a}\right). \end{aligned} \tag{31}$$

The last expression in (31) can be transformed using $\frac{1}{1+\frac{1}{a}+b} = \frac{a}{1+a+ab} = \frac{a}{1+\frac{1}{c}+a}$ and $\frac{1}{1+\frac{1}{b}+c} = \frac{1}{c(ab+a+1)} = \frac{\frac{1}{c}}{1+\frac{1}{c}+a}$. Thus

$$\begin{aligned} & \sqrt{2}\left(\frac{1}{1 + \frac{1}{a} + b} + \frac{1}{1 + \frac{1}{b} + c} + \frac{1}{1 + \frac{1}{c} + a}\right) \\ = \quad & \sqrt{2} \cdot \frac{1 + \frac{1}{c} + a}{1 + \frac{1}{c} + a} = \sqrt{2}. \end{aligned}$$

The equality can never hold.

8. Denote $T = S/2$. One of the triples (a,b,c) and (x,y,z) has the property that at least two of its members are greater than or equal to T. Assume that (a,b,c) is the one, and choose $\alpha = a - T$, $\beta = b - T$, and $\gamma = c - T$. We then have $x = T - \alpha$, $y = T - \beta$, and $z = T - \gamma$. Now the required inequality is equivalent to

$$(T + \alpha)(T - \beta) + (T + \beta)(T - \gamma) + (T + \gamma)(T - \alpha) < 4T^2.$$

After simplifying we get that what we need to prove is

$$-(\alpha\beta+\beta\gamma+\gamma\alpha)<T^2. \tag{32}$$

We also know that at most one of the numbers α, β, γ is negative. If all are positive, there is nothing to prove. Assume that $\gamma<0$. Now (32) can be rewritten as $-\alpha\beta-\gamma(\alpha+\beta)<T^2$. Since $-\gamma<T$ we have that $-\alpha\beta-\gamma(\alpha+\beta)<-\alpha\beta+T(\alpha+\beta)$ and the last term is less than T since $(T-\alpha)(T-\beta)>0$.

9. Starting from $\frac{(a-b)^2}{b}=\frac{a^2}{b}-2a+b$ and similar equalitites for $(b-c)^2/c$ and $(c-a)^2/a$ we get the required inequality is equivalent to

$$(a+b+c)\left(\frac{(a-b)^2}{b}+\frac{(b-c)^2}{a}+\frac{(c-a)^2}{b}\right)\geq 4(a-b)^2. \tag{33}$$

By the Cauchy-Schwartz inequality we have that the left-hand side of (33) is greater than or equal to $(|a-b|+|b-c|+|c-a|)^2$. (33) now follows from $|b-c|+|c-a|\geq|a-b|$.

10. Note that

$$\begin{aligned} & x_1^2+x_2^2+x_3^2+x_4^2+x_5^2 \\ = \quad & \left(x_1^2+\frac{x_2^2}{3}\right)+\left(\frac{2x_2^2}{3}+\frac{x_3^2}{2}\right)+\left(\frac{x_3^2}{2}+\frac{2x_4^2}{3}\right)+\left(\frac{x_4^2}{3}+x_5^2\right). \end{aligned}$$

Now applying the inequality $a^2+b^2\geq 2ab$ we get

$$x_1^2+x_2^2+x_3^2+x_4^2+x_5^2\geq\frac{2}{\sqrt{3}}(x_1x_2+x_2x_3+x_3x_4+x_4x_5).$$

This proves that $a\geq\frac{2}{\sqrt{3}}$. In order to prove the other inequality it is sufficient to notice that for $(x_1,x_2,x_3,x_4,x_5)=(1,\sqrt{3},2,\sqrt{3},1)$ we have

$$x_1^2+x_2^2+x_3^2+x_4^2+x_5^2=\frac{2}{\sqrt{3}}(x_1x_2+x_2x_3+x_3x_4+x_4x_5).$$

