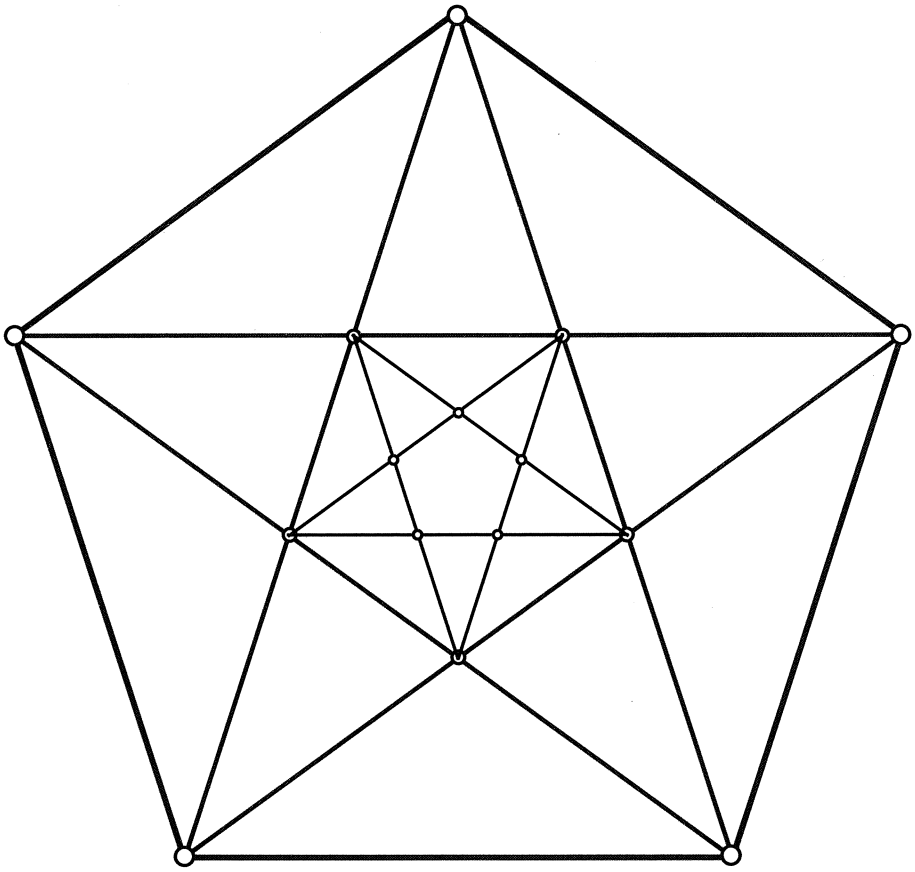


CWI Syllabi

Managing Editors

K.R. Apt (CWI, Amsterdam)
M. Hazewinkel (CWI, Amsterdam)
J.K. Lenstra (Eindhoven University of Technology)

Editorial Board

W. Albers (Enschede)
P.C. Baayen (Amsterdam)
R.C. Backhouse (Eindhoven)
E.M. de Jager (Amsterdam)
M.A. Kaashoek (Amsterdam)
M.S. Keane (Delft)
H. Kwakernaak (Enschede)
J. van Leeuwen (Utrecht)
P.W.H. Lemmens (Utrecht)
M. van der Put (Groningen)
M. Rem (Eindhoven)
H.J. Sips (Delft)
M.N. Spijker (Leiden)
H.C. Tijms (Amsterdam)

CWI
P.O. Box 94079, 1090 GB Amsterdam, The Netherlands
Telephone 31-20 592 9333, telex 12571 (mactr nl),
telefax 31-20 592 4199

CWI is the nationally funded Dutch institute for research in Mathematics and Computer Science.

Vakantiecursus 1993
Het reële getal

ISBN 90 6196 423 7
NUGI-code: 811

Copyright © 1993, Stichting Mathematisch Centrum, Amsterdam
Printed in the Netherlands

TEN GELEIDE,

Waar de vacantiecursus in het jaar 1992 de deelnemers uit hun dagelijkse werkomgeving voerde naar het terrein van de toepassingen van de wiskunde -i.c. de systeemtheorie-, richt de cursus 1993 zich op wat men wel mag noemen het dagelijkse gereedschap van iedere wiskundige: het reële getal. Hiervan zullen- in goede traditie- velerlei aspecten worden belicht. Om te beginnen de eerste aanzetten in de Griekse oudheid, waar de ontdekking van onmeetbare verhoudingen een immense schok teweeg bracht, die door Tannery werd gekarakteriseerd als een "scandale logique"; daarna de exacte fundering van het irrationale getal met de constructie van Dedekind, die zo nauw aansluit bij de gedachten van de Grieken.
Van de fundamentele vragen die het reële getal oproept krijgt de continuumhypothese bijzondere aandacht. Bestaat er een deelverzameling van de verzameling $\mathbb{R}$ van alle reële getallen met een "aantal" (kardinaalgetal) dat ligt tussen dat van de rationale getallen en dat van $\mathbb{R}$? Met het antwoord kunt U twee kanten uit: U kunt het geloven of niet!
Een andere fundamentele vraag is: "hoe zien de constructief ingestelde intuitionisten het reële getal?" Ook aan dit probleem is een voordracht gewijd.
Uiteraard komt het reële getal als dagelijks gereedschap expliciet aan de orde en dan staan we met beide benen op de grond: de meeste rekenmachines van alledag werken uitsluitend met rationale getallen en dan rijst de vraag of de irrationale getallen wel zo "reëel" zijn. Daarbij is er volop gelegenheid een aantal interessante practische vraagstukken te proberen: hoeveel is $\sqrt{2}.\sqrt{3} - \sqrt{3}.\sqrt{2}$?
Bij dit alles vraagt een rechtgeaarde leraar zich af: "Hoe vertel ik het mijn kinderen?" Welnu, ook op deze vraag zal worden ingegaan.
Een vacantiecursus zou niet compleet zijn als er niet ook aandacht en plaats zou zijn ingeruimd voor alternatieven en perspectieven. Ook daarin is voorzien. Naast "onze" reële getallen worden ook de zgn. p-adische getallen ten tonele gevoerd. Ook daarmee kan men analyse bedrijven.
De slotvoordracht plaatst het geheel in een ruimer kader. Natuurlijke getallen, rationale getallen, irrationale getallen, "oneindig kleine" en "oneindig grote" getallen (de zgn. non-standaard getallen) worden in één definitieschema samengevat.

Evenals in voorgaande jaren koestert de voorbereidingscommissie de hoop en de verwachting dat ook deze cursus velen een frisse kijk zal geven op bekende zaken en dat de cursus tot verdere studie zal aanzetten en wellicht een extra impuls zal zijn bij het doceren.

Tot slot nog dit: ieder jaar nemen de sprekers zich voor nu eens zeer tijdig de copy van hun voordracht in te leveren. Gelukkig blijkt het ieder jaar weer dat, indien het niet iedereen gelukt dit goede voornemen te realiseren, toch weer de uitmuntende staf van medewerksters en medewerkers van het CWI erin slaagt de syllabus keurig op tijd en keurig uitgevoert te produceren. Daarom, zoals ieder jaar weer, mijn zeer hartelijke dank daarvoor. Uiteraard geldt deze dank ook degene die de registratie en de ontvangst van de deelnemers (m/v) verzorgden

Rest mij slechts zowel de deelnemers als de sprekers twee genoeglijke dagen toe te wensen.

A.W. Grootendorst

Inhoud

Ten geleide
A.W. Grootendorst

Eudoxus en Dedekind 1
A.W. Grootendorst

P-adische getallen 23
W.H. Schikhof

De tussenwaardestelling in MAVO-3 35
A.J. Goddijn

De continuum-hypothese 55
J.M. Aarts

Een intuitionistische kijk op het reële getal 67
A.S. Troelstra

Bekende reële getallen 83
F. van der Blij

On Numbers and Games, chapters 0,1 & 2 101
J.H. Conway

Eudoxus en Dedekind

A.W. Grootendorst

1. *Het irrationale in de Griekse wiskunde vóór Euclides.*

1.1 Er zijn meerdere manieren om het reële getal in te voeren. In deze voordracht zal in hoofdzaak aandacht geschonken worden aan de wijze waarop Richard Dedekind (1831-1916) via zijn "Schnitte" het lichaam van de rationale getallen uitbreidde met de irrationale getallen tot het lichaam van de reële getallen [1.1].[1] De reden dat juist deze methode gekozen is als onderwerp van deze voordracht, ligt daarin dat hier een voorbeeld voor handen is hoe de diepe betekenis van een geniale gedachte uit de geschiedenis van de wiskunde, nl. de redentheorie van Eudoxus (ca. 400-347) na ruim 2200 (!) jaren werd ingezien en op even geniale wijze werd uitgewerkt door Dedekind [1.2].

