platform wiskunde nederland

NIEUWE TIJDEN

Syllabus Vakantiecursus 2014

Eindhoven, 22 en 23 augustus

Amsterdam, 29 en 30 augustus

platform wiskunde nederland

NIEUWE TIJDEN

Syllabus Vakantiecursus 2014

(uitgebreide en gecorrigeerde web-versie)

Eindhoven, 22 en 23 augustus

Amsterdam, 29 en 30 augustus

Programmacommissie

prof. dr. Frits Beukers (UU)

drs. Joke Blom (CWI)

drs. Marian Kolleveld (NVvW)

prof. dr. Jan van Maanen (RUG)

prof. dr. ir. Wil Schilders (PWN, TU/e)

dr. ir. Ionica Smeets (ionica.nl)

dr. Jeroen Spandaw (TUD)

dr. Marco Swaen (UvA)

dr. Benne de Weger (TU/e)

prof. dr. Jan Wiegerinck (UvA) (voorzitter)

e-mail: `vakantiecursus@platformwiskunde.nl`

Platform Wiskunde Nederland

Science Park 123, 1098 XG Amsterdam

Telefoon: 020-592 4006

Website: `http://www.platformwiskunde.nl`

Vakantiecursus 2014

De Vakantiecursus Wiskunde voor leraren in de exacte vakken in HAVO, VWO, HBO en andere belangstellenden is een initiatief van de Nederlandse Vereniging van Wiskundeleraren, en wordt georganiseerd door het Platform Wiskunde Nederland. De cursus wordt sinds 1946 jaarlijks gegeven op het Centrum Wiskunde en Informatica te Amsterdam en aan de Technische Universiteit Eindhoven.

Deze cursus wordt mede mogelijk gemaakt door een subsidie van de Nederlandse Organisatie voor Wetenschappelijk Onderzoek (NWO), en een bijdrage van 3TU.AMI, het toegepaste wiskunde-instituut van de 3 Nederlandse technische universiteiten. Organisatie vindt plaats in nauwe samenwerking met het Centrum voor Wiskunde en Informatica (CWI) en de Technische Universiteit Eindhoven (TU/e).

De presentaties van de sprekers zullen zo veel mogelijk beschikbaar komen op de PWN-website:
`http://www.platformwiskunde.nl/onderwijs_vakantiecursus_wiskunde.htm`.

Met dank aan

Ondersteuning PWN en CWI: Coby van Vonderen, Sjoukje Talsma.
Ondersteuning TU/e: Anita Klooster, Wil Kortsmit.

Historie

De eerste vakantiecursus wordt in het jaarverslag 1946 van het Mathematisch Centrum als volgt vermeld:

> *Op 29 en 31 Oct. '46 werd onder auspiciën van het M.C. een druk bezochte en uitstekend geslaagde vacantiecursus gehouden voor wiskundeleeraren in Nederland. Op 29 October stond de wiskunde, op 31 October de didactiek van de wiskunde op de voorgrond. De sprekers waren: Prof.Dr. O. Bottema, "De prismoide", Dr. A. Heyting, "Punten in het oneindige", Mr. J. v. IJzeren, "Abstracte Meetkunde en haar betekenis voor de Schoolmeetkunde.", Dr. H.D. Kloosterman, "Ontbinding in factoren", Dr. G. Wielenga, "Is wiskunde-onderwijs voor alpha's noodzakelijk?", Dr. J. de Groot, "Het scheppend vermogen van den wiskundige" en Dr. N.L.H. Bunt, "Moeilijkheden van leerlingen bij het beginnend onderwijs in de meetkunde".*
>
> *Aan het einde van de vacantiecursus werden diverse zaken besproken die het wiskunde-onderwijs in Nederland betroffen. Een Commissie werd ingesteld, die het M.C. over de verder te organiseren vakantiecursussen van advies zou dienen. Hierin namen zitting een vertegenwoordiger van de Inspecteurs van het V.H. en M.O. benevens vertegenwoordigers van de lerarenverenigingen Wimecos en Liwenagel.*
>
> *Ook werd naar aanleiding van "wenschen" die tijdens de cursus naar voren gekomen waren ingesteld: "een colloquium over moderne Algebra, een dispuut over de didactiek van de wiskunde, beiden hoofdzakelijk bedoeld voor de leeraren uit Amsterdam en omgeving, terwijl tevens vanwege het M.C. een cursus over Getallenleer werd toegezegd te geven door de heeren v.d. Corput en Koksma. (Colloquium, dispuut en cursus zijn in 1947 gestart en verheugen zich in blijvende belangstelling).*

Docenten

drs. Theodora Achourioti
Amsterdam University College / Institute for Logic, Language and Computation, Universiteit van Amsterdam
e-mail: t.achourioti@uva.nl

dr. Alessandro Di Bucchianico
Faculteit Wiskunde en Informatica, Technische Universiteit Eindhoven
e-mail: a.d.bucchianico@tue.nl

dr. Hans van Ditmarsch
CNRS & LORIA, Nancy, Frankrijk
e-mail: hans.van-ditmarsch@loria.fr

prof. dr. Barteld Kooi
Theoretische Filosofie & Kunstmatige Intelligentie, Faculteit Wijsbegeerte, Rijksuniversiteit Groningen
e-mail: b.p.kooi@rug.nl

prof. dr. Paul Levrie
Faculteit Toegepaste Ingenieurswetenschappen, Universiteit Antwerpen
e-mail: paul.levrie@uantwerpen.be

prof. dr. Rudi Penne
Faculteit Toegepaste Ingenieurswetenschappen, Universiteit Antwerpen
e-mail: rudi.penne@uantwerpen.be

dr. Wouter Verkerke
FOM / Nikhef, Amsterdam
e-mail: verkerke@nikhef.nl

dr. Steven Wepster
Mathematisch Instituut, Universiteit Utrecht
e-mail: s.a.wepster@uu.nl

prof. dr. Mark van de Wiel
Epidemiologie & Biostatistiek, VU Medisch Centrum en Afdeling Wiskunde, Vrije Universiteit Amsterdam
e-mail: mark.vdwiel@vumc.nl

dr. Wessel van Wieringen
Epidemiologie & Biostatistiek, VU Medisch Centrum en Afdeling Wiskunde, Vrije Universiteit Amsterdam
e-mail: w.vanwieringen@vumc.nl

Programma

Vrijdag 22 / 29 augustus 2014

15.00–15.25		*Ontvangst, koffie*
15.25–15.35	Wiegerinck	Introductie “Nieuwe tijden”
15.35–16.20	Levrie en Penne	Priem!
16.20–16.45		*Pauze*
16.45–17:30	Wepster	Analytische meetkunde
17.30–18.30		*Diner*
18.30–19.15	Achourioti	Some basics of classical logic
19.15–19.45		*Pauze*
19.45–20.30	Van Ditmarsch, Kooi	Honderd gevangenen en een gloeilamp

Zaterdag 23 / 30 augustus 2014

Speciale dag over Big Data

10.00-10.30		*Ontvangst, koffie*
10.30-11.15	Verkerke	De ontdekking van het Higgs deeltje
11.15-12.00	Di Bucchianico	Practicum 1
12.00-13.00		*Lunch*
13.00-13.45	Van de Wiel, Van Wieringen	Statistiek voor gen-gen-interactienetwerken van de cel
13.45-14.30	Di Bucchianico	Practicum 2
14.30		Afsluiting

Ten geleide

Jan Wiegerinck

Universiteit van Amsterdam

De nieuwe examenprogramma's voor wiskunde worden per 2015 ingevoerd op HAVO en VWO. Naast deze 'nieuwe tijden' in onderwijs zijn er ook nieuwe tijden in wiskundig onderzoek, en de vakantiecursus van dit jaar zal daarom ook om dit thema draaien. De focus zal liggen op onderwerpen welke ook in de klas gebruikt kunnen worden. Big Data van de bonuskaart tot gepersonaliseerde marketing, Kennislogica, Meetkunde in de architectuur, Analytische meetkunde in historisch perspectief en Priemtweelingen zijn de onderwerpen welke gepresenteerd zullen worden. Ook zal er tijd zijn voor zelfwerkzaamheid; het afgelopen jaar is deze opnieuw geïntroduceerd in de vakantiecursus, en werd alom als zeer nuttig en fijn ervaren. Dit jaar draait de zelfwerkzaamheid om het thema "Data Science", ook wel "de 4e discipline" genoemd naast theorie, experiment en simulatie.

Wij hopen u te mogen verwelkomen op wederom een zeer interessante vakantiecursus, inmiddels een traditie van 65 jaar, maar absoluut niet pensioengerechtigd gezien de nog immer grote belangstelling. De vakantiecursus wil ook als nascholing serieus genomen worden, en dat betekent dat we verder gaan dan het aanbieden van een aantal hoorcolleges. Naast de zelfwerkzaamheidsessies zijn er ook altijd interessante gesprekken aan te knopen met degenen die de voordrachten geven, en zijn het diner en de lunch uitstekende gelegenheden voor het leggen van contacten en het opdoen van additionele kennis. Daarnaast is de vakantiecursus ook als professionaliseringsactiviteit aangemeld bij het Lerarenregister.

Ik hoop en verwacht dat u een leerzame en leuke cursus tegemoet kunt zien!

Inhoudsopgave

1 Priem!

Paul Levrie en Rudi Penne

Faculteit TI, Universiteit Antwerpen

Voorwoord

De titel van dit artikel is, zoals de lezer kan merken, eerder ruim. Toen ons gevraagd werd om voor de syllabus van deze vakantiecursus een bijdrage te leveren over een onderwerp waar we net een boek [5] over hadden geschreven, leek het ons onzinnig om dubbel werk te leveren. Daarom grijpen we liever de kans om een nieuw hoofdstuk te breien aan dit boek. Ons verhaal over de priemgetallen is een mix van wiskunde, geschiedenis, en leuke weetjes. Omdat het doelpubliek anders is dan dat van het boek, mag het wat meer wiskunde zijn, al zullen we de lezer onderweg verrassen met een heuse snit-en-naad-toepassing. Het begint allemaal met dit:

> Een priemgetal is een natuurlijk getal groter dan 1
> dat enkel deelbaar is door 1 en door zichzelf.

En we gaan het hebben over een stelling die ons vertelt welke priemgetallen te schrijven zijn als een som van twee kwadraten.
Maar eerst een beetje geschiedenis:

1634 De minder bekende Franse wiskundige Albert Girard (1595-1632) schrijft een opmerking (rechts) in de Franse vertaling van het werk van Simon Stevin.

ALB. GIR. *Determinaiſon d'un nombre qui ſe peut diviſer en deux quarrez entiers.*

I. Tout nombre quarré.
II. Tout nombre premier qui excede un nombre quaternaire de l'unité.
III. Le produict de ceux qui ſont tels.
IV. Et le double d'un chaſcun d'iceux.

1640 Pierre de Fermat (1601-1665) schrijft op kerstdag in een brief naar collega Marin Mersenne (1588-1648):

1° Tout nombre premier, qui surpasse de l'unité un multiple du quaternaire, est une seule fois la somme de deux quarrés, et une seule fois l'hypoténuse d'un triangle rectangle.

Fermat beweerde dat hij een onweerlegbaar bewijs had van deze stelling (zo kennen we hem).

1742 Leonhard Euler (1707-1783) schrijft op 30 juni in een brief aan Christian Goldbach:

seyn, welches aber nicht ist. Denn alle diese Zahlen sind in dieser formula $4m + 1$ enthalten, welche so oft sie ein numerus primus ist, unfehlbar in duo quadrata, hocque unico modo, resolviret werden kann. So oft aber $4m + 1$ kein

1749 Euler schrijft op 12 april naar Goldbach:

Nunmehr habe ich endlich einen bündigen Beweis gefunden, dass ein jeglicher numerus primus von dieser Form $4n + 1$ eine summa duor. quadr. ist. Es sey $\boxed{2}$ das Zeichen der

1754 Euler publiceert een *Demonstratio theorematis Fermatiani omnem numerum primum formae 4n+1 esse summam duorum quadratorum.*

1848 Charles Hermite (1822-1901) en Joseph-Alfred Serret (1819-1885) vinden tegelijkertijd een efficiënt algoritme om een priemgetal van de vorm $4n + 1$ te schrijven als de som van twee kwadraten.

1855 Henry John Stephen Smith (1826-1883) geeft een eerste eenvoudig bewijs van de stelling.

1867 Édouard Lucas (1842-1891) ziet een toepassing van de de stelling van Fermat, die ook wel eens Fermat's Kerstmis-stelling wordt genoemd, bij het weven van satijn.

1940 Godfrey Harold Hardy (1877-1947) schrijft in zijn *A Mathematician's Apology* dit: →

Another famous and beautiful theorem is Fermat's 'two square' theorem. The primes may (if we ignore the special prime 2) be arranged in two classes; the primes

5, 13, 17, 29, 37, 41, ...

which leave remainder 1 when divided by 4, and the primes

3, 7, 11, 19, 23, 31, ...

which leave remainder 3. All the primes of the first class, and none of the second, can be expressed as the sum of two integral squares: thus

$$5 = 1^2 + 2^2, \quad 13 = 2^2 + 3^2,$$
$$17 = 1^2 + 4^2, \quad 29 = 2^2 + 5^2;$$

but 3, 7, 11, and 19 are not expressible in this way (as the reader may check by trial). This is Fermat's theorem, which is ranked, very justly, as one of the finest of arithmetic. Unfortunately, there is no proof within the comprehension of anybody but a fairly expert mathematician.

Uit de laatste zin blijkt dat Hardy het bewijs van zijn landgenoot Smith, die evenals hijzelf de Savilian Chair of Geometry bekleedde in Oxford, niet kende.

1984 Roger Heath-Brown (1952) geeft een nieuw, kort bewijs van de stelling.

1990 Don Zagier (1951) publiceert zijn *A One-Sentence Proof That Every Prime $p \equiv 1 \pmod 4$ Is a Sum of Two Squares.*

1.1 Additieve getaltheorie en de stelling van Fermat

Dat getaltheorie verslavend kan zijn, kan je nu ook lezen in de laatste editie van het beroemde boek *An introduction to the theory of numbers*, van Hardy en Wright [3]. Je vindt er namelijk in de introductie:

> Thus chs. XII–XV belong to the 'algebraic' theory of numbers, Chs. XIX–XXI to the 'addictive', and Ch. XXII to the 'analytic' theories; while Chs. III, XI, XXIII, and XXIV deal with matters usually classified under the headings of 'geometry of numbers' or 'Diophantine approximation'.

Bedoeld wordt natuurlijk 'additive'. De additieve getaltheorie behandelt eigenschappen van getallen waarbij men deze probeert te schrijven als som van andere speciale getallen. Een van de mooiste resultaten op dit gebied is de stelling van Fermat, die in haar meest ruwe vorm zegt:

> *Elk priemgetal van de vorm $4n+1$ is te schrijven als de som van twee kwadraten.*

Iets meer verfijnd hebben we dit:

> *Een getal kan geschreven worden als de som van twee kwadraten als alle priemfactoren van de vorm $4n+3$ in de priemfactorisatie van het gegeven getal voorkomen met een even macht.*

Dat de stap van de ruwe naar de verfijnde versie niet zo groot is, volgt uit twee dingen.

Eerst en vooral is het eenvoudig in te zien dat een priemgetal van de vorm $4n+3$ niet te schrijven is als een som van twee kwadraten. Indien dit wel het geval zou zijn, dan moet het gaan om een kwadraat van een even getal en een kwadraat van een oneven getal. Een even getal kunnen we voorstellen door $2k$ bijvoorbeeld, en een oneven getal door $2m+1$. We hebben dan voor de som van de kwadraten:

$$(2k)^2 + (2m+1)^2 = 4k^2 + (4m^2 + 4m + 1) = 4(k^2 + m^2 + m) + 1$$

en dit geeft dus steeds een viervoud plus 1, en nooit plus 3.

Verder kunnen we gebruik maken van de wonderbaarlijke eigenschap die ons toelaat een product van twee sommen van kwadraten te schrijven als een som van kwadraten:

$$(a^2 + b^2)(c^2 + d^2) = (ac - bd)^2 + (ad + bc)^2$$

(eenvoudig te zien door uitwerking).

Voor een willekeurig getal dat voldoet aan de voorwaarden van de verfijnde stelling gaan we nu als volgt te werk:

- de factoren 2 schrijven we als 1^2+1^2 en de priemfactoren van de vorm $4n+1$ schrijven we als som van twee kwadraten, en met de vorige eigenschap kunnen we dan het product van deze getallen voorstellen als som van twee kwadraten: A^2+B^2;
- de andere factoren hebben een even macht, en zijn samen te herschrijven in de vorm C^2, en die kunnen we dan gewoon op de volgende manier 'binnenbrengen':

$$(A^2+B^2)C^2 = (AC)^2 + (BC)^2.$$

Onze bedoeling is om in de rest van dit artikel het bewijs te geven van Henry Smith uit 1855 voor de ruwe vorm van de stelling van Fermat. We volgen in essentie de lijnen van [1], hoewel Smith gebruik maakte van *continuanten*, dingen waarvan de meeste mensen nooit gehoord hebben. In plaats daarvan zullen wij werken met *kettingbreuken*, die we o.a. kennen van [4]. Kettingbreuken hebben te maken met lineaire recursiebetrekkingen, en hierover is er een inleiding te vinden in appendix A. Het verband met de kettingbreuken lees je dan weer in appendix B. Dus als je alle details wil weten, lees dan de appendices. Soms verwijzen we naar een formule in een appendix, je herkent die verwijzingen aan de vorm: (A.#) of (B.#). Maar je kan het verhaal ook volgen zonder de details van de appendices.

Eerst vertellen we wat meer over kettingbreuken, en in de volgende paragraaf geven we dan onze eigen versie van het bewijs van Smith. Daarna gaan we op zoek naar de twee kwadraten, wiskundig, met de kettingbreuken, maar ook grafisch, via de toepassing beschreven door Édouard Lucas in 1867 (zie geschiedkundig overzicht).

1.2 Eindige kettingbreuken

Voor elk niet-geheel rationaal getal kan je wat men noemt een eindige kettingbreuk vinden. We laten dit zien aan de hand van het voorbeeld $\frac{39}{16}$. We bepalen het grootste geheel getal dat in de breuk in kwestie past, en schrijven

$$\frac{39}{16} = 2 + \frac{7}{16} = 2 + \frac{1}{\frac{16}{7}}.$$

Met de noemer van de tweede term in het rechterlid gaan we dan op dezelfde manier verder:

$$\frac{39}{16} = 2 + \frac{1}{\frac{16}{7}} = 2 + \frac{1}{2 + \frac{2}{7}} = 2 + \frac{1}{2 + \frac{1}{\frac{7}{2}}} = 2 + \cfrac{1}{2 + \cfrac{1}{3 + \cfrac{1}{2}}}.$$

Dit proces eindigt altijd en de laatste noemer is steeds een geheel getal ≥ 2 (waarom kan dit niet gelijk zijn aan 1?).

De kettingbreuk van een positief niet-geheel rationaal getal heeft dus de volgende vorm:

$$b_1 + \cfrac{1}{b_2 + \cfrac{1}{\ddots + \cfrac{1}{b_k}}} \tag{1.1}$$

waarbij alle b's positieve gehele getallen zijn. We zullen de notatie $T_{1\to k}$ en $N_{1\to k}$ gebruiken voor de teller en de noemer van het rationaal getal dat we krijgen indien we deze kettingbreuk uitwerken:

$$b_1 + \cfrac{1}{b_2 + \cfrac{1}{\ddots + \cfrac{1}{b_{k-1} + \cfrac{1}{b_k}}}} = \frac{T_{1\to k}}{N_{1\to k}} \tag{1.2}$$

(dus $1 \to k$ betekent dat de b's geordend zijn van b_1 tot b_k).

Elke stap in deze uitwerking is van de vorm

$$b_j + \frac{1}{N} = \frac{T}{N}$$

met T, N gehele getallen, waarbij T en N geen gemeenschappelijke factor kunnen hebben. In onze notatie veronderstellen we dus dat $T_{1\to k}$ en $N_{1\to k}$ onderling ondeelbaar zijn.
Indien we de b's in de omgekeerde volgorde zetten, dan noteren we het resultaat zo:

$$b_k + \cfrac{1}{b_{k-1} + \cfrac{1}{\ddots + \cfrac{1}{b_2 + \cfrac{1}{b_1}}}} = \frac{T_{k\to 1}}{N_{k\to 1}}. \tag{1.3}$$

Merkwaardig genoeg is in beide gevallen de teller gelijk, d.w.z. $T_{1\to k} = T_{k\to 1}$. We kijken even terug naar het voorbeeld aan het begin van deze paragraaf. Als we de volgorde van de b's hierin omkeren, dan krijgen we dit:

$$2 + \cfrac{1}{3 + \cfrac{1}{2 + \cfrac{1}{2}}} = \frac{39}{17}$$

met inderdaad dezelfde teller. Deze eigenschap van kettingbreuken blijkt essentieel te zijn, en is een gevolg van (B.1) en (B.2).

Meer algemeen zullen we de notaties $T_{i\to j}$ en $N_{i\to j}$ gebruiken voor de teller en de noemer van de (uitgewerkte) kettingbreuk waarbij de b's lopen van b_i tot b_j en waarbij $T_{i\to j}$ en $N_{i\to j}$ onderling ondeelbaar zijn.

1.3 Het bewijs van de stelling van Fermat

De stelling die we willen bewijzen zegt dit:

Elk priemgetal van de vorm $p = 4n + 1$ is te schrijven als de som van 2 kwadraten.

De essentie van het bewijs is het feit dat de tellers in (1.2) en (1.3) aan elkaar gelijk zijn, in combinatie met de volgende merkwaardige formule:

$$T_{1\to k} = T_{j\to 1} \cdot T_{j+1\to k} + N_{j\to 1} \cdot N_{j+1\to k}. \tag{1.4}$$

Dit is formule (B.5) uit de appendix. We kunnen ze ook zo schrijven (als $j + 2 \le k$):

$$T_{1\to k} = T_{j\to 1} \cdot T_{j+1\to k} + N_{j\to 1} \cdot T_{j+2\to k} \tag{1.5}$$

want we hebben natuurlijk dat

$$\frac{T_{j+1\to k}}{N_{j+1\to k}} = b_{j+1} + \cfrac{1}{b_{j+2} + \cfrac{1}{\ddots + \cfrac{1}{b_k}}} = b_{j+1} + \frac{1}{\frac{T_{j+2\to k}}{N_{j+2\to k}}}$$

en na uitwerking van het rechterlid volgt hieruit door het onderling ondeelbaar zijn van T en bijhorende N dat $N_{j+1\to k} = T_{j+2\to k}$.

We vertrekken van een priemgetal p dat één meer is dan een viervoud, en we stellen de kettingbreuk op voor breuken waarvan de teller gelijk is aan p. Stel dat zo'n kettingbreuk opgebouwd is met de getallen $b_1, b_2, \ldots, b_k$, dan worden (1.4) en (1.5) dus:

$$\boxed{\begin{aligned} p &= T_{j\to 1} \cdot T_{j+1\to k} + N_{j\to 1} \cdot N_{j+1\to k} \\ &= T_{j\to 1} \cdot T_{j+1\to k} + N_{j\to 1} \cdot T_{j+2\to k} \end{aligned}}$$

We beperken ons hierbij tot het geval dat $b_1 \geq 2$, dat blijkt voldoende te zijn voor het bewijs. Een gevolg van deze veronderstelling is dat de beide breuken in het rechterlid van (1.2) en (1.3) een noemer hebben die ligt tussen 2 en $2n$ (herinner je dat $p = 4n + 1$). Inderdaad, we hebben dan dat

$$\frac{p}{\text{noemer}} > 2 \quad \Rightarrow \quad p = 4n + 1 > 2 \cdot \text{noemer}.$$

Noemer 1 sluiten we overigens uit, want dan hebben we geen breuk.

We bekijken de verschillende breuken eens voor $p = 13$, dus $n = 3$. Het gaat dan om breuken met teller 13 en noemer respectievelijk 2, 3, 4, 5, 6. Dit zijn de bijhorende kettingbreuken:

$$\frac{13}{2} = 6 + \frac{1}{2}, \quad \frac{13}{3} = 4 + \frac{1}{3}, \quad \frac{13}{4} = 3 + \frac{1}{4}, \quad \frac{13}{5} = 2 + \cfrac{1}{1 + \cfrac{1}{1 + \cfrac{1}{2}}}, \quad \frac{13}{6} = 2 + \frac{1}{6}$$

Let op het verband tussen de kettingbreuken die horen bij de breuken met noemers 2 en 6: de volgorde van de b's is omgewisseld. Ook die met noemers 3 en 4 komen op deze wijze met elkaar overeen. De overblijvende breuk, met noemer 5, valt op door het feit dat als je de volgorde van de b's in de kettingbreuk omkeert, dat deze niet verandert ten gevolge van een soort palindromisch effect.

Dit zal ook meer algemeen zo zijn: met elk getal uit $\{2, 3, 4, \ldots, 2n\}$ (=de mogelijke noemers) kunnen we een ander getal uit deze verzameling associëren op deze manier. Neem als getal j, bepaal de kettingbreuk voor $\frac{p}{j}$, keer de volgorde van de b's om zoals in (1.2) en (1.3), en dan is de noemer van de resulterende breuk opnieuw een getal uit $\{2, 3, 4, \ldots, 2n\}$.

Omdat het aantal elementen in de verzameling $\{2, 3, 4, \ldots, 2n\}$ gelijk is aan $2n-1$, een oneven getal, en omdat we deze dus kunnen verdelen in groepjes

van 2 noemers die bij elkaar horen, zal er altijd 1 noemer overblijven, die dan ook deze palindromische eigenschap zal moeten hebben:

$$b_1 = b_k,\ b_2 = b_{k-1}, \ldots .$$

Hiermee werken we verder. We onderscheiden nu twee gevallen.

Ofwel is het aantal b's voor deze ene speciale breuk even, stel dus $k = 2j$. Dan ziet de breuk er zo uit:

$$b_1 + \cfrac{1}{\ddots + \cfrac{1}{b_j + \cfrac{1}{b_{j+1} + \cfrac{1}{\ddots + \cfrac{1}{b_k}}}}} = b_1 + \cfrac{1}{\ddots + \cfrac{1}{b_j + \cfrac{1}{b_j + \cfrac{1}{\ddots + \cfrac{1}{b_1}}}}}$$

want inderdaad geldt:

$$b_{j+1} + \cfrac{1}{\ddots + \cfrac{1}{b_k}} = b_j + \cfrac{1}{\ddots + \cfrac{1}{b_1}}.$$

Nu kunnen we dit zo schrijven:

$$\frac{T_{j+1\to k}}{N_{j+1\to k}} = \frac{T_{j\to 1}}{N_{j\to 1}} \tag{1.6}$$

en hieruit volgt wegens het onderling ondeelbaar zijn van tellers en noemers dat $T_{j+1\to k} = T_{j\to 1}$, noem dit geheel getal a, en ook $N_{j+1\to k} = N_{j\to 1}$, noem dit b.

De eerste vergelijking in het kadertje wordt dan:

$$p = T_{j\to 1} \cdot T_{j+1\to k} + N_{j\to 1} \cdot N_{j+1\to k} = (T_{j\to 1})^2 + (N_{j\to 1})^2$$

of nog

$$p = a^2 + b^2.$$

In dit eerste geval is de zaak dus bewezen.

Ofwel is het aantal b's voor de speciale breuk oneven, stel dus $k = 2j + 1$. Dan ziet de breuk er zo uit:

$$\begin{aligned} & b_1 + \cfrac{1}{\ddots + \cfrac{1}{b_j + \cfrac{1}{b_{j+1} + \cfrac{1}{b_{j+2} + \cfrac{1}{\ddots + \cfrac{1}{b_k}}}}}} \\ = {} & b_1 + \cfrac{1}{\ddots + \cfrac{1}{b_j + \cfrac{1}{b_{j+1} + \cfrac{1}{b_j + \cfrac{1}{\ddots + \cfrac{1}{b_1}}}}}} \end{aligned}$$

en dus geldt:

$$b_{j+2} + \cfrac{1}{\ddots + \cfrac{1}{b_k}} = b_j + \cfrac{1}{\ddots + \cfrac{1}{b_1}}.$$

Dit laatste kunnen we nu zo schrijven:

$$\frac{T_{j+2\to k}}{N_{j+2\to k}} = \frac{T_{j\to 1}}{N_{j\to 1}}$$

en hieruit volgt dan onmiddellijk dat $T_{j+2\to k} = T_{j\to 1}$. Uit het tweede deel van de vergelijking in het kadertje volgt dan:

$$p = T_{j\to 1} \cdot T_{j+1\to k} + N_{j\to 1} T_{j\to 1} = T_{j\to 1} \cdot (T_{j+1\to k} + N_{j\to 1})$$

waarbij $T_{j\to 1}$ niet gelijk kan zijn aan 1. Uit deze laatste uitdrukking volgt dat p niet priem is, wat in strijd is met het gestelde.

We zijn er dus nu zeker van dat p te schrijven is als de som van twee kwadraten, en dat k in dit geval even is. Hiermee is het bewijs klaar.

1.4 Op zoek naar de twee kwadraten

We willen natuurlijk ook graag weten van welke twee kwadraten het priemgetal $p = 4n + 1$ de som is. Hiervoor proberen we meer te weten te komen over de noemer in de palindromische kettingbreuk, we stellen deze noemer voor door d. Hierbij gebruiken we de formule (B.4) uit de appendix:

$$N_{1\to k} \cdot T_{1\to k-1} - T_{1\to k} \cdot N_{1\to k-1} = (-1)^{k-1} = -1 \tag{1.7}$$

die we hier toepassen met k even. Hierbij is $T_{1\to k} = p$ en $N_{1\to k} = d$. Uit (1.2):

$$\frac{p}{d} = \frac{T_{1\to k}}{N_{1\to k}} = b_1 + \cfrac{1}{b_2 + \cfrac{1}{\ddots + \cfrac{1}{b_{k-1} + \cfrac{1}{b_k}}}} = b_1 + \frac{1}{\frac{T_{2\to k}}{N_{2\to k}}}$$

en uit het onderling ondeelbaar zijn van T en bijhorende N volgt dat d de teller is van de volgende kettingbreuk:

$$\frac{T_{2\to k}}{N_{2\to k}} = b_2 + \cfrac{1}{b_3 + \cfrac{1}{\ddots + \cfrac{1}{b_{k-1} + \cfrac{1}{b_k}}}}.$$

Maar we hebben vroeger gezien dat de teller van zo'n kettingbreuk niet verandert als we de volgorde van de b's omkeren. We krijgen dan de volgende kettingbreuk, die we anders kunnen schrijven door gebruik te maken van het palindromisch karakter van de b's:

$$b_k + \cfrac{1}{b_{k-1} + \cfrac{1}{\ddots + \cfrac{1}{b_3 + \cfrac{1}{b_2}}}} = b_1 + \cfrac{1}{b_2 + \cfrac{1}{\ddots + \cfrac{1}{b_{k-2} + \cfrac{1}{b_{k-1}}}}}$$

vermits $b_k = b_1$, $b_{k-1} = b_2$ enzoverder. De teller van deze laatste kettingbreuk is dus ook gelijk aan d. Dit betekent dat $T_{1\to k-1} = d$. De vergelijking (1.7) wordt dan:

$$\begin{aligned} d^2 - pN_{1\to k-1} = -1 \quad &\Leftrightarrow \quad d^2 + 1 = pN_{1\to k-1} \\ &\Leftrightarrow \quad d^2 + 1 \text{ is deelbaar door } p \end{aligned}$$

Het is niet moeilijk om in te zien dat er maar 1 getal d kan zijn tussen 2 en $2n$ met deze eigenschap. Stel namelijk dat er zo twee zijn, d_1 en d_2, met $d_1 > d_2$, met zowel $d_1^2 + 1$ als $d_2^2 + 1$ deelbaar door p. Dan volgt onmiddellijk dat ook het verschil $d_1^2 - d_2^2 = (d_1 + d_2)(d_1 - d_2)$ deelbaar zal zijn door het priemgetal p, maar beide factoren zijn kleiner dan p, dus dat kan niet.

Voor het bepalen van de unieke noemer d gaan we dan als volgt tewerk: bepaal het kwadraat van de getallen $2, 3, \ldots$ en tel er 1 bij op. Indien het resultaat deelbaar is door p, dan heb je d gevonden. Voor $p = 29$ vinden we voor die kwadraten plus 1 het volgende, op veelvouden van 29 na:

$$\begin{array}{l|l} 1 \to 2 & 7 \to 21 \\ 2 \to 5 & 8 \to 7 \\ 3 \to 10 & 9 \to 24 \\ 4 \to 17 & 10 \to 14 \\ 5 \to 26 & 11 \to 6 \\ 6 \to 8 & 12 \to 0 \end{array}$$

We besluiten dat $d = 12$. We bepalen nu de kettingbreuk voor $29/12$:

$$\frac{29}{12} = 2 + \cfrac{1}{2 + \cfrac{1}{2 + \cfrac{1}{2}}}$$

(palindromisch!) en we berekenen dit deel ervan:

$$\cfrac{1}{2 + \cfrac{1}{2}} = \frac{2}{5}$$

en met de teller en de noemer kunnen we 29 schrijven als een som van twee kwadraten:

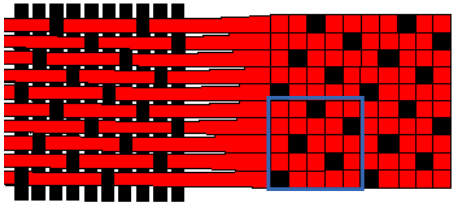
$$29 = 2^2 + 5^2.$$

Dit laatste volgt uit (1.6), want we hebben bewezen dat teller en noemer van het rechterlid de twee getallen zijn die we zoeken.