11. Since $xy+yz+zx-2xyz=(x+y+z)(xy+yz+zx)-2xyz=T[2,1,0]+\frac{1}{6}T[1,1,1]$ the left part of the inequality follows immediately. In order to prove the other part notice that

$$\frac{7}{27}=\frac{7}{27}(x+y+z)^3=\frac{7}{27}\left(\frac{1}{2}T[3,0,0]+3T[2,1,0]+T[1,1,1]\right).$$

After multiplying both sides by 54 and cancel as many things as possible we get that the required inequality is equivalent to:

$$12T[2,1,0]\leq 7T[3,0,0]+5T[1,1,1].$$

This inequality is true because it follows by adding up the inequalities $2T[2,1,0]\leq 2T[3,0,0]$ and $10T[2,1,0]\leq 5T[3,0,0]+5T[1,1,1]$ (the first one is a consequence of the Muirhead's and the second one of the Schur's theorem for $\alpha=\beta=1$).

12. The expressions have to be homogenous in order to apply the Muirhead's theorem. First we divide both left and right-hand side by $(abc)^{\frac{4}{3}} = 1$ and after that we multiply both sides by $a^3b^3c^3(a+b)(b+c)(c+a)(abc)^{\frac{4}{3}}$. The inequality becomes equivalent to

$$2T\left[\frac{16}{3},\frac{13}{3},\frac{7}{3}\right]+T\left[\frac{16}{3},\frac{16}{3},\frac{4}{3}\right]+T\left[\frac{13}{3},\frac{13}{3},\frac{10}{3}\right]\geq 3T[5,4,3]+T[4,4,4].$$

The last inequality follows by adding the following three which are immediate consequences of the Muirhead's theorem:

1. $2T\left[\frac{16}{3},\frac{13}{3},\frac{7}{3}\right]\geq 2T[5,4,3]$,
2. $T\left[\frac{16}{3},\frac{16}{3},\frac{4}{3}\right]\geq T[5,4,3]$,
3. $T\left[\frac{13}{3},\frac{13}{3},\frac{10}{3}\right]\geq T[4,4,4]$.

The equality holds if and only if $a=b=c=1$.

13. The left-hand side can be easily transformed into $\frac{a^3(b+c)}{b^3+c^3}+\frac{b^3(c+a)}{c^3+a^3}+\frac{c^3(a+b)}{a^3+b^3}$. We now multiply both sides by $(a+b+c)(a^3+b^3)(b^3+c^3)(c^3+a^3)$. After some algebra the left-hand side becomes

$$\begin{aligned}L= \quad & T[9,2,0]+T[10,1,0]+T[9,1,1]+T[5,3,3]+2T[4,4,3]\\ & +T[6,5,0]+2T[6,4,1]+T[6,3,2]+T[7,4,0]+T[7,3,1],\end{aligned}$$

while the right-hand side transforms into

$$D=3(T[4,4,3]+T[7,4,0]+T[6,4,1]+T[7,3,1]).$$

According to Muirhead's theorem we have:

1. $T[9,2,0]\geq T[7,4,0]$,
2. $T[10,1,0]\geq T[7,4,0]$,
3. $T[6,5,0]\geq T[6,4,1]$,
4. $T[6,3,2]\geq T[4,4,3]$.

The Schur's inequality gives us $T[4,2,2]+T[8,0,0]\geq 2T[6,2,0]$. After multiplying by abc, we get:

5. $T[5,3,3]+T[9,1,1]\geq T[7,3,1]$.

Adding up $1,2,3,4,5$, and adding $2T[4,4,3]+T[7,4,0]+2T[6,4,1]+T[7,3,1]$ to both sides we get $L\geq D$. The equality holds if and only if $a=b=c$.