1.2 Men vermoedt dat in de school van Pythagoras (560-480) het irrationale ontdekt is, waarbij het niet zeker is welke irrationaliteit het eerst gevonden werd: de onderlinge onmeetbaarheid van de lengte van zijde en diagonaal in het vierkant of in de regelmatige vijfhoek. In ieder geval kwam deze ontdekking als een grote schok aan, immers bij de Pythagoreërs gold de suprematie van het (natuurlijke) getal, zoals verwoord is door de Pythagoreër Philolaus [1.3].
Inderdaad heeft alles wat men kan kennen, een getal, want het is niet mogelijk iets te begrijpen of te kennen zonder het getal.
Traditioneel wordt de ontdekking van het irrationale toegeschreven aan Hippasus van Metapontum (ca. 520 - ca. 480), de eerste belangrijke wiskundige uit de school van Pythagoras. Volgens de overlevering [1.4] zou hij als straf voor de openbaarmaking van deze "gruwelijke" ontdekking de dood op zee gevonden hebben gevonden.

1.3 De onderlinge onmeetbaarheid van zijde en diagonaal in een regelmatige vijfhoek (zie afb. 1.2) kan worden bewezen met behulp van de stelling in El.X.2.[2]
Hier wordt de bekende Euclidische algoritme ("delen met rest") - officieel anthyphaeresis of antanairesis genoemd - ingevoerd. We lezen daar:
Indien men van twee ongelijke grootheden steeds afwisselend de kleinste van de grootste aftrekt en de rest nooit af te passen is op de voorgaande, dan zullen deze grootheden onderling onmeetbaar zijn.

[1] Verwijzingen naar de aantekeningen zijn tussen [ ] geplaatst.
[2] El.X.2 verwijst naar hoofdstuk 2 van boek X van de Elementen van Euclides.

In moderne taal betekent dit dat a_0 en a_1 onderling onmeetbaar zijn indien het volgende schema (Euclidische algoritme) nooit afbreekt.

$$\begin{aligned} a_0 &= a_1 q_0 + r_1 \quad 0 < r_1 < a_1 \\ a_1 &= r_1 q_1 + r_2 \quad 0 < r_2 < r_1 \\ r_1 &= r_2 q_2 + r_3 \quad 0 < r_3 < r_2 \\ &\vdots \end{aligned}$$

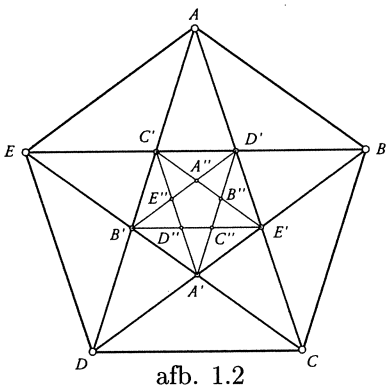
Zou dit schema wel eindigen, bijv. met

$$r_{n-1} = r_n q_n$$

dan zien men eenvoudig in dat r_n de g.g.d. van a_0 en a_1 is.
Dit passen we toe op de regelmatige vijfhoek in afb. 1.2.

afb. 1.1: Euclides

Men ziet eenvoudig in: $AE = ED' = DC$ en $BD' = D'A' = A'D$.
Stelt men in vijfhoek $ABCDE$ de diagonaal op d_1 en de zijde op a_1 en in vijfhoek $A'B'C'D'E'$ de diagonaal op d_2 en de zijde op a_2 en gaat men zo voort met steeds weer de "binnenste" vijfhoek, dan zien we:



afb. 1.2

$$EB = ED' + D'B = AE + D'A'$$

dus

$$d_1 = 1 \cdot a_1 + d_2 \quad ; \text{met} \quad 0 < d_2 < a_1$$

en analoog

$$d_2 = 1 \cdot a_2 + d_3 \quad \text{met} \quad 0 < d_3 < a_2$$

en men ziet dat deze keten niet afbreekt, dus zijde en diagonaal van een regelmatige vijfhoek zijn onderling ondeelbaar.
Tenslotte volgt uit de gelijkvormigheid van de driehoeken AEB en $A'DC$ dat $EB : AE = DC : A'D$
maar $DA' = A'D' = D'B = d_1 - a_1$, dus

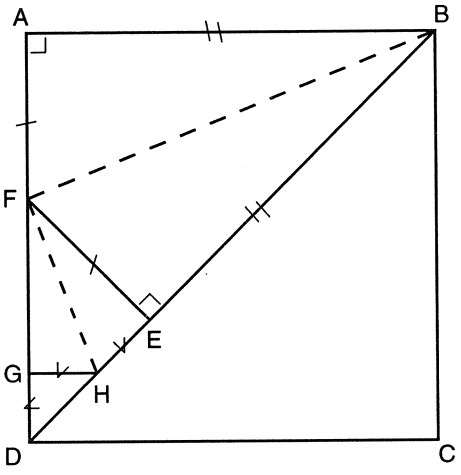
$$d_1 : a_1 = a_1 : (d_1 - a_1)$$

$$a_1^2 + d_1 a_1 - d_1^2 = 0$$

$$a_1 = \frac{-1+\sqrt{5}}{2} \cdot d_1$$

dus $\frac{-1+\sqrt{5}}{2}$ en dus ook $\sqrt{5}$ is irrationaal. De verdeling van d_1 in a_1 en $d_1 - a_1$ is de bekende verdeling in uiterste en middelste reden.

1.4 De methode van de oneindig voortlopende Euclidische algoritme kunnen we ook gebruiken om de onderlinge ondeelbaarheid van zijde en diagonaal in een vierkant aan te tonen. Zie daarvoor afb. 1.3.



afb. 1.3

We beginnen met de zijde AB af te passen op de diagonaal BD. De rest ED moet nu worden afgepast op AB, maar daarvoor kunnen we ook AD nemen. Aangezien kennelijk geldt: $AF = FE = ED$, zien we dat we ED tweemaal kunnen afpassen op AD met rest GD. Echter $GD = GH$, dus het komt er op neer dat we GH moeten afpassen op ED. Omdat echter $GH = HE$, rest ons dat we GH nog moeten afpassen op DH, maar dan zijn we in dezelfde situatie als in het begin, waaruit volgt dat dit proces nooit eindigt. Hieruit blijkt dat de verhouding $BD : AB = \sqrt{2}$ irrationaal is. Dit zullen we nog op 3 andere manieren aantonen.

1.5 De eerste daarvan is alom bekend en stamt uit een supplement op El. X 115 [1.5]. In moderne notatie:
Stel $\sqrt{2} = \frac{t}{n}$ met natuurlijke t en n en $(t, n) = 1$.
Dan geldt: $t^2 = 2n^2$, dus t^2 en derhalve t even, bijv. $t = 2s$. Hieruit volgt dan $n^2 = 2s^2$ dus ook n^2 en n even, in strijd met de onderstelling dat $(t, n) = 1$.

1.6 Een fraai bewijs valt te ontlenen aan Aristoteles [1.6]; het vereist de mogelijkheid van ontbinding in priemfactoren en is uit te breiden tot de stelling dat een getal dat geen kwadraat is van een natuurlijk getal, ook niet het kwadraat is van een rationaal getal.
Stel $D \neq m^2 (m \in \mathbb{N})$, dan bevat D een oneven aantal priemfactoren. Als nu $D = (\frac{t}{n})^2$ met natuurlijke t en n, dan zou $n^2 D = t^2$. In het linkerlid staat dan een oneven aantal priemfactoren, in het rechterlid een even aantal. Dit levert een contradictie.