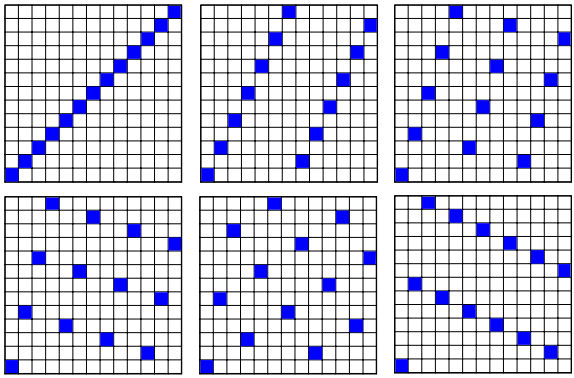
1.5 De methode van Édouard Lucas

We kunnen de getallen a en b ook op een andere manier bepalen. Édouard Lucas bracht in 1867 de stelling van Fermat in verband met het weven

van draden tot textiel. Hierbij worden verticale en horizontale draden met elkaar vervlochten, zoals aan de linkerkant van de figuur:



Het gaat hier om een zogenaamde satijnbinding. Bij satijn kruisen wat men noemt de ketting- en de inslagdraden elkaar op een speciale manier. Als de zwarte kettingdraad op een bepaalde plaats boven ligt, dan ligt bij minstens 4 van de volgende kruisingen de rode inslagdraad boven. Op de figuur zijn het er precies vier. Het patroon herhaalt zich, en kan daarom voorgesteld wordt zoals rechts in de figuur, met wat in het blauwe vierkant te zien is. Een vierkant opgebouwd uit 5×5 kleine vierkanten, waarvan er als volgt een aantal zwart gekleurd zijn: kleur het vierkantje linksonder in, ga dan 1 kolom opzij en 2 rijen omhoog en kleur daar het vierkantje in, ga weer naar rechts en 2 omhoog enzovoort. Kom je uit boven de 5, dan tel je van onderen te beginnen verder.
Omdat de zijde van het vierkant een priemgetal is, zal het toepassen van deze methode er steeds voor zorgen dat er op elke rij en kolom precies 1 vierkant is ingekleurd, één per rij en één per kolom.

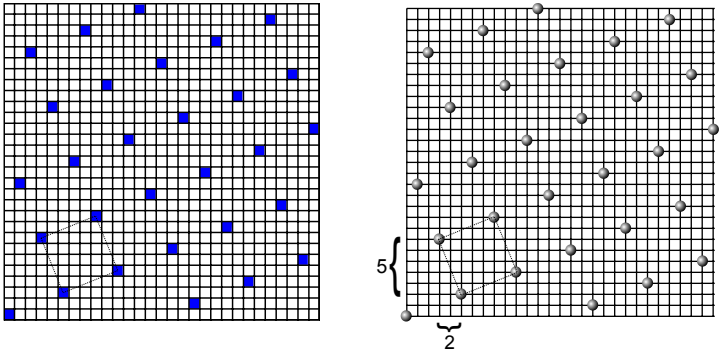


Hier is een voorbeeld in het geval we niet 4 maar 12 opeenvolgende kruisingen hebben met de inslagdraad boven, het gaat dan om een vierkant van 13×13. Dit kan op verschillende manieren, zoals je kan zien op de figuur hierboven. In de verschillende delen van de figuur gaat men respectievelijk

met stappen van 1, 2, 3, 4, 5 en 6 naar boven, steeds startend linksonder in een vierkant. Zo ontstaan patronen. Lucas merkte op dat het patroon in het midden onderaan het mooiste is omdat er allerlei vierkanten in te zien zijn. Bovendien is het het enige dat niet verandert als je het een kwartslag draait. Merk op dat we bij dit patroon stijgen in stappen van 5. En 5 was precies de noemer die hoorde bij het priemgetal 13 waarvan de kettingbreuk palindromisch was!

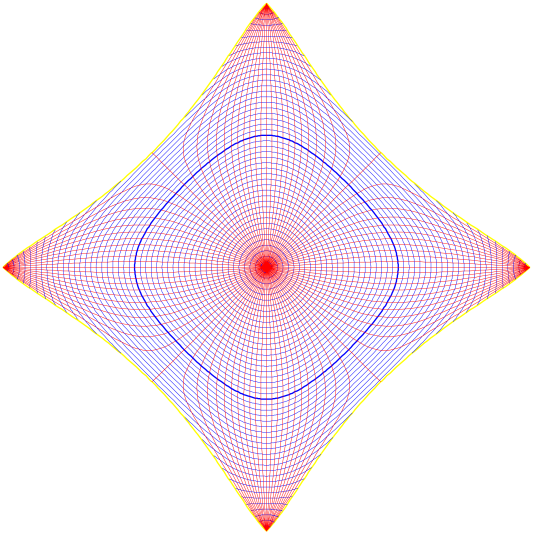
Lucas bewees ook dat een dergelijke symmetrie niet voorkomt indien het aantal vierkantjes per zijde een priemgetal is van de vorm $4n + 3$.

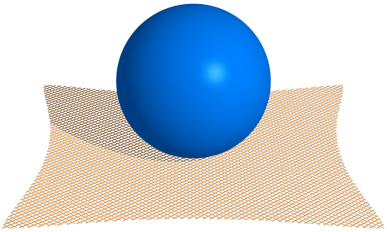
We doen het nog even opnieuw voor $p = 29$, en de bijhorende noemer 12. Kleur in een vierkant van 29×29 het vierkantje linksonder in, ga dan 1 kolom opzij en 12 rijen omhoog en kleur daar het vierkantje in, ga weer naar rechts en 12 omhoog enzovoort: (figuur links)

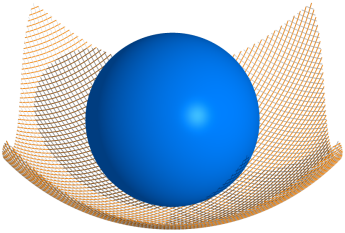


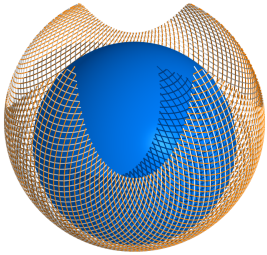
Opnieuw ontstaan er vierkanten, waarvan er één is aangeduid op de figuur. Het wordt duidelijker indien we de blauwe vierkantjes vervangen door roosterpunten in een rooster van 29×29 (figuur rechts). Via de stelling van Pythagoras vinden we dan dat het kwadraat van de zijde van het aangeduide vierkant te schrijven is als $5^2 + 2^2$.

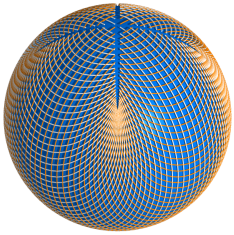
Bijna 250 jaar na de eerste formulering van de stelling van Fermat door Girard, vond Lucas [6, 7] er zo een echte praktische toepassing van!
Voor Lucas werd het zelfs een heel nieuwe tak van de meetkunde, die hij *Géométrie plane des tissus* (Vlakke meetkunde van de stoffen) doopte. Hij vond hierbij navolging bij andere wiskundigen, waarvan de meest bekende Pafnuty Chebyshev (1821-1894) is, met zijn ongepubliceerd artikel *Sur la coupe des vêtements* uit 1878 waarin hij berekent welke vorm je uit een stuk stof moet uitknippen om er een boloppervlak volledig mee te kunnen 'aankleden'.











Voor de geïnteresseerde lezer, je doet het zoals op de figuren. Bovenaan zie je de vorm die je uit de stof moet snijden, onderaan het aankleden van de bol.
(Met dank aan Étienne Ghys [2] voor de eerste figuur, en aan Jos Leys voor de andere.)

Appendix A. Lineaire recursiebetrekkingen

Een lineaire recursiebetrekking (van de tweede orde) is een vergelijking die het verband geeft tussen 3 opeenvolgende termen van een oneindige rij getallen $y_{-1}, y_0, y_1, \ldots, y_n, \ldots$ Zo'n vergelijking neemt de volgende vorm aan:

$$y_k = b_k \cdot y_{k-1} + a_k \cdot y_{k-2} \quad \text{voor } k = 1, 2, 3, \ldots$$

De getallen $a_1, a_2, a_2, \ldots$ en $b_1, b_2, b_3, \ldots$ zijn hierbij gegeven.

Het bekendste voorbeeld van een lineaire recursiebetrekking is dit:

$$y_k = y_{k-1} + y_{k-2} \quad \text{voor } k = 1, 2, 3, \ldots$$

waarbij alle a's en b's gelijk zijn aan 1.

In feite staan hier oneindig veel vergelijkingen, waarvan de eerste gegeven zijn door:

$$y_1 = y_0 + y_{-1}\ , \quad y_2 = y_1 + y_0\ , \quad y_3 = y_2 + y_1\ , \ldots .$$

Merk op dat je met deze vergelijkingen inderdaad een oneindige rij getallen genereert, op voorwaarde dat je de eerste twee getallen y_{-1} en y_0 vastlegt. Voor deze recursiebetrekking is de standaardkeuze $y_{-1} = y_0 = 1$ (we spreken van beginvoorwaarden). We krijgen dan de volgende rij getallen:

$$y_{-1} = 1, y_0 = 1, y_1 = 2, y_2 = 3, y_3 = 5, y_4 = 8, y_5 = 13, y_6 = 21, \ldots .$$

Deze rij is beroemd en wordt de rij van Fibonacci genoemd.

Als we andere beginvoorwaarden kiezen, dan krijgen we een andere rij getallen.

Een mooie eigenschap van dit soort recursiebetrekkingen is dat als je al twee oplossingen berekend hebt, dat alle andere oplossingen dan te vinden zijn uit die twee. Het is ideaal als je eerst deze twee specifieke oplossingen berekent:

$$\begin{array}{lll} \text{oplossing } A_k & \text{met} & A_{-1} = 1, A_0 = 0 \\ \text{oplossing } B_k & \text{met} & B_{-1} = 0, B_0 = 1 \end{array}$$

In de volgende tabel zie je de eerste paar termen van beide rijen. Door de speciale keuze van de beginvoorwaarden kunnen we nu elke andere oplossing y_k van de recursiebetrekking schrijven m.b.v. deze twee. Je ziet in de tabel dat als je de oplossing A_k vermenigvuldigt met y_{-1} en daar B_k, vermenigvuldigd met y_0, bij optelt, dat je dan de eerste 2 waarden van y_k krijgt.

$k:$	-1	0	1	2	$\ldots$	n	$\ldots$
$A_k:$	1	0	a_1	b_2a_1	$\ldots$	A_n	$\ldots$
$B_k:$	0	1	b_1	$b_2b_1+a_2$	$\ldots$	B_n	$\ldots$
$y_k:$	y_{-1}	y_0	y_1	y_2	$\ldots$	y_n	$\ldots$

Door de lineariteit van de recursiebetrekking is het dan eenvoudig in te zien dat voor elke k geldt:

$$y_k = y_{-1}A_k + y_0B_k.$$

We hebben de volgende eigenschap nodig, eveneens een gevolg van de lineariteit:

$$A_kB_{k-1} - A_{k-1}B_k = (-1)^{k-1} \cdot a_1 \cdot a_2 \cdot \ldots \cdot a_k. \tag{A.1}$$

Dit kunnen we bewijzen per inductie, en daarvoor gebruiken we eigenschappen van determinanten:

$$\begin{vmatrix} A_k & A_{k-1} \\ B_k & B_{k-1} \end{vmatrix} = \begin{vmatrix} b_kA_{k-1} + a_kA_{k-2} & A_{k-1} \\ b_kB_{k-1} + a_kB_{k-2} & B_{k-1} \end{vmatrix} = -a_k \cdot \begin{vmatrix} A_{k-1} & A_{k-2} \\ B_{k-1} & B_{k-2} \end{vmatrix}$$

met

$$\begin{vmatrix} A_0 & A_{-1} \\ B_0 & B_{-1} \end{vmatrix} = -1.$$

Vanaf nu zullen we enkel nog werken met recursiebetrekkingen waarbij $a_k = 1$ voor elke waarde van k:

$$y_k = b_k \cdot y_{k-1} + y_{k-2} \quad \text{voor } k = 1, 2, 3, \ldots.$$

We definieren een aantal basisoplossingen zoals in de volgende tabel. De positie van de twee opeenvolgende waarden 0 en 1 karakteriseert de verschillende basisoplossingen waarvan je de naam kan lezen in de eerste kolom. Om te beginnen kan je opmerken dat $B_k^{(1)} = B_k$ en $B_k^{(2)} = A_k$, en dus wordt eigenschap (A.1):

$$B_k^{(2)}B_{k-1}^{(1)} - B_{k-1}^{(2)}B_k^{(1)} = (-1)^{k-1}. \tag{A.2}$$

Verder is er natuurlijk voor elke waarde van j voldaan aan de recursiebetrekking:

$$B_k^{(j)} = b_k B_{k-1}^{(j)} + B_{k-2}^{(j)}. \tag{A.3}$$

$k:$	-1	0	1	2	$\dots$	$j-1$	j	$j+1$	$\dots$	n	$\dots$
$B_k^{(1)}:$	0	1	b_1	$B_2^{(1)}$	$\dots$	$B_{j-1}^{(1)}$	$B_j^{(1)}$	$B_{j+1}^{(1)}$	$\dots$	$B_n^{(1)}$	$\dots$
$B_k^{(2)}:$	1	0	1	b_2	$\dots$	$B_{j-1}^{(2)}$	$B_j^{(2)}$	$B_{j+1}^{(2)}$	$\dots$	$B_n^{(2)}$	$\dots$
$B_k^{(3)}:$		1	0	1	$\dots$	$B_{j-1}^{(3)}$	$B_j^{(3)}$	$B_{j+1}^{(3)}$	$\dots$	$B_n^{(3)}$	$\dots$
$\vdots$											
$B_k^{(j)}:$					$\dots$	1	b_j	$B_{j+1}^{(j)}$	$\dots$	$B_n^{(j)}$	$\dots$
$B_k^{(j+1)}:$					$\dots$	0	1	b_{j+1}	$\dots$	$B_n^{(j+1)}$	$\dots$
$B_k^{(j+2)}:$					$\dots$	1	0	1	$\dots$	$B_n^{(j+2)}$	$\dots$

Uit deze tabel kunnen we twee interessante relaties tussen deze oplossingen aflezen. We baseren ons hiervoor op de waarden van $B_k^{(j+1)}$ en $B_k^{(j+2)}$ in de blauwe rechthoek. Om te beginnen kijken we naar $B_k^{(1)}$, en je ziet dan dat deze oplossing de volgende lineaire combinatie is van de twee onderste oplossingen:

$$B_k^{(1)} = B_j^{(1)} B_k^{(j+1)} + B_{j-1}^{(1)} B_k^{(j+2)}. \tag{A.4}$$

Voor $B_n^{(j)}$ vinden we op dezelfde manier:

$$B_n^{(j)} = b_j B_n^{(j+1)} + B_n^{(j+2)}. \tag{A.5}$$

Deze recursiebetrekking verbindt drie onder elkaar staande getallen in de tabel, en heeft als onmiddellijk gevolg dat twee opeenvolgende elementen van een kolom geen gemeenschappelijke factor kunnen hebben. Hadden ze wel een gemeenschappelijke factor, dan volgt uit (A.5) dat deze factor de hele kolom deelt, wat niet mogelijk is door het voorkomen van het getal 1 in elke kolom.

Appendix B. Verband met kettingbreuken

Met de resultaten uit de vorige appendix kunnen we laten zien dat we het volgende resultaat hebben (zie verder):

$$\frac{T_{1\to k}}{N_{1\to k}} = b_1 + \cfrac{1}{b_2 + \cfrac{1}{\ddots + \cfrac{1}{b_{k-1} + \cfrac{1}{b_k}}}} = \frac{B_k^{(1)}}{B_k^{(2)}} \tag{B.1}$$

waarbij de b's de coëfficiënten zijn in de recursiebetrekking in appendix A.

Uit het feit dat beide breuken (links en rechts) niet te vereenvoudigen zijn, volgt dan dat $T_{1\to k} = B_k^{(1)}$ en $N_{1\to k} = B_k^{(2)}$.
Verder hebben we ook:

$$\frac{T_{k\to 1}}{N_{k\to 1}} = b_k + \cfrac{1}{b_{k-1} + \cfrac{1}{\ddots + \cfrac{1}{b_2 + \cfrac{1}{b_1}}}} = \frac{B_k^{(1)}}{B_{k-1}^{(1)}} \tag{B.2}$$

voor de kettingbreuk waarin de b's in de omgekeerde volgorde staan. Dus $T_{k\to 1} = B_k^{(1)}$ en $N_{k\to 1} = B_{k-1}^{(1)}$.

Formule (B.1) kunnen we veralgemenen tot:

$$\frac{T_{j+1\to k}}{N_{j+1\to k}} = b_{j+1} + \cfrac{1}{b_{j+2} + \cfrac{1}{\ddots + \cfrac{1}{b_k}}} = \frac{B_k^{(j+1)}}{B_k^{(j+2)}}. \tag{B.3}$$

De formules (B.1) en (B.3) volgen uit (A.5). We herschrijven deze vergelijking als volgt:

$$B_k^{(j)} = b_j B_k^{(j+1)} + B_k^{(j+2)} \quad \Rightarrow \quad \frac{B_k^{(j)}}{B_k^{(j+1)}} = b_j + \frac{1}{\frac{B_k^{(j+1)}}{B_k^{(j+2)}}}.$$

We starten nu met $j = 1$:

$$\frac{B_k^{(1)}}{B_k^{(2)}} = b_1 + \frac{1}{\frac{B_k^{(2)}}{B_k^{(3)}}} = b_1 + \frac{1}{b_2 + \frac{1}{\frac{B_k^{(3)}}{B_k^{(4)}}}}$$

$$= b_1 + \cfrac{1}{b_2 + \cfrac{1}{\ddots + \cfrac{1}{b_{k-1} + \cfrac{1}{\frac{B_k^{(k)}}{B_k^{(k+1)}}}}}} = b_1 + \cfrac{1}{b_2 + \cfrac{1}{\ddots + \cfrac{1}{b_{k-1} + \cfrac{1}{b_k}}}}.$$

De laatste stap volgt uit de beginvoorwaarden in combinatie met de recursiebetrekking. We hebben namelijk dat $B_k^{(k)} = b_k$ en $B_k^{(k+1)} = 1$.

Voor (B.2) vertrekken we van de vergelijking (A.3) met $j = 1$ en herschrijven deze als volgt:

$$B_k^{(1)} = b_k B_{k-1}^{(1)} + B_{k-2}^{(1)} \quad \Rightarrow \quad \frac{B_k^{(1)}}{B_{k-1}^{(1)}} = b_k + \frac{1}{\frac{B_{k-1}^{(1)}}{B_{k-2}^{(1)}}}.$$

We vervangen dan in de linkse vergelijking k door $k - 1$ en herhalen het procedé.

Formule (A.2) kunnen we nu herschrijven als:

$$N_{1\to k} \cdot T_{1\to k-1} - T_{1\to k} \cdot N_{1\to k-1} = (-1)^{k-1}. \tag{B.4}$$

Formule (A.4) wordt:

$$T_{1\to k} = T_{j\to 1} \cdot T_{j+1\to k} + N_{j\to 1} \cdot N_{j+1\to k}. \tag{B.5}$$

Bibliografie

[1] Clarke, F.W., Everitt, W.N., Littlejohn, L.L., Vorster, S.J.R., *H.J.S. Smith and the Fermat two squares theorem*, Amer. Math. Monthly 106 (1999), 7, p. 652–665.

[2] Ghys, É., *Sur la coupe des vêtements: variation autour d'un théorème de Tchebychev.* [*On the cutting of garments: variation on a theme of Chebyshev*], Enseign. Math. (2) 57 (2011), no. 1-2, p. 165–208.

[3] Hardy, G.H. en Wright, E.M., *An Introduction to the Theory of Numbers*, Edited by Roger Heath-Brown, Joseph Silverman, and Andrew Wiles, Oxford University Press, 2008 .

[4] Kindt, M. en Lemmens, P., *Ontwikkelen met kettingbreuken*, Zebrareeks 33, Epsilon Uitgaven Amsterdam, 2011.

[5] Levrie, P. en Penne, R., *De pracht van priemgetallen*, Prometheus Amsterdam, 2014.

[6] Lucas, É., *Application de l'arithmétique à la construction de l'armure des satins réguliers*, Paris: G. Retaux, 1867.

[7] Lucas, É., Les principes fondamentaux de la géométrie des tissus, *Congrès de l'Association française pour l'avancement des sciences*, 40, 2 (1911), p. 72–88.

2 Wat is er Cartesiaans aan analytische meetkunde?

Steven Wepster

Mathematisch Instituut, Universiteit Utrecht

Voorwoord

Analytische meetkunde komt terug in het examenprogramma. Dat is een mooie gelegenheid om nog eens kritisch te kijken naar de rol die aan Descartes wordt toegeschreven bij de "uitvinding" of "introductie" van dat gebied in de wiskunde. Het doel van deze lezing is om te benadrukken dat die rol heel wat bescheidener is dan vaak wordt aangenomen[1]. Om dat te onderbouwen kijken we naar het werk van Descartes, en minstens zo interessant in deze context zijn tijdgenoot Fermat. Maar we moeten ook een beeld hebben van wat we onder analytische meetkunde verstaan. Laten we met dat laatste beginnen.

Ik vraag u om nu na te denken over wat u onder analytische meetkunde verstaat, voordat u doorleest.

2.1 Wat is analytische meetkunde?

Wie vraagt wat een priemgetal is, krijgt meestal prompt een definitie voorgeschoteld. Bij analytische meetkunde ligt dat anders: het is geen precies gedefinieerd begrip. In leerboeken over het vakgebied staat zelden of nooit wat het *is*, ze *doen* het gewoon. We moeten blijkbaar kijken naar de inhoud van de leerboeken om te weten waar analytische meetkunde zich mee bezig houdt. Of, gezien de hervorming van het examenprogramma

[1] Dit is reeds eerder betoogd; zie bijv. [7], [10], of [3, p. 95–96].

die de aanleiding vormde voor deze zoektocht: laten we beginnen bij dat examenprogramma zelf.

Het examenprogramma Wiskunde B voor vwo bevat een domein E over meetkunde, dat weer is onderverdeeld in vier subdomeinen. Hierin lezen we het volgende[2].

E1 **Meetkundige vaardigheden**
De kandidaat kan meetkundige eigenschappen van objecten onderzoeken en bewijzen en kan daarbij gebruik maken van meetkundige en algebraïsche technieken en van ICT.

E2 **Algebraïsche methoden in de vlakke meetkunde**
De kandidaat kan eigenschappen en onderlinge ligging van punten, lijnen, cirkels en andere geschikte figuren onderzoeken met behulp van algebraïsche voorstellingen, kan in een gegeven of zelfgekozen coördinatenstelsel algebraïsche voorstellingen van figuren opstellen en kan algebraïsche voorstellingen gebruiken om meetkundige problemen op te lossen.

E3 **Vectoren en inproduct**
De kandidaat kan met behulp van vectoren en inproducten eigenschappen van figuren in het vlak afleiden en berekeningen uitvoeren.

E4 **Toepassingen**
De kandidaat kan de aangegeven technieken toepassen in geschikte natuurwetenschappelijke en technische situaties.

Wat opvalt is dat het woord "analytisch" in het meetkundedomein niet in voorkomt. Wel is er sprake van het gebruik van algebra om meetkundige problemen op te lossen. Er worden coördinatenstelsels gebruikt, er is sprake van de eigenschappen en onderlinge ligging van lijnen, cirkels, en "andere geschikte figuren". We zien ook vectoren en inproduct en tenslotte is er sprake van toepassen. Kortom, we zien een aantal elementen die we herkennen als analytische meetkunde, en een aantal die we niet of minder herkennen. De scheidslijn is niet helder.

Laten we eens kijken hoe zo'n programma in de praktijk uitpakt. Ik geef twee voorbeelden. Het eerste is uit een Wiskunde D-module van de Wageningse Methode[3]:

> Gegeven is de lijn k: $x = -1$ en het punt $E : (1, 0)$. De kromme K bestaat uit alle punten P waarvoor

[2] Examenprogramma's geraadpleegd op http://www.examenblad.nl/
[3] [21, p. 21].

$d(P, E) = e \cdot d(P, k)$ voor zeker getal $e > 0$.

a. Laat zien dat voor punten $P(x, y)$ op K geldt: $(1-e^2)x^2 - (2 + 2e^2)x + 1 - e^2 + y^2 = 0$.

b. Laat zien dat de coëfficiënten van x^2 en y^2 in de vergelijking van K hetzelfde teken hebben, precies dan als $0 < e < 1$.

Hierin herkennen we duidelijk het verband tussen meetkunde en algebra: de punten van de kromme voldoen aan een meetkundige eis die wordt vertaald in een algebraïsche relatie tussen de coördinaten van zulke punten. De coëfficiënten in de formule hebben een bepaalde eigenschap (nl. hetzelfde teken hebben) precies dan als de meetkundige eis aan een bepaalde voorwaarde voldoet ($0 < e < 1$).

Het tweede voorbeeld komt uit een experimenteel programma voor havo en betreft een toepassing.

> Mobiele telefoons hebben soms geen bereik. Dat betekent dat er geen mast met antenne dicht genoeg in de buurt is om mee in verbinding te kunnen staan. Stel je voor dat zo'n antenne een bereik heeft van 30 km. Op 15 km van de snelweg A1 staat een mast met zo'n antenne. Gedurende hoeveel km op de A1 kun je via die antenne met je mobiele telefoon verbinding maken?[4]

De opgave geeft een context die aanleiding geeft tot een cirkel doorsneden door een rechte (want een snelweg is natuurlijk altijd recht) en de vraag is dan hoe lang de afgesneden koorde is. Als toepassing van analytische meetkunde is dit eigenlijk wat minder gelukkig, omdat de vraag makkelijker beantwoord kan worden met de stelling van Pythagoras. Toch vinden we er wel vertrouwde "analytische" elementen in terug.

Een van de klassiekere leerboeken uit de eerste helft van de vorige eeuw is Rutgers' *Inleiding tot de Analytische Meetkunde*[5]. Zo'n gespecialiseerd boek kan veel verder gaan dan een hedendaagse wiskundemethode. Al bladerend komen we er een veelheid aan onderwerpen in tegen: uiteraard natuurlijk coördinaten, punten, lijnen; krommen die door vergelijkingen worden voorgesteld. Zo'n kromme wordt dan vaak als *locus* of meetkundige plaats beschouwd: als verzameling punten die voldoen aan een of andere meetkundige eigenschap. Dat zagen we in het voorbeeld uit de Wageningse Methode ook al, alleen werd dat daar niet expliciet zo genoemd.

[4] [6, p. 21]
[5] [18].

Tot zover lezen we weinig nieuws bij Rutgers, maar dat verandert ras. Eerst komen er wat voorbeelden van meetkundige plaatsen, zoals natuurlijk kegelsneden, maar ook de conchoide en andere krommen van graad 3 of 4. Daarna komt het plotten van een kromme bij gegeven vergelijking, gevolgd door diverse coördinatentransformaties. Rutgers vervolgt met een meer algemene behandeling van algebraïsche krommen en hun symmetrieën, dubbelpunten en dergelijke. Het grootste deel van het boek tenslotte is gewijd aan een uitputtende behandeling van (bundels van) kegelsneden. Her en der staan uitstapjes naar projectieve meetkunde. Zo komen homogene coördinaten en dubbelverhouding aan bod.

Ter illustratie volgt hier Rutgers' behandeling van het vraagstuk om de raaklijnen met gegeven richting aan een gegeven ellips te vinden. Wie bij de woorden "raaklijn" en "analytisch" denkt aan differentiëren, doet hier een mooie ontdekking[6].

> **Raaklijnen aan de ellips, evenwijdig aan een gegeven rechte**[7] worden gevonden door de voorwaarde af te leiden, waaraan voldaan moet worden, opdat de rechte $y = mx+n$, (m de richtingscoëfficient der gegeven rechte) *twee samenvallende snijpunten* met de ellips gemeen heeft.
>
> Dit is het geval, indien b.v. de *abscissen* der snijpunten aan elkaar gelijk zijn, daar dan ook de *ordinaten* aan elkaar gelijk zijn [...]
>
> Na eliminatie van y uit $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$ en $y = mx+n$ vindt men als voorwaarde voor gelijke abscissen $n^2 = a^2m^2 + b^2$, zoodat de vergelijkingen der raaklijnen zijn
>
> $$y = mx \pm \sqrt{a^2m^2 + b^2},$$
>
> terwijl men voor de coördinaten der raakpunten vindt:
> $x = -\frac{a^2m}{n}$, $y = \frac{b^2}{n}$, waarin $n = \pm\sqrt{a^2m^2 + b^2}$.

Om ons beeld van de analytische meetkunde verder te ontwikkelen, kijken we naar de Nederlandse Wikipedia[8]. Daar lezen we:

> De **analytische meetkunde**, ook wel bekend als Cartesiaanse meetkunde, is de studie van meetkunde die de principes

[6] [18, p. 146–147].

[7] [18, p. 146], cursivering is door mij.

[8] nl.wikipedia.org/wiki/Analytische_meetkunde geraadpleegd op 19/8/2014; hoewel niet de meest betrouwbare bron is het wel een voor de hand liggende bij bijv. een profielwerkstuk.

van algebra gebruikt. [...]
Gewoonlijk wordt het Cartesisch coördinatenstelsel toegepast om vergelijkingen voor vlakken, lijnen, krommen en cirkels te manipuleren, vaak in twee of drie, maar in principe in willekeurig veel dimensies. Sommigen zijn van mening dat de introductie van analytische meetkunde door René Descartes het begin van moderne wiskunde was.

De aan Descartes toegeschreven introductie, hier als onomstreden *feit* geponeerd, zullen we verderop nader onderzoeken. De wikipagina vervolgt met een opsomming van onderwerpen uit de analytische meetkunde die grotendeels eerder in de lineaire algebra thuishoren. Onder het kopje "geschiedenis" staat vervolgens Descartes' *Géométrie* uit 1637 genoemd en daarover lezen we dan:

> Dit in het Frans geschreven werk en de er aan ten grondslag liggende filosofische principes, legden het fundament voor de infinitesimaalrekening in Europa.

Dat het in het Frans geschreven is klopt, maar het fundament voor de infinitesimaalrekening vind ik het niet, en ik vrees dat het een poging van de wiki-auteur is om het woord *analytisch* in deze tak van meetkunde te duiden. Daarover straks meer; we verplaatsen ons nu eerst naar de 18e eeuw en raadplegen het *Volkoomen Wiskundig Woordenboek* van Stammetz.

In Stammetz' woordenboek komt de aanduiding "analytische meetkunde" niet voor, er staat wel een lemma over "hogere meetkunde":

> De Hoogere Geometrie [...] welk van de Krommen Linien, en de door haar voortgebragte Lighamen handeld. [...] Gelyk nu egter niet alleen de Algebra van Vieta en Carthesius in eenen zulken stand gebragt is, dat zy met veel nut in de Geometrie konde aangebragt worden, maar ook de Heer van Leibnitz zyne Differentiaal- en Integraal-Rekening uitgevonden [...] zoo heeft de Hoogere Geometrie ook een geheel ander Aanzien gekreegen als zy eertyds gehad heeft[9].

Deze al zeer lang bestaande meetkunde heeft dus, volgens Stammetz, door het *algebraische* werk van François Viète en René Descartes, en later door de *infinitesimaalrekeing* van Leibniz, belangrijke vernieuwingen ondergaan. Stammetz plaatst, in mijn ogen terecht, Descartes bij de algebraici en legt geen verbinding tussen Descartes en infinitesimaalrekening

[9] [20, p. 175–176].

of analyse.

2.2 Algebraïsch, analytisch en synthetisch

Hoe zit dat nu met die woorden algebra, analyse, synthese? Laten we daar eerst wat licht op werpen voordat we in Descartes' werk duiken en ook in dat van Fermat. We laten Stammetz wederom aan het woord:

> Analysis, Oplossings- of Ontleedings-Kunst, heet de Weetenschap, welke leerd, hoe men uit eenige bekende waarheeden andere nog onbekende vinden en dus verborgene Vraagen oplossen zal. Hierin heeft men het zeer ver gebragt: Maar de middelen, door dewelke men daartoe komt, zyn de Letter ReekenKunst, de Algebra en heerlijke Differentiaal- en Integraal-Reekening van den Heer Leibnitz ofte Newton welke Uitvindings-Kunsten, by een genoomen, deze Oplossings-Kunst uitmaaken[10].

Kort gezegd: volgens Stammetz is analyse dus een *uitvindkunst*. Een vergelijkbare verklaring staat te lezen in de welbekende *Encyclopédie* van Diderot en d'Alembert:

> Analyse is precies de methode om wiskundige problemen op te lossen, door ze te reduceren tot vergelijkingen. Om problemen op te lossen gebruikt de analyse de algebra, of de rekening van grootheden in het algemeen: de woorden Analyse en Algebra worden dan ook vaak als synonym gezien[11].

Op autoriteit van een van de meest gezaghebbende wiskundigen uit die tijd, namelijk d'Alembert, worden algebra en analyse aan elkaar gelijk verklaard. Gedurende het grootste deel van de 18e eeuw was dat inderdaad ook zo. In de tweede helft van die eeuw is de aanduiding "analytische meetkunde" in zwang geraakt, en we begrijpen nu dat dat ook "algebraïsche meetkunde" had kunnen zijn. In de huidige tijd zijn allebei de benamingen in gebruik waarbij de eerste kan worden gezien als een wat eenvoudiger variant van de tweede; het verschil zit in het soort vragen die gesteld worden en de technieken die nodig zijn om die vragen te beantwoorden.

De betekenis van analyse als uitvindkunst gaat terug op de klassieke Griekse tijd. We vinden dit helder verwoord o.a. bij Pappus (ca. 300 n. Chr.):

[10] [20, p. 24].
[11] [1, v.1, p. 400].

> In analyse nemen we aan wat gezocht is alsof het er al is, we zoeken naar waar het uit volgt, en dan waar dat weer uit volgt, todat we terug zijn bij iets dat we al weten.
>
> In synthese daarentegen, beginnen we bij het laatste punt van de analyse dat we al wisten, en daarna zetten we de eerdere gevolgtrekkingen in hun natuurlijke volgorde, als precedenten, en we sinteren ze aaneen totdat we het doel bereikt hebben[12].

Wie meetkundige proposities zoals in Euclides' *Elementen* bestudeert, ziet alleen de synthese: vanuit de gegeven condities werkt Euclides stap voor stap het gevraagde bewijs of de constructie uit, maar je hebt geen idee hoe hij aan de juiste stappen komt. Volgens Pappus kun je die stappen vinden door te doen alsof je het eindresultaat al kent en te onderzoeken waar dat eindresultaat een direct gevolg van is, enzovoort totdat je bij de gegeven uitgangspunten bent aangeland. Er zijn maar erg weinig Griekse teksten bewaard waarin zo'n analytische vindmethode voorkomt; synthetische deductieve bouwwerken (zoals Euclides' *Elementen*) zijn er echter te over.

Aan het eind van de 16e, begin van de 17e eeuw waren de Europese wiskundigen op zoek naar "vindmethoden" om meetkundige problemen op te lossen. De algebra had zich, o.a. door het werk van Viète, voldoende ver ontwikkeld om bepaalde problemen waarin bijvoorbeeld een onbekende lengte gevonden moest worden, op te lossen. We zullen nu naar het werk van Descartes en Fermat gaan kijken hoe zij hierop verderbouwen.