14. Multiplying the both sides with the common denominator we get

$$T_{5,5,5}+4T_{7,5,0}+T_{5,2,2}+T_{9,0,0}\geq T_{5,5,2}+T_{6,0,0}+2T_{5,4,0}+2T_{4,2,0}+T_{2,2,2}.$$

By Schur's and Muirhead's inequalities we have that $T_{9,0,0}+T_{5,2,2}\geq 2T_{7,2,0}\geq 2T_{7,1,1}$. Since $xyz\geq 1$ we have that $T_{7,1,1}\geq T_{6,0,0}$. Therefore

$$T_{9,0,0}+T_{5,2,2}\geq 2T_{6,0,0}\geq T_{6,0,0}+T_{4,2,0}.$$

Moreover, Muirhead's inequality combined with $xyz\geq 1$ gives us $T_{7,5,0}\geq T_{5,5,2}$, $2T_{7,5,0}\geq 2T_{6,5,1}\geq 2T_{5,4,0}$, $T_{7,5,0}\geq T_{6,4,2}\geq T_{4,2,0}$, and $T_{5,5,5}\geq T_{2,2,2}$. Adding these four inequalities to (1) yields the desired result.

15. Let $a_i = e^{x_i}$ and let $(m_1,\ldots,m_n)$, $(k_1,\ldots,k_n)$ be two permutations of $(1,\ \ldots,\ n)$ for which the sequences $(3x_{m_1} - x_{m_1+1},\ldots,3x_{m_n} - x_{m_n+1})$ and $(2x_{k_1},\ \ldots,\ 2x_{k_n})$ are non-increasing. As above we assume that $x_{n+1} = x_n$. Similarly as in the problem 11 from the section 5 we prove that $(2x_{k_i}) \prec (3x_{m_i} - x_{m_i+1})$. The function $f(x) = e^x$ is convex so the Karamata's implies the required result.

16. Hint: Choose x_i such that $a_i = e^{x_i}$. Sort the sequences $(2x_1 - x_2,\ldots,2x_n - x_1)$ and $(x_1,\ldots,x_n)$ in non-increasing order, prove that the first majorizes the second, and apply Karamata's inequality with the convex function $f(x) = 1 + e^x$.

17. Applying the Chebyshev's inequality first we get

$$\frac{a^n}{b+c} + \frac{b^n}{c+a} + \frac{c^n}{a+b} \geq \frac{a^n+b^n+c^n}{3} \cdot \left(\frac{1}{a+b} + \frac{1}{b+c} + \frac{1}{c+a}\right).$$

The Cauchy-Schwartz inequality gives:

$$2(a+b+c)\left(\frac{1}{a+b} + \frac{1}{b+c} + \frac{1}{c+a}\right) \geq 9,$$

and the inequality $M_n \geq M_2$ gives

$$\frac{a^n+b^n+c^n}{3} \geq \left(\frac{a+b+c}{3}\right)^n.$$

In summary

$$\begin{aligned}\frac{a^n}{b+c} + \frac{b^n}{c+a} + \frac{c^n}{a+b} &\geq \left(\frac{a+b+c}{3}\right)^n \left(\frac{1}{a+b} + \frac{1}{b+c} + \frac{1}{c+a}\right) \\ &\geq \frac{1}{3}\cdot\frac{1}{2}\cdot\left(\frac{2}{3}s\right)^{n-1}\cdot 9 = \left(\frac{2}{3}\right)^{n-2} s^{n-1}.\end{aligned}$$

18. It is enough to prove that

$$\begin{aligned}&\left(\sqrt{x_1} + \frac{1}{\sqrt{x_1}}\right) + \left(\sqrt{x_2} + \frac{1}{\sqrt{x_2}}\right) + \cdots + \left(\sqrt{x_n} + \frac{1}{\sqrt{x_n}}\right) \\ \geq\ & n\left(\frac{1}{\sqrt{x_1}} + \frac{1}{\sqrt{x_2}} + \cdots + \frac{1}{\sqrt{x_n}}\right),\end{aligned}$$

or equivalently

$$\begin{aligned}&\left(\frac{1+x_1}{\sqrt{x_1}} + \cdots + \frac{1+x_n}{\sqrt{x_n}}\right)\left(\frac{1}{1+x_1} + \frac{1}{1+x_2} + \cdots + \frac{1}{1+x_n}\right) \\ \geq\ & n\cdot\left(\frac{1}{\sqrt{x_1}} + \frac{1}{\sqrt{x_2}} + \cdots + \frac{1}{\sqrt{x_n}}\right).\end{aligned}$$