1.7 Tot slot het bewijs dat Dedekind gaf van de irrationaliteit van $\sqrt{D}$ als D een natuurlijk getal is dat niet zelf het kwadraat is van een natuurlijk getal. [1.7]
Daar dit bewijs weinig bekend is, wordt het hier weergegeven. Het verloopt aldus:
Als D het kwadraat is van een rationaal getal, dan zijn er twee natuurlijke getallen t en n met $(t, n) = 1$, waarvoor geldt

$$t^2 - Dn^2 = 0. \qquad (*)$$

Laat nu n_0 het kleinste natuurlijke getal zijn dat hieraan voldoet. Voor zekere $\lambda \in \mathbb{N}$ geldt voor de bij die n_0 behorende t_0:

$$\lambda n_0 < t_0 < (\lambda + 1)n_0$$

dus, als we stellen

$$n_1 = t_0 - \lambda n_0$$

dan geldt:

$$0 < n_1 < n_0.$$

Stelt men verder

$$t_1 = Dn_0 - \lambda t_0,$$

dan geldt $t_1 > 0$ en

$$t_1^2 - Dn_1^2 = (\lambda^2 - D)(t_0^2 - Dn_0^2) = 0.$$

Daar echter $0 < n_1 < n_0$ levert dit een tegenspraak met de onderstelling dat n_0 het kleinste natuurlijke getal is dat aan (*) voldoet.

2. *Verhouding en evenredigheid van getallen bij Euclides: Eudoxus.*

2.1 Het zevende boek van de Elementen van Euclides (ca. 300 v. C.) is het eerst van de drie getallentheoretische boeken (VII, VIII, IX) die als een merkwaardige enclave voorkomen in het meetkundige werk van Euclides. Met getal wordt daarin steeds - in goede Pythagorische traditie - bedoeld: natuurlijk getal. Dit blijkt al direct uit de eerste twee definities van dit boek [2.1]:

ζ′

OPOI

α′. Μονάς ἐστιν, καθ' ἣν ἕκαστον τῶν ὄντων ἓν λέγεται.

β′. Ἀριθμὸς δὲ τὸ ἐκ μονάδων συγκείμενον πλῆθος.

γ′. Μέρος ἐστὶν ἀριθμὸς ἀριθμοῦ ὁ ἐλάσσων τοῦ μείζονος, ὅταν καταμετρῇ τὸν μείζονα.

δ′. Μέρη δέ, ὅταν μὴ καταμετρῇ.

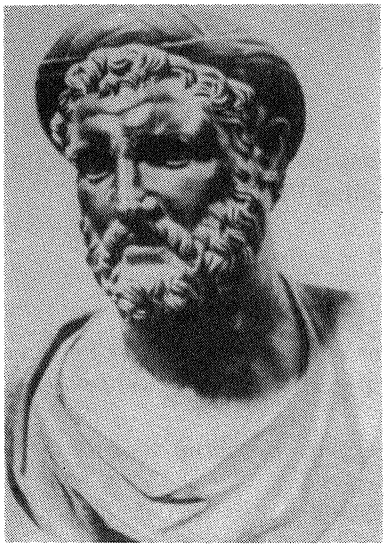
ε′. Πολλαπλάσιος δὲ ὁ μείζων τοῦ ἐλάσσονος, ὅταν καταμετρῆται ὑπὸ τοῦ ἐλάσσονος.

ς′. Ἄρτιος ἀριθμός ἐστιν ὁ δίχα διαιρούμενος.

ζ′. Περισσὸς δὲ ὁ μὴ διαιρούμενος δίχα ἢ [ὁ] μονάδι διαφέρων ἀρτίου ἀριθμοῦ.

afb. 2.1

Eenheid is datgene op grond waarvan elk van de dingen "één" genoemd wordt.
en
Een getal (arithmos, staat daar) *is een hoeveelheid, samengesteld uit eenheden.*



afb. 2.2: Pythagoras

In de school van Pythagoras (ca. 560 - ca. 480 v. C.), die duidelijk zijn stempel op de Elementen drukte, gold 1 niet als getal, maar als grondslag daarvan, zoals een steen, waaruit een muur is opgebouwd zelf geen muur is. Sommige Pythagoreërs sloten zelfs 2 uit als getal. Getal was in eerste instantie een aantal. Velen schrijven de kerngedachten van het zevende boek van de Elementen toe aan Theaetetus (circa 417-369). Sommigen doen dat met meer zekerheid dan anderen. Vast staat echter dat hij beschikte over de leer van de evenredigheden uit de Pythagorische school en bekend was met het werk van Hippocrates van Chios (circa 450 v. C.) en Archytas van Tarente (circa 375 v. C.).

Theaetetus stond in de oudheid bekend als in alle opzichten briljant. Hij is vermoedelijk de ontdekker van het regelmatige achtvlak en twintigvlak,

gold als grondlegger van de stereometrie en ontdekte de irrationaliteiten in de rij $\sqrt{3}, \sqrt{5}, \ldots, \sqrt{17}$. Ook was hij de hoofdfiguur in de gelijknamige dialoog van Plato die gewijd is aan de kennisleer.
Hij vond de dood in 369 v. C. ten gevolge van zijn deelname aan een veldslag bij Corinthe.

2.2 Na zorgvuldige definities van even, oneven, deelbaarheid, priem etc. volgt dan als twintigste definitie die van evenredigheid. Let wel niet van verhouding, maar in feite van de gelijkheid van niet-gedefinieerde verhoudingen.
VII Def. 20
Getallen zijn evenredig (analogon) wanneer het eerste van het tweede en het derde van het vierde evenzoveel maal veelvoud is of hetzelfde deel of dezelfde delen.
De eerste twee delen van deze definitie spreken voor zichzelf:
$A : B = C : D$ indien $A = tB$ en $C = tD$ of $B = tA$ en $D = tC$.
Voor een goed begrip van het derde deel moeten we iets verder in het boek VII lezen en dan blijkt dat met "dezelfde delen" bedoeld wordt

$$A = ad_1; B = bd_1 \quad \text{en} \quad C = ad_2; D = bd_2$$

waarbij d_1 de g.g.d is van A en B en d_2 de g.g.d. van C en D is.
Men ziet eenvoudig in dat dit geval beide voorgaande impliceert.

2.3 Met behulp van deze definitie worden dan de bekende eigenschappen afgeleid, zoals

$$a : b = c : d \iff (a \pm b) : (c \pm d) = a : b$$

$$a : b = c : d \iff ad = bd \text{ etc.}$$

Uiteraard is deze notatie een anachronisme. In de Elementen gaat het geheel verbaal toe. Zo leest men in stelling VII.11 in plaats van

$$a : b = c : d \implies (a - c) : (b - d) = a : b,$$

de fraaie volzin:
Indien het geheel tot het geheel staat zoals een afgenomen stuk tot een afgenomen stuk, dan zal de rest staan tot de rest zoals het geheel tot het geheel [2.2].
Met behulp van def. 20 zijn al deze eigenschappen op voor de hand liggende wijze af te leiden.
Voor ons doel is echter het belangrijkste dat men uit def. 20 een eigenschap kan afleiden die de band legt met de definitie van evenredigheid van "grootheden" (lengten, oppervlakken, inhouden) zoals Eudoxus (ca. 400 - ca. 347) die gaf, wellicht geïnspireerd door de definitie van Theaetetus.
De bedoelde eigenschap van de evenredigheid van getallen luidt:

Er geldt: $A : B = C : D$ (*)
dan en slechts dan als voor willekeurige natuurlijke getallen m en n geldt:

$$mA \gtreqless nB \iff mC \gtreqless nD \qquad (**)$$

Het bewijs is eenvoudig. Uit (*) volgt immers

$$A = ad_1,\ B = bd_1 \quad \text{en} \quad C = ad_2,\ D = bd_2$$

en dus

$$mA \gtreqless nB \iff mad_1 \gtreqless nbd_1 \iff mad_2 \gtreqless nbd_2 \iff mC \gtreqless nD.$$

Dat uit (**) ook (*) volgt, blijkt als we alleen letten op het gelijkteken in (**); dan zien we immers:

$$mA = NB \iff mC = nD \qquad (***)$$

Nu zijn A en B getallen, dus we mogen deze als resp. n en m kiezen. Aangezien $BA = AB$, volgt met (***):
$BC = AD$ en met één van de bovengenoemde stellingen hebben we dan $A : B = C : D$.