2.3 Descartes' Géométrie

Descartes' *Géométrie* verscheen in 1637 als appendix bij zijn filosofische verhandeling *Discours de la méthode*[13]. Descartes verbleef in die tijd in de Nederlanden, en was goed bevriend met Frans van Schooten Jr., die op zijn beurt weer de spil was van een netwerk van veelal Nederlandse wiskundigen waar Descartes' *Géométrie* dan ook gretig aftrek vond. Jantien Dopper heeft in haar onlangs verschenen proefschrift[14] aangetoond dat de bewondering van Van Schooten voor Descartes zo ver ging dat de laatste in de ogen van de eerste werkelijk niets fout kon doen – zelfs niet als die

[12] [17, p. 82]
[13] [8]; er is daarnaast ook een recente Nederlandse uitgave met een verdienstelijke vertaling, zie [9].
[14] [11].

fouten door andere wiskundigen glashard werden aangetoond. Maar dit terzijde.

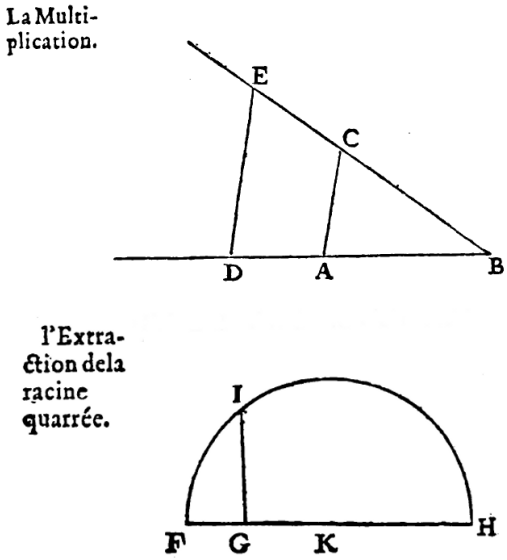
Het doel van Descartes in de *Géométrie* is om zekere en exacte methoden te ontwikkelen voor het oplossen van meetkundige problemen.[15]. Dit doel bereikt hij niet, maar het middel dat hij gebruikt in zijn poging het doel te bereiken is een eigen leven gaan leiden. De openingszin van het boek luidt:

> Alle meetkundige problemen kunnen eenvoudig worden herleid tot zulke termen, dat het daarna alleen nodig is de lengte van enkele rechten te kennen, om ze te construeren[16].

De *herleiding* waar Descartes het hier over heeft, blijkt te zijn: vertaal het probleem in een (of meer) algebraïsche vergelijking(en) en los die op. De oplossing van de vergelijking stelt de lengte van een lijnstuk voor en het lijnstuk zelf, eenmaal geconstrueerd, geeft de oplossing van het oorspronkelijke probleem. Meetkundige problemen vragen om een oplossing in de vorm van een constructie.

Om dat laatste – het construeren van algebraïsch gegeven lengten – mogelijk te maken, geeft Descartes eerst enkele basisconstructies. Optellen en aftrekken corresponderen op natuurlijke wijze met het achter elkaar of langs elkaar leggen van lijnstukken. Verder is er de vermenigvuldiging (figuur 2.1 boven): gevraagd wordt een lijnstuk met lengte gelijk aan het product van twee gegeven lijnstukken BD en BC. Hiertoe zet Descartes de twee lijnstukken uit op de benen van een willekeurige hoek met hoekpunt B. Op een van deze benen zet hij ook een eenheidslijnstuk AB af. Hij trekt een lijnstuk DE door D en evenwijdig aan het lijnstuk AC en laat E het snijpunt met BC verlengd zijn. Vanwege gelijkvormigheid geldt nu dat $BE : BD = BC : BA$ oftewel $BE = BC \cdot BD$, immers $BA = 1$. Met dezelfde figuur is een deling van twee lijnstukken mogelijk. Het belang van deze constructie ligt erin dat het product van twee lijnstukken een lijnstuk oplevert, en niet een oppervlakte. Zodoende kan Descartes een voorheen ongekende stap zetten: vergelijkingen hoeven niet langer dimensioneel homogeen te zijn. Voorheen was een algebraïsche vergelijking slechts betekenisvol als alle termen precies dezelfde "dimensie" hadden (bijv. ax^2 en abx zijn beide van dimensie 3). Bovendien ontdoet Descartes zich van mogelijk commentaar dat zijn vergelijkingen betekenisloos zouden zijn als er hogere dimensies dan 3 in zouden voorkomen.

[15] Hierover is belangrijk onderzoek gedaan door Bos waarover hij o.a. sprak bij de Vakantiecursus 1989; zie [3] of het uitvoerige [2].
[16] [8, p. 297].

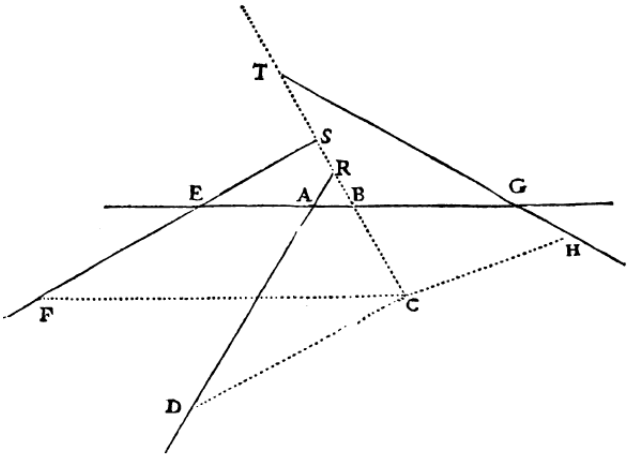


Figuur 2.1: Meetkundige constructies overeenkomende met algebraïsche operaties.

Voor het trekken van vierkantswortels dient de onderste afbeelding in figuur 2.1, waar boog FIH een halve cirkel is op middellijn FH. Als $FG = 1$ en $GH = k$, dan geldt dat $GI = \sqrt{k}$. Algemener: er geldt $FG : GI = GI : GH$ oftewel zoals men zegt GI is middenevenredige tussen FG en GH. Het bewijs hiervan loopt via de gelijkvormigheid van driehoeken FGI, IGH en FIH. Deze constructie was van fundamenteel belang in de klassieke meetkunde en verdient eigenlijk een vaste plaats in het voortgezet onderwijs maar ik geloof niet dat het daar veel voorkomt.

Nadat Descartes op bovenbeschreven wijze een verbinding legt tussen aritmetiek en algebra enerzijds en meetkunde anderzijds, wendt hij zich tot het zgn. *Pappusprobleem van vier lijnen.* Pappus beschreef het probleem in zijn reeds geciteerde *Collectio* en besprak de vorderingen van de Griekse meetkundigen vóór hem. Voor Descartes is het een uitstekende kapstok om de kracht van zijn nieuwe methode te demonstreren. Het probleem is als volgt. Er zijn vier lijnen "in positie" gegeven. In figuur 2.2 zijn het de doorgetrokken lijnen AB, AD, EF en GH. Bij elke lijn is tevens een hoek gegeven maar zonder de plaats ervan op de lijn vast te leggen; in de figuur zijn ze ingetekend als respectievelijk $\angle ABC$, $\angle ADC$, $\angle EFC$,

$\angle GHC$. Gevraagd wordt naar de meetkundige plaats van alle punten C waarvoor geldt dat $(CB \cdot CF) : (CD \cdot CH) = r$ voor een of andere gegeven verhouding r. In zijn gekunsteldheid, ingewikkelde formulering en willekeur (waarom niet $(CB \cdot CD) : (CF \cdot CH)$ etc) vind ik het een welhaast pervers probleem en we zullen hier niet langer bij stilstaan dan nodig is om te zien wat Descartes ermee doet.



Figuur 2.2: Het Pappusprobleem voor vier lijnen.

De algemene techniek om meetkundeproblemen op te lossen heeft Descartes dan al eerder uiteengezet. In het kort komt het neer op het volgende. Je doet of het probleem al opgelost is, en je geeft namen aan alle lijnen die nodig lijken te zijn om de oplossing te construeren, bekende zowel als onbekende. Dan zoek je relaties tussen de lijnen net zo lang tot het lukt om een enkele grootheid op twee manieren uit te drukken. Je hebt dan een vergelijking. Je moet proberen net zoveel vergelijkingen te krijgen als er onbekende grootheden zijn. Als dat niet lukt dan is het probleem onderbepaald en heb je een aantal vrije keuzes (en bijgevolg als oplossing een meetkundige plaats zoals een kromme of oppervlak). Voor het overige druk je alle onbekende grootheden uit in bekende, waarna je de oplossing van het probleem kunt construeren middels de hierboven al genoemde basisconstructies.

En passant introduceert Descartes hier een gewoonte die we nog steeds aanhangen: bekende grootheden kies je uit het begin van het alfabet en onbekende uit het eind ervan.

Om het Pappusprobleem op te lossen schrijft Descartes[17]:

> Allereerst onderstel ik dat de zaak geklaard is, en om me te redden uit de wirwar van rechten beschouw ik een van de gegeven rechten, en een van die welke gezocht moeten worden, bijvoorbeeld AB en CB, als de belangrijkste waarop ik alle andere zal proberen te betrekken. Noem het lijnstuk AB tussen de punten A en B x en noem BC, y.

Het valt niet zo erg op maar hier kiest Descartes zijn "co"-ordinatensysteem! Vergeleken met een modern rechthoekig assenstelsel zou je kunnen zeggen dat A de oorsprong is, en AB de richting van de positieve x-as. Een y-as is er niet (laat staan dat die loodrecht op de x-as is) maar het punt C kan hij wel aangeven met de abscis x en ordinaat y. Dichter bij ons moderne coördinatenstelsel komt hij niet en het is nogal vreemd dat we iets Cartesiaans noemen wat bij Descartes zelf niet voorkomt.[18]

We slaan de langdradige details over en pakken de draad weer op als Descartes de algebraïsche oplossing van het probleem te pakken heeft. Hij heeft afgeleid dat

$$y = m - \frac{n}{z}x + \sqrt{m^2 + ox - \frac{p}{m}x^2}.$$

Waar de overige letters precies voor staan doet er niet toe; het zijn grootheden die terug te voeren zijn op bekende lengten in de figuur. Descartes merkt nu op dat deze formule bij elke keuze van $x = AB$, precies één lengte voor $y = BC$ geeft. Bovendien is deze lengte te construeren middels de eerder beschreven constructiestappen. Descartes merkt ook op dat hij aan de formule ziet dat alle punten C op een kegelsnede liggen. Dat de locus een kegelsnede is was bekend bij Pappus; Descartes gaat eraan voorbij om zijn lezer uit te leggen hoe je dat aan de vergelijking ziet.

Concluderend zien we dat Descartes het Pappusprobleem kan oplossen door de gevraagde meetkundige plaats puntsgewijs te construeren; hij herkent de oplossing bovendien als kegelsnede. Hij heeft een volwassen en goed bruikbare algebraïsche notatie ingevoerd. Hij propageert het gebruik van algebra als vindmethode om meetkundige problemen op te lossen, maar dat is niet helemaal nieuw: we kunnen dat ook terugvinden bij o.a. Viète, Van Ceulen, of Oughtred. Wat wel nieuw is, is dat Descartes erkent dat

[17] [8, p. 28–29].

[18] Aan de andere kant is het een veel voorkomend verschijnsel in de wiskunde om begrippen, stellingen e.d. te vernoemen naar personen die er weinig mee te maken hebben.

één vergelijking met twee onbekenden een locus geeft door verschillende waarden aan één van de onbekenden te geven.

Wat we niet terugvinden bij Descartes is het "Cartesiaanse" assenstelsel. Daarbij moeten we gelijk ook opmerken dat de grootheden in Descartes' formules altijd een meetkundige betekenis hebben: ze zijn daarom altijd positief. Als Descartes al een assenstelsel had gehad dan zou hij dus alleen in het eerste kwadrant kunnen werken. Voor punten C (figuur 2.2) die corresponderen met een abscis AB ter linkerzijde van A zou hij opnieuw moeten beginnen en zo heeft hij heel veel gevalsonderscheidingen nodig.

Tenslotte merken we nog een belangrijk verschil tussen Descartes en moderne analytische meetkunde op: de algebra levert Descartes wel een techniek voor *analyse* op, maar een meetkundig probleem vereist toch een *synthese*, een meetkundige constructie als oplossing.

We verlaten nu de *Géométrie* (en doen het boek daarmee eigenlijk tekort) en besluiten de bespreking ervan met een citaat van Dijksterhuis:

> [wanneer men in de *Géométrie*] vergeefs naar het Cartesiaansche assenstelsel zoekt en [...] noch van de rechte lijn, noch van een der kegelsneden de vergelijking ziet afleiden en niettemin een optredende vergelijking van den tweeden graad op juiste wijze als kegelsnede ziet interpreteeren, wanneer men dan ten slotte nog het derde [deel] vrijwel geheel over de theorie der algebraische vergelijkingen ziet handelen en het nieuw gelegde verband van meetkunde en algebra evenzeer in de meetkundige oplossing van algebraische vergelijkingen ziet toepassen als in de algebraische behandeling van meetkundige problemen, ... dan kan men zich moeilijk onttrekken aan den indruk, dat, wat men als analytische meetkunde heeft omschreven, in het werk van Descartes, dat als haar geboorteplaats bekend staat, deels ontbreekt, deels als volkomen bekend wordt behandeld en gebruikt en dat de titel *Géométrie* den sterk algebraisch getinten inhoud van het werk maar zeer volkomen dekt. [19]

Waarom wordt de analytische meetkunde dan toch zo vaak aan Descartes toegeschreven? Henk Bos gaf daar bij de vakantiecursus precies 25 jaar geleden het volgende antwoord op:

> ... omdat latere lezers van het boek zich vooral door de algebraïsche technieken lieten inspireren en weinig interesse toon-

[19] [10, p. 59–60].

den in de methodologische problemen die voor Descartes zo belangrijk waren geweest. [20]

Het gebruik van algebra om meetkundige vraagstukken op te lossen was voor Descartes geen doel maar middel, en bijna een vanzelfsprekendheid. Hij besteedde dan ook weinig woorden aan de verduidelijking ervan en liet veel details over aan de lezer.

2.4 Fermat

Ongeveer in dezelfde tijd dat Descartes de *Géométrie* uitbracht, schreef Fermat een klein pamfletje van slechts 8 pagina's, waar hij de Latijnse titel "Ad locos planos et solidos isagoge" aan gaf, te vertalen met *Inleiding in de vlakke en ruimtelijke plaatsen*. Het is pas in 1679 postuum in zijn verzamelde werk gepubliceerd, maar circuleerde kort na productie wel in manuscriptvorm in wiskundige kringen te Parijs[21]. Het is als inleiding in analytische meetkunde veel beter verteerbaar dan de beroemde appendix van Descartes waar onterecht zoveel aan is toegeschreven. Het begint zo:

> Dat de ouden zeer veel over plaatsen geschreven hebben is zonder twijfel. Daarvan getuigt Pappus, die aan het begin van zijn 7e boek verzekert dat Apollonius over vlakke, en Aristaios over ruimtelijke plaatsen geschreven heeft. Maar als wij ons niet vergissen, viel hun het onderzoek van plaatsen niet bepaald licht. [...] Wij onderwerpen daarom deze wetenschap aan een eigen en bijzondere analyse, opdat er voortaan een algemene weg tot de locus(problemen) openstaat[22].

Ook Fermat sluit dus aan bij de klassieken, Pappus voorop. Hij vervolgt:

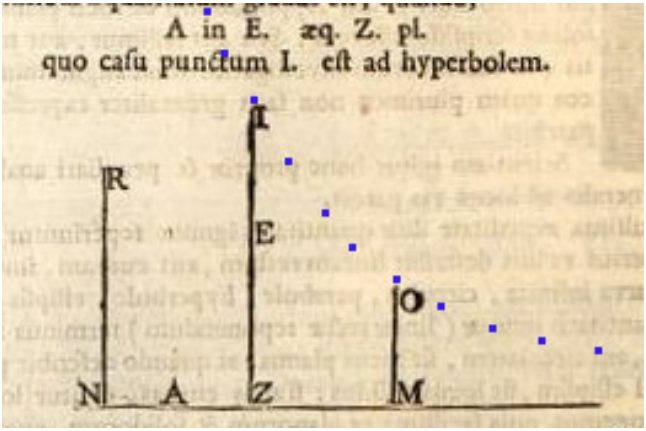
> Wanneer in een uiteindelijke vergelijking twee onbekende grootheden voorkomen, hebben we een locus, en het eindpunt van een van hen beschrijft een rechte of kromme lijn. [...] Vergelijkingen kunnen gemakkelijk aangevat worden wanneer wij de twee onbekende grootheden onder een vaste hoek, die we meestal recht nemen, samenbrengen, met het eindpunt van één van hen in positie gegeven. Indien geen van de onbekende

[20] [3, p. 96].
[21] Zie [13], Duitse vertaling in [14], Engelse vertaling in [19, 389–396].
[22] [13, p. 1]

grootheden een kwadraat te boven gaat, is de plaats vlak of ruimtelijk[23].

Hierbij moeten we weten dat in de Griekse traditie "vlakke" plaatsen overeenkomen met constructies met uitsluitend passer en latje, terwijl voor "ruimtelijke" plaatsen bovendien ook kegelsneden zijn toegestaan. Fermat legt dus uit dat een vergelijking met twee onbekenden correspondeert met een kromme door het uitzetten van die onbekenden onder een (meestal rechte) hoek, waarbij vergelijkingen van maximaal graad 2 overeenkomen met lijnen, cirkels en kegelsneden.

Als voorbeeld zullen we laten zien hoe Fermat de vergelijking $xy = a^2$ behandelt, die hij zelf weergeeft in de notatie van Viète "A in E aeq. Z plano" (Viète gebruikte klinkers voor onbekenden, medeklinkers voor bekenden, en hield vast aan de dimensionele homogeniteit waardoor Z, naast het product van A en E, een oppervlak (*plan*) moet zijn).



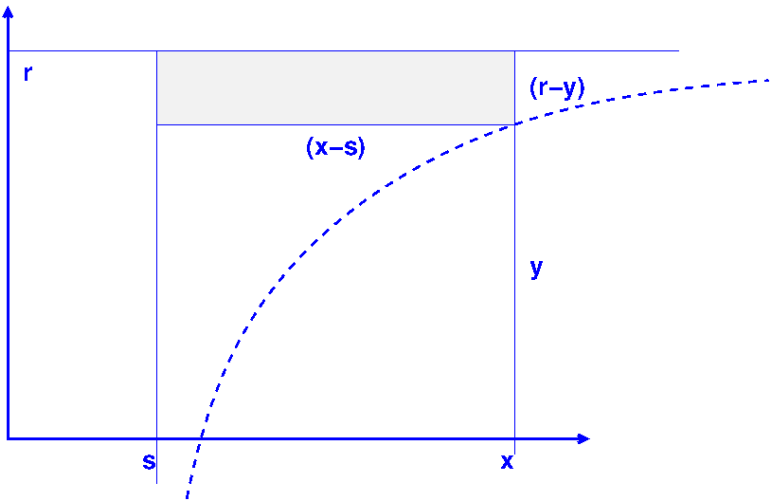
Figuur 2.3: Fermats constructie van de hyperbool bij de vergelijking $xy = z^2$ (met enkele punten van de kromme door de auteur toegevoegd).

Zie figuur 2.3. Fermat kiest N als "oorsprong" en zet op de rechte NM de eerste variabele A uit, zeg $A = NZ$. Vanuit Z komt de tweede variabele $E = ZI$, onder een vaste, in dit geval rechte hoek. Vanuit N trekt hij ook een tweede rechte NR evenwijdig aan ZI. Je kunt nu zeggen dat voor verschillende combinaties van $A = NZ$ en $E = ZI$, het eindpunt I van ZI

[23] [13, p. 1]

steeds op een hyperbool met asymptoten NM en NR moet liggen. Om de hyperbool vast te leggen heb je, behalve de asymptoten, maar één punt nodig, dat je kunt vinden met een zelfgekozen abscis NM en bijbehorende ordinaat MO zodanig dat $NM \cdot MO$ even groot is als het oppervlak Z.

Zo werkt Fermat een aantal standaardvormen van vergeijkingen van graad 1 en 2 af. Hij laat ook zien dat je andere vormen tot deze kunt herleiden. Bijvoorbeeld (we gebruiken moderne notatie): de vergelijking $rx + sy = d^2 + xy$ kun je herleiden tot $(x - s)(r - y) = d^2 - rs$, door aftrekken van $xy+rs$ aan beide zijden. In het resultaat kun je links wederom het product van twee variabele grootheden herkennen, en rechts het verschil van twee vaste grootheden. De bijbehorende hyperbool staat in figuur 2.4.



Figuur 2.4: De hyperbool bij de vergelijking $rx+sy = d^2+xy$ (figuur door de auteur).

Fermat illustreert zijn nieuwe methode met het volgende meetkundige vraagstuk. Gegeven zijn twee punten N en M op onderlinge afstand b; gevraagd de locus van alle punten I waarvoor geldt dat de vierkanten op IN en IM samen in een gegeven verhouding r staan tot driehoek INM. Dit probleem geeft aanleiding tot een vergelijking met twee onbekenden $2x^2 + b^2 - 2bx + 2y^2 = rby$ die van één van de types is die hij al besproken heeft en daarmee is de locus bekend en het probleem opgelost.

In acht bladzijden geeft Fermat zo een systematische, leesbare introductie in de “analytische meetkunde” van krommen tot en met graad 2. Het is

echt een aanrader om te lezen, zeker vergeleken met de *Géométrie* van Descartes die de lezer vaak allerlei lastige details zelf laat uitzoeken.

Als we nu Descartes en Fermat vergelijken dan zien we bij Fermat, net als bij Descartes, een één-assig "co"-ordinatensysteem, dat niet noodzakelijk rechthoekig is. We zien dat beide auteurs ten volle het verband zien tussen locusproblemen en onbepaalde vergelijkingen met twee onbekenden. Beide sluiten aan en bouwen voort op Viète, Pappus, en vroegere klassieke meetkundigen. Hoewel ik het bij Descartes niet heb laten zien gebruiken beide auteurs transformaties. Opvallend is dat ook Fermat, aan het eind van zijn betoog, refereert aan het Pappusprobleem (voor willekeurig aantal lijnen), weliswaar zonder er een uitgebreide behandeling van te geven als Descartes. Tot slot: beide auteurs overschatten de kracht van algebra, want zij denken dat oplossingen van vergelijkingen van willekeurig hoge graad gevonden en geconstrueerd kunnen worden. Ook al staat Descartes in feite elke algebraïsche kromme met een lagere graad als constructiemiddel toe, het oplossen van de vergelijkingen zelf zou in het algemeen onmogelijk blijken.

2.5 Slotopmerkingen

Wat zijn de verschillen tussen de analytische meetkunde van nu en de versies die we bij Fermat en Descartes hebben aangetroffen? Allereerst natuurlijk het coördinatensysteem: de zogenaamd "cartesische" coördinaten doen pas hun intrede bij Cramer in 1750[24]. Daarnaast gebruikten Fermat en Descartes wel algebra bij het oplossen van meetkundeproblemen, maar zij bleven daarbij binnen de meetkundige context. Dit betekende enerzijds dat zij alleen positieve waarden van de variabelen erkenden, en anderzijds dat de algebra slechts een *vindmiddel* was voor een meetkundige constructie: een meetkundige vraag vereiste een meetkundig antwoord. Vergelijk dat eens met het voorbeeld van de raaklijn aan een ellips bij Rutgers: algebra en meetkunde staan daar veel meer op gelijke voet, en voor Rutgers is het vraagstuk opgelost zodra hij n in de lijnvergelijking $y = mx + n$ heeft uitgerekend.

Er zijn meer verschillen te noemen, zoals het doen van meetkunde in drie of meer dimensies, of het gebruik van pool- en andere coördinatensystemen, maar ik denk dat hierboven de meest kenmerkende verschillen staan.

[24] De eerste plaats waar nadrukkelijk sprake is van *twee* assen en een paar coördinaten is, bij mijn weten, [5, p. 3–4]. Interessant is dat Cramer zelf naar Descartes verwijst terwijl er toch duidelijke verschillen zijn tussen de twee.

Tot besluit geef ik u het volgende citaat van Maxwell uit 1871:

> Position and form, which were formerly supposed to be in the exclusive possession of geometers, were reduced by Descartes to submit to the rules of arithmetic by means of that ingenious scaffolding of co-ordinate axes which he made the basis of his operations[25].

We leren hieruit twee dingen: ten eerste dat Maxwell de *Géométrie* van Descartes onmogelijk gelezen kan hebben, want dan zou hij wel beter weten, en ten tweede dat het verdichtsel van het *cartesische* coördinatensysteem in Maxwells tijd zich al volledig voltrokken had.

Bibliografie

[1] D. Diderot en J. le Rond d'Alembert (eds.) *Encyclopédie ou Dictionnaire raisonné des sciences, des arts et des mtiers, par une Société de Gens de lettres*, Parijs 1751 (online: http://encyclopedie.uchicago.edu/)

[2] H.J.M. Bos, *Redefining Geometrical Exactness*, Springer 2001.

[3] H.J.M. Bos, *Descartes en het begin van de Analytische Meetkunde*, CWI Syllabus 25, Vakantiecursus 1989.

[4] C.F. Boyer, *History of Analytic Geometry*, New York 1956.

[5] G. Cramer, *Introduction à l'analyse des lignes courbes algébriques*, Genève 1750.

[6] cTWO (opdrachtgever), *Domein Meetkunde havo B: 1 Analytische meetkunde*, versie voor proefscholen met ingedikt programma, Utrecht 2012.

[7] Hk. de Vries, *De "Geometrie" van Descartes, en de "Isagoge" van Fermat*, Historische Studiën, Deel I, Hst. VII, Groningen 1926.

[8] D.E. Smith en M.L. Latham (eds), *The Geometry of René Descartes: with a facsimile of the first edition*, 1925 (Dover reprint 1954).

[9] W.W. Wilhelm (ed), *Meetkunde / René Descartes: Vertaald en ingeleid door Wim W. Wilhelm*, Delft 2009.

[10] E.J. Dijksterhuis, *Descartes als wiskundige*, Euclides 9 (1932), 56–76.

[25] [16, p. 258].

[11] J.G. Dopper, *A life of learning in Leiden: the mathematician Frans van Schooten (1615–1660)*, proefschrift, Utrecht 2014.

[12] L. Euler, *Introductio in analysin infinitorum: liber secundus, continens Theoriam Linearum curvarum, una cum appendice de Superficiebus*, Lausanne 1748 (facsimile Brussel 1967).

[13] P. de Fermat, *Ad locos planos et solidos isagoge*, in: Varia Opera, Toulouse 1679, 1–8; Engelse vertaling in [19, 389–396].

[14] H. Wieleitner (vert.), *Einführung in die ebenen und körperlichen Örter von Pierre de Fermat (um 1636)*, Ostwald's Klassiker der exakten Wissenschaften, Leipzig 1923

[15] M.S. Mahoney, *The mathematical career of Pierre de Fermat (1601–1665)*, Princeton 1973

[16] J.C. Maxwell, *On the mathematical classification of physical quantities*, in: Proceedings of the London Mathematical Society, III:34, 1871

[17] Pappus Alexandrinus, A. Jones (ed.) *Book 7 of the Collection*, Sources in the history of mathematics and the physical sciences 8, New York 1986 (2 delen)

[18] J.G. Rutgers, *Inleiding tot de analytische meetkunde: eerste deel*, Groningen 1928 (tweede druk)

[19] D.E. Smith, *A Source Book in Mathematics*, 1929 (Dover reprint 1959).

[20] J.L. Stammetz, *Volkoomen wiskundig woordenboek, daar in alle kunstwoorden en zaaken [. . .] duidelyk verklaart worden*, Leiden 1740.

[21] L. van den Broek e.a., *Analytische meetkunde*, Verbeterde experimentele uitgave voor wiskunde D vwo, Ede 2010 (online beschikbaar: www.wageningse-methode).

3 Some Basics of Classical Logic

Theodora Achourioti

AUC/ILLC, University of Amsterdam

Introduction

In this lecture we recall some of the basics of logic in general and of classical logic in particular. Logic is traditionally regarded to lie in the intersection between mathematics and philosophy, yet it finds applications in many other areas such as computer science, linguistics, artificial intelligence, psychology etc. While in these areas one often sees the appearance of non-classical logics, the logic underlying mathematical reasoning is predominantly classical with the only serious contestant historically being intuitionistic logic, the logic of intuitionistic mathematics, first developed by the Dutch mathematician Brouwer.

In contrast to mathematics, language is central to logic. The logician reflects on mathematical results inside a formal system in which these results can be represented; so the expressive power needed for such representation becomes an essential part of the formalism. Thus, the first step in defining a logic is to define a language. We first recall the syntax of the language of classical propositional and predicate (or first-order) logic.

3.1 Syntax

The language of classical propositional logic consists of:

- Propositional variables: $p, q, r...$
- Logical operators: $\neg$, $\wedge$, $\vee$, $\rightarrow$
- Parentheses: (,)

Recursive definition of a well-formed formula (wff) in classical propositional logic:

Definition 1.

1. *A propositional variable is a wff.*
2. *If ϕ and ψ are wffs, then $\neg\phi$, $(\phi \wedge \psi)$, $(\phi \vee \psi)$, $(\phi \rightarrow \psi)$ are wffs.*
3. *Nothing else is a wff.*

Clause 1 gives the *atomic* propositions. The Greek letters ϕ, ψ etc. are used as meta-variables to denote arbitrary formulas.

First-order logic extends the language of propositional logic with terms, relations and quantifiers. Formally, the language of first-order logic consists of:

- Terms: a) constants $(a, b, c...)$, b) variables $(x, y, z...)$[1]
- n-place predicates $P, Q, R...$
- Logical operators: a) propositional connectives, b) quantifiers $\forall$, $\exists$
- Parentheses: (,)

Recursive definition of a wff in classical first-order logic:

Definition 2.

1. *If R is a n-place predicate and $t_1, ...t_n$ are terms, then $R(t_1...t_n)$ is a wff.*
2. *If ϕ and ψ are wffs, and t is a variable, then $\neg\phi$, $(\phi \wedge \psi)$, $(\phi \vee \psi)$, $(\phi \rightarrow \psi)$, $\forall t\phi$ and $\exists t\phi$ are wffs.*
3. *Only strings of symbols constructed by the previous clauses are wffs.*

First-order formulas with *free* (as opposed to *bound*) occurrences of variables are called *open* formulas; otherwise they are *closed* (also called *sentences*). A variable x is bound if it falls under the scope of a quantifier $\forall x$ or $\exists x$.

[1] Functions are also typically included in first order theories.

The recursive nature of the above definitions is important in that it allows inductive proofs on the entire logical language.
Exercise 1: Prove that all wffs of classical first-order logic have an even number of brackets.

A special binary predicate '=', is often added to the language of first-order logic, especially if the language is employed to talk about mathematics. This is different from other predicates which generally belong to the non-logical part of the vocabulary; it is typically taken to express the identity relation which has well-known logical properties, in particular, it is reflexive, symmetric and transitive. We see in the next section that '=' admits a special semantics.

3.2 Semantics

The semantics of a logic specifies the meaning of expressions that are syntactically well-formed and does so in an unambiguous way (as opposed to the meaning of expressions in natural language which is often subject to ambiguities). The usual semantics for a logic is truth-conditional, that is, the meaning of a wff is given by the conditions under which this formula is true. In the case of classical logic, this semantics is based on the principle of *bivalence*, according to which the truth-value of a formula is exactly one of *true* (1) or *false* (0). The meaning of complex expressions is *compositionally* defined, that is, its truth-value is fully determined by the truth-values of its atomic components and the semantics of the logical operators. This is achieved by the *valuation* function, a total function that assigns truth-values to wffs as follows:

Definition 3. *The valuation function V is defined as the function that assigns to each wff one of* 1 *or* 0 *in accordance with the following clauses:*

- $V(\phi) = 1$ *or* 0 *and not both, for* ϕ *atomic*
- $V(\neg\phi) = 1$ *iff* $V(\phi) = 0$
- $V(\phi \wedge \psi) = 1$ *iff* $V(\phi) = 1$ *and* $V(\psi) = 1$
- $V(\phi \vee \psi) = 1$ *iff either* $V(\phi) = 1$ *or* $V(\psi) = 1$
- $V(\phi \rightarrow \psi) = 1$ *iff either* $V(\phi) = 0$ *or* $V(\psi) = 1$

Given this semantics, the language of propositional logic is truth-functionally

complete also with smaller sets of logical operators; e.g. $\{\neg, \wedge\}$, $\{\neg, \vee\}$, and $\{\neg, \rightarrow\}$. A set of logical operators is truth-functionally complete when it is sufficient in order to express all possible meanings. For the set $\{\neg, \rightarrow\}$ for example, this means that for some assignment of truth-values to atomic propositions, any wff can be translated to one containing only $\neg$ and $\rightarrow$ as logical operators while preserving its truth-value.

The *duality* of $\wedge$ and $\vee$ as given by the so-called De Morgan laws, is a good example of this truth-preserving syntactic translation between formulas:

De Morgan laws:
$\neg(\phi \wedge \psi) \equiv \neg\phi \vee \neg\psi$
$\neg(\phi \vee \psi) \equiv \neg\phi \wedge \neg\psi$

The sign '$\equiv$' belongs to the meta-language (the language used to talk about the object language, in this case, the language of propositional logic) and is used to denote *logical equivalence*; that is, the formulas on the left and right side obtain the same truth-value under the same assignment of truth-values to their atomic components. Again, this means that, in the presence of negation, one can safely remove one of $\wedge$ or $\vee$ from the language without loss in the semantics (given the right translations are performed). Admitting all four operators, however, makes the language easier to work with and to use for translating expressions from natural language.

Some more terminology is in order. *Tautologies* and *contradictions* are wffs that are true or false respectively under all possible valuations. For example, any instance of the form $\phi \vee \neg\phi$, or $\phi \rightarrow \phi$ is a tautology and any instance of $\phi \wedge \neg\phi$ is a contradiction. *Satisfiable* are wffs that are true under some valuation. It follows that contradictions are not satisfiable formulas.

For first-order logic, the semantics that has prevailed in mathematical logic is the so-called model-theoretic semantics due to Tarski. This semantics is grounded on set theory.