Consider the function $f(x) = \sqrt{x} + \frac{1}{\sqrt{x}} = \frac{x+1}{\sqrt{x}}, x \in (0,+\infty)$. It is easy to verify that f is non-decreasing on$(1,+\infty)$ and that $f(x) = f\left(\frac{1}{x}\right)$ for every $x > 0$. Furthermore from the given

conditions it follows that only x_1 can be less than 1 and that $\frac{1}{1+x_2} \leq 1 - \frac{1}{1+x_1} = \frac{x_1}{1+x_1}$. Hence $x_2 \geq \frac{1}{x_1}$. Now it is clear that (in both of the cases $x_1 \geq 1$ and $x_1 < 1$):

$$f(x_1) = f\left(\frac{1}{x_1}\right) \leq f(x_1) \leq \cdots \leq f(x_n).$$

This means that the sequence $\left(\frac{1+x_k}{x_k}\right)_{k=1}^{n}$ is non-decreasing. Thus according to the Chebyshev's inequality we have:

$$\begin{aligned} & \left(\frac{1+x_1}{\sqrt{x_1}} + \cdots + \frac{1+x_n}{\sqrt{x_n}}\right)\left(\frac{1}{1+x_1} + \frac{1}{1+x_2} + \cdots + \frac{1}{1+x_n}\right) \\ \geq \quad & n \cdot \left(\frac{1}{\sqrt{x_1}} + \frac{1}{\sqrt{x_2}} + \cdots + \frac{1}{\sqrt{x_n}}\right). \end{aligned}$$

The equality holds if and only if $\frac{1}{1+x_1} = \cdots = \frac{1}{1+x_n}$, or $\frac{1+x_1}{\sqrt{x_1}} = \cdots = \frac{1+x_n}{\sqrt{x_n}}$, which implies that $x_1 = x_2 = \cdots = x_n$. Thus the equality holds if and only if $x_1 = \cdots = x_n = n-1$.

19. Denote by S the set of all members of the society, by A the set of all pairs of friends, and by N the set of all pairs of enemies. For every $x \in S$, denote by $f(x)$ number of friends of x and by $F(x)$ number of pairs of friends among enemies of x. It is easy to prove:

$$q = |A| = \frac{1}{2}\sum_{x \in S} f(x);$$

$$\sum_{\{a,b\} \in A} (f(a) + f(b)) = \sum_{x \in S} f^2(x).$$

If a and b are friends, then the number of their common enemies is equal to $(n-2) - (f(a) - 1) - (f(b) - 1) = n - f(a) - f(b)$. Thus

$$\frac{1}{n}\sum_{x \in S} F(x) = \frac{1}{n}\sum_{\{a,b\} \in A} (n - f(a) - f(b)) = q - \frac{1}{n}\sum_{x \in S} f^2(x).$$

Using the inequality between arithmetic and quadratic mean on the last expression, we get

$$\frac{1}{n}\sum_{x \in S} F(x) \leq q - \frac{4q^2}{n^2}$$

and the statement of the problem follows immediately.

20. Consider the partition of plane π into regular hexagons, each having inradius 2. Fix one of these hexagons, denoted by γ. For any other hexagon x in the partition, there exists a unique translation τ_x taking it onto γ. Define the mapping $\varphi : \pi \to \gamma$ as follows: If A belongs to the interior of a hexagon x, then $\varphi(A) = \tau_x(A)$ (if A is on the border of some hexagon, it does not actually matter where its image is).

The total area of the images of the union of the given circles equals S, while the area of the hexagon γ is $8\sqrt{3}$. Thus there exists a point B of γ that is covered at least $\frac{S}{8\sqrt{3}}$ times, i.e.,

such that $\varphi^{-1}(B)$ consists of at least $\dfrac{S}{8\sqrt{3}}$ distinct points of the plane that belong to some of the circles. For any of these points, take a circle that contains it. All these circles are disjoint, with total area not less than $\dfrac{\pi}{8\sqrt{3}}S \geq 2S/9$.

klaar voor de oefeningen