3. *Verhouding en evenredigheid van grootheden bij Euclides: Eudoxus.*

3.1 In het vijfde boek van de Elementen van Euclides vinden we een volledige redentheorie, d.w.z. een volledige theorie van evenredigheden voor grootheden, die - toegepast op gehele getallen - dezelfde resultaten geeft als de zojuist genoemde uit El.VII en dit dus in feite overbodig maakt. Merkwaardig is dat er geen antieke bronnen zijn die verband leggen tussen boek V en VII.
Men is het er algemeen over eens dat Eudoxus van het Kleinaziatische Cnidus (ca. 400 - ca. 347) de auctor intellectualis is van dit boek en niet alleen van dit deel van de Elementen. In een scholion (toelichting) schreef Proclus (410-485), na de opsomming van een aantal wiskundigen [3.1]:
Niet veel jonger dan deze is Euclides, die de Elementen samenstelde, veel van de resultaten van Eudoxus samenvatte, veel voltooide wat Theaetetus was begonnen en de minder strenge bewijzen van zijn voorganger in een niet te weerleggen vorm bracht. Met betrekking tot boek V zegt hij expliciet [3.2]
Sommigen zeggen dat dit boek de vinding is van Eudoxus, een leerling van Plato.
Van het vele werk dat volgens antieke getuigen van zijn hand verscheen rest ons nog slechts een aantal fragmenten [3.3].
Hij behoort echter tot de belangrijkste wiskundigen van de oudheid. Opgeleid o.a. door Archytas van Tarente en Plato (427-347) ontwikkelde hij

zich tot astronoom en wiskundige die niet alleen belangrijke ontdekkingen deed op het gebied van de wiskunde, maar deze ook toepaste op het gebied van de astronomie. Zo ontwikkelde hij een geocentrisch wereldbeeld waarin de aarde het middelpunt was van een stel concentrische bollen waarop de planeten zich bevonden en waarmee hij hun schijnbare bewegingen wist te verklaren. Hierbij toonde hij zich een meester in stereometrie, vooral waar het de meetkunde van de bol betrof.
Van zijn wiskundewerk staat in deze voordracht zijn redentheorie centraal. Daarnaast echter vermelden we dat Archimedes (287-212) aan hem de exhaustie methode toeschreef. Deze komt neer op het volgende: als men van een gegeven oppervlakte, zeg A, de helft of meer wegneemt, van de rest eveneens de helft of meer en dit procédé voortzet, houdt men op de duur minder over dan een vooraf gegeven oppervlakte. In moderne notatie: $\lim_{n\to\infty} A(1-r)^n = 0$ ($\frac{1}{2} \leq r < 1$). Hiermee toonde hij o.a. aan dat de oppervlakten van twee cirkels zich verhouden als de kwadraten van hun stralen; van hem stamt ook het bewijs van de formule voor de inhoud van een rechte cirkelkegel. Ook zou de axiomatische methode en de systematische presentatie, die zo kenmerkend zijn voor de Elementen van Euclides, van hem afkomstig zijn.

3.2 De grote betekenis van de nu te bespreken theorie ligt daarin dat deze niet beperkt blijft tot (natuurlijke) getallen, maar zich uitstrekt tot continu veranderlijke grootheden ook in het geval dat deze onderling geen gemene maat hebben. In feite hebben we hier een theorie waarin de definitie en de eigenschappen van het onmeetbare getal besloten liggen. De ontdekking daarvan zou echter voorbehouden blijven aan Richard Dedekind (1831-1916), die met grote genialiteit de theorie van El.V 22 eeuwen later - in 1858 - oppakte.

3.3 Boek V van de Elementen van Euclides zet in met de definities van deler en veelvoud van grootheden. Het begrip grootheid wordt daarbij niet gedefinieerd, maar in een scholion op boek V [3.4] leest men:
Een grootheid is dat wat in het oneindige kan worden vermeerderd en gedeeld. Er zijn daarvan drie soorten: lengten, oppervlakten, inhouden.
Daarna volgt (bij Euclides) de "definitie" van verhouding [3.5].
Een verhouding is een zekere betrekking tussen gelijksoortige grootheden inzake hun afmeting.
Met deze "definitie" valt uiteraard niet veel te beginnen. Het enige dat er uit blijkt is dat een verhouding geen getal is en zeker ook niet het quotiënt van twee getallen, want aan lengten etc. werden geen getallen toegekend. (Hoe zou dan toen ook gekund hebben?)
Ook het begrip "gelijksoortig" wordt hier niet verklaard, maar de scholiast [3.6] wijst er op dat
Niemand een lengte met een oppervlakte vergelijke.
en dat laatste blijkt ook uit Def.V,4:
Men zegt dat die grootheden een verhouding kunnen hebben die na vermenigvuldiging elkaar kunnen overtreffen.

Het gaat dus om verhoudingen van lengten onderling, oppervlakten onderling, inhouden onderling. In deze definitie wordt geformuleerd datgene wat later bekend geworden is als het axioma van Eudoxus/Archimedes en dat aldus luidt:
Bij twee gelijksoortige grootheden A en B kan men steeds natuurlijke getallen n en m vinden z.d.d.

$$nA > B \quad \text{en} \quad mB > A.$$

Het gaat in feite om het uitsluiten van "oneindige kleine" grootheden.

3.4 Hoewel dus het begrip verhouding van grootheden niet op operationele wijze is gedefinieerd, wordt in de beroemde Def.V,5 het begrip "gelijkheid" van verhoudingen vastgelegd:
Grootheden worden gezegd dezelfde verhouding te hebben, de eerste tot de tweede en de derde tot de vierde, wanneer willekeurige, gelijke veelvouden van de eerste en de derde tegelijkertijd groter zijn dan, gelijk zijn aan, of kleiner zijn dan willekeurige gelijke veelvouden van de tweede en de vierde, in dezelfde volgorde genomen.

ε'. Ἐν τῷ αὐτῷ λόγῳ μεγέθη λέγεται εἶναι πρῶτον πρὸς δεύτερον καὶ τρίτον πρὸς τέταρτον, ὅταν τὰ τοῦ πρώτου καὶ τρίτου ἰσάκις πολλαπλάσια τῶν τοῦ δευτέρου καὶ τετάρτου ἰσάκις πολλαπλασίων κατ' ὁποιονοῦν πολλαπλασιασμὸν ἑκάτερον ἑκατέρου ἢ ἅμα ὑπερέχῃ ἢ ἅμα ἴσα ᾖ ἢ ἅμα ἐλλείπῃ ληφθέντα κατάλληλα.

Deze volzin zouden wij aldus in formule samenvatten.
Voor de grootheden A, B, C, D geldt dat:

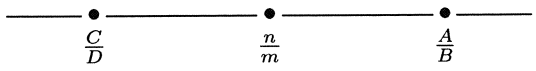
$$A : B = C : D$$

dan en slechts dan als voor alle $m, n \in \mathbb{N}$ geldt:

$$mA \gtreqless nB \iff mC \gtreqless nD.$$

Zoals gezegd hadden de Grieken voor de verhouding van twee grootheden A en B geen notatie; als wij die dan toch aangeven met $A : B$, dan mogen we - in hun gedachtengang - zeker niet aan een getal denken!
De genoemde definitie ligt in het verlengde van (is geïnspireerd door?) de eigenschap die we in §2.3 afleidden voor natuurlijke getallen en die volgde uit de definitie van evenredigheid van natuurlijke getallen.
Def.V,6 zegt dan dat grootheden die dezelfde verhouding hebben "evenredig" (analogon) genoemd worden.
Direct daarna wordt de ordening van verhoudingen ingevoerd door Def.V,7, die in onze taal zou luiden:
$A : B > C : D$ *dan en slechts dan als er* $n, m \in \mathbb{N}$ *bestaan z.d.d.* $mA > nB$ *en tevens* $mC < nD$.
Deze definitie lijkt in eerste instantie ondoorzichtig, maar als we even brutaalweg bij $A : B$ en $C : D$ toch aan "getallen" denken (later worden zij dat

toch) dan staat hier te lezen dat $A : B > C : D$ wanneer er een rationaal getal $\frac{n}{m}$ tussen $\frac{C}{D}$ en $\frac{A}{B}$ ligt, immers $mA > nB$ betekent dan $\frac{n}{m} < \frac{A}{B}$ en $mC < nD$ betekent dan $\frac{n}{m} > \frac{C}{D}$.