Just as the truth-value of a wff in the propositional language is relative to the assignment of truth-values to the atomic propositions (first clause in the definition of the valuation function) the truth-value of a wff in predicate logic is relative to a set-theoretic structure called a *model* $\mathcal{M}$, and to an *assignment* g. A model $\mathcal{M}$ consists of a domain $\mathcal{D}$, that is, a non-empty set, and an interpretation function $\mathcal{I}$: $\mathcal{M} = \{\mathcal{D}, \mathcal{I}\}$. The interpretation $\mathcal{I}$ is a total function that fixes the meaning of the non-

logical expressions of the language, that is, constants and predicates. The assignment g is a total function that assigns objects to the variables of the language. Assignments are necessary in order to fix the meaning of formulas that contain free variables. There are in principle more than one assignments possible relative to a particular model $\mathcal{M}$.

The interpretation function $\mathcal{I}$ is subject to the following constraints:

Definition 4.

- $\mathcal{I}(t) \in \mathcal{D}$, *if* t *is a constant*
- $\mathcal{I}(R)$ *is a set of ordered n-tuples from the domain* $\mathcal{D}$, *for* R *an n-ary relation*

Before giving the valuation function, we define the auxiliary notion of *denotation* $[t]_{\mathcal{M},g}$ of a term t as follows:

Definition 5.

- $[t]_{\mathcal{M},g} = \mathcal{I}(t)$ *if* t *is a constant*
- $[t]_{\mathcal{M},g} = g(t)$ *if* t *is a variable*

We are now in a position to define the valuation function $V_{\mathcal{M},g}$ that fixes the truth-value of a wff in first-order logic relative to a model $\mathcal{M}$ and an assignment g:

Definition 6. *For any wffs* ϕ, ψ, *n-ary predicate* R *and terms* t_1, ... , t_n:

- $V_{\mathcal{M},g}(R(t_1, ..., t_n)) = 1$ *iff* $< [t_1]_{\mathcal{M},g}, ..., [t_n]_{\mathcal{M},g} > \in \mathcal{I}(R)$
- $V_{\mathcal{M},g}(\neg\phi) = 1$ *iff* $V_{\mathcal{M},g}(\phi) = 0$
- $V_{\mathcal{M},g}(\phi \wedge \psi) = 1$ *iff* $V_{\mathcal{M},g}(\phi) = 1$ *and* $V_{\mathcal{M},g}(\psi) = 1$
- $V_{\mathcal{M},g}(\phi \vee \psi) = 1$ *iff either* $V_{\mathcal{M},g}(\phi) = 1$ *or* $V_{\mathcal{M},g}(\psi) = 1$
- $V_{\mathcal{M},g}(\phi \rightarrow \psi) = 1$ *iff either* $V_{\mathcal{M},g}(\phi) = 0$ *or* $V_{\mathcal{M},g}(\psi) = 1$
- $V_{\mathcal{M},g}(\forall x\phi(x)) = 1$ *iff for all assignments* g,$V_{\mathcal{M},g}(\phi(x)) = 1$
- $V_{\mathcal{M},g}(\exists x\phi(x)) = 1$ *iff for some assignment* g, $V_{\mathcal{M},g}(\phi(x)) = 1$

Finally, we add a clause for the special predicate '=' expressing the identity relation:

- $V_{\mathcal{M},g}(t_1 = t_2) = 1$ *iff* $[t_1]_{\mathcal{M},g}$ *and* $[t_2]_{\mathcal{M},g}$ *are the same*

Tautologies and *contradictions* are now formulas that are true or false respectively in relation to all possible models $\mathcal{M}$ and assignments g. Note that it is a consequence of the non-empty domains restriction of the model theory that the formula $\exists x(x = x)$ is a tautology. So-called free logics lift this restriction.

We now give the definition of *satisfiability* for a set of formulas Γ.

Definition 7. *A set of formulas Γ is satisfiable iff there is model $\mathcal{M}$ and assignment g such that for all $\gamma \in \Gamma$, $V_{\mathcal{M},g}(\gamma) = 1$.*

As a final remark, it is interesting to point out that the semantics of first-order logic is *extensional*: intersubstituting terms with the same denotation does not change the truth-value of formulas under some valuation relative to a model $\mathcal{M}$ and assignment g. According to the clauses above, for any predicate P and terms t_1 and t_2 such that $t_1 = t_2$, $V_{\mathcal{M},g}(P(t_1)) = 1$ iff $V_{\mathcal{M},g}(P(t_2)) = 1$. This kind of semantics does not account for expressions that create *intensional* or so-called *opaque* contexts; to use Frege's famous example, I may believe that what I see in the sky is the evening star, but not that it is the morning star, if I do not also know that these two names refer to the same entity, namely, planet Venus. Note that truth-functionality breaks down in these contexts. The truth-value of the sentence 'I believe that what I see in the sky is not the morning star' does not depend on whether what I see is indeed the morning star or not. Epistemic logics, the logics of knowledge and belief develop semantics to account precisely for such contexts (see next lecture).

3.3 Derivability

Besides its expressive part, a logic is characterised by its normative part, in particular, its definition of derivability and semantic validity. Derivability '$\vdash$' is a syntactic or proof-theoretic notion. A conclusion ϕ is derivable inside a formal system from a set of premises Γ if it can be syntactically generated in accordance with the rules (and axioms) of that system in a finite number of steps. Formalised proofs are finite objects:

$\Gamma \vdash \phi$ iff there exists finite $\Gamma' \subseteq \Gamma$ such that $\Gamma' \vdash \phi$ in a finite number of steps.

Premises and conclusion constitute an 'argument'. When the set of pre-

mises is the empty set, the conclusion is a called a *theorem*.

In order to get a good grasp of classical derivability we look at the proof system of Natural Deduction (ND) invented in its current form by Gentzen in 1934. The system is called 'natural' in that it aims to capture mathematical reasoning with logical expressions. ND contains rules which allow to introduce or eliminate logical operators (*Introduction* and *Elimination* rules). We now give the ND rules for classical first-order logic.

Conjunction [2]

n	ϕ	
$\vdots$	$\vdots$	
m	ψ	
$\vdots$	$\vdots$	
l	$\phi \wedge \psi$	$\wedge$I, n, m

n	$\phi \wedge \psi$	
$\vdots$	$\vdots$	
m	ϕ	$\wedge$E, n

Disjunction

n	ϕ	
$\vdots$	$\vdots$	
m	$\phi \vee \psi$	$\vee$I, n

n	$\phi \vee \psi$	
$\vdots$	$\vdots$	
m	ϕ	
$\vdots$	$\vdots$	
s	χ	
t	ψ	
$\vdots$	$\vdots$	
l	χ	
r	χ	$\vee$E, n, m–s, t–l

[2] We have to add that ψ may also be the conclusion for the Elimination rule for Conjunction as well as the premise for the Introduction rule for Disjunction.

Implication

$$\begin{array}{lll} n & \phi & \\ \vdots & \vdots & \\ m & \psi & \\ s & \phi \to \psi & \Rightarrow\text{I}, n, m \end{array} \qquad \begin{array}{lll} n & \phi & \\ \vdots & \vdots & \\ m & \phi \to \psi & \\ \vdots & \vdots & \\ s & \psi & \Rightarrow\text{E}, n, m \end{array}$$

Negation [3]

$$\begin{array}{lll} n & \phi & \\ \vdots & \vdots & \\ m & \bot & \\ \vdots & \vdots & \\ s & \neg\phi & \neg\text{I}, n, m \end{array} \qquad \begin{array}{lll} n & \neg\neg\phi & \\ \vdots & \vdots & \\ m & \phi & \neg\neg\text{E}, n \end{array}$$

Some of the rules make use of assumptions which need then to be discharged.

We can now derive the *law of excluded middle* (LEM): $\phi \vee \neg\phi$. LEM is a characteristic law of classical logic.
Exercise 2: Perform the derivation.

The rule of *double negation elimination* (DNE) is not an intuitionistically acceptable rule. All other rules are contained in both classical and intuitionistic ND. Since DNE is used essentially in the derivation of LEM, this law is not intuitionistically derivable. The exclusion of DNE reflects the intuitionistic philosophical idea that there are no mind independent mathematical facts.
Example of a mathematical proof that is not intuitionistically acceptable:

There exist irrational numbers a, b such that a^b is rational.

[3] $\bot$ is used as short for $\phi \wedge \neg\phi$.

Proof

- Consider $\sqrt{2}$, which we know to be irrational.
- $\sqrt{2}^{\sqrt{2}}$ is either rational or irrational. (By excluded middle)

  1. If $\sqrt{2}^{\sqrt{2}}$ is rational: take $a = b = \sqrt{2}$. Then a^b is rational.
  2. If $\sqrt{2}^{\sqrt{2}}$ is irrational: take $a = \sqrt{2}^{\sqrt{2}}$, and $b = \sqrt{2}$.
     Then, $a^b = (\sqrt{2}^{\sqrt{2}})^{\sqrt{2}} = \sqrt{2}^2 = 2$, which is rational. □

Note that the converse of DNE, that is, $\phi \vdash \neg\neg\phi$, is both classically and intuitionistically derivable.
Exercise 3: Perform the derivation.

The rule of *ex falso quodlibet sequitur* (meaning 'from a contradiction everything follows') is also characteristic of classical logic, as it distinguishes it from other logics, crucially, paraconsistent logics.

Ex falso

$$\begin{array}{l|ll} \vdots & \vdots & \\ n & \bot & \\ \vdots & \vdots & \\ m & \psi & \neg\text{E}, n, m \end{array}$$

This rule need not be added to the classical calculus as all of its instances are classically derivable from the existing rules.
Exercise 4: Prove that $\bot \vdash \psi$ for arbitrary ψ in classical ND.

Note that the derivation for exercise 4 uses the DNE rule essentially which means that *ex falso* is not intuitionistically derivable. It is, however, intuitionistically valid, hence added to the system as a separate rule.

We proceed with the rules for the quantifiers.

Quantifiers

$$\begin{array}{l|ll} n & \phi(a) & \\ \vdots & \vdots & \\ m & \forall x\phi(x) & \forall\text{I},\ n \end{array} \qquad \begin{array}{l|ll} n & \forall x\phi(x) & \\ \vdots & \vdots & \\ m & \phi(a) & \forall\text{E},\ n \end{array}$$

Where in the rule $\forall I$, a cannot appear in any premise or assumption active at line m.

$$\begin{array}{l|ll} n & \phi(a) & \\ \vdots & \vdots & \\ m & \exists x\phi(x) & \exists\text{I},\ n \end{array} \qquad \begin{array}{ll|ll} n & \exists x\phi(x) & & \\ \vdots & \vdots & & \\ m & & \phi(a) & \\ \vdots & & \vdots & \\ s & & \chi & \\ r & \chi & & \exists\text{E},\ n,\ m\text{–}s \end{array}$$

Where in the rule $\exists E$, a cannot appear earlier in the derivation, and it cannot appear in the formula χ derived at line r.

A different proof theoretic approach is found in the older so-called axiomatic, or Hilbert type systems. Whereas ND consists only of rules, an axiomatic system lists a number of axioms which are seen as the foundations on which the rest of the theorems are grounded.[4] Axiomatic systems typically have very few rules, and do not make use of assumptions. This means that they do not support conditional proofs and any line in an axiomatic proof is itself a theorem. This is also the reason why they are less 'natural' in describing mathematical reasoning. For a brief comparison, we present an axiomatic system for propositional classical logic.

Axiomatic System for Classical Propositional Logic:

- Rule: Modus Ponens[5]

[4]The invention of ND was partly motivated by dissatisfaction with axiomatic systems.
[5]Modus Ponens is the same as the Elimination rule for implication on page 7.

- Axioms: The result of *substituting* wffs for ϕ, ψ, and χ in any of the following schemas is an axiom.
  1. $\phi \to (\psi \to \phi)$
  2. $(\phi \to (\psi \to \chi)) \to ((\phi \to \psi) \to (\phi \to \chi))$
  3. $(\neg\psi \to \neg\phi) \to ((\neg\psi \to \phi) \to \psi)$

Exercise 5: Give an axiomatic derivation of $p \to p$ and one in ND.

As the derivation for exercise 5 shows, proofs in ND are usually shorter than axiomatic proofs. They are also much more informative than axiomatic proofs. Yet, for meta-mathematical results that require induction over the entire proof system, it is much easier to work with a simple axiomatic system. Since the systems are proof theoretically equivalent, the logician can switch between proof theories depending on the task at hand.

Finally, an important proof-theoretic notion is that of *consistency*. We define it via its negation:

Definition 8. *A set of formulas* Γ *is inconsistent iff* $\Gamma \vdash \bot$.

This means that there is a set $\Gamma' \subseteq \Gamma$ such that for some formula ϕ, both ϕ and its negation $\neg\phi$ are derivable from Γ' in a finite number of steps.

3.4 Entailment

Whereas *derivability* '$\vdash$' in a formal system is the proof theoretic way to make precise that a conclusion ϕ follows logically from a given set of premises Γ (what is also called 'logical consequence'), *entailment* '$\models$' is the semantic or model theoretic way. One can also talk about *syntactic* and *semantic validity* in order to distinguish the two. The main idea behind entailment is *truth-preservation*:

Definition 9. *An argument* $< \Gamma, \phi >$ *is semantically valid iff whenever all premises* $\gamma \in \Gamma$ *are true the conclusion* ϕ *is also true.*

For propositional logic, this means that for any assignment of truth-values to the atomic components of the formulas such that the valuation function

decides the premises to be true, the conclusion is also decided to be true. For predicate logic it means that:

Definition 10. $\Gamma \vDash \phi$ *iff for every model* $\mathcal{M}$ *and assignment* g*, if* $V_{\mathcal{M},g}(\gamma) = 1$ *for all* $\gamma \in \Gamma$*, then* $V_{\mathcal{M},g}(\phi) = 1$*.*

In the special case that Γ is the empty set, ϕ is said to be a *logical truth.*[6]

Although derivability is a procedural notion, entailment is better thought of as a relational notion –where the relation holds between the semantic values of the formulas– in the sense that one need not worry about how to get from premises to conclusion: the conclusion need not be syntactically generated from the premises by some formal procedure.[7]

Truth-preservation basically comes down to the absence of counter-examples, that is, valuations that make the premises true and the conclusion false. It is often easier to check the validity of an argument by constructing a counter-example.[8] This seems especially obvious for first order logic where one would have to survey an infinite number of models (including infinite models) which make the premises true in order to find out whether they also make the conclusion true. In contrast, a single counter-model, i.e. a model that satisfies the premises but not the conclusion, is sufficient to prove an argument invalid. However, finding whether such a counter-model exists may turn out to be as challenging (see discussion of undecidability in the next section).

We now list some basic properties of classical entailment.

Theorem 1. *For any* $\gamma \in \Gamma$*,* $\Gamma \vDash \gamma$*.*

The following property is *transitivity.* It is present, for example, in the use of lemmas in mathematical proofs.

Theorem 2. *If* $\Gamma \vDash \delta$ *and* $\delta \vDash \phi$*, then* $\Gamma \vDash \phi$*.*

Classical entailment is *monotonic.*[9] Monotonicity ensures the stable growth

[6] Logicians sometimes use the term 'valid' for such formulas; to avoid confusion we use this term here only for arguments.

[7] Note, though, that there are operational approaches to logical consequence, where the latter is seen as an operation which outputs sets of possible conclusions.

[8] This is indeed the idea behind a proof method we have not discussed here, namely, semantic tableaux invented by the Dutch mathematician Beth.

[9] Tarski took these three properties to be characteristic of logical consequence in gene-

of mathematical knowledge in the sense that new axioms should not invalidate existing theorems.

Theorem 3. *If* $\Gamma \subseteq \Delta$ *and* $\Gamma \models \phi$*, then* $\Delta \models \phi$*.*

It is a corollary of monotonicity that:

Theorem 4. *If* ϕ *is a tautology, then* $\Gamma \models \phi$ *for any* Γ*.*

Notice that $\Gamma \models \phi$ iff the set $\{\Gamma, \neg\phi\}$ is not satisfiable, which also means that:

Theorem 5. *If* Γ *is not satisfiable, then* $\Gamma \models \phi$ *for any* ϕ*.*

This last one can be seen as the semantic counterpart of the ex falso rule. The proof is a straightforward application of the definition of entailment. **Exercise 6**: Prove monotonicity for first order classical entailment.

3.5 Meta-theory

Some of the most important meta-theorems of classical logic have to do with the relation between derivability and entailment, or put somewhat vaguely, the relation between syntax and semantics.

Soundness guarantees that all provable formulas are logical truths. A sound proof system is one that does not generate falsehoods; however, it may undergenerate by proving less than what is true according to the given semantics. It is *completeness* that guarantees of all logical truths that they can be proven. Of course a complete proof system with respect to a given semantics may overgenerate by proving more than logical truths. An inconsistent proof system, for instance, is trivially complete.

The proof systems we have presented are both sound and complete with respect to the classical semantics we have given. Completeness for first order logic was first established by Gödel in 1929, although it is a later and more simplified proof due to Henkin that is mostly used today. The combination of soundness and completeness shows that semantic validity and provability coincide:

ral. However, there are prominent logics today that violate one or the other. E.g. monotonicity fails in logics for belief revision or some logic programming languages.

Theorem 6. $\vdash \phi$ *iff* $\vDash \phi$.

This means that we can safely switch between proof theory and semantics. For example, we can use model theoretic arguments to establish facts about provability, as when showing that $\Gamma \nvdash \phi$ by constructing a counter-model, or proof theory to establish facts about the semantics, as when showing that $\Gamma \vDash \phi$ by means of a deductive proof. In fact, the proof theoretic notion of consistency and the semantic notion of satisfiability also coincide:

Theorem 7. *A set of formulas* Γ *is satisfiable iff it is consistent.*

Proof hint: Soundness is used to prove the left to right direction, completeness for the converse.

Exercise 7: Complete the proofs.

Theorem 8 can now be used to prove another fundamental meta-theoretic result about first order logic, so-called *compactness*:

Theorem 8. *A set of formulas* Γ *is satisfiable iff every finite* Γ' *such that* $\Gamma' \subseteq \Gamma$ *is satisfiable.*

Proof
Left to right direction. Take model $\mathcal{M}$ and assignment g such that they satisfy Γ. Then $\mathcal{M}$, g also satisfy every Γ' such that $\Gamma' \subseteq \Gamma$.
Right to left direction. Assume that Γ is not satisfiable. It follows that Γ is inconsistent, that is, there is finite $\Gamma' \subseteq \Gamma$ such that $\Gamma' \vdash \bot$. It follows that Γ' is not satisfiable. $\square$

The following theorem is an interesting corollary of compactness. It points to the limitations of the expressive power of first order logic:

Theorem 9. *Finiteness cannot be expressed in first order logic.*

Proof
Assume there exists a first order formula ϕ such that it is satisfiable only in finite models. Take the infinite set $\{\phi, \psi_1, \psi_2, ...\}$, where ψ_1 expresses 'there exists at least one element', ψ_2 expresses 'there exist at least two elements' etc. This set violates compactness. Hence, there is no such formula ϕ. $\square$

Compactness can also be used to prove another very interesting result, this time about the first order theory of Peano arithmetic (PA):

Theorem 10. *There exist non-standard models of PA with infinite numbers.*

Proof
Take constant c and the infinite set of sentences $\{c \geq n_1, c \geq n_2, ...\}$ for all natural numbers n. Every finite subset of this set has a model, therefore, by compactness, the set itself has a model. It follows that c must be larger than any natural number, hence, an infinite number. □

Finally, an important difference between propositional and first order logic is that the first is *decidable* whereas the second is not. A logic is decidable if there exists an effective mechanical procedure that decides for arbitrary formulas whether they are logical truths or not.[10] For propositional logic such an effective procedure can be given: one may consider all possible valuations and since there is a finite number of those it can be expected that the process terminates. This will obviously not do for first order logic where infinite models would have to be considered.[11] One can take the proof theory to provide a decision procedure in the following sense: if the formula is indeed a logical truth, we know that it will be generated at some point, given completeness. But at any given point that such a proof is not present there is no way of knowing whether it will be had some time in the future (a sound proof system proves only truths). In order to mark this asymmetry between deciding formulas that are logical truths and formulas that are not, first order logic is said to be *semi-decidable*.[12]

But the reader should not be left with the impression that all is simple with propositional logic. Even though propositional logic is decidable, questions of efficiency naturally arise since the time needed for such a decision process grows exponentially with the size of the formulas. It is an open question today whether there exists an algorithm to terminate this decision process in polynomial time. Problems for which this question is unanswered form the so-called NP-complete class. Problems for which such a polynomial-time algorithm exists are called P problems. Hence the famous, open to this day, question: is it that $P = NP$?

[10] The question of whether first order logic is decidable was first asked by Hilbert in 1928 as his famous Entscheidungsproblem.

[11] Note that infinite models are necessary for proving completeness of first order logic.

[12] This is a general result about first-order logic. Fragments of the logic may be decidable as, for example, monadic logic is (e.g. the Aristotelian theory of syllogisms).

3.6 Further Reading

(in reccommended order)

Shapiro, S., "Classical Logic", The Stanford Encyclopedia of Philosophy (Winter 2013 Edition), Edward N. Zalta (ed.), URL = `http://plato.stanford.edu/archives/win2013/entries/logic-classical/`

Van Benthem, J., van Ditmarsch, H., van Eijck, J. Jaspars, J., *Logic in Action* (Winter 2014 Edition), URL = `http://logicinaction.org`

Gamut, L.T.F., *Logic, Language and Meaning*, Vol. 1, University of Chicago, 1990.

Boolos, G.S., Burgess, J.P., Jeffrey, R. C., *Computability and Logic* (5th Edition 2007), Cambridge University Press.

Van Dalen, D., *Logic and Structure* (5th Edition 2012), Springer.

4 Honderd gevangenen en een gloeilamp

Hans van Ditmarsch en Barteld Kooi

CNRS & LORIA, Nancy, Frankrijk / Theoretische Filosofie & Kunstmatige Intelligentie, RUG Groningen

Een groep van 100 *gevangenen, gezamenlijk in de gevangeniskantine, wordt medegedeeld dat ze allemaal in isolatiecellen geplaatst zullen worden en daarna één voor één ondervraagd, in een kamer met een lamp met een aan-uitschakelaar. De gevangenen kunnen met elkaar communiceren door de lamp aan of uit te doen (en dat is de enige manier waarop ze kunnen communiceren). De lamp is aan het begin uit. Er is geen vaste volgorde van ondervraging, er is geen standaard tijdsduur tussen de ondervragingen, en dezelfde gevangene kan best meerdere keren achter elkaar ondervraagd worden. We mogen wel aannemen dat op ieder moment iedere gevangene ooit nog eens ondervraagd zal worden. Als een gevangene ondervraagd wordt, kan deze: niets doen, de lamp uitdoen, de lamp aandoen, of verklaren dat iedereen (ten minste een keer) ondervraagd is. Als dit waar is, worden alle gevangenen vrijgelaten. Anders worden ze allemaal opgehangen. Kunnen de gevangenen, zolang ze nog bij elkaar zijn in de kantine en niet naar de isolatiecellen gebracht zijn, een protocol overeenkomen waardoor ze vrijgelaten worden?*

4.1 Tot 100 tellen met één bit?

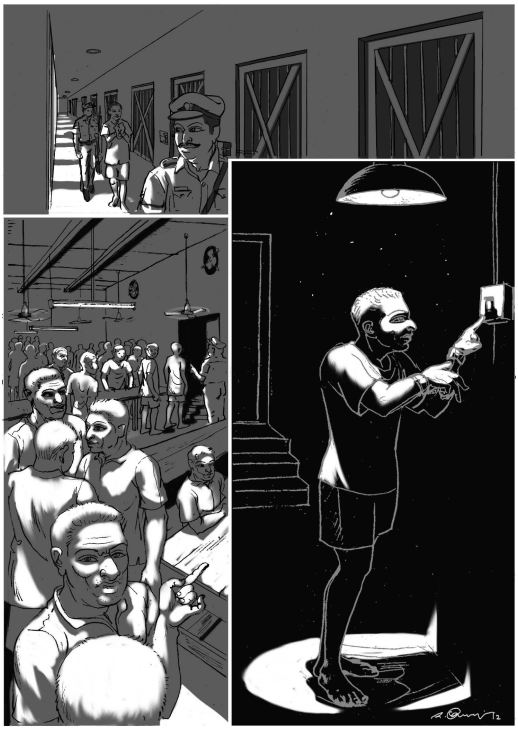
Het raadsel lijkt onoplosbaar. Er is immers maar één bit beschikbaar voor de informatieoverdracht: de lamp kan aan, of uit. Maar er zijn 100 gevangenen. Het getal 100 ligt tussen 64 en 128 en vergt dus 9 bits om het te representeren. En dan hebben we het nog niet over het protocol om het raadsel op te lossen. Hoe kan één bit nu voldoende zijn om dat allemaal te doen?

Een struikelblok *lijkt* de wiskundige intuïtie. De natuurlijke inductie waarmee we vaak problemen te lijf gaan, suggereert een methode om de oplossing te zoeken, die helaas in dit geval niet ergens toe leidt. Laten we het even hardop doen. Oplossen voor één gevangene, oplossen voor twee gevangenen, oplossen voor meer gevangenen.

4.2 Eén gevangene

We beginnen met de basis. Stel er is één gevangene: Anna. De eerste keer dat Anna ondervraagd wordt, weet zij dat iedereen ondervraagd is. Daarvoor is de lamp niet eens nodig. Dus:

Protocol 1: *Als je ondervraagd wordt, is iedereen ondervraagd.*



Figuur 4.1: Wie komt met het goede protocol? Illustratie met dank aan Elancheziyan

4.3 Twee gevangenen

Protocol 1 werkt natuurlijk niet als er meer dan een gevangene is. Maar misschien kunnen we het aanpassen. Stel er zijn twee gevangenen: Anna en Bert. De eerste van de twee die ondervraagd wordt, doet de lamp aan. Stel, zonder verlies aan algemeenheid van de oplossing, dat dit wederom Anna is. Nu is het even afwachten wie er daarna ondervraagd wordt. Als het Bert is, dan ziet Bert dat de lamp aan is, en weet daarom dat Anna reeds ondervraagd is (alleen gevangenen mogen de lamp aan en uit doen). Bert kan dan dus naar waarheid zeggen dat iedereen ondervraagd is. En Anna en Bert gaan vrijuit. Als daarentegen de volgende ondervraagde opnieuw Anna is, is de lamp nog aan, en 'denkt ze dus' dat Bert nog niet ondervraagd is, anders was ze al vrijgelaten. We zijn nu wat vaag: Anna's argument is gebaseerd op wat ze in alle redelijkheid vermoedt dat Bert gedaan heeft. Dit kan in dit geval wel zonder afspraak, maar aan de andere kant is er niets op tegen dat Anna en Bert dit met elkaar afspreken: hier komt het protocol om de hoek kijken. Dit kunnen we als volgt verwoorden:

Protocol 2: *Als je ondervraagd wordt en de lamp is uit, doe hem dan aan; als je ondervraagd wordt, en jij hebt de lamp eerder aangedaan en de lamp is nog steeds aan, doe dan niets; als je ondervraagd wordt, en de lamp is aan maar jij hebt hem niet aangedaan, verklaar dan dat iedereen ondervraagd is.*

4.4 Een protocol voor drie gevangenen?

Stel er zijn drie gevangenen, Anna, Bert en Caroline ... Nu wordt het moeilijker. Anna, die weer als eerste ondervraagd wordt, kan de lamp aandoen. Als daarna bijvoorbeeld Bert ondervraagd wordt, kan Bert de lamp weer uitdoen. Stel dat daarna Anna weer ondervraagd wordt, dan weet Anna dat ten minste twee gevangenen nu ondervraagd zijn. Dat schiet op! Als daarna Caroline ondervraagd wordt, weet Caroline echter niet of een andere gevangene al eerder ondervraagd is: het kan dan ook best zo zijn dat zij als eerste ondervraagd wordt (de tijdsduur tussen ondervragingen is niet bekend)! Wat moet Caroline nu doen? Dat is eigenlijk dezelfde vraag als wat Anna de eerste keer moet doen. Maar laten we nog even doorfantaseren – want het is namelijk fantaseren, we komen er zo echt niet uit: het zou handig zijn als Anna eerst de lamp aan kan doen, daarna zoals hiervoor de lamp weer uit kan zien, en daarna de lamp nog een

keer aandoet. De volgende keer als ze de lamp weer uit ziet, zou dan dus iedereen ondervraagd moeten zijn? Laten we het formaliseren: het protocol wordt nu:

Protocol 3: *Als je ondervraagd wordt en de lamp is uit, doe hem dan aan – houd bij hoe vaak je dat doet; als je ondervraagd wordt, en jij hebt de lamp eerder aangedaan en de lamp is nog steeds aan, doe dan niets; als je ondervraagd wordt en de lamp is aan maar jij hebt hem niet aangedaan, doe hem dan uit. Als je de lamp twee keer hebt uitgedaan, en je wordt opnieuw ondervraagd en de lamp is aan, verklaar dat alle drie gevangenen ondervraagd zijn.*

Dit protocol levert niet wat we wensen: soms gaat het goed, maar soms niet. Het gaat goed als de ondervragingsvolgorde is: Anna – Bert - Anna – Caroline – Anna. Het gaat fout als de ondervragingsvolgorde is: Anna – Bert – Anna – Bert – Anna. Maar hoe weet Anna nu of Bert dan wel Caroline, die allebei ook zo'n protocol volgen, de gevangene was die op zijn beurt de lamp voor de tweede keer uitdoet? Met andere worden, voor 100 gevangenen, als de eerste twee gevangenen om de beurt ondervraagd worden en verder niemand, dan zegt Anna na de 199ste keer ten onrechte dat iedereen ondervraagd is en worden ze allemaal opgehangen. Niet goed ... Met andere woorden, het probleem van dit protocol is dat het soms wel en soms niet werkt. De gevangenen kunnen best een geïformeerd *gokje* wagen dat iedereen ondervraagd is. Maar om *kennis* te verkrijgen dat iedereen ondervraagd is, is meer nodig.

4.5 Geen trucjes

Allerlei trucjes waar de lezer wellicht aan gedacht heeft, zijn niet toegestaan, en het raadsel is ook geen strikvraag in welke zin dan ook: voelen of de lamp warm of koud is, is niet toegestaan (als de lamp uit is maar warm, weet je dat er iemand voor jou ondervraagd is – maar wat schiet je ermee op?); je mag de lamp ook niet stukslaan als je deze voor de tweede keer aan ziet (waarna de gevangene die een stukgeslagen lamp ziet maar nog niet zelf een lamp heeft uitgedaan, kan verklaren dat iedereen ondervraagd is – dit geeft een oplossing voor drie gevangenen). En het interval tussen de ondervragingen is variabel: dus de tijd bijhouden heeft geen zin. De ondervragingskamer kan niet worden gezien vanuit de isolatiecellen (daarom zijn het ook *isolatiecellen*), en er is geen geheime klikverbinding tussen de schakelaar van de lamp en de isolatiecellen; je kunt er ook niet de schakelaar horen overgaan ... De gevangenen mogen allemaal verschillende

namen hebben, of met opeenvolgende getallen van 1 tot 100 geïdentificeerd worden, maar dat is allemaal lood om oud ijzer.

4.6 Een beter protocol voor drie gevangenen

De oplossing komt snel naderbij als we ons realiseren dat een protocol niet aan iedere gevangene dezelfde *rol* hoeft toe te bedelen: aangezien de gevangenen het protocol kunnen afspreken zolang ze nog in de kantine zijn, kunnen ze elkaar verschillende taken geven. De telling die Anna bijhoudt in het voorbeeld hiervoor, werkt, zolang iedereen maar weet dat *alleen* Anna dit doet. Dit brengt ons tot het volgende protocol.

Protocol 4: *De gevangenen wijzen onder elkaar een teller aan. Voor de teller geldt: Als je ondervraagd wordt en de lamp is uit doe dan niets; als je ondervraagd wordt en de lamp is aan, doe hem dan uit – houd bij hoe vaak je dat doet; als je de lamp $n-1$ keer hebt uitgedaan, verklaar dan dat alle n gevangenen ondervraagd zijn. Voor de overige gevangenen geldt: Als je ondervraagd wordt en de lamp is uit en je hebt hem nog niet eerder aangedaan, doe hem dan aan; als je ondervraagd wordt en de lamp is uit en je hebt hem al eerder aangedaan, doe dan niets; als je ondervraagd wordt en de lamp is aan, doe dan niets.*

Het zal de lezer duidelijk zijn dat dit protocol inderdaad werkt. Voor drie gevangenen, waarbij Anna als teller is aangewezen, zijn de volgende drie uitvoeringen van protocol 4 succesvol:

1. Bert-Anna-Caroline-Anna
2. Anna-Bert-Caroline-Anna-Bert-Anna-Caroline-Caroline-Bert-Bert-Anna
3. Bert-Anna-Bert-Caroline-Bert-Anna

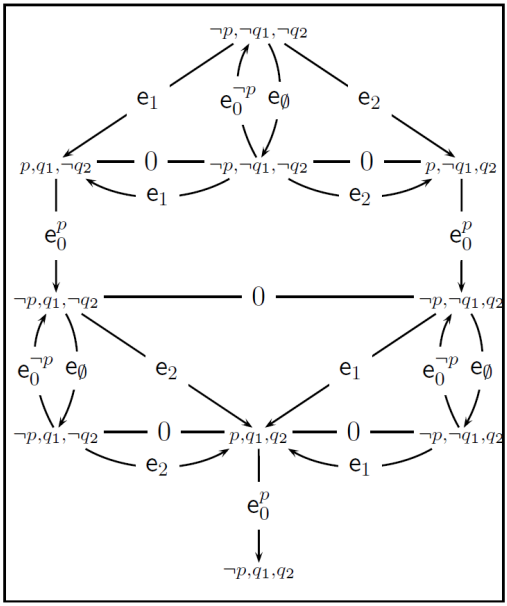
4.7 Uitstapje naar meer formele logica

De stappen uit het nieuwe protocol 4 voor drie gevangenen, en het effect van die stappen afhankelijk van de volgorde van ondervragen, kunnen mooi in logische notatie weergegeven worden. Op deze wijze kun je het totale proces 'doorrekenen'. Dat werkt zo:

- Stel gevangene 0 ('onze' Anna) is de teller. Gevangenen 1 en 2 zijn niet-tellers.