afb. 3.2

3.5 Met deze definitie van gelijkheid en ordening van verhoudingen worden dan de gebruikelijke eigenschappen afgeleid. We geven hiervan als voorbeeld stelling V,18, maar dan in moderne notatie.
De stelling luidt:

$$(A + B) : B = (C + D) : D \iff A : B = C : D$$

Bewijs:
$\Longrightarrow$ Uit het gegeven volgt voor willekeurige $m, n \in \mathbb{N}$:

$$m(A + B) \gtreqless (m + n)B \iff m(C + D) \gtreqless (m + n)D$$

en dus

$$mA \gtreqless nB \iff mC \gtreqless nD$$

maar dit betekent juist

$$A : B = C : D$$

$\Longleftarrow$ Dit volgt door de redenering "van onder naar boven" te volgen.
N.B. alle genomen kleine tussenstappen, zoals $m(A+B) = mA+mB$ zijn nauwkeurig verantwoord in de Elementen.
Als tweede voorbeeld, de

Stelling: $A > B \implies A : C > B : C$.

Bewijs:
We onderscheiden
a. $A - B < B$
b. $A - B \geq B$.
a. $A - B < B$. Kies $m \in \mathbb{N}$ z.d.d. $m(A - B) > C$ en daarna $n \in \mathbb{N}$ z.d.d.

$$(n - 1)\, C \leq mB < nC. \qquad (*)$$

dan geldt

$$C < mA - mB$$

$$nC - C \leq mB$$

dus

$$nC < mA \qquad (**)$$

en tevens (*)

$$mB < nC \qquad (*)$$

Uit (*) en (**) volgt dan, per definitie

$$A : C > B : C$$

b. $B \leq A - B$
Kies nu $m \in \mathbb{N}$ z.d.d.

$$mB > C$$

en daarna $n \in \mathbb{N}$ z.d.d.

$$(n-1)C \ \leq \ m(A-B) < nC, \qquad (***)$$

dan volgt uit $B \leq A - B$
en dus

$$mB \ \leq \ m(A-B),$$

met (***)

$$mB < nC.$$

Verder volgt uit:

$$C < mB$$

met

$$nC - C \ \leq \ mA - mB \ < nC \qquad (***)$$

dat

$$nC \ < \ mA.$$

Dus we hebben

$$mA > nC \quad \text{en} \quad mB < nC$$

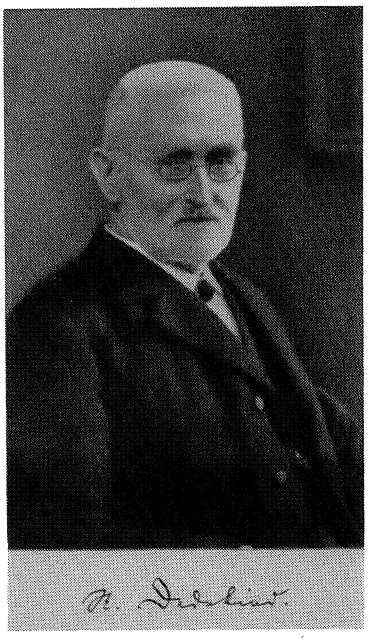
hetgeen betekent

$$A : C > B : C.$$

4. *Het irrationale getal bij Dedekind*

4.1 De kroon op het werk, waarvoor Eudoxus de basis legde, werd gezet door Julius Wilhelmus Dedekind. Deze grote, bescheiden wiskundige werd in 1831 geboren in Braunschweig en studeerde vanaf 1850 in Göttingen o.a. bij Gauss, onder wiens leiding hij in 1852 promoveerde op een proefschrift over integralen van Euler, d.w.z. integralen van het type

$$\int_0^1 x^{a-1}(1-x)^{b-1}dx \text{ en } \int_0^1 x^{t-1}e^{-x}dx.$$



afb. 4.1: Dedekind

Achtereenvolgens was hij privaatdocent in Göttingen, hoogleraar aan het Polytechnikum in Zürich (de voorloper van de Eidgenössische Technische Hochschule) en vanaf 1862 tot zijn emeritaat in 1894, hoogleraar aan het Polytechnikum in Braunschweig.
Van zijn vele bijdragen tot de wiskunde worden hier slechts genoemd zijn fundamentele werk op het gebied van de algebraïsche getallentheorie - met als hoogtepunt de ideaaltheorie met de ontbindbaarheid van een ideaal in priemfactoren - en zijn strenge definitie van het irrationale getal.
Ook mag zijn "editorial" werk niet onvermeld blijven: met Weber gaf hij het verzamelde werk van Riemann (1826-1866) uit, hij werkte mee aan de uitgave van het getallentheoretische werk van Gauss (1777-1855) en publiceerde na de dood van Dirichlet (1805-1859) diens "Vorlesungen über Zahlentheorie".

4.2 Hier houden we ons bezig met Dedekinds constructie van de irrationale getallen uit de rationale.
Bij zijn onderwijs in de differentiaal- en integraalrekening voelde hij behoefte aan een goede fundering daarvan. Zo schrijft hij:
Man sagt so häufig, die Differentialrechnung beschäftige sich mit den stetigen Grössen, und doch wird nirgends eine Erklärung von dieser Stetigkeit gegeben.... [4.1]
Verderop leest men op dezelfde bladzijde, waar hij spreekt over de stelling dat iedere naar boven begrensde verzameling van reële getallen een kleinste bovengrens heeft:
Es kam nur noch darauf an, seinen eigentlichen Ursprung in den Elementen der Arithmetik zu entdecken und hiermit zugleich eine wirkliche Definition von dem Wesen der Stetigkeit zu gewinnen. Dies gelang mir am 24 November 1858.
In het voorwoord tot " Was sind und was sollen die Zahlen?" (litt. (3),(4)) noemt hij - schrijvende over irrationale getallen - als zijn bron van inspiratie de hierboven behandelde def. V.5 in de Elementen van Euclides:
... So ist diese Art ihrer Bestimmtheit schon auf das deutlichste in der berühmten Definition ausgesprochen, welche Euklid (Elemente V.5) für die Gleichheit der Verhältnisse aufstellt. Eben diese uralte Überzeugung ist nun gewiss die Quelle meiner Theorie.

4.3 Wij herhalen deze definitie hier in moderne notatie:
Men zegt dat $A : B = C : D$ dan en slechts dan als voor alle $m, n \in \mathbb{N}$ geldt:

$$mA \gtreqless nB \Leftrightarrow mC \gtreqless nD.$$

Dedekind zag de diepere betekenis hiervan in. Deze houdt nl. in dat de definitie die Eudoxus gaf van de gelijkheid van verhoudingen, een evenredigheid dus, in de verzameling $\mathbb{Q}^+$ van alle positieve rationale getallen een verdeling in twee klassen A_1 en A_2 induceert, z.d.d.
voor alle $\frac{n}{m} \in A_1$, geldt $mA \geq nB$
en voor all $\frac{n}{m} \in A_2$ geldt $mA < nB$,
Als we even brutaal zijn en $A : B$ schrijven als een "breuk" (Wat dat dan ook moge beteken), dan zou gelden voor $\frac{n}{m} \in A_1 : \frac{n}{m} \leq \frac{A}{B}$ en voor $\frac{n}{m} \in A_2 : \frac{n}{m} > \frac{A}{B}$. Dit verklaart waarom we A_1 "onderklasse" noemen en A_2 "bovenklasse".

4.4 Deze klassen blijken de volgende eigenschappen te bezitten:

1. $A_1 \neq \phi$ en $A_2 \neq \phi$.

2. $A_1 \cup A_2 = \mathbb{Q}^+$ en $A_1 \cap A_2 = \phi$.

3. als $\frac{n_1}{m_1} \in A_1$ en $\frac{n_2}{m_2} \in A_2$, dan geldt $\frac{n_1}{m_2} < \frac{n_2}{m_2}$.

Bewijs:

1. Volgens het axioma van Eudoxus- Archimedes bestaan er natuurlijke getallen m en z.d.d. $mA \geq B$ en $nB > A$ dus $\frac{1}{m} \in A_1$ en $n \in A_2$; derhalve $A_1 \neq \phi$ en $A_2 \neq \phi$.
2. Steeds geldt voor $\frac{n}{m}$: $mA \geq nB$ òf $mA < nB$ en nooit beide tegelijkertijd.
3. Stel $\frac{n_1}{m_1} \in A_1$ en $\frac{n_2}{m_2} \in A_2$, dan geldt:

   $$m_1 A \geq n_1 B \text{ en } m_2 A < n_2 B$$

   dus $n_1 m_2 A < n_1 n_2 B \leq n_2 m_1 A$
   dus $n_1 m_2 < n_2 m_1$
   oftewel $\frac{n_1}{m_1} < \frac{n_2}{m_2}$.

4.5 Verder merken we op dat m.b.t. een grootste en een kleinste element in A_1 resp. A_2, er 3 mogelijkheden zijn:

1. A_1 heeft een grootste element.
2. A_2 heeft een kleinste element.
3. A_1 heeft geen grootste element en A_2 geen kleinste.

De vierde mogelijkheid A_1 heeft een grootste element, zeg m_1 en A_2 heeft een kleinste element, zeg m_2 kan zich niet voordoen, want dan zou daar $m_1 < m_2$ voor $\mu = \frac{1}{2}(m_1 + m_2)$ gelden: $\mu > m_1$ en dus $\mu \notin A_1$ dus $\mu \in A_2$ maar tevens $\mu < m_2$ en dus $\mu \notin A_2$ dus $\mu \in A_1$ maar $A_1 \cap A_2 = \phi$ dus tegenspraak.
Uiteraard kunnen we het altijd zo inrichten dat A_2 geen kleinste element heeft door dat dit eventuele element aan A_1 toe te kennen.

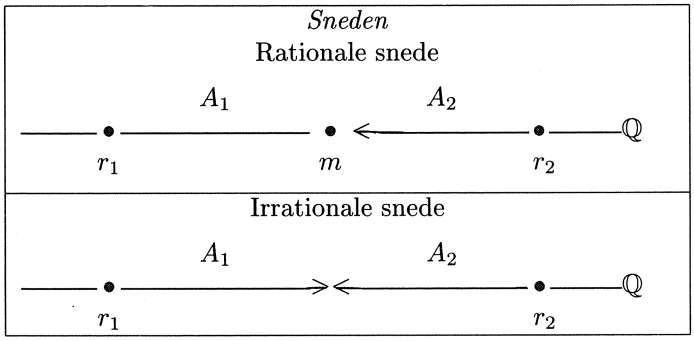
4.6 Indien A_1 een grootste (rationaal) element, zeg r, heeft dan correspondeert met de verdeling, d.w.z. met dit paar klassen (A_1, A_2) het rationale getal r. Heeft A_1 geen grootste element (en A_2 geen kleinste), dan kunnen we het paar (A_1, A_2) laten corresponderen met "iets nieuws". Dit bracht Dedekind op de gedachte om in $\mathbb{Q}$, het lichaam van *alle* rationale getallen, in abstracto een verdeling in twee klassen A_1 en A_2 te definiëren, welke verdeling *per definitie* de karakteristieke eigenschappen heeft die hierboven voortvloeiden voor $\mathbb{Q}^+$ uit de definitie van Eudoxus.
Het paar (A_1, A_2) noemde Dedekind een snede (Schnitt) en is als "snede van Dedekind" de geschiedenis ingegaan. Deze snede stelde Dedekind in staat om het lichaam $\mathbb{Q}$ van de rationale getallen uit te breiden met de irrationale getallen tot het lichaam $\mathbb{R}$ van de reële getallen.

4.7 De zuinigste definitie van een snede van Dedekind luidt aldus:
Een snede is een geordend paar deelverzamelingen (A_1, A_2) van het lichaam $\mathbb{Q}$ van de rationale getallen en wel zó dat.

1. $A_1 \neq \phi$, $A_2 \neq \phi$.

2. $A_1 \cup A_2 = \mathbb{Q}$; $A_1 \cap A_2 = \phi$.
3. Als $r_1 \in A_1$ en $r_2 \in A_2$ dan geldt: $r_1 < r_2$.
4. A_2 heeft geen kleinste element.

In het geval dat A_1 een grootste element, zeg m, heeft laat men met de snede (A_1, A_2) het rationale getal m corresponderen. Men spreekt dan van een *rationale* snede. In het geval dat A_1 geen grootste element heeft spreken we van een *irrationale* snede.



afb. 4.2

De verzameling sneden in $\mathbb{Q}$ is nu wat Dedekind de verzameling van de reële getallen noemde.
De rationale getallen die één-éénduidig corresponderen met de rationale sneden zijn hierin ingebed.
De irrationale sneden noemt men irrationale getallen. Dit zijn de nieuwe aanwinsten, maar nu formeel ingevoerd.
Van deze collectie reële getallen kan bewezen worden dat zij - bij geschikt gekozen operaties en definities - een geordend lichaam vormen met een continuïteitseigenschap die we nog zullen preciseren.

4.8 Het bekendste voorbeeld van een irrationale snede is wel de snede waarmee $\sqrt{D}$ wordt gedefinieerd in het geval dat D behoort tot $\mathbb{N}$, maar niet het kwadraat is van een geheel getal.
In §1.6 en §1.7 zagen we al dat in dit geval D ook niet het kwadraat is van een rationaal getal. De bijbehorende snede wordt als volgt gedefinieerd:

$$A_2 = \{r \in \mathbb{Q} | r > 0 \wedge r^2 > D\} \text{ en } A_1 = \mathbb{Q} - A_2.$$

We zien dan eenvoudig in dat voldaan is aan de eisen (1), (2) en (3) van de definitie.
Blijft nog te bewijzen dat A_1 geen grootste element heeft en A_2 geen

kleinste.
Voor het bewijs nemen we $r \in A_1$, met $r > 0$ (er geldt dan $r^2 < D$) en laten zien dat er in A_1 nog een groter element is. Daartoe kiezen we allereerst een rationale h z.d.d. $0 < h < 1$. Dan geldt:

$$(r+h)^2 = r^2 + 2rh + h^2 < r^2 + 2rh + h = r^2 + (2r+1)h.$$

Bepalen we h nader z.d.d. $h < \frac{D-r^2}{2r+1}$, dan geldt $(r+h)^2 < D$, dus $r+h \in A_1$ dus A_1 heeft geen grootste element.
Analoog: zij $r \in A_2$ (er geldt dan $r > 0$ en $r^2 > D$). Kies nu $0 < h < r$. Er geldt dan $(r-h)^2 = r^2 - 2r + h^2 > r^2 - 2rh > D$, indien we ook nog zorgen dat $h < \frac{r^2-D}{2r}$. Voor $r-h$ geldt: $r-h \in A_2$ en $r-h < r$. A_2 heeft dus geen kleinste element. [4.2]

4.9 Na de introductie van het begrip "snede", gaat Dedekind de "orde op zaken stellen", d.w.z. hij definieert voor deze sneden de begrippen "gelijk aan" $(=)$ "groter dan" $(>)$ en "kleiner dan" $(<)$.
Als α en β twee sneden zijn, waarbij $\alpha = (A_1, A_2)$ en $\beta = (B_1, B_2)$ dan definieert men $\alpha = \beta$ d.e.s.d, als $A_1 = B_1$ en tevens $A_2 = B_2$.
Indien $A_1 \neq B_1$ dan onderscheiden we twee gevallen:

1. A_1 en B_1 hebben slechts één rationaal getal niet gemeen
2. A_1 en B_1 hebben minstens twee rationale getallen niet gemeen.