- p staat voor 'het licht is aan'.
- q_1 staat voor 'gevangene 1 heeft het licht aangedaan'.
- q_2 staat voor 'gevangene 2 heeft het licht aangedaan'.
- $\neg p$ staat voor 'niet p', en $p \to q$ staat voor 'p impliceert q'.

gebeurtenis	*beschrijving van de situatie*
$\mathsf{e}_\emptyset$	er gebeurt niets
e_1	p wordt $q_1 \to p$, en q_1 wordt $p \to q_1$
e_2	p wordt $q_2 \to p$, en q_2 wordt $p \to q_2$
$\mathsf{e}_0^{\neg p}$	indien $\neg p$, er gebeurt niets
e_0^{p}	indien p, p wordt onwaar



Figuur 4.2: Doorwerking van protocol 4 voor drie gevangenen

Hoe doen de gebeurtenissen zich voor aan gevangene 0:

- hij kan geen onderscheid maken tussen e_1, e_2, en $\mathsf{e}_\emptyset$
- hij kan e_0^p onderscheiden van de overige gebeurtenissen

- hij kan $\mathsf{e}_0^{\neg p}$ onderscheiden van de overige gebeurtenissen

De doorwerking van het protocol voor 100 gevangenen zou een aanzienlijk gecompliceerder diagram opleveren, maar het principe is niet wezenlijk anders. Elke niet-teller wordt op een gegeven moment ondervraagd, en doet dan (en ook alleen maar dan) de lamp aan. De teller doet, elke keer als zij ondervraagd wordt, de lamp uit. Bij de 99ste keer dat ze de lamp uitdoet, weet ze dat alle 99 mede-gevangenen evenals zijzelf ondervraagd zijn, en kan ze dus met zekerheid verklaren dat iedereen ondervraagd is.

4.8 Oplossing voor 100 gevangenen

Voor 100 gevangenen werkt protocol 4 dus op dezelfde manier als voor drie gevangenen. Je moet wel met het volgende rekening houden:

Eerlijk roosteren

Stel dat Anna de teller is en dat van de 100 gevangenen alleen om de beurt Anna en Bert ondervraagd worden. Dan zal het nooit gebeuren dat Anna kan zeggen dat iedereen ondervraagd is. Maar dan werkt het protocol toch niet? Wel, in dat geval komt de oplossing inderdaad nooit dichterbij. Maar dat is dan omdat nooit alle gevangenen ondervraagd worden! Een voorwaarde voor de terminatie van protocol 4 is de zogenaamde 'fairness' (eerlijkheid) of 'liveness' van het ondervragingsrooster: we hebben dit in het raadsel verwoord als de eis 'We mogen wel aannemen dat op ieder moment iedere gevangene ooit nog eens ondervraagd zal worden.' Onder die aanname zal de teller ooit naar waarheid kunnen beweren dat iedereen ondervraagd is.
We kunnen ons ook nog voorstellen dat eerst alle gevangenen eenmaal na elkaar ondervraagd worden, en daarna tot in de eeuwigheid alleen Anna en Bert om de beurt. Dan is wel iedereen ondervraagd, maar komt de oplossing toch nooit naderbij. Maar ook dan wordt niet voldaan aan 'We mogen wel aannemen dat op ieder moment iedere gevangene ooit nog eens ondervraagd zal worden': er is immers een moment waarna alleen nog Anna en Bert ondervraagd worden. Uit de eis van eerlijk roosteren volgt dat alle gevangenen oneindig vaak ondervraagd zullen worden: Neem een zeker moment. Daarna wordt Anna nog eens ondervraagd. Neem dat moment. Daarna wordt Anna nog eens ondervraagd. Neem dat moment. Enzovoort.

4.9 Varianten

Bert weet het voordat Anna het weet

In uitvoering 1 is iedereen ondervraagd na drie ondervragingen, en weet Anna dat bij de vierde ondervraging. In uitvoering 2 is iedereen eveneens ondervraagd na drie ondervragingen, maar duurt het nog acht ondervragingen voordat Anna, als teller, dat ook weet en bekend kan maken. Kan dit sneller?
In uitvoering 3 weet Bert vóór Anna dat iedereen ondervraagd is, ook al heeft Bert niet de rol van teller. De eerste keer dat hij ondervraagd wordt doet hij de lamp aan. De tweede keer dat hij ondervraagd wordt is de lamp uit en doet hij niets, maar hij weet daarom dat Anna nu ook ondervraagd is en de lamp weer heeft uitgedaan. De derde keer dat hij ondervraagd wordt ziet hij dat de lamp weer aan is: dat kan dus alleen door Caroline gedaan zijn. En hij kan dan naar waarheid verklaren dat iedereen ondervraagd is, *voordat* teller Anna dat weet en dat kan verklaren.
Net zo, weet Caroline in uitvoering 2 bij de tweede keer dat ze ondervraagd wordt, dat iedereen ondervraagd is. De eerste keer dat ze wordt ondervraagd ziet ze dat de lamp aan is: dat moet dus door Bert gedaan zijn. De tweede keer dat ze wordt ondervraagd ziet ze dat de lamp uit is: dat moet dus door Anna gedaan zijn. Dus iedereen is ondervraagd.
De gevangenen kunnen de ondervragingen kennelijk met slimme trucs wellicht beperken, en zich eerder in vrijheid stellen. Als niets bekend is over de tijdsduur tussen ondervragingen, weten we geen slimmere trucs (ook niet voor 100 gevangenen) dan dat iedere niet-teller voor zichzelf bijhoudt hoeveel keer hij het licht uit en aan ziet gaan, net als Bert en Caroline hiervoor.

Eén ondervraging per dag

Hoe lang duurt het gemiddeld voordat Anna bekend kan maken dat iedereen ondervraagd is? Die vraag is zinloos, want de tijdsduur tussen ondervragingen is onbekend. Hoeveel ondervragingen duurt het dan gemiddeld voordat Anna dat bekend kan maken? Die vraag heeft wel zin. Dit hangt natuurlijk af van de verroostering van de ondervragingen door de gevangenisbewaarders. Als de verroostering puur op basis van toeval is ('random') kunnen we dit bepalen. Het antwoord is leuker als we eveneens aannemen dat er iedere dag een ondervraging is.

Puzzel 1: *Stel dat er iedere dag precies één ondervraging is. Hoe lang duurt het gemiddeld voordat de honderd gevangenen vrijgelaten worden?*

Dit duurt dus wel een tijdje. Omdat de gevangenen weten dat er precies

één ondervraging per dag is, kan het echter (gemiddeld) ook veel sneller. We onderzoeken eerst weer de situatie voor drie gevangenen: stel dat teller Anna niet op de eerste dag ondervraagd wordt. Dan weet ze meteen al dat daarom Bert en Caroline ondervraagd moeten zijn, zonder zelf de lamp aan gezien te hebben. Kunnen we dit uitbuiten?

Protocol 5: *Het protocol bestaat uit twee stadia. Het eerste stadium duurt n dagen en verloopt als volgt. De gevangene die voor het eerst twee keer ondervraagd wordt, doet het licht aan. Stel dit is op dag m. Op dag n van het eerste stadium: als het licht uit is verklaart de gevangene die dan ondervraagd wordt dat iedereen ondervraagd is; anders doet die gevangene het licht uit. Het tweede stadium verloopt als volgt; er zijn nu drie verschillende rollen.*
(i) De teller is de gevangene die in stadium één voor het eerst twee keer ondervraagd is. Voor de teller geldt: Als je ondervraagd wordt en de lamp is uit doe dan niets; als je ondervraagd wordt en de lamp is aan, doe hem dan uit – houdt bij hoe vaak je dat doet; als je de lamp $n - m$ keer hebt uitgedaan, verklaar dan dat alle n gevangenen ondervraagd zijn.
(ii) De gevangenen die in het eerste stadium het licht alleen uit gezien hebben, doen niets in het tweede stadium.
(iii) Voor de overige gevangenen geldt: Als je ondervraagd wordt en de lamp is uit en je hebt hem nog niet eerder aangedaan, doe hem dan aan; als je ondervraagd wordt en de lamp is uit en je hebt hem al eerder aangedaan, doe dan niets; als je ondervraagd wordt en de lamp is aan, doe dan niets.

Als m maar groter dan 2 is, levert dit protocol tijdwinst op ten opzichte van protocol 4. Het gemiddelde aantal dagen voordat iemand voor de tweede keer ondervraagd wordt is 13. Dit betekent dat de hiermee aangewezen teller al van 11 andere gevangenen weet dat ze ondervraagd zijn voordat hij met tellen begint. (En die 11 doen in het tweede stadium niets meer!) Hij hoeft dus nog maar tot 88 te tellen, in plaats van tot 99. De tijdwinst die hiermee bereikt wordt is een jaar of vier. Dit komt voornamelijk omdat de teller 11 keer minder 100 dagen hoeft te wachten totdat hij het licht weer uit kan doen.

Puzzel 2: *Stel dat niet bekend is of het licht in eerste instantie aan of uit is. En, als voorheen, niets is bekend over de tijdsduur tussen de ondervragingen. Wat is dan een protocol om het probleem op te lossen?*

4.10 Herkomst

Op `http://domino.watson.ibm.com/Comm/wwwr_ponder.nsf/challenges/July2002.html` van IBM Research wordt in 2002 vermeld dat "this puzzle has been making the rounds of Hungarian mathematicians' parties" in een versie voor 23 gevangenen. Het raadsel deed de ronde in de VS vanaf 2001. William Wu, destijds van Stanford University, was vanaf dit vermoedelijke begin bij de ontwikkelingen en oplossingen van dit raadsel betrokken, zie `http://wuriddles.com`. Het raadsel wordt uitvoerig behandeld in het tijdschrift *Mathematical Intelligencer* door Dehaye, Ford, en Segerman (2003), in 'Mathematical Puzzles: A Connoisseur's Collection' door Winkler (2004), in de *Nieuwe Wiskrant* door Van Ditmarsch (2007), en in 'One hundred prisoners and a lightbulb – logic and computation' door Van Ditmarsch, van Eijck, en Wu (2010).
Onder de aanname dat iedere dag een ondervraging plaatsvindt, valt het protocol nog op allerlei manieren te optimaliseren. Het is niet bekend wat de kleinste verwachtingswaarde is voor de terminatie van een protocol om het probleem op te lossen, het huidige 'record' staat op ongeveer negen jaar. Dat is al een stevige verbetering ten opzichte van de 28 en een half jaar die als oplossing uit puzzel 1 rolt! De snellere algoritmes onderscheiden nog meer dan twee rollen voor de gevangenen, en meerdere fasen in het protocol waarin gevangenen van rol kunnen wisselen (net als hiervoor, in protocol 5, maar dan nog ingewikkelder). Zie voor details het antwoord op puzzel 2, hieronder, en (Van Ditmarsch et al., 2010).

Antwoorden op de puzzels

Antwoord op puzzel 1

Eerst moet een niet-teller ondervraagd worden, om het licht aan te doen. De kans hierop is 99/100. Dan moet de teller ondervraagd worden, om het licht weer uit te doen. De kans daarop is 1/100. Daarna moet een niet-teller die nog niet ondervraagd is, het licht weer aandoen. Die kans is nu 98/100. Daarna de teller weer, dat blijft altijd 1/100. Enzovoorts. De verwachtingswaarde in dagen is net andersom. Het verwachte aantal dagen voordat de eerste niet-teller ondervraagd wordt is 100/99, ongeveer een dag dus, de verwachting dat daarna de teller ondervraagd wordt is 100/1, honderd dagen dus, enzovoorts. De gemiddelde tijd die verstrijkt

voordat de teller kan verklaren dat iedereen ondervraagd is, is daarom:

$$\frac{100}{99} + 100 + \frac{100}{98} + 100 + \ldots + \frac{100}{2} + 100 + \frac{100}{1} + 100.$$

De deelreeks $\frac{100}{99} + \frac{100}{98} + \ldots + \frac{100}{2} + \frac{100}{1}$ telt op tot ongeveer 518 dagen en voor de rest hebben we 99 keer 100 dagen: 9.900 dagen. Dit sommeert tot 10.418 dagen wat ruim $28,5$ jaar is. Dat valt toch wat tegen. Het is maar goed dat niet alle gevangenen wiskundigen zijn, dan waren ze er vast niet eens aan begonnen.

Antwoord op puzzel 2

Protocol 4 werkt nu niet, en ook niet als we de teller één verder door laten tellen. Bijvoorbeeld, voor drie gevangenen, kunnen we de volgende uitvoeringen niet van elkaar onderscheiden:

1. (licht aan)-Anna-Caroline-Anna-Bert-Anna
2. (licht uit)-Bert-Anna-Caroline-Anna-Bert-Anna

Als Anna slechts tot twee telt, verklaart ze bij 1 na de tweede keer ondervraagd te zijn ten onrechte dat iedereen ondervraagd is. Als we Anna daarentegen tot drie laten tellen, dan komt ze bij 2 nooit aan die verklaring toe, hoe vaak ook Anna, Bert, en Caroline ondervraagd zullen blijven worden in de toekomst: zowel Bert als Caroline hebben het licht precies één keer aangedaan, en Anna doet daarna dat licht weer uit, en daar blijft het bij. We kunnen protocol 4 echter aanpassen - de truc is dat we moeten compenseren voor de onzekerheid dat Anna één keer teveel het licht moet uitdoen, omdat het initieel aan had kunnen zijn.

Protocol 6: *De gevangenen wijzen onder elkaar een teller aan. Voor de teller geldt: Als je ondervraagd wordt en de lamp is uit doe dan niets; als je ondervraagd wordt en de lamp is aan, doe hem dan uit – houd bij hoe vaak je dat doet; als je de lamp* $2n-2$ *keer hebt uitgedaan, verklaar dan dat alle* n *gevangenen ondervraagd zijn. Voor de overige gevangenen geldt: Als je ondervraagd wordt en de lamp is uit en je hebt hem nog niet twee keer aangedaan, doe hem dan aan; als je ondervraagd wordt en de lamp is uit en je hebt hem al twee keer eerder aangedaan, doe dan niets; als je ondervraagd wordt en de lamp is aan, doe dan niets.*

We leggen uit waarom dit werkt voor het geval van drie gevangenen. Er moet dan vier keer ondervraagd worden. In het ergste geval was het licht aan het begin aan, 1; heeft een van Bert of Caroline het licht al twee keer aangedaan, stel dit is Bert, 2; en heeft Caroline het licht pas één keer aangedaan, 1. Anna hoeft dus niet te wachten totdat Caroline het licht twee

keer aangedaan heeft. Maar drie keer is te weinig: want dat krijgen we al als het licht initieel aan is en Bert het twee keer heeft aangedaan, zonder dat Caroline ooit ondervraagd is.
In het algemeen hoeft de teller niet te wachten totdat de laatste niet-teller die het licht nog niet twee keer heeft aangedaan, dat voor de tweede keer heeft gedaan. Maar alle overige niet-tellers moeten dat wel doen. We krijgen dus: $1 + 2(n-2) + 1 = 2n - 2$. En $2n - 3$ is echt niet genoeg: want dat bereiken we al als iedereen behalve één niet-teller het licht al twee keer heeft aangedaan, en het licht initieel aan was.

Referenties

Dehaye, P., D. Ford, en H. Segerman (2003). One hundred prisoners and a lightbulb. *Mathematical Intelligencer 25(4)*, 5361.

van Ditmarsch, H. (2007). Honderd gevangenen en een gloeilamp. *Nieuwe Wiskrant 27(1)*, 15–18.

van Ditmarsch, H., J. van Eijck, en W. Wu (2010). One hundred prisoners and a lightbulb logic and computation. In F. Lin, U. Sattler, en M. Truszczynski (Eds.), *Proceedings of KR 2010 Toronto*, pp. 90100.

van Ditmarsch, H. en B. Kooi (2013). *Honderd gevangenen en een gloeilamp.* Amsterdam: Epsilon Uitgaven (Epsilon # 73).

Winkler, P. (2004). *Mathematical Puzzles: A Connoisseurs Collection.* Wellesley, MA : Peters.

5 Zoeken naar het Higgs deeltje

Wouter Verkerke

FOM/Nikhef

Voorwoord

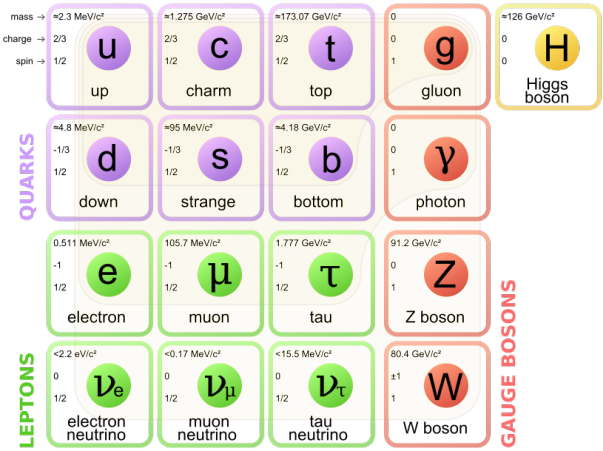
In 2012 is het bestaan van het Higgs deeltje aangetoond door middel van een statistische analyse van proton-proton botsingen die werden geproduceerd door de 'Large Hadron Collider' op het CERN laboratorium in Genève[1]. In deze zoektocht zijn biljoenen botsingen geanalyseerd die vele petabytes aan data hebben opgeleverd, zodat de moderne deeltjesfysica een echte 'big data' onderneming is geworden. In de bewijsvorming zijn statistische analyse technieken essentieel geweest en ik zal de basis van de technieken die zijn gebruikt in het bewijs voor het Higgs deeltje toelichten, en ook enkele van de geavanceerde technieken die zijn gebruikt om deze spreekwoordelijke speld in de hooiberg te vinden illustreren.

5.1 De natuurkunde van elementaire deeltjes

5.1.1 Het Standaard Model van elementaire deeltjes

Het doel van elementaire deeltjes fysica is om een beschrijving van de natuur te maken in termen van elementaire bouwstenen van materie, en de krachten die op deze deeltjes kunnen werken. Sinds de ontdekking van het atoom in de 19e eeuw, en het bijbehorende periodieke systeem van elementen (de atomen), is grote vooruitgang geboekt in het beschrijven van atomen in kleinere elementaire deeltjes, gedreven door voortschrijdende versneller technologie die effectief fungeert als een steeds krachtigere microscoop. In ons huidig begrip van de natuur zijn protonen en neutronen waaruit de atoomkernen zijn opgebouwd geen elementaire deeltjes, maar

zijn opgebouwd uit zogenaamde *up-quarks* en *down-quarks*. Tezamen met de electronen en het onzichtbare neutrino vormen zij de bouwstenen van alle zichtbare materie. Daarnaast blijkt dat er van deze bouwstenen nog twee andere 'generaties' bestaan waaruit exotische materie kan worden gebouwde, zodat er in totaal 12 bouwstenen voor materie bestaan, voor zover wij nu weten. Een overzicht van deze bouwstenen is weergegeven in Fig. 1.



Figuur 5.1: De bouwstenen van het Standaard Model van elementaire deeltjes.

Een van de grote doorbraken in de beschrijving van de natuur op het niveau uit de jaren vijftig is de ontwikkeling van zogenaamde 'ijktheorieën', die ons in staat stellen om drie van de vier de bekende fundamentele natuurkrachten (elektromagnetisme, de zwakke kernkracht, en de sterke kernkracht) in te beschrijven door middel van de uitwisseling van zogenaamde boodschapperdeeltjes. Alhoewel de eigenschappen van deze drie krachten erg verschillend zijn, alsmede de eigenschappen van de bijbehorende deeltjes blijkt dat deze eigenschappen, zoals wij ze in natuur vinden, in de theorie kunnen worden gerealiseerd, door te eisen dat de theorie invariant is onder een set van symmetriegroepen (SU3, SU2 en U1). Deze wonderbaarlijke eigenschap van deze ijktheorieën, waar de ons bekende fundamentele natuur krachten 'automatisch' geïntroduceerd worden door een geschikte keuze van symmetrieën maakt dat deze klasse van theoretische modellen favoriet zijn als beschrijving van de natuur. Met de keuze van de bovengenoemde

symmetrie SU(3)xSU(2)xU(1) ontstaan in de theorie het foton als boodschapper van elektromagnetisme, de W en Z bosonen, als boodschappers van de zwakke kernkracht, en de gluonen in 8 'kleuren', allen waargenomen door moderne experimenten.

Een belangrijke tekortkoming de ijktheorieën is dat ze alleen massaloze elementaire deeltjes kunnen beschrijven, waardoor zie in hun standaard vorm ongeschikt zijn om de natuur beschrijven, waarin zowel de bouwsteendeeltjes als boodschapper-deeltjes wel aantoonbaar een massa hebben. Een oplossing van dit probleem is voorgesteld door Higgs, Brout en Englert in de jaren zestig[2]: in hun model is massa geen intrinsieke eigenschap van deeltjes, maar een dynamische eigenschap die word 'opgelopen' doordat deeltjes interactie hebben met een alomtegenwoordig *Higgs veld*, dat zich ook als een deeltje kan manifesteren, het zogenaamde Higgs deeltje. De huidige alomvattende theorie van de natuur, het Standaard Model[3], is een ijktheorie die gebruikt maakt van het Higgs mechanisme om deeltjes om massa te geven. Deze theorie is spectaculair succesvol gebleken in het beschrijven van de waargenomen resultaten van meer dan 40 jaar aan experimentele inspanningen, met uitzondering van het het zogenaamde Higgs deeltje zelf, dat pas in 2012 na zeer grote inspanningen van duizenden experimentele fysici wereldwijd is waargenomen.

5.2 Kwantummechanika en kansrekening

De natuurkundige processen waar we naar kijken zijn op een zodanig kleine schaal dat alledaagse wetten van de natuur, zoals de bewegingswetten van Newton, niet meer van toepassing zijn. Door de enorm kleine afstandschalen van de processen, tot een miljoen keer kleiner dan een atoomkern, en de zeer hoge energieën van de botsingen in de experimenten zijn de wetten van de kwantummechanica en de speciale relativiteitstheorie leidend[1]. Een van de consequenties van kwantummechanica is dat de natuurkundige processen in een experiment, bijvoorbeeld een proton-proton botsing bij de Large Hadron Collider, geen deterministische uitkomst hebben zoals bij klassieke mechanica, maar uitsluitend beschreven kunnen worden door kansverdelingen over vele mogelijk uitkomst, waarbij nooit met zekerheid vastgesteld kan worden wat in een bepaalde botsing gebeurd is.

De uitkomsten van moderne deeltjesfysica experimenten kunnen dus ook

[1] Het gangbare wiskundig framework waarin zowel de wetten van de kwantummechanica als de relativiteitstheorie verwerkt is de kwantumveldentheorie, waarvan de ijktheorieën een speciale subset vormen.

alleen beschreven worden in kansmodellen, waardoor statische data analyse een hoofdrol speelt in experimentele bewijsvoering. Alvorens met kansmodellen voor de Large Hadron Collider data aan de slag te gaan zal ik eerst aan de hand van eenvoudig fictief experiment de basisconcepten invoeren.

5.3 Knikkers in een ton

Een van de eenvoudigste experimenten die je kan verzinnen die berusten op een toevalsproces is het willekeurig trekken van knikkers uit een ton. In zo'n experiment is een ton gevuld met N_W witte knikkers en N_Z zwarte knikkers. Een meting bestaat uit het (blind) trekken van een knikker uit de ton en bepalen van de kleur van de knikker. Een vraagstelling die gelijkenis heeft met de Higgs ontdekking is dan de vraag of het bestaan van zwarte knikkers wel of niet aangetoond kan worden op basis van het trekken van een eindig aantal knikkers uit de ton.

5.3.1 Hoe bewijs je dat zwarte knikkers niet bestaan?

Deze vraag is in zijn vorm het tegenovergestelde van die van het Higgs deeltje, waarvan we het bestaan juist *wel* willen aantonen, maar heeft in zijn statistische analyse grote overeenkomsten. Een vraagstelling over het (niet) bestaan van een deeltje of knikker begint met het opstellen van twee strijdige hypotheses over de natuur. Bijvoorbeeld: hypothese H_0: er zijn geen zwarte knikkers en duizend witte knikkers, hypothese H_1: er zijn 200 zwarte knikkers en 800 witte knikkers. Voor beide hypotheses kunnen we een kansmodel opstellen dat de mogelijke experimentele uitkomsten beschrijft. Bijvoorbeeld voor de voor de kleur van een enkele getrokken knikker geldt:

$$H_0 \quad : \quad p(zwart) = 0 \; ; \; p(wit) = 1. \tag{5.1}$$

$$H_1 \quad : \quad p(zwart) = 0.2 \; ; \; p(wit) = 0.8. \tag{5.2}$$

Het zal duidelijk zijn dat het trekken van een enkele knikker in de veel gevallen onvoldoende zal zijn om uitsluitsel te geven over het (niet) bestaan van zwarte knikkers. Een realistisch experiment zal daarom n knikkers trekken (en weer terugleggen na elke trekking). Op basis van Vgl 5.1 en 5.2 kan een kans worden toegekend aan elke mogelijke trekkingsserie, door het product van de kansen te nemen, bijvoorbeeld

$$
\begin{aligned}
p(wwww)_{H_0} &= 1 * 1 * 1 * 1 = 1 \\
p(wwww)_{H_1} &= 0.8 * 0.8 * 0.8 * 0.8 = \frac{4096}{10000} \\
p(wzwz)_{H_0} &= 1 * 0 * 1 * 0 = 0 \\
p(wzwz)_{H_1} &= 0.8 * 0.1 * 0.8 * 0.1 = \frac{64}{10000}
\end{aligned}
$$

Het is evident dat het apart berekenen van de kansen voor elke trekkingsserie weinig zinvol is aangezien alle permutaties van trekkingsuitkomsten dezelfde kans hebben (bv p(wzwz)=p(wwzz)). Een betere manier om de trekkingsuitslagen te classificeren is bijvoorbeeld om het aantal witte knikkers in een trekking als gemeten grootheid te nemen. De kans op k witte knikkers in een trekking van n knikkers is dan

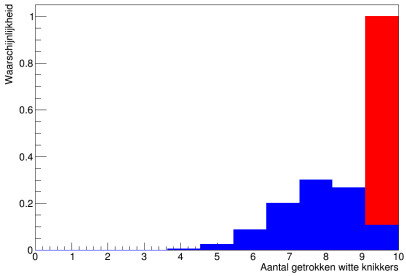
$$
p(n_w) = \frac{n!}{k!(n-k)!} \cdot p_w^k \cdot (1 - p_w)^{n-k} \tag{5.3}
$$

In deze uitdrukking, bekend als de binomiaalverdeling, geeft de term $p_w^k \cdot (1 - p_w)^{n-k}$ de kans op k witte knikkers en $n - k$ zwarte knikkers in de trekking, en de term $n!/k!(n-k)!$ het aantal permutaties in de trekkingsvolgorde dat mogelijk is voor deze trekkingsuitslag. Met behulp van Vgl. 5.3 kunnen we nu de kans berekenen voor elke mogelijke trekkingsuitslag onder hypothese H_0 en hypothesis H_1. In Figuur 5.2 is een grafiek voor de kansen van de mogelijke uitkomsten onder hypotheses H_0 en H_1 voor trekkingen van 10 knikkers.

Stel dat we nu ons experiment doen, en we trekken 10 witte knikkers, wat weten we dan? We de kans berekening op deze trekkingsuitslag onder beide hypotheses:

$$
\begin{aligned}
p(n_w = 10|H_0) &= 1 \cdot 1 \cdot 1 = 1 \\
p(n_w = 10|H_1) &= 1 \cdot 1 \cdot (0.8)^{10} \approx 0.11
\end{aligned}
$$

Is dit voldoende om te concluderen dat zwarte knikkers niet bestaan? Nee, want de kans op deze uitslag on de de hypothese H_1, waar zwarte knikkers wel in voorkomen, is immers nog 10%, en op zijn best een aanwijzing. Uiteraard kunnen met deze statistische analyse, nooit met *volledige* zekerheid zeggen dat zwarte knikkers niet bestaan, maar kunnen wel aantonen,



Figuur 5.2: Binomiaalverdeling van waarschijnlijkheden voor het aantal witte knikkers in een trekking van 10 knikkers uit een ton onder hypothese dat 80% van de knikkers in de ton wit zijn (blauw) of de hypothese dat 100% van de knikkers in de ton wit zijn (rood).

als we meer data hebben, dat de kansen onder hypothese H_1 zo klein zijn, dat we H_0 als de nieuwe wetenschappelijke consensus kunnen aannemen. Bijvoorbeeld, als in een herhaald experiment, 70 witte knikkers zouden trekken, dan worden de kansen

$$p(n_w = 70|H_0) = 1 \quad (5.4)$$
$$p(n_w = 70|H_1) = (0.8)^{10} \approx 1.6 \cdot 10^{-7} \quad (5.5)$$

en is deze uitkomst zo onwaarschijnlijk onder hypothese H_1 dat dit een verantwoorde onderbouwing kan zijn voor de claim van de ontdekking van hypothese H_0, dat wil zeggen, we claimen de ontdekking dat knikkers alleen wit zijn. Er zijn geen harde regels bij welke waarschijnlijke kansen een ontdekkingsclaim wetenschappelijk geaccepteerd is.

Binnen elementaire deeltjesfysica word algemeen een kans van kleiner dan $2.87 \cdot 10^{-7}$ voor hypothese H_0 gepaard een kans van $O(1)$ voor hypothese H_1 geaccepteerd als bewijs[2] voor hypothese H_1, maar consensus in de 'a priori' geloofwaardigheid van de theorie H_1 speelt ook een belangrijke rol. Zo is het bestaan van het Higgs deeltje, dat theoretisch goed gemotiveerd

[2]De gekozen waarde van ca $2.87 \cdot 10^{-7}$ komt overeen met een kans op een fluctuatie van $+5$ of groter in een normale (Gaussische) verdeling, en wordt om die reden vaak aangeduid als 5σ.

was, geaccepteerd op basis van een kans van $2 \cdot 10^{-7}$ voor de H_0 hypothese (dwz de kans dat de geobserveerde data het product is van toeval in een natuur zonder Higgs), terwijl het statistisch veel sterkere bewijs voor neutrinos die sneller vlogen dan de snelheid van het licht *niet* geaccepteerd werd omdat er zo veel theoretische haken en ogen aan modellen met superluminale neutrinos zitten dat het moeilijk was om dit resultaat te accepteren. De meting van superluminale neutrinos is later ook teruggetrokken nadat een bekabelingsfout in het experiment was gevonden.

5.3.2 Bayesiaanse kansrekening

Het is mogelijk de 'a priori' geloofwaardigheid van de hypotheses expliciet mee te nemen in de kansrekening en zo een statement te formuleren over de geloofwaardigheid van de een hypothese gegeven *na* de meeting, met inachtneming van zowel deze *a priori* geloofwaardigheid als het resultaat van de meting. Deze formulering komt neer op het inverteren van de conditionaliteit van de kans zoals we hem tot nu toe hebben uitgedrukt:

$$p(n_w = 70|H_{0,1}) \rightarrow p(H_{0,1}|n_w = 70)$$

De relatie tussen deze twee van zulke conditionele waarschijnlijkheden word algemeen gegeven door de stelling van Bayes[4]

$$p(A|B) = \frac{p(B|A)P(A)}{P(B)} \tag{5.6}$$

Als we dan voor A de observatie $n_w = 70$ nemen en voor B de hypothese H_0 dan krijgen we

$$p(H_0|n_w = 70) = \frac{p(n_w = 70|H_0)P(H_0)}{P(n_w = 70)} \tag{5.7}$$

De betekenis van $P(n_w = 70)$ in de deler van vgl 5.7 lijkt in eerste instantie onduidelijk, maar is eenvoudig de som van kansen op deze uitkomst onder *alle* hypotheses die beschouwd worden. In het huidige voorbeeld zijn dat H_0 en H_1 zodat

$$P(n_w = 70) = P(n_w = 70|H_0)P(H_0) + P(n_w = 70|H_1)P(H_1) \tag{5.8}$$

Het invullen van Vgl 5.8 in Vgl 5.7 geeft dan

$$p(H_0|n_w = 70) = \frac{p(n_w = 70|H_0)P(H_0)}{P(n_w = 70|H_0)P(H_0) + P(n_w = 70|H_1)P(H_1)} \tag{5.9}$$

In deze vergelijking representeren de kansen $p(H_0)$ en $p(H_1) = 1 - p(H_0)$ het 'a priori' geloof in beide hypotheses.

Stel dat ons geloof in beide hypotheses voor het knikkerexperiment even groot was, dwz. $p(H_0) = p(H_1) = 0.5$), dan kunnen we nu berekenen dat ons geloof in deze hypothese *na* het experimentele resultaat is:

$$\begin{aligned} p(H_0|n_w = 70) &= \frac{p(n_w = 70|H_0)P(H_0)}{P(n_w = 70|H_0)P(H_0) + P(n_w = 70|H_1)P(H_1)} \\ &= \frac{1 \cdot 0.5}{1 \cdot 0.5 + 1.6 \cdot 10^{-7} \cdot 0.5} \approx 0.99999984 \\ p(H_1|n_w = 70) &= \frac{p(n_w = 70|H_1)P(H_1)}{P(n_w = 70|H_0)P(H_0) + P(n_w = 70|H_1)P(H_1)} \\ &= \frac{1.6 \cdot 10^{-7} \cdot 0.5}{1 \cdot 0.5 + 1.6 \cdot 10^{-7} \cdot 0.5} \approx 0.00000016 \end{aligned}$$

Een ander handige manier om dezelfde informatie te representeren is om naar verhouding van het geloof in beide hypotheses te kijken voor en na het experiment,

$$\frac{p(H_0|n_w = 70)}{p(H_1|n_w = 70)} = \frac{p(n_w = 70|H_0)P(H_0)}{p(n_w = 70|H_1)P(H_1)} \tag{5.10}$$

waarbij de deler van vgl 5.9 handig wegvalt. Voor ons knikker voorbeeld-experiment drukt dit uit dat voor een scenario waar ons 'geloof' in de hypotheses H_0 en H_1 voor het knikkerexperiment gelijk waren (1:1), dat deze na de experimentele uitkomst van 70 getrokken witte knikkers 625000:1 is in het voordeel van H_0 (waarin zwarte knikkers niet bestaan).

Deze Bayesiaanse vorm van kansrekeningen wordt in veel takken van wetenschap gebruikt, echter minder in de elementaire deeltjesfysica omdat het is moeilijk zinvolle numerieke kansen toe te dichten aan een 'a priori' geloof in fundamentele natuurkundige theorien zonder subjectieve persoonlijke argumenten te introduceren. Bijvoorbeeld, het antwoord op de vraag 'Wat is de kans dat het Higgs deeltje bestaat' in is moeilijk procenten uit te drukken. Derhalve wordt de experimentele bewijsvoering in de elementaire deeltjes fysica meestal gedaan op basis van de kansen op gemeten uitkomsten $p(H_0|n_w = 70)$ en $p(H_1|n_w = 70)$. De uiteindelijke beslissing of een experimentele uitkomst kwalificeert als een 'ontdekking' is dan nog steeds onderhevig aan variabele criteria die afhangen van het geloof in de geteste theorie, maar de vraag een concrete set meetresultaten meting voldoende of onvoldoende bewijs is voor een ontdekkingsclaim is in de

praktijk makkelijker te beantwoorden dan er getalsmatige 'a priori' kans aan toe te kennen aan de de geteste theorieën.