1. Stel $B_1 \neq A_1$ en a_1' is het enige rationale getal dat wel in A_1 maar niet in B_1 ligt. We kunnen dan bewijzen dat $B_1 \subset A_1$ en dat α en β beide met het rationale getal a_1' corresponderen, dus in wezen gelijk zijn (zie aantekening [4.3]).
2. Stel $B_1 \neq A_1$ en er zijn minstens twee rationale getallen a_1' en a_1'' die wel A_1 maar niet in B_1 liggen. In dit geval zijn er uiteraard oneindig veel rationale getallen in A_1 die niet tot B_1 behoren en we noemen de sneden α en β wezenlijk verschillend.
   We zeggen dan $\alpha > \beta$ en $\beta < \alpha$.
   In bovenstaande redenering kunnen uiteraard α en β verwisseld worden, zodat we alle gevallen te pakken hebben.
   Het is niet moeilijk te bewijzen dat deze ordening *totaal* is, d.w.z. als $\alpha \in \mathbb{R}$, $\beta \in \mathbb{R}$, dan geldt

   $$\alpha = \beta \text{ of } \alpha > \beta \text{ of } \alpha < \beta.$$

   Verder geldt de transitieve wet:

   $$\text{als } \alpha < \beta \text{ en } \beta < \gamma \text{ dan } \alpha < \gamma.$$

   en:
   Tussen twee verschillende reële getallen liggen oneindig veel reële getallen.
   Ook deze eigenschappen zijn niet moeilijk te bewijzen.

4.10 Een belangrijke consequentie van deze ordening van $\mathbb{R}$ is het continuiteitsbeginsel, dat de samenhang van $\mathbb{R}$ bepaalt en dat als volgt geformuleerd kan worden:
Indien de verzameling $\mathbb{R}$ van alle reële getallen in twee niet lege klassen K_1 en K_2 wordt verdeeld zodanig dat voor $\alpha_1 \in K_1$ en $\alpha_2 \in K_2$ geldt $\alpha_1 < \alpha_2$, dan is er precies één $\alpha_0 \in \mathbb{R}$ zodanig dat voor elke $\beta \in \mathbb{R}$ met $\beta < \alpha_0$, geldt: $\beta \in K_1$ en voor elke $\beta \in \mathbb{R}$ met $\beta > \alpha_0$, geldt $\beta \in K_2$.
Voor het bewijs leze men aantekening [4.4]. De betekenis van deze stelling is nauwelijks te overschatten: zij bevat de kern van het begrip continuiteit. In het voorwoord tot zijn "Stetigkeit und irrationale Zahlen" vergelijkt Dedekind de rechte lijn met de verzameling rationale getallen en constateert bij deze laatste collectie "Lückenhaftigkeit, Unvollstandigkeit oder Unstetigkeit" en vraagt zich af waarin dan wel de "Stetigkeit" van de rechte lijn bestaat. Hij wil dit probleem streng aanpakken:
Mit vagen Reden über den ununterbrochenen Zusamenhang in den kleinsten Teilen ist naturlich nichts erreicht.
Hij merkt dan op dat een blik op de rechte lijn leert dat een gegeven punt p de rechte verdeelt in twee stukken, zódat ieder punt van het ene deel links van ieder punt van het andere deel ligt en gaat dan als volgt verder:
Ich finde nun das Wesen der Stetigkeit in der Umkehrung, also in dem folgenden Prinzip: Zerfallen alle Punkte der Geraden in Zwei Klassen von der Art, dass ieder Punkt der ersten Klasse links van jedem Punkte der zweiten Klassen liegt, so existiert ein und nur ein Punkt, welcher diese Einteilung aller Punkte in Zwei Klassen, diese Zerschneidung der Geraden in zwei Stücke hervorbringt [4.5].
Voor de door hem ingevoerde reële getallen lukt het hem dit te bewijzen en dat is dan ook een antwoord op zijn "cri de cœur" die we in §4.2 weergaven.

4.11 Na de ordening worden de algebraïsche bewerkingen - optelling, vermenigvuldiging en hun inversen - voor sneden ingevoerd.
Bij de definitie van de optelling zijn er weinig problemen.
Als (A_1, A_2) en (B_1, B_2) twee sneden zijn, dan definiëren wij de som $(S_1, S_2) = (A_1, A_2) + (B_1, B_2)$ daarvan, door:

$$S_1 = \{(a_1 + b_1) | a_1 \in A_1; b_1 \in B_1\} \text{ en } S_2 = \{(a_2 + b_2) | a_2 \in A_2; b_2 \in B_2\}$$

Men kan dan rechttoe, rechtaan bewijzen dat hierdoor inderdaad een snede is gedefinieerd en dat de sneden met deze definitie een geordende Abelse groep vormen,waarbij de snede die $0 \in \mathbb{N}$ bevat het nulelement is [4.6].
Met betrekking tot de vermenigvuldiging is de situatie gecompliceerder. Allereerst bezien we twee niet-negatieve sneden (A_1, A_2) en (B_1, B_2). Hiervoor definieert men het produkt $(P_1, P_2) = (A_1, A_2) \cdot (B_1, B_2)$ door:

$$P_1 = \{a_1 b_1 | a_1 \in A_1; b_1 \in B_1\}$$

en

$$P_2 = \{a_2 b_2 | a_2 \in A_2; b_2 \in B_2\}.$$

Dit product is inderdaad een snede. De vermenigvuldiging is associatief, commutatief, distributief over hier boven gedefinieerde opstelling en heeft als één-element de snede die $1 \in \mathbb{N}$ bevat.
Elke positieve snede blijkt dan een multiplicatieve inverse te bezitten. Niet zonder veel moeite kan men de vermenigvuldiging uitbreiden tot de negatieve sneden. Voor de bewijzen hiervan zij ook hier verwezen naar de zoeven genoemde litteratuur [4.5].
Het resultaat is echter dat de door Dedekind ingevoerde sneden een geordend lichaam ($\mathbb{R}$) vormen waarin het lichaam $\mathbb{Q}$ van de rationale getallen isomorf kan worden ingebed.
Dit lichaam $\mathbb{R}$ is - zoals we zagen - totaal geordend en heeft de in §4.10 genoemde continuiteitseigenschap.

4.12 Tot slot vermelden we nog de constructie van reële getallen met behulp van inkrimpende intervallen met rationale eindpunten. Zo'n "nest" van intervallen heeft dan een aan alle intervallen gemeenschappelijk rationaal getal of niet. Op deze verzameling intervallen kan men een equivalentierelatie definieren. De bijbehorende klassen vormen dan met geschikt gekozen operaties een lichaam, het lichaam van de reële getallen. De klassen van intervallen met (rationaal) lege doorsnede brengen dan de irrationale getallen in.
De gedachtengang van inkrimpende intervallen is al zeer oud en gaat terug tot de Babyloniërs en de Grieken. Men denke aan de benadering van de oppervlakte van een cirkel met behulp van in-en omgeschreven regelmatige veelhoeken door Archimedes.

Aantekeningen

N.B. tussen () geplaatste getallen verwijzen naar de litteratuurlijst

1.1 Dedekind (3) en (4)

1.2 Deze voordracht is een nader uitwerking van en een aanvulling op de opmerking die ik maakte in mijn lezing over de meetkundige algebra bij de Euclides, in het kader van de vacantiecursus van het CWI in 1991. Zie (20) pag. 12 t/m 15.