5.4 Kansberekeningen voor botsingen bij de Large Hadron Collider

Experimenten met proton-proton botsingen bij de Large Hadron Collider in Genève zijn ingewikkelder dan het trekken van knikkers uit een ton, maar in zijn eenvoudigste vorm wel erg vergelijkbaar.

5.4.1 Experimenteren bij de Large Hadron Collider

De experimentele setup van de Large Hadron Collider bestaat uit een ringvormige deeltjesversneller die protonen in pakketjes versnelt die beide richten draaien. Op vier plekken in de ring botsen deze pakketjes protonen op elkaar waar gemiddeld tussen de 10 en 50 protonen op elkaar botsen met zeer hoge energie per kruising van twee pakketjes. De energie van deze botsingen is zodanig hoog dat nieuwe deeltjes kunnen worden gemaakt uit energie, waar Einsteins bekende vergelijking $E = mc^2$ aangeeft hoe energie E benodigd is om een deeltje met massa m te maken (de constante c hier is de lichtsnelheid). De energie van de botsingen is zodanig hoog dat bij een volledige voltreffer er voldoende energie in een proton-proton botsing is om nieuwe deeltjes met een gezamenlijke massa van maximaal 7000 maal de proton massa te maken. Omdat protonen zelf geen elementaire deeltjes zijn, maar samengestelde deeltjes, is een proton-proton botsing in de praktijk een botsing tussen een onderdeel van een protonen (een quark of een gluon) met een onderdeel van een ander proton, waardoor effectieve botsingsenergie per botsing zeer verschilt, en in de praktijk vaak veel lager is dan het maximaal haalbare van 7000 maal de proton massa.

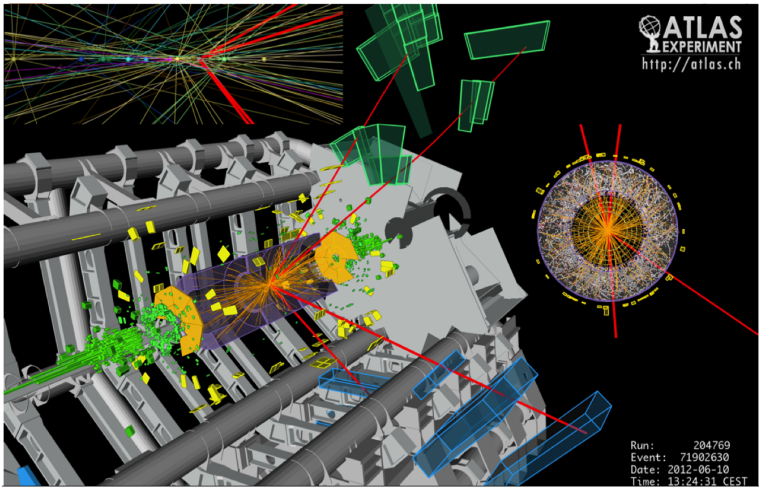
Op de vier vaste plaatsen op de ring waar deze botsingen plaatsvinden zijn grote detectie instrumenten opgesteld die de botsingsproducten waarnemen. Deze producten zijn niet alleen de restanten van de originele protonen van de botsingen, maar ook honderden tot duizenden nieuwe deeltjes die uit de botsingsenergie zijn produceert volgende de wetten van de relativistische quantummechanica. Van al deze deeltjes die vrijkomen wordt - voor zover mogelijk - die richting bepaalt, alsmede de impuls en energie, en identiteit (aan de hand van de waargenomen massa).

De overgrote meerderheid van deze botsingen is oninteressant, maar heel af en toe wordt er in deze botsingen een enkel zeer zwaar exotisch deeltje produceert (bijvoorbeeld het zojuist ontdekte Higgs deeltje), in plaats van oninteressante lichte deeltjes. Het is niet mogelijk het botsingsproces te 'sturen' om zoveel mogelijk interessante botsingen op te leveren, maar door de botsingen zeer vaak uit te voeren, en zeer stringent te selecteren kan toch een grote hoeveelheid interessante botsingen worden verzameld. Zo is in de jaren 2011 en 2012 25 miljoen keer per seconde gebotst gedurende ca. 100 dagen per jaar, resulterend in $25 \cdot 10^6 \cdot 3600 \cdot 24 \cdot 100 \approx 2 \cdot 10^{14}$ botsingen per jaar. Een geavanceerd drietraps *real-time* selectie systeem, gebaseerd op speciaal hier voortgebouwde snelle elektronica, alsmede grote computer farms, beslist om 4999 van elke 5000 botsingen meteen weg te gooien, en voor de overgebleven botsingen de gehele detector uit te lezen voor verdere analyse. Op deze wijze zijn per jaar ca. 10 miljard botsingen bewaard voor verdere analyse, goed voor 10 petabyte aan data per jaar. De 'big data' analyse uitdaging is dan om zeldzame botsingen met bijzonder signatuur te vinden, waarvan de verwachte aantallen uiteen lopen een van enkele tientallen to enkele honderden in de gehele dataset. Figuur 5.3 laat een computerreconstructie zien van een waargenomen 'interessante' proton-proton botsing in de ATLAS detector aan de LHC.

De kwantummechanische natuur van botsingsproces zorgt hier ook weer voor een extra uitdaging in de analyse van de data: het is fundamenteel onweetbaar of een individuele geselecteerde botsing met een interessant signatuur daadwerkelijk het gevolg is van de productie en verval van een exotisch zwaar deeltje (zoals het Higgs deeltje), of toch een botsing is waarin uitsluitend oninteressante processen plaatsvinden dat 'toevallig' erop lijkt. Daardoor is weer een hoofdrol weggelegd voor een statistische analyse van alle bewaarde botsingen om bewijs te formuleren over het bestaat van nieuwe deeltjes.

5.4.2 Een voorbeeld van statistische analyse: Higgs productie en verval in vier leptonen

Het bepalen van een signatuur van een proton-proton botsing dat typerend is voor de productie van een nieuw soort deeltje is in eerste instantie een taak voor getrainde experimentele fysici, hierin bijgestaan door uitgebreide computersimulatie van zowel natuurkundige processen als van de detector. Van het Higgs deeltje is bekend dat het zeer kort leeft, als het gemaakt wordt in een proton-proton botsing, en dat dus alleen de vervalsproducten kunnen worden waargenomen door de detector. Daarnaast kan het Higgs



Figuur 5.3: Computerreconstructie van een waargenomen proton-proton botsing in de ATLAS detector aan de LHC versneller. De vier rode sporen zijn zogenaamde *leptonen*, zeldzame geproduceerde deeltjes, die wanneer ze in meervoud voorkomen een aanwijzing kunnen zijn dat er een Higgs deeltjes is produceert in de botsing, dat vervolgens is vervallen in deze leptonen.

deeltje op vele verschillende manieren vervallen. Sommige vervalsmogelijken zijn goed te herkennen, maar uiterst zeldzaam (bijvoorbeeld naar vier leptonen, via een twee Z bosonen), andere zijn minder zeldzaam maar vrijwel onmogelijk te onderscheiden van andere botsingen (Higgs verval naar twee b quarks).

Het eerste type verval – een zeer zeldzaam process, maar met een relative makkelijk te onderscheiden signatuur – is een geschikt voorbeeld om de eenvoudigste vorm van statistische analyse in de Higgs zoektocht te illustreren. Voor het zoeken naar Higgs deeltjes die vervallen in 4 *leptonen* is het afdoende de botsingen selecteren die aan slechts een criterium te hoeven voldoen: dat deze 4 leptonen gevonden zijn (met een aantal aanvullende eisen op hun richting en snelheid). Zodoende kan de statistische analyse voor de zoektocht in dit kanaal vereenvoudigd worden tot een 'tel-experiment' analoog aan het experiment met de knikkers in de ton. Als het Higgs deeltje *niet* bestaat (hypothese H_0) zal een klein aantal van deze botsingen met vier leptonen worden waargenomen door toevalsprocessen

in proton-proton botsingen. Als het Higgs deeltje *wel* bestaat (H_1), zullen deze toevalsprocessen ook plaats vinden, maar zullen er daarnaast ook botsingen voorkomen waarin Higgs deeltjes vervallen in vier leptonen: het aantal waargenomen botsingen met deze signatuur zal daardoor gemiddeld groter zijn.

Om de kans op een behaald experimenteel resultaat onder twee hypotheses (wel/geen Higgs) te kunnen bereken, moeten we eerst kansmodellen opstellen die dat de kans op elke mogelijke teluitkomstsen voor beide hypotheses formuleren. Voor het experiment met de knikkers in de ton was eenvoudig af te leiden dat het kansmodel een zogenaamde binomiaalverderling volgde, voor tel-experimenten bij de LHC is de verdeling anders, maar is ook eenvoudig af te leiden.

Bij een binomiale verdeling is het aantal trekkingen vastgelegd in het experiment, zoals bij het knikkerexperiment. Bij botsingen in de LHC is het totaal aantal botsingen *niet* vast gelegd, dit is een toevalsprocess, maar de *tijdsperiode* waarover deze geregistreerd worden is wel vast. Het is dus alsof we het aantal clicks van een Geigerteller bij een radioactieve bron proberen te modelleren, voor een vastgesteld tijdsinterval.

De vraag is dus: stel dat we in het tijdsinterval *gemiddeld* k clicks verwachten, wat is dan de kans om daadwerkelijk r clicks te meten? We kunnen dit process benaderen door het tijdsinterval te verdelen in n gelijke intervals, waarvan we aannemen dat die elk of één click *of* geen click bevatten. De verdeling van intervals die wel click bevatten in een 'vaste trekking' van n intervals volgt dan weer de bekende binomiaalverdeling, waar de kans om een click te zien in één tijdinterval dan k/n is, zodat

$$p(r|k/n, n) = \frac{k^r}{n^r}(1 - \frac{k}{n})^{n-r}\frac{n!}{r!(n-r)!} \tag{5.11}$$

De aanname dat elke tijdinterval maximaal een click bevat is uiteraard alleen correct in de limiet $n \to \infty$ te nemen waardoor de tijdsduur van elke interval naar nul gaat. Gebruik makend van

$$\lim_{n\to\infty} \frac{n!}{r!(n-r)!} = n^r \tag{5.12}$$

$$\lim_{n\to\infty} (1 - \frac{\lambda}{n})^{n-r} = e^{-\lambda} \tag{5.13}$$

vinden we dan

$$p(r|\lambda) = \frac{e^{-\lambda}\lambda^r}{r!} \tag{5.14}$$

bekend als de Poisson verdeling, genoemd naar Simeon de Poisson die 1837 onderzoek deed naar de frequentie van gerechtelijke dwalingen in het Franse justitiële systeem[5]. De Poisson verdeling heeft slechts één parameter: het gemiddelde verwachte aantal botsingen in het tijdsinterval.

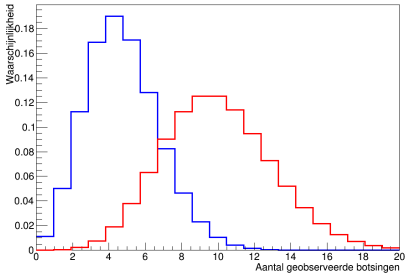
Met deze kennis kunnen we nu, als we weten hoeveel botsingen we *gemiddeld* verwachten met een 4-lepton signatuur onder de hypothese dat het Higgs deeltje wel/niet bestaat, een waarschijnlijkheidsmodel beschrijven voor beide hypotheses en de statistische analyse van het knikkerexperiment herhalen zonder verdere wijzingen.

Voor het beantwoorden van de vraag "*hoeveel 4-lepton botsingen verwachten we gemiddeld?*" is kennis van de onderliggende natuurkunde onontbeerlijk, en deze vraag in de praktijk beantwoord door middel van een uitgebreide keten van computersimulaties van zowel natuurkundige processen die plaatsvinden van de botsing, als een computersimulatie van de detector, die rekening houdt met de geometrie van de detector, zijn (in)efficiencties en eventuele miscalibraties. Deze computer-simulaties zijn zeer tijdrovend, en nemen vaak ruim 10 minuten computer-tijd per botsing in beslag. Voor tel-experimenten, hoeven deze echter slechts minimaal gebruik gemaakt te worden, omdat het niet nodig is de kans van elke mogelijke tel-uitkomst per computer te berekenen, maar slechts het gemiddeld verwachte aantal botsingen.

Uit deze berekeningen blijkt dat we in twee jaar data gemiddeld 4.5 botsingen met een 4-lepton signatuur verwachten als het Higgs deeltje *niet* bestaat (H_0), en gemiddeld 10 botsingen verwachten als het Higgs deeltje *wel* bestaat (H_1). In figuur 5.4 zijn de bijhorende Poisson kansverdelingen voor de hypotheses uitgezet.

In de geobserveerde data van het ATLAS experiment hebben 13 van zulke botsingen geobserveerd. De kans op 13 (of meer) geobserveerde botsingen kan nu eenvoudig worden uitgerekend:

$$\begin{aligned} p(N \geq 13|k = 4.5) &= \sum_{i=13,\infty} \text{Poisson}(i, 4.5) \approx 0.08\% \\ p(N \geq 13|k = 10) &= \sum_{i=13,\infty} \text{Poisson}(i, 10) \approx 21\% \end{aligned}$$



Figuur 5.4: Poisson verdeling van waarschijnlijkheden om het aantal geobserveerde LHC proton-proton botsingen met vier leptonen te observeren in twee jaar data onder hypothese dat het Higgs deeltje niet bestaat (blauw) en dat het Higgs deeltje wel bestaat (rood). Het aantal daadwerkelijk geobserveerde botsingen van de type is 13.

Alhoewel deze observatie van 13 botsingen zeer compatibel is met de H_1 hypothese, is de kans onder de H_0 hypothese van 0.08% nog altijd ver verwijderd van de '5 σ' drempel van $2.87 \cdot 10^{-7}$. Daarom is de observatie in dit kanaal nog geen overtuigend bewijs van het bestaan van het Higgs deeltje, maar wel een interessant hint.

5.5 Beter zoeken met Machine Learning

De zoektocht naar het Higgs deeltje kan worden uitgebreid door er ook ander vervalsmogelijkheden van het deeltje bij te betrekken. Deze andere vervalsmogelijkheden zijn vaak minder zeldzaam van het 4-lepton verval, maar zijn minder gemakkelijk te onderscheiden van andere botsingen, waardoor een selectie van botsingen op basis van een of twee criteria ontoereikend zal zijn.

Het is echter nog steeds mogelijk om zo'n voldoende zuivere verzameling te definiëren aan de hand van een groter aantal snedes op meerdere meetbare botsingeneigenschappen. Het process om deze optimale combinatie van snedes te vinden is echter tijdrovend: het aantal combinaties van mogelijke snedes wordt enorm groot als er meer dan slechts een handjevol gemeten grootheden wordt beschouwd. Een van de technieken om de

meest efficiënte combinatie van selectie-criteria te ontwikkelen die bij voorkeur botsingen selecteren met andere Higgs signatuur is 'machine learning'. Hier wordt op basis van natuurkundige overwegingen een lange lijst van meetbare grootheden in proton-proton botsingen opgesteld die elk enig potentieel hebben om botsingen met een Higgs verval te kunnen onderscheiden van de vele overige botsingen. Deze worden vervolgens aan een computer algoritme gegeven om de optimale selectie te definiëren.

Er zijn in de afgelopen decennia vele algoritmes ontwikkeld om dit te doen. De techniek die momenteel populair voor elementaire deeltjes fysica is de zogenaamde 'Boosted Decision Tree'.

5.5.1 Decision Trees

Een 'Decision Tree'[6] is wiskundige implementatie van een series snedes op gemeten grootheden die een botsing classificeert als 'signaal' of 'achtergrond', en laat zich het makkelijkst uitleggen aan de hand van een voorbeeld.

Gegeven een serie met meetbare grootheden aan een proton-botsing, vaak met exotische namen zoals $n(lep), H_T, E_T, m(\gamma\gamma)$, en twee zuivere verzamelingen met gesimuleerde signaal botsingen (waarin een Higgs deeltje wordt geproduceerd dat op de gewenste wijze vervalt) en gesimuleerde achtergrond botsingen (een mix van alle andere botsingen die kunnen voorkomen), gaat een *Decision Tree learner* als volgt te werk:

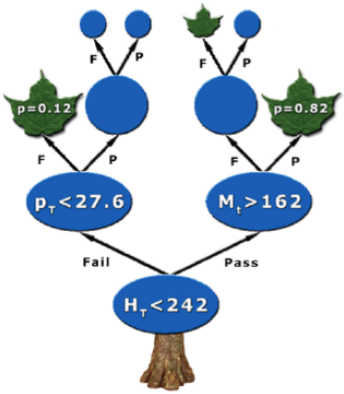
1. Bepaal voor elke botsings-grootheid $(n(lep), H_T, E_T, m(\gamma\gamma))$ welke snede de beste scheiding tussen signaal en achtergrond oplevert, aan de hand van de classificatie van de aangeleverde zuivere verzamelingen met signaal en achtergrond botsingen.

2. Gebruik de de snede op de botsings-grootheid die de beste scheiding oplevert om een eerste classificatie van botsingen te maken in twee deelruimtes S_0 (voornamelijk signaal botsingen) en B_0 (voornamelijk achtergrond botsingen) Deze classificatie zal zeer waarschijnlijk imperfect zijn (dwz er zullen signaal botsingen als achtergrond worden geclassificeerd en vice-versa), maar is de best mogelijke classificatie op basis van één criterium.

3. Itereer nu stappen 1) en 2) om de deelruimtes S_0 and B_0 elk weer in tween te splitsen met behulp van een nieuw gekozen 'beste' criterium in beide deelverzamelingen: $S_0 \to S_1(S_0), S_0 \to B_1(S_0), B_0 \to S_1(B_0), B_0 \to B_1(B_0)$. Het bestwerkende criterium voor de splitsing

van S_0 en B_0 hoeft uiteraard niet hetzelfde te zijn.

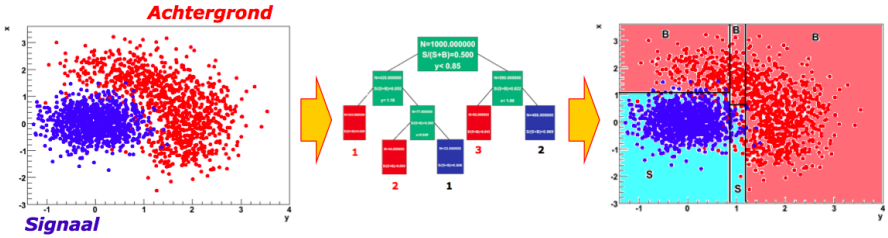
4. Itereer de procedure net zo lang totdat alle gesimuleerde botsingen perfect geclassificeerd zijn, of er geen criteria meer zijn die een statistisch significantie verbeterde zuiverheid in de deelverzamelingen resulteren.



Figuur 5.5: Grafische illustratie van een 'Decision Tree'. Elke vertakking in de boom splits een verzameling in twee deel-verzamelingen, met als doen te eindigen met een serie deelverzamelingen elk voornamelijk uit signaal of achtergrond botsingen bestaan.

Het resultaat is een beslissingsboom ('Decision Tree') die elke botsing aan de hand van zijn gemeten criteria classificeert als signaal of achtergrond. Figuur 5.5 illustreert grafisch de structuur van een beslissingsboom zoals hij gebouwd wordt. Figuur 5.6 illustreert de werking van een decision tree aan de hand van een fictieve verzameling signaal en achtergronden punten verdeeld in een ruimte van twee gemeten grootheden.

In vrijwel alle realistische toepassingen zal een Decision Tree de data nooit perfect kunnen classificeren (en in het geval van elementaire deeltjes fysica is dat vanuit het oogpunt van de kwantummechanica ook theoretisch niet mogelijk). Daarom is het gebruikelijk om de rapporteerde uitkomst van een Decision Tree niet simpelweg een Boolean te laten zijn, maar de signaalzuiverheid van de deelverzameling van trainingsbotsingen te rapporteren, die hoort bij het gekozen pad in de Decision Tree hoort. Bijvoorbeeld, als de Decision Tree van Fig 5.5 voor een bepaalde botsing uitkomt in het meest rechtse 'blad', zal de uitkomst score van de Decision Tree een waarde



Figuur 5.6: Een Decision Tree in actie op een voorbeeld probleem met een verzamelijk signaal (blauw) en achtergrond (rood) punten die elk worden gekarakteriseerd door twee observabelen. De rechter grafiek illustreert hoe de mogelijke beslissingen van de getrainde Decision Tree uitpakken.

van 0.82 hebben, in plaats van een label 'signaal'.

De mogelijke uitkomstwaarden van een Decision Tree, en het aantal mogelijke uitkomstwaarden, hangen van de vorm van de Decision Tree af: grote complexe Decision Trees kunnen veelvoud aan signaal-zuiverheids scores opleveren, zeer eenvoudige Decision Trees slechts twee. Een snede op de Decision Tree score bepaalt dan de uiteindelijke botsing-selectie, die vervolgens gebruikt kan worden voor een tel-experiment zoals uitgelegd in sectie 5.4.2. Het rapporteren van een zuiverheidscore in plaats van een Boolean geeft de gebruiker die dan de mogelijkheid om de efficiëntie van de selectie aan te passen door de snede op de BDT score te variëren (een snede op een lage score zal meer signaal en achtergrond selecteren dan een snede op een hoge score).

5.5.2 Boosted Decision Trees

De trainings procedure voor Decision Trees kan worden verbeterd door een de *boosting* techniek. Het idee achter deze techniek is dat classificatie uiteindelijk wordt bepaald door een ensemble van Decision Trees[7]. De Decision Trees in dit ensemble onderscheiden zich van elkaar doordat ze van elkaars fouten leren in het trainingsalgoritme.

Het construeren van een *Boosted* Decision Tree begint met het trainen van een gewone Decision Tree T_0, waar weer voor elke beslissingspad wordt bijgehouden wat de classificatiezuiverheid is op basis van trainings samples.

Na de training wordt de performance van de Decision Tree berekend als de gemiddelde misclassificatie frequentie $\langle \epsilon_k \rangle$:

$$\langle \epsilon_k \rangle = \frac{\sum_{i=1,N} w_i^k \cdot \epsilon_k(i)}{\sum_{i=1,N} w_i^k} \tag{5.15}$$

waar de sommatie over alle trainings botsingen loopt, w_i^k het gewicht is van deze botsingen in het trainingsproces (voor de eerste Decision Tree met $k = 0$ zijn deze gewichten initieel 1), en $\epsilon_k(i)$ is een Boolean functie is die aangeeft of trainings botsing i wel of niet correct geclassificeerd is door de Decision Tree. De gemiddelde misclassificatie rate $\langle \epsilon_k \rangle$ heeft een waarde tussen de nul (voor een nutteloze tree) en een (voor een perfecte tree).

Deze informatie wordt vervolgens op twee manieren gebruikt. Ten eerste zal het trainingssample voor de *volgende* Decision tree (T_{k+1}) worden herwogen voor botsingen waar T_k een vergissing maakte zodat deze een belangrijkere rol spelen in de training van deze volgende decision tree. De gewichtsfactor die wordt toegekend aan de gemisclassificeerde botsingen in T_k is gerelateerd aan de performance van deze vorige tree:

$$w_i^{k+1} = w_i \cdot \frac{1 - \langle \epsilon \rangle_k}{\langle \epsilon_k \rangle} \tag{5.16}$$

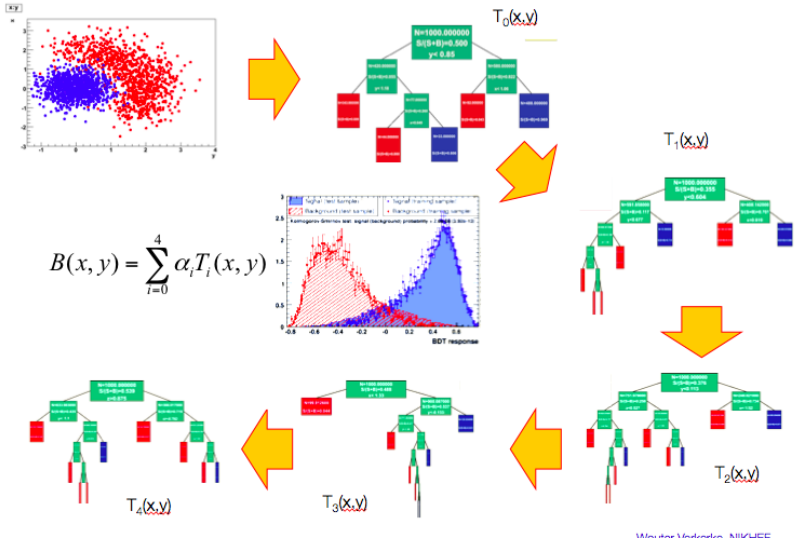
Bijvoorbeeld, als een Decision Tree T_k matig was ($\langle \epsilon_k \rangle = 40\%$) dan worden de botsingen die waren gemisclassificeerd door T_k herwogen met een factor 1.5. Fouten gemaakt door matige Decision Trees hebben dus relatief weinig invloed op het leerproces. Daarentegen, als de een Decision Tree T_k uitstekend was ($\langle \epsilon_k \rangle = 5\%$) dan worden de botsingen die waren gemisclassificeerd door T_k herwogen met een factor 19. Een volgende Decision Tree zal dus zeer sterk focusseren op fouten gemaakt door uitstekende Decision Tree.

De uiteindelijke beslissing van de Boosted Decision Tee is een gewogen gemiddelde van de 'bomen in het bos', waar de weging wordt bepaald door de gemiddelde misclassificatie $\langle \epsilon_k \rangle$

$$BDT(\vec{x}) = \sum_{k=1}^{N_{tree}} \log \left(\frac{1 - \langle \epsilon_k \rangle}{\langle \epsilon_k \rangle} \right) T_k(\vec{x}) \tag{5.17}$$

Figuur 5.7 laat de serie Decision Trees zien die getraind worden als de boos-

ting procedure wordt toegepast op het voorbeeld van figuur 5.5. Naast een verbeterde classificatie performance van boosted decision trees ten opzichte van enkele decision trees, hebben BDT als karakteristieke eigenschap dat de distributie van BDT classificatie uitkomsten op een botsingssample een bijna continue distributie is geworden als gevolg door het wegingsproces.

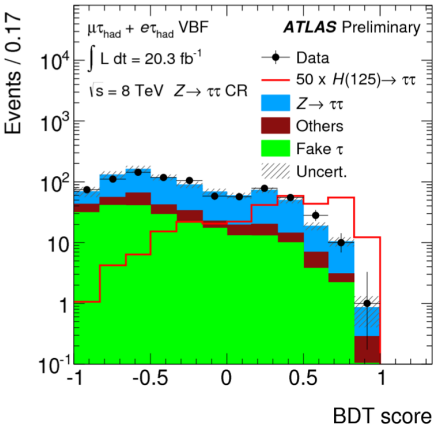


Figuur 5.7: Illustratie van het proces van boosting van Decision Trees. Een serie van vijf Decision Trees wordt getraind op een verzameling trainings events dat telkens herwogen wordt op basis van de foutieve classificatie door de vorige DT in de serie. De grafie in het midden laat zien wat de verdeling van gewogen gemiddeld beslissing is van het 'bos' van Decision Trees.

5.5.3 Boosted Decision Trees voor Higgs identificatie

Voor diverse van de 'moeilijke' signaturen van het Higgs deeltje zijn Boosted Decision Trees gebruikt om de selectie van botsingen te definiëren. Figuur 5.8 laat de distributie van BDT scores zien voor één van deze moeilijke kanalen waar het Higgs deeltje vervalt in twee exotische τ mesonen[8]. Signaal events populeren voornamelijk de rechterkant van het histogram, corresponderend met hoge BDT scores, achtergrond events populeren voor-

namelijk de linkerkant van het histogram corresponderend met lage BDT scores. Het is evident uit de geringe scheiding tussen gesimuleerde signaal en achtergrondsbotsingen dat identificeren van signaal-botsingen in deze vervalsmodus zeer moeilijk is.



Figuur 5.8: Voorbeeld van het gebruik van botsingsklassificatie door middel van Boosted Decision Trees in de zoektocht naar het Higgs deeltje. In deze grafiek zijn de voorspelde en geobserveerde verdeling van BDT klassificaties uitgezet voor de notoir moeilijke Higgs vervalssignatuur in twee exotische τ deeltjes.

5.6 Van tellen naar likelihood-ratio tests

De eenvoudigste manier om een statistische analyse uit te voeren op een Higgs vervalskanaal waar de botsingen zijn geselecteerd door een Boosted Decision Tree is om een snede te introduceren op de BDT score en een statistische analyse voor een tel-experiment uit te voeren op de geselecteerde events, analoog aan het voor beeld van de 4-lepton tel analyse in sectie 5.4.2.

Voor moeilijke kanalen zoals $H \rightarrow \tau\tau$, waar de scheiding tussen signaal en achtergrond matig is, zoals te zien in Fig. 5.8, is een tel-experiment echter evident sub-optimaal omdat veel signaal events met lage BDT scores worden associeert en dus niet meer meetellen in analyse omdat ze een

selectie niet overleven. Een voor de hand liggende uitbreiding van de statistische data analyse die dit euvel verhelpt is is om eenvoudigweg *alle* botsingen te tellen door BDT scores geclassificeerd zijn, maar wel gebruikt te maken van de variërende signaalzuiverheid als functie van de BDT score.

5.6.1 Een kansmodel voor alle BDT-geklassificeerde botsingen

Een eenvoudige manier om dit te doen is om de data te verdelen is klasses $i = 0...n$ gebaseerd op intervals in de BDT score, zoals al reeds gedaan is het histogram in Fig 5.8. Het geobserveerde aantal botsingen in elke individuele data klasse i kan vervolgens beschreven worden met een Poisson kansverdeling, aangezien we eenvoudigweg het aantal botsingen in een selectie tellen.

$$p_\mu(n_i) = \text{Poisson}(n_i, \mu \cdot s_i + b_i) \tag{5.18}$$

waar n_i het aantal geobserveerde botsingen in bin i van het BDT histogram van Fig 5.8 en s_i en b_i de verwachte hoeveelheid signaal en achtergronden botsingen die gemiddeld worden verwacht in elke klasse. In vgl 5.18 is een parameter μ gentroduceerd om dezelfde uitdrukking als kansmodel zowel voor hypothese H_0 te kunnen gebruiken ($\mu = 0$) als voor hypothese H_1 ($\mu = 1$). Het kansmodel voor *alle* BDT-geklassificeerde botsingen is dan het product van al deze kansmodellen voor elke klasse

$$p_\mu(\vec{n}) = \prod_{i=0,n_{bin}} \text{Poisson}(n_i, \mu \cdot s_i + b_i) \tag{5.19}$$

waar de $\vec{n}$ de vector van het geobserveerde aantal botsingen in BDT klasses is.

5.6.2 Rekenen aan kansmodellen met multi-dimensionale observabelen

Alhoewel $p(\vec{n})$ eenvoudig is uit te rekenen, is het minder eenvoudig om de grafiek of tabel te maken van distributie van mogelijke uitkomsten, zoals gedaan in Fig 5.4 voor het Poisson tel-experiment, omdat de observatie niet langer een scalaire grootheid is die op de x-as kan worden uitgezet, maar een vector van observaties is van dimensie n. Afgezien van het visualisatie probleem, is de hogere dimensionaliteit van de observatie ook een

fundamenteel probleem: om een kans op een waarneming uit te drukken in de vorm '*de kans om 13 of meer botsingen te zien onder de hypothese wel/niet bestaat*' is het noodzakelijke alle mogelijke observaties te kunnen ordenen van 'minder signaal' naar 'meer signaal'. Als de observatie simpelweg een aantal botsingen N is, is deze ordening triviaal, maar als de observatie een vector van aantallen $\vec{N}$ is deze ordening niet meer voor de hand liggend. Bijvoorbeeld: bevat de observatie $\vec{N} = (10, 20, 30)$ meer signaal of minder signaal dan de observatie $\vec{N} = (10, 21, 29)$?

Een uitkomst voor dit probleem biedt het Neyman-Pearson lemma[9] dat stelt dat informatie bevat in de vector $\vec{N}$ over onderscheid tussen twee hypotheses H_0 en H_1 (bv signaal en achtergrond) kan worden altijd kan worden samengevat in een scalaire grootheid zonder verlies van informatie:

$$\lambda(\vec{N}) = \frac{P(\vec{N}|H_0)}{P(\vec{N}|H_1)} \tag{5.20}$$

Dit lemma stelt ons in staat om kansmodellen voor multi-dimensionale observaties altijd te reduceren tot kansmodellen voor scalair observaties. Het invullen van vgl 5.19 in vgl 5.20 geeft de volgende uitdrukking:

$$\lambda(\vec{n}) = \frac{\prod_{i=0,n_{bin}} \text{Poisson}(n_i, b_i)}{\prod_{i=0,n_{bin}} \text{Poisson}(n_i, b_i + s_i)} \tag{5.21}$$

De waarde van deze nieuw geconstrueerde samenvatting van de geobserveerde data is eenvoudig te interpreteren: een waarde van $\lambda < 1$ geeft aan dat de geobserveerde data waarschijnlijker is onder de achtergrondshypothese dan de signaalhypothese (aangezien de noemer kleiner is dan de deler), een waarde van $\lambda > 1$ geeft daarentegen aan dat de data waarschijnlijker is onder signaalhypothese. Een hoge gemeten waarde van λ in een experiment is dus een aanwijzing voor signaal in de data, terwijl een lage gemeten waarde een aanwijzing is voor het ontbreken van signaal. De waarde van λ voor de observeerde data getoond in figuur 5.8 is bijvoorbeeld 1.30, die een kleine hoeveelheid signaal suggereert.

5.6.3 Kansmodellen voor likelihood ratios

Om een gemeten waarde van λ te vertalen naar het gebruikelijke statement over de kans van deze waarneming onder een hypothese ('*de kans om de*

observeerde waarde van $\lambda = 1.30$ of hoger te meten is X% onder de hypothese van wel/geen Higgs') is het wederom noodzakelijk on een kansmodel voor nieuwe observabele λ hebben voor beide hypotheses. In tegenstelling tot kansmodellen voor de multidimensionale tel-experimenten, waar de distributie analytisch bekend is als een (product van) Poisson verdelingen, is de distributie van een likelihood-ratio observable λ in het algemeen *niet* analytisch berekenbaar. Deze distributies kunnen wel 'brute-force' door middel van computersimulaties kunnen worden benaderd, maar deze deze computer berekeningen zeer tijdsintensief zijn, en zijn dit daardoor onpraktisch. Het blijkt echter dat met een kleine aanpassing[3] in de deler van de formulering van de likelihood ratio van vgl 5.21

$$\lambda_0(\vec{n}) = \frac{\prod_{i=0,n_{bin}} Poisson(n_i, b_i}{\prod_{i=0,n_{bin}} Poisson(n_i, b_i + \hat{\mu} \cdot s_i)} \tag{5.22}$$

de distributie van deze vorm van λ_0 *wel* analytisch bekend is in de limiet van oneindige statistiek[4]. Deze convergentie naar een analytische form is bekend als Wilks theorema[10]:

$$p(\lambda_0) = p_{\chi^2}(\lambda_0, k)_{k=1} = \left(\frac{\lambda_0^{(k/2-1)} \cdot e^{-\lambda_0/2}}{2^{k/2}\Gamma(k/2)} \right)_{k=1} \tag{5.23}$$

waar $p_{\chi^2}(\lambda, k)$ de bekende χ^2 verdeling is voor k vrijheidsgraden. De convergentie van de distributie van $p(\lambda_0)$ naar de asymptotische vorm blijkt voor tel-experiment al gerealiseerd te worden bij enkele tientallen geobserveerde botsingen[11], en is dus in zeer veel gevallen praktisch toepasbaar. De kans is op een observatie λ_0^{obs} of hoger onder hypothese H_0 kan nu evenvoudig worden uitgedrukt als een integraal over vgl 5.23:

$$p(\lambda_0 \geq \lambda_0^{obs}) = \int_{\lambda_0^{obs}}^{\infty} p_{\chi^2}(\lambda_0, k)_{k=1} \tag{5.24}$$

Zo is bijvoorbeeld de kans op de geobserveerde waarde van $\lambda = 1.30$ (behorende bij fig 5.8) of hoger, onder de hypothese H_0 (er bestaat geen Higgs

[3]Het verschil tussen λ en λ_0 is dat in in de deler de vaste aanname van het nominale signaal vervangen door een aanname van een hoeveelheid signaal die best past bij de geobserveerde data, hier symbolisch genoteerd $\hat{\mu} \cdot s_i$ waar s_i de nominaal verwachte hoeveelheid signaal is, en $\hat{\mu}$ een schaalfactor voor het nominale signaal die zo gekozen is dat de het kansmodel in de deler van vgl 5.22 maximaal is. Deze procedure om de deze speciale waarde van μ te vinden staat bekend als de 'maximum likelihood fit'

[4]En onder enkele aanvullende regulariteitsvoorwaarden.

deeltje) 25.4%.

Voor de ontdekkingsclaim van het Higgs deeltje is het asymptotische kansmodel expliciet geverifieerd door middel een brute-force simulatie berekening van de distributie $p(\lambda_0)$ *zonder* asymptotische aannames. Deze berekening heeft meer dan 100.000 uur CPU-tijd gekost en heeft aangetoond dat voor de likelihood ratio die gebruikt is voor de Higgs ontdekking de asymptotische benadering zeer goed was.

5.7 Statistische analyse van alle kanalen

De techniek van likelihood ratios om multi-dimensionale datasets te reduceren tot een enkele scalaire observable kan niet alleen gebruikt worden om de informatie van alle dataklasses op basis van BDT scores te combineren, maar kan op eenzelfde wijze gebruikt worden om kansmodellen van Higgs analysis in verschillende vervalskanalen te combineren. Analoog aan vgl 5.22 kan het kansmodel voor *alle* Higgs-gerelatieerde waarnemingen worden geschreven als het product van de kansmodellen voor elke Higgs vervalsignatuur signatuur

$$p_{comb}(n_W, n_Z, n_\gamma, n_\tau, n_b) = p_W(n_W) \cdot p_Z(n_Z) \cdot p_\gamma(n_\gamma) \cdot p_\tau(n_\tau) \cdot p_b(n_b) \tag{5.25}$$

waar $n_W, n_Z, n_\gamma, n_\tau, n_b$ de geobserveerde data voor de diverse Higgs vervalsmogelijkheden representeren. Alhoewel het gecombineerde kansmodel een zeer lange uitdrukking zal worden, die in praktijk alleen als een geprogrammeerde computer functie kan worden geschreven, is de procedure voor de kansberekeningen voor de data onder hypotheses H_0 en H_1 voor dit gecombineerde model precies hetzelfde als in het voorgaande hoofdstuk is beschreven, en kunnen de asymptotische benaderingen van sectie 5.6.3 evenzo worden toegepast.

5.8 Omgaan met onzekerheden

De statistische analyse procedures die tot nu toe zijn behandeld zijn voornamelijk gefocused op het gebruiken en classificeren van zo veel mogelijk informatie. Een uit wetenschappelijk oogpunt zeer belangrijk aspect is tot nu toe echter geheel onbelicht gebleven: veel van de informatie die ge-

bruikt wordt om de kansmodellen voor hypothese H_0 en H_1 te formuleren is onzeker. De bronnen van onzekerheden in de Higgs analyse zijn veelvoudig en variërend van de eindige precisie van theoretische berekeningen van voorspelde botsingsprocessen tot onzekerheden detectorcalibraties die leiden tot onzekerheden in detectie efficiëntie schattingen.

Als deze onzekerheden niet worden meegenomen in de statistische analyse is het eenvoudig de significantie van een observatie te overschatten. Neem bijvoorbeeld een observatie van 13 botsingen voor een proces waar slechts 2 achtergrondsbotsingen worden verwacht. Een Poisson model voor een tel-experiment leert dat de kans om 13 of meer botsingen waar te nemen, waar slechts 2 botsingen verwacht worden, een kans heeft van slechts $2 \cdot 10^{-7}$, wat onder normale omstandigheden tot de claim van een ontdekking van signaal zou leiden. Als de schatting van de gemiddelde verwachte achtergrond een aanzienlijke onzekerheid heeft, bijvoorbeeld $n_b = 2 \pm 2$ in plaats van exact 2, dan is de kans om 13 botsingen waar te nemen onder de achtergrondshypothese een stuk groter , en is er waarschijnlijk geen sprake meer van voldoende bewijs voor een ontdekkingsclaim.

Hoe kan je dergelijke onzekerheden meenemen in de statistische analyse? Door het waarschijnlijkheidsmodel uit te breiden met een extra parameters die de onzekere grootheden beschrijven, en kansmodellen toe te voegen die metingen die tot de onzekere schattingen van deze grootheden hebben geleid. Stel dat de gemiddeld verwachte achtergrond voor ons tel-experiment niet precies 2 is maar, maar 2 ± 2. Dit betekent dat er (ergens) een meting heeft plaatsgevonden die tot het resultaat $n_b = 2 \pm 2$ heeft geleid. We kunnen deze onzekere informatie nu meenemen door het kansmodel voor de achtergronds meting op te nemen in ons totale kansmodel, mogelijk in vereenvoudigde vorm. Zo breiden we het originele Poisson kansmodel voor ons (Higgs) tel-experiment

$$p(n_{obs}) = \text{Poisson}(n_{obs}|2)$$

uit met het kansmodel voor de achtergronden schatting, hier met de gesimplificeerde aanname dat het een normale (Gaussische) verdeling betreft:

$$p(n_{obs}|b) = \text{Poisson}(n_{obs}|b) \cdot \text{Gaussian}(2, b, 2)$$

Het bestaan van onzekere parameters in het kansmodel moet worden onderkend in de Likelihood Ratio berekening vgl 5.21 die moet worden aangepast in een zogenaamde Profile Likelihood Ratio, die waardes van deze

onzekere parameters in de Likelihood Ratio berekening vastgelegd op de waardes die waarschijnlijkheid van kansmodellen de noemer, respectievelijke deler maximaliseren. Dit proces van kansmaximalisatie in de onzekere parameters zal de waarde van de likelihood ratio voor observaties veranderen, waardoor het effect van deze onzekerheden effectief wordt gepropageerd in de berekening.

Op deze wijze kan berekend worden dat voor het voorbeeld tel-experiment met een onzekere achtergrond van $n_b = 2 \pm 2$ de kans op 13 of meer botsingen onder de achtergrondshypothese 2.7% is. Dit is een aanzienlijk verschil met de kans van $2 \cdot 10^{-7}$ voor het originele model zonder onzekerheid op de gemiddelde verwachte achtergrond. Het belang van het beschijven van onzekerheden in kansmodellen is dus evident, en alhoewel slechts kort belicht in deze syllabus vormt het overgrote deel van het mensenwerk achter de statistische analyse in de deeltjes fysica. Het computer-geprogammeerde gecombineerde kansmodel dat is gebruikt voor de Higgs ontdekkingsclaim heeft maat liefst 1600 parameters die corresponderen met een even groot aantal modelonzekerheden.

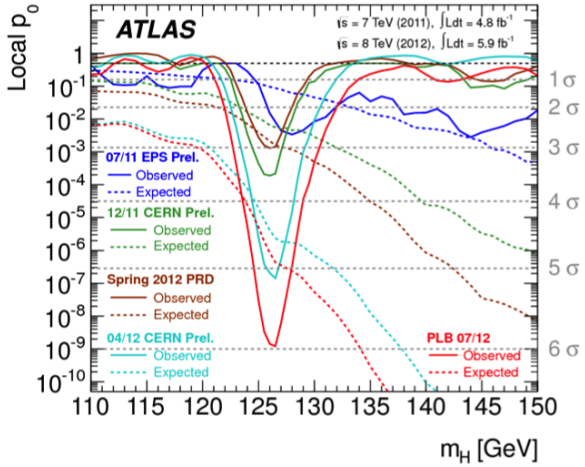
5.9 De ontdekking van het Higgs deeltje

Als afsluiting van dit onderwerp voeg ik een figuur uit het ontdekkingen paper van het ATLAS experiment toe (fig.5.9), die de resultaten van de statistische analyse samenvat die alle hier beschreven technieken gebruikt heeft. Op de verticale as van de grafiek is uitgezet de waarschijnlijkheid om de geobserveerde hoeveelheid signaal, of groter, or waar te nemen onder de hypothese dat er *geen* Higgs deeltje bestaat. Het gecombineerde kansmodel voor deze analyse beschrijft 192 verschillende signaal selecties die gemaakt zijn van de LHC botsingsdata.

Omdat de massa van het Higgs deeltje voor de ontdekking niet bekend was, en de selectiecriteria voor signaalbotsingen sterk afhangen van deze (onbekende) massa, is de statistische analyse herhaald voor een wijd scala aan aangenomen waarden voor de Higgs massa. Deze aangenomen waarde voor de Higgs massa in de statistische analyse is uitgezet op de horizontale as.

De gekleurde curves laten zien hoe de waarschijnlijkheid van de geobserveerde data onder de geen-Higgs hypothese steeds kleiner werd naar mate er meer data genomen is. Bij de laatste curve, corresponderend met die van de data genomen tot Juli 2012 is die waarschijnlijkheid kleiner gewor-

den dan $2.87 \cdot 10^{-7}$ voor een aangenomen Higgs massa rond de 125 GeV. Op basis van deze '5σ' observatie van het ATLAS experiment, tezamen met een geheel onafhankelijk waargenomen 4.9σ observatie door het CMS experiment, is de ontdekking van het Higgs deeltje gezamenlijk geclaimd door de ATLAS en CMS experiment op CERN, zoals aangekondigd op de live-uitzending van de wetenschappelijke presentatie op 4 juli 2012 door de directeur-generaal van CERN.



Figuur 5.9: Samenvatting van de statistische analyse die heeft geleid tot de ontdekkingsclaim van het Higgs deeltje. Op de verticale as is uitgezet de kans om het geobserveerde Higgs signaal, of groter, waar te nemen (in alle bestudeerrde kanalen), onder de hypothese dat het Higgs deeltje *niet* bestaat. Zie de de tekst van sectie 5.9 voor een gedetailleerde uitleg van de betekenis van deze plot

Bibliografie

[1] "Observation of a New Particle in the Search for the Standard Model Higgs Boson with the ATLAS Detector at the LHCÓ", ATLAS Collaboration (Georges Aad et al.), Phys. Lett. B 716 (2012) 1-29.

"Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHCÓ", CMS Collaboration, Phys. Lett. B 716 (2012) 30-61.

[2] "Broken Symmetry and the Mass of Gauge Vector Mesons", F. Englert, R. Brout (1964). Physical Review Letters 13 (9): 321323. "Broken Symmetries and the Masses of Gauge Bosons", P.W. Higgs (1964). Physical Review Letters 13 (16): 508509.

[3] [6] "Partial-symmetries of weak interactions", S.L. Glashow (1961), Nucl. Phys. 22 (4): 579588 "A Model of Leptons", S. Weinberg (1967)., Physical Review Lett 19 (21): 12641266

[4] "An Essay towards solving a Problem in the Doctrine of Chance. By the late Rev. Mr. Bayes, communicated by Mr. Price, in a letter to John Canton, A. M. F. R. S.". Bayes, Thomas, and Price, Richard (1763), Philosophical Transactions of the Royal Society of London 53 (0): 370418

[5] "Probabilité des jugements en matire criminelle et en matire civile, précédées des règles générales du calcul des probabilitiés", S.D. Poisson, (Paris, France: Bachelier, 1837), p. 206

[6] "Classification and Regression Trees", L. Breiman et al., Chapman & Hall (1984)

[7] "A decision-theoretic generalization of on-line learning and an application to boosting". AUTHORS, Journal of Computer and System Sciences 55. 1997

[8] "Evidence for Higgs Boson Decays to the tau+tau- Final State with the ATLAS Detector", ATLAS Collaboration, ATLAS-CONF-2013-108.

[9] "On the Problem of the Most Efficient Tests of Statistical Hypotheses", J. Neyman, E. Pearson, Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences 231 (694706): 289337 (1933).

[10] "The Large-Sample Distribution of the Likelihood Ratio for Testing Composite Hypotheses", S. S. Wilks, The Annals of Mathematical Statistics 9: 6062 (1938).

[11] "Asymptotic formulae for likelihood-based tests of new physics", G. Cowan, K. Cranmer, E. Gross, O. Vitells, arXiv:1007.1727v2

6 Reconstruction of the gene-gene interaction network

Wessel N. van Wieringen en Mark A. van de Wiel

Vrije Universiteit Amsterdam

Op de volgende bladzijden vindt u de presentatie.

Reconstruction of the gene-gene interaction network

Wessel N. van Wieringen
w.vanwieringen@vumc.nl
Amsterdam (30-08-2014)

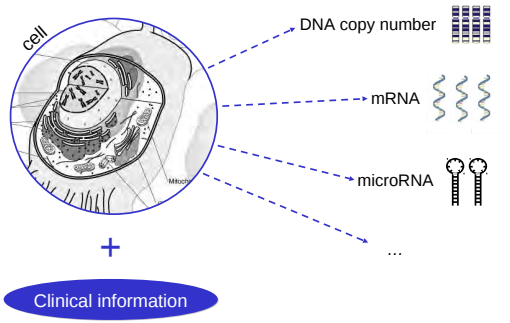
Mark A. van de Wiel
mark.vdwiel@vumc.nl
Eindhoven (23-08-2014)

Department of Epidemiology and Biostatistics, VUmc
& Department of Mathematics, VU University
Amsterdam, The Netherlands

Platform Wiskunde Nederland, Vakantiecursus 2014, "Nieuwe tijden"

A bare minimum of molecular biology

Context



Molecular biology

Molecular biology aims to understand the molecular processes that occur in the cell. That is, which molecules present in the cell interact, and how is this coordinated?

For many cellular processes, it is unknown which genes play what role.

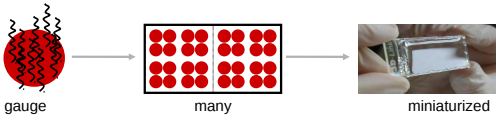
Solution

Simply measure (the expression of) all genes.

High-throughput

Microarrays

Measure, e.g., the expression of many genes simultaneously (which genes are expressed and to what extent).

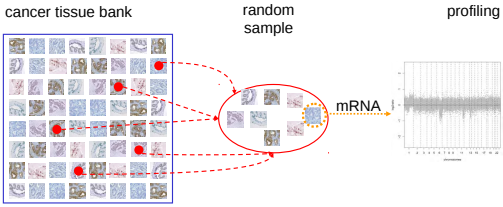


gauge many miniaturized

expression profile

gene 1	-0.21968
gene 2	-0.08376
gene 3	-1.08336
gene 4	0.04333
gene 5	1.16542
gene 6	-0.29687
gene 7	1.76249
gene 8	0.46387
...	...

Experimental design



- in vivo
- cross-sectional

Data

To reconstruct which genes interact, we have available:
→ expression profiles of *n* samples,
→ each profile comprises *p* genes.

	sample 1	sample 2	sample 3	sample 4	sample 5
gene 1	-0.21968	-0.42796	0.26441	-5.74971	-0.96908
gene 2	-0.08376	-7.21648	-3.86460	0.77440	-3.18557
gene 3	-1.08336	-1.14688	-1.22544	-2.36134	0.19293
gene 4	0.04333	-0.46377	0.12756	-0.39535	-0.20215
gene 5	1.16542	0.86248	1.16049	1.23941	0.51927
gene 6	-0.29687	0.28602	-0.69624	-1.19779	0.19546
gene 7	1.76249	1.07556	1.46201	1.16876	1.29921
gene 8	0.46387	0.21271	0.49455	0.58267	-0.44349
gene 9	-1.27492	3.95515	-0.26441	-2.95037	-0.77896
gene 10	0.25603	1.33263	0.91686	0.27558	1.44118
gene 11	3.55160	-3.66900	-1.56704	-5.44092	-3.71877
gene 12	-2.63154	-4.47593	-4.11052	-0.46330	-2.21023
gene 13	0.04514	0.82428	0.23972	-0.87675	0.41242
gene 14	0.04152	-0.10269	-0.74541	-0.25346	-0.39054
gene 15	1.31159	0.88571	1.36031	1.09076	0.55596
gene 16	-1.29186	2.93666	-2.06078	-6.91695	6.82540
gene 17	-0.10796	-1.39422	-0.54767	0.97816	0.16774
gene 18	0.98841	0.35618	0.68502	1.37255	1.52144
gene 19	-1.01269	-0.92959	-1.46227	-0.76152	-0.30942
gene 20	-2.86074	-5.50497	2.21568	-8.09685	-3.51475
...	...	...	...	...	...

Pathway

Pathway = chain of chemical reactions (that processes a signal)

≈ a set of genes believed to carry one function

Knowledge of pathways is incomplete and biased towards a few well-studied pathways.

Pathways are loosely defined using repositories, such as:
- KEGG
- BioCarta
- GennMapp
- Reactome
- GO

Pathway

Pathways are represented by a **graph** or **network**.

node or *vertex*, representing a gene.

edge or *arrow*, representing an interaction between two genes.

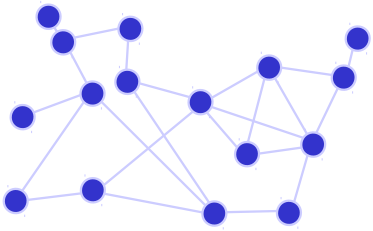
undirected and *directed* edges

undirected

directed

Pathway

Goal: reconstruct the topology, i.e. the edges present, of a pathway from gene expression data.



An edge ≈ interaction (formally: conditional dependence).

Two-gene pathway & correlation

Two-gene pathway & correlation

Two-gene pathways comprise two genes, and ignore the possibility there may be more.

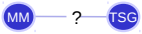
Cancer research example

Y_2 : gene expression measurements of a tumor suppressor gene

Y_1 : gene expression of a methylation marker

Question

Does the methylation marker (MM) influence the expression of the tumor suppressor gene (TSG)?



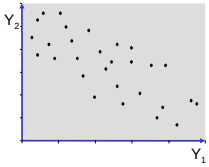
Two-gene pathway & correlation

Consider a pathway comprising of two genes.

Expression levels of genes 1 and 2 are *dependent*: $Y_1 \not\perp Y_2$

Graph: 1 — 2

Data:



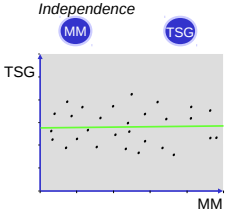
If a low value for Y_1 is observed, it is more likely to observe a high value for Y_2.

And vice versa.

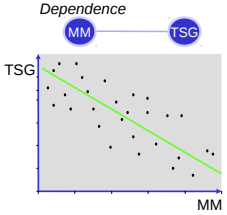
Two-gene pathway & correlation

TSG suppresses tumor formation. Ideally, its expression levels are high. If the expression levels of MM and TSG are dependent, we may aim to control those of TSG via MM.

Independence

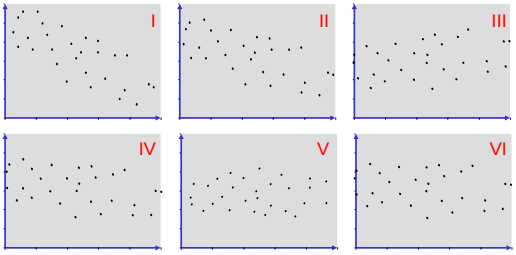


Dependence



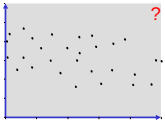
Two-gene pathway & correlation

Scatterplots of data on two random variables.
Which show association?



Two-gene pathway & correlation

→ Assess association between two random variables graphically.
→ Not very exact and in boundary cases no consensus.



Ideally, a measure of interrelatedness of the two variables.

Covariance measures whether a positive deviation from the mean in one variable systemically coincides with a positive (or negative) deviation from the mean in another variable.

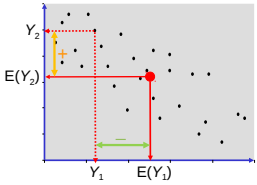
Two-gene pathway & correlation

Illustration

Let Y_1 and Y_2 be random variables and their covariance:

$$\mathrm{Cov}(Y_1, Y_2) = E\{[Y_1 - E(Y_1)]\,[Y_2 - E(Y_2)]\}$$

deviation from mean



sign of covariance:
+ x – = –

Two-gene pathway & correlation

Standardizing the covariance defines the ***correlation*** coefficient:

$$\rho(Y_1, Y_2) = \mathrm{Cor}(Y_1, Y_2) = \frac{\mathrm{Cov}(Y_1, Y_2)}{\sqrt{\mathrm{Var}(Y_1)}\sqrt{\mathrm{Var}(Y_2)}}$$

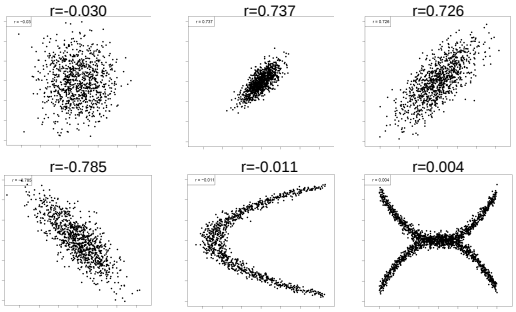
$\rho(Y_1, Y_2)$ in [-1, 1], with

→ $\rho(Y_1, Y_2) = 1$: perfect positive linear relationship,
→ $\rho(Y_1, Y_2) = -1$: perfect negative linear relationship,
→ $\rho(Y_1, Y_2) = 0$: no linear relationship.

Zero correlation implies independence of Y_1 and Y_2

Two-gene pathway & correlation

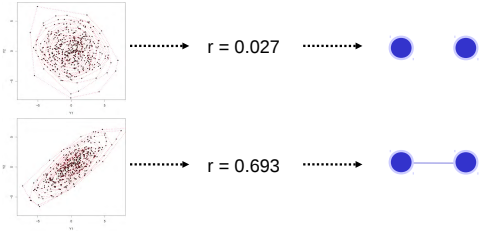
Pearson's correlation measures only linear dependence.



Two-gene pathway & correlation

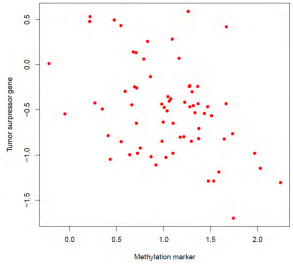
Two-gene system

Calculate correlation between any two genes. If the correlation is large (in some sense), the two genes interact.



Two-gene pathway & correlation

Expression levels of the TSG vs. MM



Question: is there dependence between TSG and MM?

Two-gene pathway & correlation

Cancer research example

```
> cov(cbind(MM, TSG))
              MM          TSG
MM   0.23897377 -0.09787409
TSG -0.09787409  0.25388099

> cor(cbind(MM, TSG))
           MM         TSG
MM   1.000000 -0.397354
TSG -0.397354  1.000000

> rho <- cor(MM, TSG)
> T <- log((1+rho)/(1-rho))/2
> sd <- sqrt(1/(length(MM)-3))
> pvalue <- 2*pnorm(T, sd=sd)
> pvalue
[1] 0.0006984108
```

Conclusion

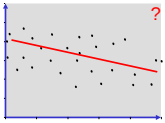
Significant association between of MM on TSG.

Note: the edge is undirected, the two variates are (causally) on a par.

Two-gene pathway & regression

Two-gene pathway & regression

Instead of using some measure to assess the dependence between two variables, one may explicitly model their relationship.



Regression analysis is a statistical method to estimate the relation among variables, e.g.:

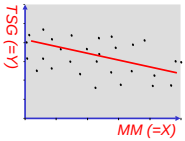
$Y_{tsg} = f(Y_{mm}) + \text{error}$

where f() is some function deemed appropriate. Commonly, f() is taken to be linear (as a first order approximation).

Two-gene pathway & regression

More formally, the simple linear regression model:

$$\underset{TSG}{Y_i} = \beta_0 + \beta_1 \underset{MM}{X_i} + \varepsilon_i$$



Some nomenclature:

$$Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i$$

response or dependent variable

regression parameter

covariate or explanatory variable

error ≈ part of Y not explained by the model

Two-gene pathway & regression

Note

The simple linear regression model:

$$Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i$$

is called *linear* because it is linear in the regression parameters.

Hence, the extension with a quadratic term:

$Y = \beta_0 + \beta_1 X + \beta_2 X^2 + \varepsilon$

Is also a linear model (in the regression parameters).

Similarly, the multivariable model

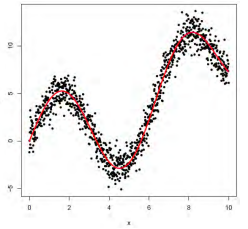
$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \varepsilon$

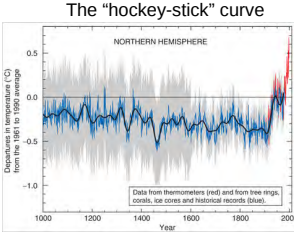
is also linear.

Two-gene pathway & regression

Examples of linear models

The "hockey-stick" curve





Two-gene pathway & regression

Note

We write:

$$Y_i \quad = \quad \beta_0 + \beta_1 X_i + \varepsilon_i$$

while it is equivalent to write:

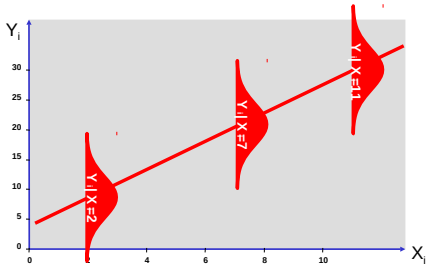
$$Y_i \mid X_i \quad \sim \quad \mathcal{N}(\beta_0 + \beta_1 X_i, \sigma^2)$$

error variance

The latter explicitly assumes that the explanatory variable X_i is (temporarily) taken as non-random. It is to be read as: Y_i *conditional* on X_i is distributed as ….

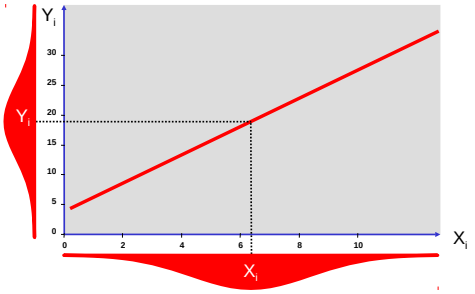
Two-gene pathway & regression

The ***conditional*** distribution of Y_i on X_i



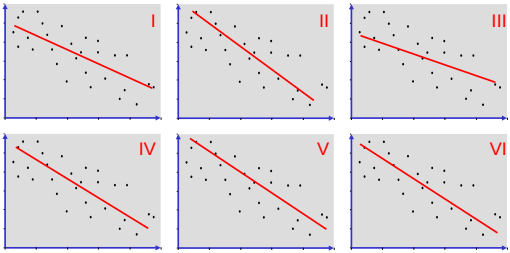
Two-gene pathway & regression

The unconditional (***marginal***) distribution of Y_i



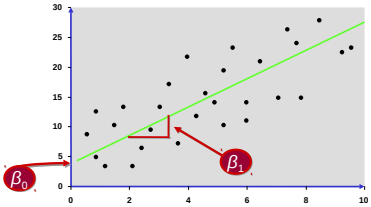
Two-gene pathway & regression

Question
What is now the best model? Best in what sense?



Two-gene pathway & regression

We search a *linear* (= straight line) relation:
$Y = \beta_0 + \beta_1 X$.
How to choose β_0 and β_1?

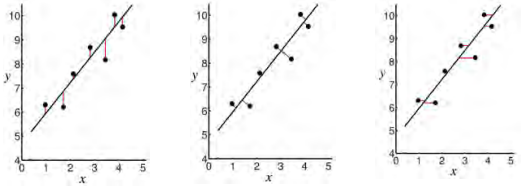


β_0: constant
β_1: slope

Two-gene pathway & regression

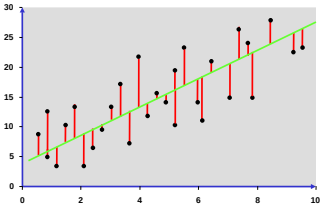
Best: smallest distance from line to observations.

Question
Which do you prefer?



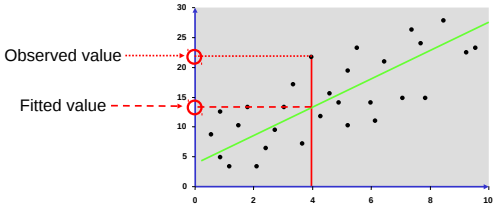
Two-gene pathway & regression

β_0 and β_1 are chosen such that the total quadratic distance of the observations to the regression line is minimal.



Two-gene pathway & regression

Residuals and fits

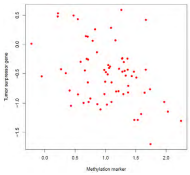


residual = observed value – fitted value

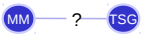
Two-gene pathway & regression

Two-gene pathways comprise two genes, and ignore the possibility there may be more.

Cancer research example
Y : expression measurements of a tumor suppressor gene
X : expression of a methylation marker

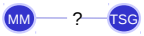


Question
Does the methylation marker (MM) influence the expression of the tumor suppressor gene (TSG)?



Two-gene pathway & regression

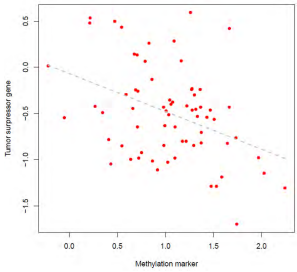
Cancer research example




```
> regRes <- lm(TSG ~ MM)
> summary(regRes)

Call:
lm(formula = TSG ~ MM)

Residuals:
    Min      1Q  Median      3Q     Max
-0.9170 -0.2957  0.0103  0.2921  1.1754

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) -0.06796    0.13419  -0.506 0.614231
MM          -0.40956    0.11643  -3.518 0.000793 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1
```

Two-gene pathway & regression

```
> plot(TSG ~ MM, ...)
> lines(regressionResults$fitted.values ~ MM, ...)
```



Conclusion
Significant effect of MM on TSG.

Thus, $\beta \neq 0$. Hence, gene expression levels of MM and TSG are related.

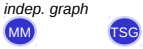
Two-gene pathway: regression & correlation

Correlation & regression

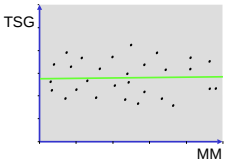
Correlation and regression are 1-1 related

TSG & MM independent

Cond. indep. graph

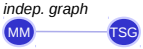


Data (ρ=0 ↔ β_1=0)

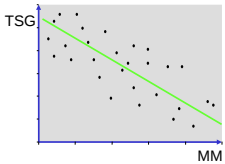


TSG & MM dependent

Cond. indep. graph



Data (ρ≠0 ↔ β_1≠0)



Correlation & regression

Undirected edges only

In regression analysis random variables Y_{MM} and Y_{TSG} are not on equal footing. The equation:

$Y_{TSG} = f(Y_{MM}) + \text{error}$

suggests MM controls TSG.

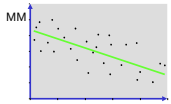
However, correlation is symmetric:

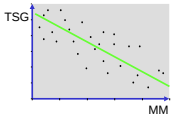
$\rho(Y_{MM}, Y_{TSG}) = \rho(Y_{TSG}, Y_{MM})$,

and β is 1-1 related to ρ.

E.g, signs of β_{MM} and β_{TSG} are same.

Thus, regression provides no clue about direction of the relationship.





Interpretation pitfall

Interpretation pitfall

Eat chocolate, win the Nobel!

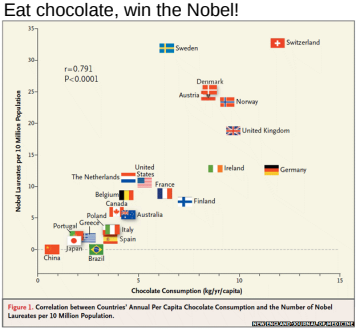
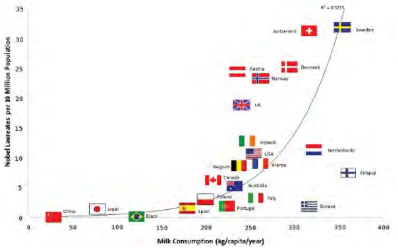


Figure 1. Correlation between Countries' Annual Per Capita Chocolate Consumption and the Number of Nobel Laureates per 10 Million Population.

Messerli, 2012.

Interpretation pitfall

Even better: drink milk, win the Nobel!

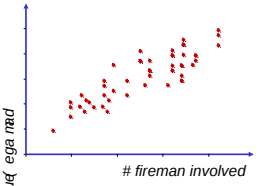


Best: drink chocolate-milk, win the Nobel?

Linthwaite, Fuller, 2013.

Interpretation pitfall

Does the involvement of more fireman result in more damage?



Possible interpretations of these data:

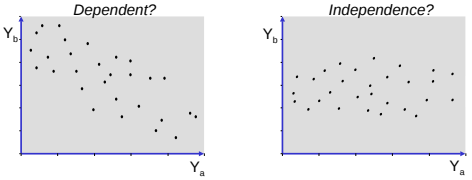
X → Y More firemen result in more damage.

X ← Y More damage results in more firemen.

A bigger fire (Z) results in more firemen and more damage.

Interpretation pitfall

What to conclude about the relation between the expression levels of gene A and B?

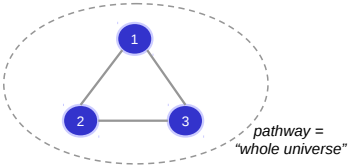


Could other genes be responsible for the observed (in)dependence?

Multi-gene pathways

Multi-gene pathways

Multi-gene pathways comprise of more than two genes, and assume no gene "lives" outside the pathway.



Two methods:
- Regression
- Correlation

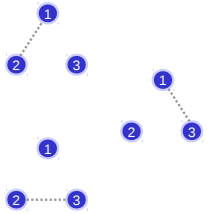
Multi-gene pathways

Correlation method
Correlate all possible pairs of genes in the pathway.

Cor(Y_1, Y_2)

Cor(Y_1, Y_3)

Cor(Y_2, Y_3)



Correlation ignores the other variables, and thus assesses the marginal dependence between two variables.

Multi-gene pathways

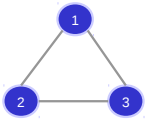
A ***partial correlation coefficient*** quantifies the correlation between two variables when conditioning on one or several other variables.

The partial correlation coefficient between Y_a and Y_b conditional on variables $\mathbf{Y}_c$ is defined as:

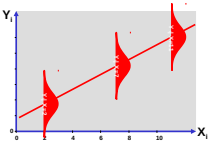
$$\rho(Y_a, Y_b \mid \mathbf{Y}_c) = \frac{\mathrm{Cov}(Y_a, Y_b \mid \mathbf{Y}_c)}{\sqrt{\mathrm{Var}(Y_a \mid \mathbf{Y}_c)}\sqrt{\mathrm{Var}(Y_b \mid \mathbf{Y}_c)}}$$

A zero partial correlation implies conditional independence.

Multi-gene pathways



How to condition on another gene?
Consider a three-gene pathway.
Suppose we wish to condition gene 1 on gene 3.



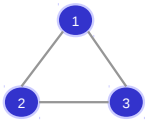
Recall

$$Y_i = \mathbf{X}_{i,*}\boldsymbol{\beta} + \varepsilon_i$$

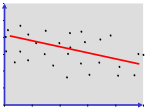
is equivalent to:

$$Y_i \mid \mathbf{X}_{i,*} \sim N(\mathbf{X}_{i,*}\boldsymbol{\beta}, \sigma^2)$$

Multi-gene pathways



How to condition on another gene?
Consider a three-gene pathway.
Suppose we wish to condition
gene 1 on gene 3.



Hence, regress gene 1 on gene 3:

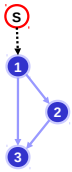
$$Y_{i,1} \quad = \quad \beta_0 + \beta_1 Y_{i,3} + \varepsilon_{i,1}$$

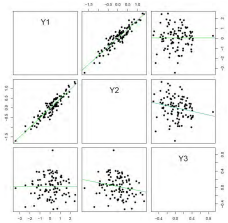
and "obtain":

$$Y_{i,1} \,|\, Y_{i,3} \quad \sim \quad \mathcal{N}(\beta_0 + \beta_1 Y_{i,3}, \sigma^2)$$

Multi-gene pathways

Example
A three-gene pathway with no
corrrelation between Y_1 and Y_3.

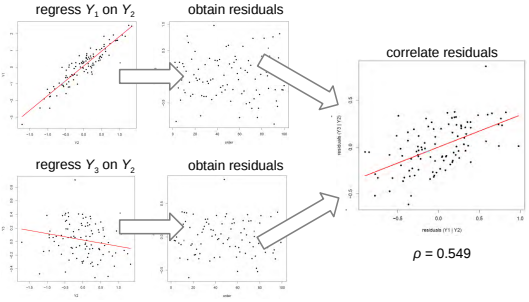




Calculate *partial* correlation between Y_1 and Y_3.

Multi-gene pathways

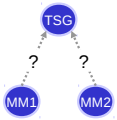
Example (continued)



Multi-gene pathways

Cancer research example
Y : gene expression measurements of a tumor suppressor gene

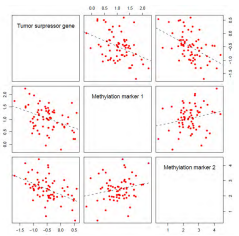
X_1 : gene expression of methylation marker 1
X_2 : gene expression of methylation marker 2



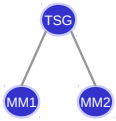
Question
Do the methylation markers (MMs) influence the expression of the tumor suppressor gene (TSG)?

Multi-gene pathways

Finally, the partial correlation method clearly indicates there is no conditional correlation between MM1 and MM2,.




```
> Sigma <- var(cbind(TSG, MM1, MM2))
> invSigma <- solve(Sigma)
> partCorMat <- cov2cor(invSigma)
> round(partCorMat, d=3)
       [,1]   [,2]   [,3]
[1,]  1.000 -0.342 -0.441
[2,] -0.342  1.000  0.032
[3,] -0.441  0.032  1.000
```



Multi-gene pathways

What about regression?
Regress the expression data of each gene on that of all other genes.

$$Y_1 = b_{01} + b_{21}Y_2 + b_{31}Y_3 + e_1$$

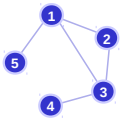
$$Y_2 = b_{02} + b_{12}Y_1 + b_{32}Y_3 + e_2$$

$$Y_3 = b_{03} + b_{13}Y_1 + b_{23}Y_2 + e_3$$

2 ⋯⋯ 3 : Y_3 response, Y_2 explanatory

Multi-gene pathways

Partial correlation vs. regression
Partial correlations and regression coefficients are one-to-one related.



Zero partial correlations indicate conditional independence.

Hence, zero regression coefficients also indicate conditional independence.

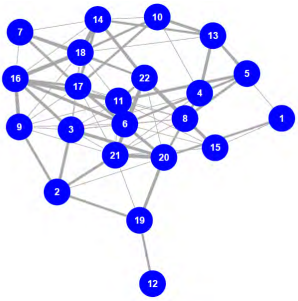
In fact, regression and partial correlations yield the same conditional independence graph (≈ network).

So what?

So what?

Study the reconstructed pathway topology

E.g. which genes interact?

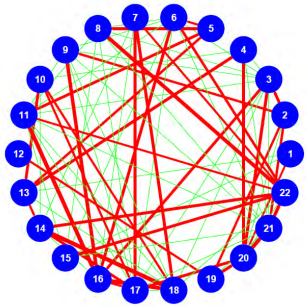


So what?

Study the reconstructed pathway topology

E.g. which genes interact?

Visualization is important!

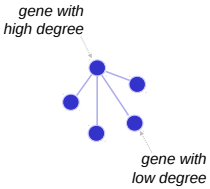


Red edge: ρ > 0
Green edge: ρ < 0

So what?

From network (properties) to biological function

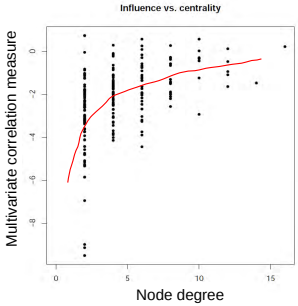
Hub ≈ gene with many connections.



Question: what is the role of the hub?

So what?

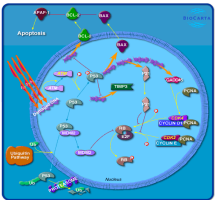
Hub has large influence on network



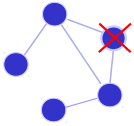
Hypothesis
Hubs are disease genes

So what?

Study the effect of changes in the regulatory system.



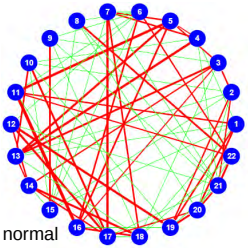
interpretation



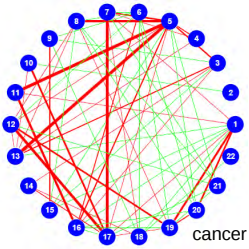
knock-out

So what?

Network differences between two conditions



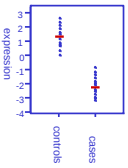
normal



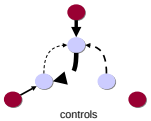
cancer

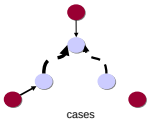
So what?

Traditional medicine
Biomarker is a gene or a locus.



Network medicine
Biomarker is an interaction between genes or some other topological feature.





High-dimensional data

High-dimensional data

Big data vs. high-dimensional data

Big data:
→ large *n*: many individuals, large sample size.
→ large *p*: information on many traits of these individuals.

Google:
→ *n* large: many people use Google software.
→ *p* large: Google registers everything these *n* do.
Similarly, Facebook, ING, et cetera.

Why:
→ many individuals available.
→ information is cheap to gain.

High-dimensional data

Big data vs. high-dimensional data

High-dimensional data:
→ small *n*: few individuals, small sample size.
→ large *p*: information on many traits of these individuals.

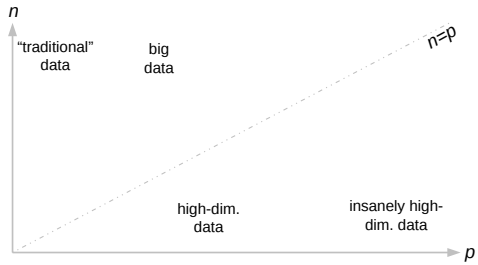
VUmc:
→ *n* small: few cancer patients.
→ *p* large: many genes.

Why:
→ individuals with particular disease not abound.
→ information is expensive.

VUmc

High-dimensional data

Big data vs. high-dimensional data



High-dimensional data

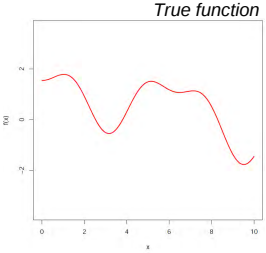
Big data vs. high-dimensional data

Nonlinear function f(x).

Estimate f(x) from:
→ big data,
→ high-dim. Data.

Form f(x) unknown.
Approximate f(x) by polynomial of degree *p* (number parameters).

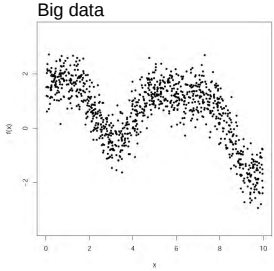
True function



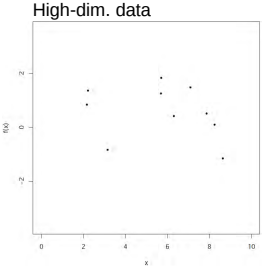
High-dimensional data

Big data vs. high-dimensional data

Big data



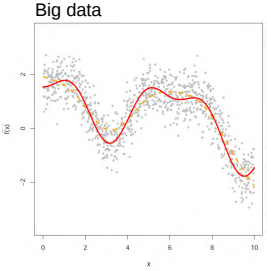
High-dim. data



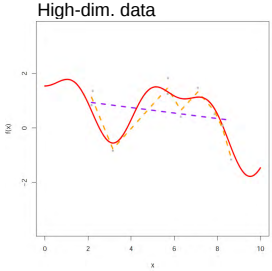
High-dimensional data

Big data vs. high-dimensional data

Big data



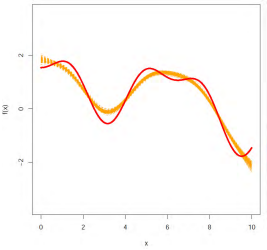
High-dim. data



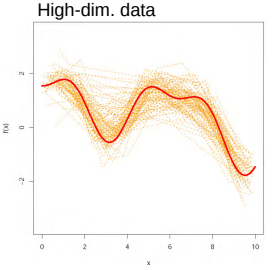
High-dimensional data

Big data vs. high-dimensional data

Big data



High-dim. data



High-dimensional data

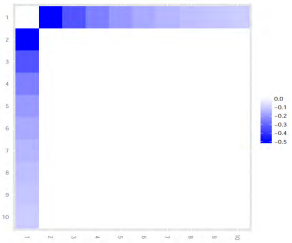
High-dimensionality & networks

In network reconstruction throughout assumed: n >> p.

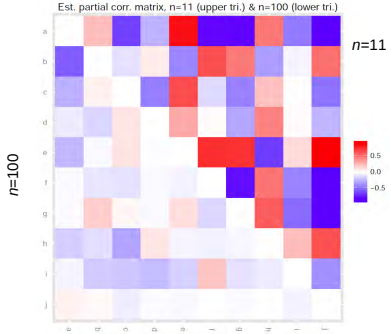
Study effect of high-dimensionality in the estimation of partial correlations.

Network of 10 genes.

Visualization of true partial correlations stucture: ⟶



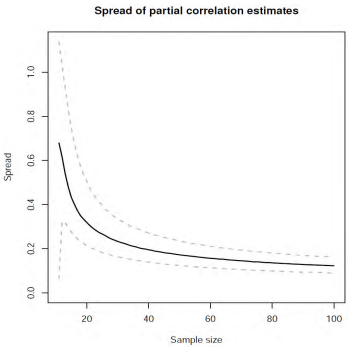
High-dimensional data



High-dimensional data

No incident

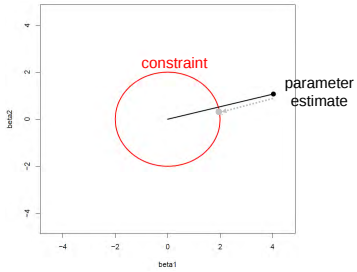
Partial correlation estimates inflate with decreasing sample size.



High-dimensional data

Solution

Penalization: discourage “large” parameter estimates.

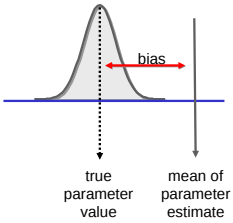


High-dimensional data

Solution
Penalization: discourage "large" parameter estimates.

What is "large"?
→ determined in
data-driven fashion.

Effect of penalization
→ well-defined estimates,
→ with reduced variance,
→ but … biased!

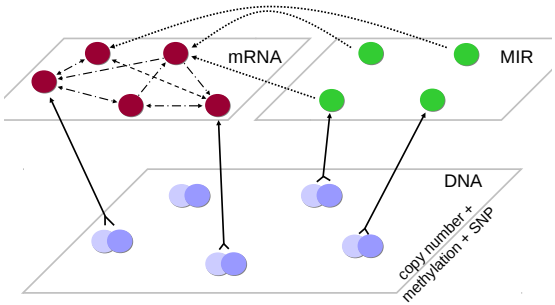


What's next?

Integration
(even larger *p*)

Integration

Include more molecular levels.

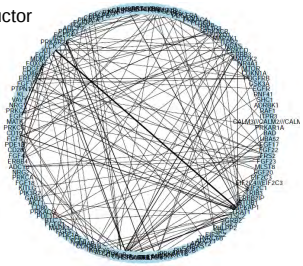


Integration

Reconstruct the ErbB signalling pathway

Available
Off the shelf from Bioconductor
4 breast cancer data sets

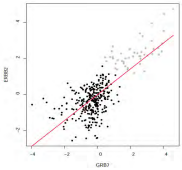
Reconstruction
Fit GGM with ridge penalty to pathway data. Post-hoc sparsification by local FDR procedure.

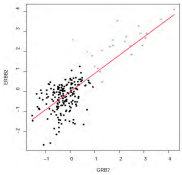


Integration

ErbB2 - GRB7 edge

→ Top ranking edge
→ ErbB2 often amplified
→ GRB7 maps to ErbB2
→ ErbB2 and GRB7 co-expressed



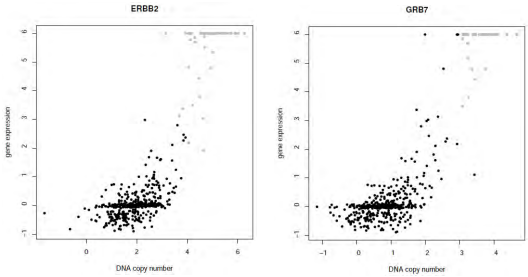


Breast cancer data set	*Marginal correlation*	*Partial correlation*
VDX	0.733	0.624
MAINZ	0.772	0.668
TRANSBIG	0.795	0.767
UPP	0.866	0.815

Integration

ErbB2 - GRB7 edge

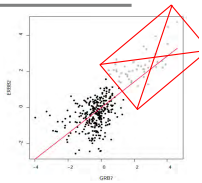
TCGA breast cancer data: copy number vs. expression



Integration

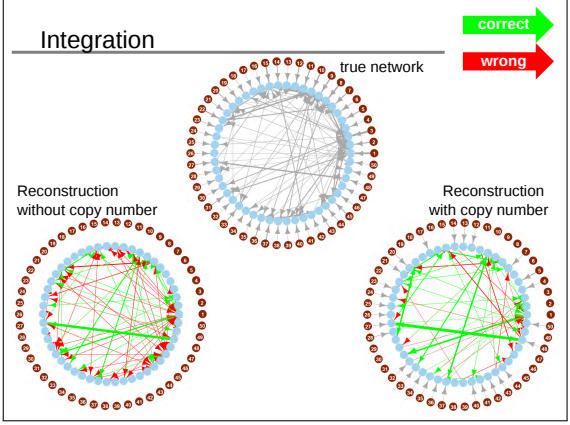
ErbB2 - GRB7 edge

→ Remove “amplified samples”
→ Reconstruct CIG of pathway
→ Amplification contributes to edge strength

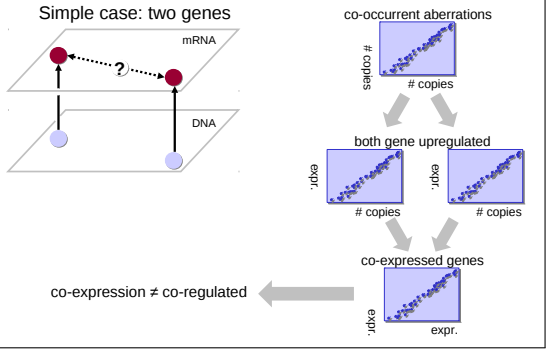


Breast cancer data set	*Marginal correlation*	*Partial correlation*	*Marg. cor. ampl. rem.*	*Part. cor. ampl. rem.*	*Rank*
VDX	0.733	0.624	0.340	0.238	69 (out of 9453)
MAINZ	0.772	0.668	0.357	0.293	521 (out of 9453)
TRANSBIG	0.795	0.767	0.428	0.314	457 (out of 9453)
UPP	0.866	0.815	0.370	0.305	237 (out of 11628)

Integration



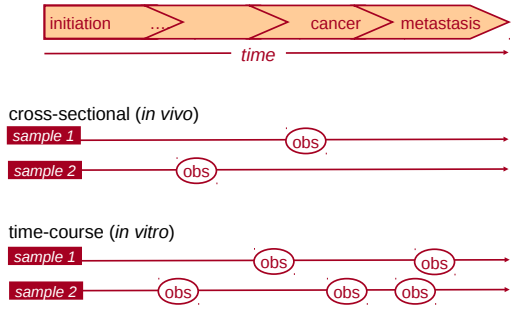
Integration



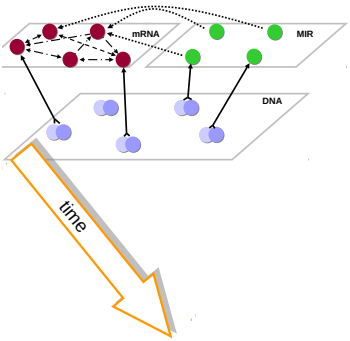
What's next?

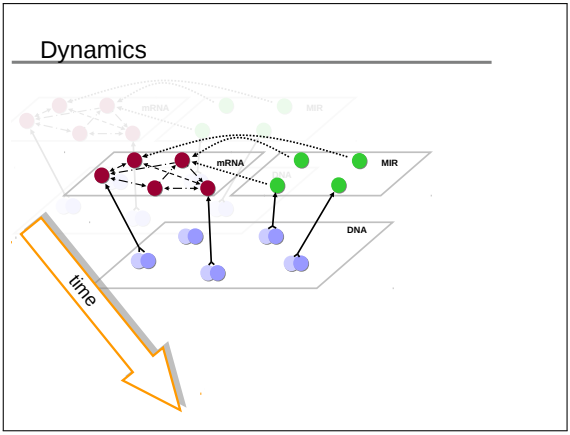
Dynamics

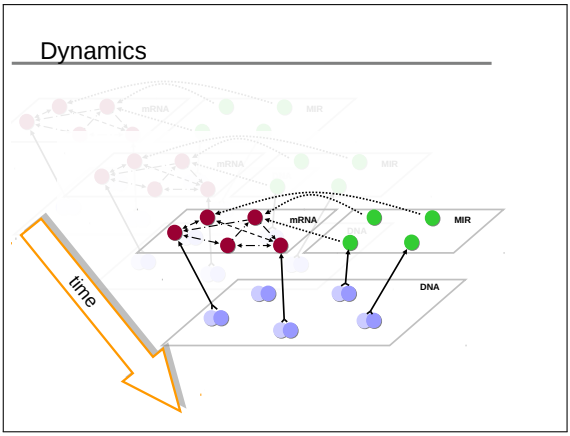
Dynamics

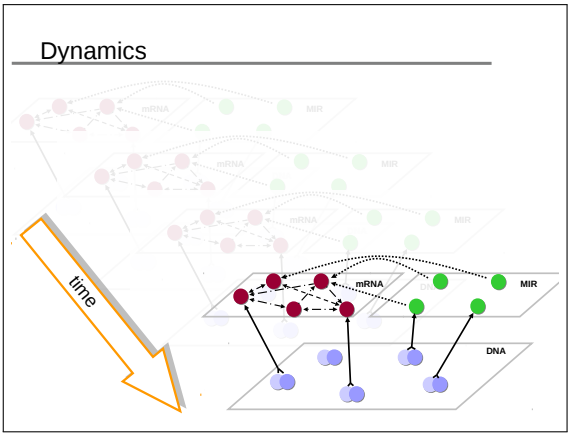


Dynamics




Dynamics
mRNA
MIR
DNA
time


Dynamics
mRNA
MIR
DNA
time


Dynamics
mRNA
MIR
DNA
time

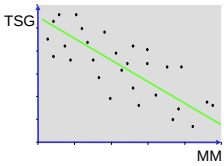
What's next?

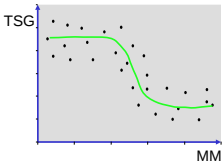
Nonlinearity

Nonlinearity

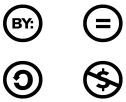
Assumed: all players in the network are linearly related.

Biologically, this is untenable. Hence, nonlinear dependencies need to be considered.





Slides by: Wessel van Wieringen
Input from: Mark van de Wiel and Carel Peeters



7 Practicum Data

Alessandro Di Bucchianico

Technische Universiteit Eindhoven

7.1 Doel van het practicum

Het doel van dit practicum is een kennismaking met verscheidene aspecten van data-analyse. Er is gekozen voor opdrachten die ook uitvoerbaar zijn door middelbare scholieren. De opdracht is een aangepaste versie van een opdracht die op de Technische Universiteit Eindhoven gegeven wordt in het eerstejaars vak Effectiveness of Mathematics. In dit vak laten we studenten kennismaken met allerlei toepassingen van wiskunde en proberen we over te brengen hoe wiskunde een bijdrage kan leveren. Omdat dit vak gegeven wordt tijdens de eerste weken van het eerstejaar, is de gebruikte wiskunde op VWO-niveau. Samen met mijn collega's Remco van der Hofstad en Eduard Belitser heb ik de statistiekopdracht van dit vak opgezet.

In het practicum komen de volgende aspecten van data-analyse aan de orde

- verzamelen van data
- representativiteit steekproef
- exploratieve data-analyse, nut van grafieken
- omgaan met software
- interpretatie van resultaten

7.2 Context

Door technologische ontwikkelingen m.b.t. computers en het internet is er een enorme toename aan data afkomstig van netwerken. Netwerken komen voor in allerlei gebieden, zoals in bijvoorbeeld sociale en professionele media (Facebook, LinkedIn, ResearchGate,...), biologie (gen-gen interactienetwerken, metabolische netwerken, ...), computernetwerken (internet, WWW, ...). Onlangs heeft NWO een grote subsidie (Zwaartekrachtsubsidie) toegekend voor wiskundig onderzoek naar netwerken, zie

`http://www.thenetworkcenter.nl/`

Dit toont niet alleen aan dat er groot belang aan dit thema wordt gehecht, maar ook dat er in Nederland veel expertise is op dit gebied.

7.3 Opzet

Netwerken worden wiskundig gemodelleerd als een graaf. Een graaf bestaat uit knopen en kanten. De knopen zijn de elementen uit het netwerk, de kanten beschrijven welke knopen met elkaar verbonden zijn. In dit practicum bestuderen we de collaboratiegraaf van wiskundigen, waarin de knopen de wiskundigen zijn, en er een kant is tussen twee wiskundigen indien zij samen een wetenschappelijk artikel hebben geschreven. Hiervoor zal U zelf data moeten verzamelen via de database van wiskunde-artikelen MathSciNet, zie

`http://www.ams.org/mathscinet/`

voor meer informatie. Op deze site kunt U met behulp van de tool op

`http://www.ams.org/mathscinet/collaborationDistance.html`

afstanden tussen twee wiskundigen berekenen. De tool geeft zelfs het pad van artikelen dat de twee wiskundigen verbindt. Een vergelijkbare functionaliteit wordt geboden door Microsoft Academic Search:

`http://academic.research.microsoft.com`.

Als men een wiskundige kiest, dan verschijnt er aan de linkerkant een menu waarin men de optie 'Co-author graph' kan kiezen. Dan ontstaat er

een grafische representatie van de collaboratiegraaf (ook interessant zijn de opties 'Co-author path' en 'Citation path').

Op de website van het 'Erdös Number Project'

`http://www.oakland.edu/enp/`

kunt U meer informatie vinden over de eigenschappen van de collaboratiegraaf tussen wiskundigen.

7.4 Wiskundige achtergrond

Om statistische uitspraken te doen, worden uitkomsten van waarnemingen gemodelleerd via kansverdelingen. Een belangrijke stap bij modelleren is de keuze van een passende kansverdeling. De keuze van een kansverdeling kan op verschillende manieren gebeuren. Zo blijkt bij analyse van netwerk data dat de graad D van een knoop (dat is het aantal kanten verbonden aan een knoop) vaak een zogenaamde 'power-law distribution' volgt, d.w.z. een kansverdeling van de vorm $P(D \geq k) = c_\tau k^{-\tau+1}$ met $\tau > 1$. Voor sommige klassen van random grafen kan deze eigenschap bewezen worden.

Zoals bekend treden binomiale verdelingen op als er men het aantal successen telt bij onafhankelijke, gelijkverdeelde experimenten met twee mogelijke uitkomsten.

De zogenaamde normale verdeling (ook wel Gaussverdeling genoemd) is een kansverdeling waarvan de kansdichtheid de bekende klokvorm heeft:

$$f(x) = \frac{e^{-(x-\mu)^2/2\sigma^2}}{\sqrt{2\pi\sigma^2}}.$$

De parameter μ is de verwachte waarde van deze verdeling, terwijl σ de standaardafwijking is. Indien $\mu = 0$ en $\sigma = 1$ spreken we van een *standaardnormale* verdeling (notatie: $Z \sim N(0, 1)$).

Het gebruik van de normale verdeling vaak wordt gerechtvaardigd door een beroep te doen op de centrale limietstelling. Een eenvoudige vorm van deze stelling zegt dat de kansverdeling van steekproefgemiddelden voor grote steekproefomvangen bij benaderingen normaal verdeeld is.

7.4.1 Steekproefomvang

Een steekproefgemiddelde geeft niet aan hoe betrouwbaar dat getal is. In de statistiek werkt men met betrouwbaarheidsintervallen waardoor we met een bepaalde kans (bijv. $0,95$) weten dat de gezochte grootheid in een interval ligt. Als data een breed betrouwbaarheidsinterval oplevert, dan geeft dit aan dat we weinig zekerheid hebben. Betrouwbaarheidsintervallen kunnen ook gebruikt worden voordat data verzameld wordt om te bepalen hoeveel data men moet verzamelen om met een gegeven nauwkeurigheid uitspraken te doen. We zullen dit illustreren aan de hand van zowel de normale verdeling als de binomiale verdeling.

Als we aannemen dat (gemiddelden van) waarnemingen normaal verdeeld zijn met een bekende standaardafwijking σ (bijv. door gebruik te maken van de centrale limietstelling), dan kunnen we afleiden dat voor deze steekproefgemiddelden $\overline{X}_n$ geldt dat :

$$P\left(\overline{X}_n - z_{\alpha/2}\frac{\sigma}{\sqrt{n}} < \mu < \overline{X}_n + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right) = 1-\alpha. \tag{7.1}$$

De waarden $z_{\alpha/2}$ zijn zogenaamde quantielen van de standaardnormale verdeling. Ze worden eenduidig vastgelegd via de relatie $P(Z > z_\alpha) = \alpha$. Deze waarden moeten numeriek benaderd worden. Ze zijn beschikbaar via tabellen, software of via applets op het WWW. Veel gebruikte waarden zijn $z_{0,025} = 1,96$ en $z_{0,05} = 1,645$. Een $100(1-\alpha)\%$-betrouwbaarheidsinterval voor μ als σ bekend wordt verondersteld, wordt dus gegeven door

$$\left(\overline{X}_n - z_{\alpha/2}\frac{\sigma}{\sqrt{n}}, \overline{X}_n + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right).$$

Als we met $100(1-\alpha)\%$ betrouwbaarheid willen hebben dat onze schatter $\overline{X}_n$ maximaal E verschilt van de echte, onbekende μ, dan moet volgens (7.1) gelden dat

$$z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \leq E \Rightarrow n \geq \left(\frac{z_{\alpha/2}\,\sigma}{E}\right)^2. \tag{7.2}$$

Als de standaardafwijking σ niet bekend is, dan geldt de formule bij benadering als men σ vervangt door de steekproefstandaardafwijking.

We gaan nu een vergelijkbare formule presenteren voor het geval men niet in een gemiddelde, maar in een fractie of proportie met een bepaald kenmerk K geïnteresseerd is. Een eenvoudig kansmodel hiervoor

is dat $X_1, \ldots, X_n$ onafhankelijk stochasten zijn met $P(X_i = 1) = p$ en $P(X_i = 0) = 1 - p$, waarbij $X_i = 1$ aangeeft dat de waarneming kenmerk K heeft. Dan geldt dat $X = \sum_{k=1}^{n} X_k$ een binomiale verdeling heeft met parameters n en p (de succeskans). De onbekende werkelijke fractie p schatten we dan met de waargenomen fractie:

$$\widehat{p} = \frac{\text{aantal waarnemingen met kenmerk } K}{\text{totaal aantal waarnemingen}} = \frac{X}{n} = \bar{X}_n,$$

Voor n voldoende groot vinden we m.b.v. de centrale limietstelling

$$P\left(\widehat{p} - \frac{z_{\alpha/2}\sqrt{\widehat{p}(1-\widehat{p})}}{\sqrt{n}} \leq p \leq \widehat{p} + \frac{z_{\alpha/2}\sqrt{\widehat{p}(1-\widehat{p})}}{\sqrt{n}}\right) \approx 1 - \alpha.$$

We hebben dus een $100(1-\alpha)\%$ betrouwbaarheidsinterval voor de onbekende fractie of proportie van de vorm

$$\left(\widehat{p} - \frac{z_{\alpha/2}\sqrt{\widehat{p}(1-\widehat{p})}}{\sqrt{n}}, \widehat{p} + \frac{z_{\alpha/2}\sqrt{\widehat{p}(1-\widehat{p})}}{\sqrt{n}}\right) \tag{7.3}$$

Als we met $100(1-\alpha)\%$ betrouwbaarheid willen hebben dat onze schatter $\overline{X}_n$ maximaal E verschilt van de echte, onbekende p, dan moet volgens (7.3) gelden dat

$$\frac{z_{\alpha/2}\sqrt{\widehat{p}(1-\widehat{p})}}{\sqrt{n}} \leq E \Rightarrow n \geq \left(\frac{z_{\alpha/2}}{E}\right)^2 \widehat{p}(1-\widehat{p}).$$

Deze benaderende formule heeft het praktische bezwaar dat we van tevoren $\widehat{p}$ niet weten omdat we nog geen data hebben. Als we redelijkerwijs een onder- en bovengrens kunnen aangeven voor de onbekende p, dan kunnen we toch een ondergrens aangeven voor de minimale steekproefomvang n. Dit is gebaseerd op elementaire eigenschappen van de functie $p \mapsto p(1-p)$ op $[0, 1]$. Deze eigenschappen kunnen we ook gebruiken voor het geval we geen grenzen voor p kunnen aangeven en veiligheidshalve van het ongunstigste geval uit willen gaan. Dit laten we om didactische redenen over aan de lezer.

7.5 Software

Het practicum is bewust zo opgezet dat de opdrachten met eenvoudige software uitgevoerd kunnen worden zoals bijv. Excel. Indien Excel gebruikt

wordt, dan is het wel handig om de Data Analysis Toolpak te activeren (dit staat bij Opties onder Invoegtoepassingen (Add-ins bij Engelstalige versies)).

Het Analysis Toolpak bevat onder Gegevensanalyse (Data Analysis bij Engelstalige versies) o.a. een optie om histogrammen te maken en eenvoudige statistische analyses uit te voeren.

Er is overigens allerlei gratis statistische software van uitstekende kwaliteit, o.a. `R` (zie `http://www.r-project.org`) of PSPP (zie `http://www.gnu.org/software/pspp/`).

Voor wie is PWN interessant?

Beroepswiskundigen
Wiskundeleraren
Bedrijven
Leerlingen en studenten
Breed publiek

Platform Wiskunde Nederland is hét landelijke loket voor alles wat met wiskunde te maken heeft.

PWN behartigt de belangen van, en fungeert als spreekbuis voor, de gehele Nederlandse wiskunde.

Platform Wiskunde Nederland | Science Park 123 | kamer L013 | 1098 XG Amsterdam | 020 592 40 06

Ga voor meer informatie naar:
www.platformwiskunde.nl

platform
wiskunde nederland