1.3 H. Diels - W. Kranz (5).

1.4 Jamblichus (13) §88 en Heiberg (12) pag. 85.

1.5 We vinden dit wel bij Thaer (19) maar niet bij Heiberg - Stamatis (11).

1.6 Zie hier voor Fowler (8) pag. 215, 304, 305.

1.7 Dedekind (3) pag. 324, 325; (4) pag. 13, 14, 15.

2.1 Heath (9), dl.II pag. 277; Heiberg (11) dl.II pag. 103; Thaer (19) pag. 141.

2.2 O.c. resp. pag. 311; pag. 118; pag. 150.

3.1 Steck (18), pag. 150.

3.2 Heiberg (12), pag. 211.

3.3 Lasserre (17).

3.4 Heiberg (12), pag. 213.

3.5 Heath (9), II pag. 126; Heiberg (11), II pag.1; Thear (9), pag. 91.

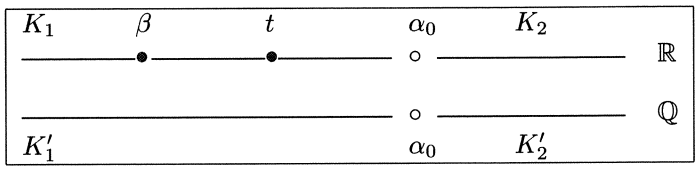
3.6 Heiberg (12), pag. 315,

4.1 Dedekind (3), pag. 316; (4) pag. 2.

4.2 Het bewijs is hier enigzins gewijzigd doordat we twee gevallen onderscheiden. Hier door komt h op iets minder gekunstelde wijze te voorschijn. Dedekind werkt met slechts één h die in beide gevallen voldoet, maar deze h "valt uit de lucht"

4.3 We gaan uit van de sneden (A_1, A_2) en (B_1, B_2) met $A_1 \neq B_1$ en wel zó dat a_1' het enige rationale getal is dat wel in A_1, maar niet in B_1 ligt en dus wel in B_2. Om dit laatste te accentueren noteren we a_1' dan ook wel als b_2'. Voor elke $b_1 \in B_1$ geldt dan $b_1 < b_2'$, dus $b_1 < a_1'$ (*). Hieruit volgt $B_1 \subset A_1$; stel nl. dat voor zekere $b_1 \in B_1$ geldt $b_1 \notin A_1$ dan $b_1 \in A_2$ en dus $b_1 > a_1'$ $(\in A_1)$. Tegenspraak. Nu is a_1' het enige element van A_1 dat niet in B_1 ligt. Voor elke andere $a_1 \in A_1$ geldt dus $a_1 \in B_1$ en dus volgens (*): $a_1 < a_1'$, m.a.w. a_1' is het grootste element van A_1 en (A_1, A_2) is de rationale snede die met het getal a_1' correspondeert. Uit het bovenstaande volgt tevens dat A_1 en B_1 slechts in a_1' verschillen.
Nu bezien we de snede (B_1, B_2) en tonen aan dat $a_1'(= b_2')$ het kleinste element van B_2 is. We zagen al dat voor elke $b_1 \in B_1$ geldt $b_1 < a_1'(= b_2')$ (zie (*)).
Voor elke $b_2 \in B_2$ met $b_2 \neq b_2'$ geldt dan $b_2 > b_2'$ want anders zou $b_2 < b_2'$, dus $b_2 < a_1'(\in A_1)$; maar omdat A_1 en B_1 alleen in a_1' verschillen, zou dan $b_2 \in B_1$ en dat geeft een tegenspraak. Hieruit volgt dat $b_2' = a_1'$ het kleinste rationale getal is van B_2 en dus correspondeert ook de snede (B_1, B_2) met het rationale getal a_1'.

4.4 Laat $\mathbb{R}$ verdeeld zijn in twee niet lege klassen K_1 en K_2, zó dat uit $\alpha_1 \in K_1$ en $\alpha_2 \in K_2$ volgt $\alpha_1 < \alpha_2$. Deze klassenindeling van $\mathbb{R}$ induceert in de verzameling (het lichaam) $\mathbb{Q}$ van alle rationale getallen een indeling in twee klassen K_1' en K_2' waarbij K_1' alle rationale getallen van K_1 bevat en K_2' die van K_2.



(K_1', K_2') is een snede van Dedekind in $\mathbb{Q}$ en bepaalt dus op éénduidige wijze een - al dan niet rationaal - reël getal α_0 Neem nu $\beta \in \mathbb{R}$ en $\beta \neq \alpha_0$. Onderscheid nu $\beta < \alpha_0$ en $\beta > \alpha_0$. Als $\beta < \alpha_0$, dan is er een *rationaal* getal t met $\beta < t < \alpha_0$. Omdat $t < \alpha_0$, geldt $t \in K_1$, anders $t \in K_2$ dus $t \in K_2'$ en dus $t > \alpha_0$ het geen een tegenspraak oplevert. Maar dan geldt ook $\beta \in K_1$, want $\beta < t$ en als $\beta \in K_2$ dan zou $\beta > t$ omdat $t \in K_1$. Het geval $\beta > \alpha_0$ verloopt analoog. Bij de gegeven inleiding in klassen K_1 en K_2 is er dus precies één reëel getal α_0 zodanig dat $\beta < \alpha_0$ als $\beta \in K_1$ en $\beta > \alpha_0$ als $\beta \in K_2$. Men kan α_0 als grootste element aan K_1 of als kleinste aan K_2 toevoegen.

4.5 Dedekind (3), pag. 322 en (4), pag. 10 en 11.

4.6 Voor de details van de bewijzen zie litt. (3), (4), (7), (16).

LITTERATUUR

1. E.M.J. BERTIN, H.J.M. BOS, A.W. GROOTENDORST ed., Two Decades of Mathematics in the Netherlands, M.C., Amsterdam, 1978.
2. R. DEDEKIND, Was sind und was sollen die Zahlen? In: Gesammelte Abhandlungen III, p. 335-391, Chelsea Publishing Company, New York, 1969.
3. R. DEDEKIND, Stetigkeit und Irrationale Zahlen, o.c.p. 315-334.
4. R. DEDEKIND, Essays on the Theory of Numbers, Dover Publications Inc., New York, 1963. (Engelse vertaling van (2) en (3)).
5. H. DIELS, W. KRANZ, Die Fragmente der Vorsokratiker, 6^e ed., Brill, Leiden, 1969.
6. E.J. DIJKSTERHUIS, De Elementen van Euclides I en II, Noordhoff, Groningen, 1929.
7. H.-D. EBBINGHAUS et al., Zahlen, 3^e Auflage, Springer, Berlin etc., 1992.
8. D.H. FOWLER, The Mathematics of Plato's Academy, Oxford, 1987.
9. T.L. HEATH, The Thirteen Books of Euclid's Elements (translated from the Text of Heiberg), Dover Publications Inc., New York, 1956.
10. T.L. HEATH, A History of Greek mathematics I, II, Clarendon Press, Oxford, 1965.
11. I.L. HEIBERG, Euclidis Elementa (ed. E.S. Stamatis), Teubner, Leipzig, 1969.
12. I.L. HEIBERG, Euclidis Elementa, Scholia in Libros VI-XIII (ed. E.S. Stamatis), Teubner, Leipzig, 1977.
13. JAMBLICHUS-Porphyrius, Leven en leer van Pythagoras, vertaling H.W.A. van Rooijen-Dijkman, Ambo, Baarn, 1987.

14. M. KLINE, Mathematical Thought from Ancient to Modern Times, Oxford University press, New York, 1972.
15. W.B. KNORR, The Evolution of the Euclidean Elements, Reidel Publishing Company, Dordrecht, 1975.
16. E. LANDAU, Grundlagen der Analysis, Wiss. Buchgesellschaft, Darmstadt, 1960.
17. F. LASSERRE, Die Fragmente des Eudoxus von Knidos, Berlin, 1966.
18. MAX STECK, Proklos Diadochos, Kommentar zum Ersten Buch von Euklids "Elementa" Halle, 1945.
19. C. THAER, Euklid, die Elemente, Wiss. Buchgesellschaft, Darmstadt, 1962, (Duitse vertaling).
20. Vacantiecursus 1991: Meetkundige Structuren, C.W.I. Amsterdam 1991.

A.W. Grootendorst

p-Adische getallen

W.H. Schikhof

§1. De 10-adische getallen

Definitie 1.1. Een 10-*adisch getal* is een oneindige rij cijfers

$$\ldots\ a_3\ a_2\ a_1\ a_0,\ a_{-1}\ a_{-2}\ \ldots$$

waarbij $a_{-n} = 0$ voor grote n (d.w.z. vanaf een zeker moment zie je naar rechts toe alleen maar nullen). Een *cijfer* is een der symbolen $0, 1, 2, 3, 4, 5, 6, 7, 8, 9$.

Voorbeelden.

$$\ldots 8427159,43100000\ldots$$
$$\ldots 0000365,00000\ldots$$

De 10-adische getallen stoppen we in een verzameling genaamd $\mathbb{Q}_{10}$. vskip0.1em De 10-adische getallen met alleen nullen achter de komma (dan schrijf je meestal noch de komma, noch die nullen op: $\ldots 318,0000\ldots$ " = " 318) vormen bij elkaar de verzameling $\mathbb{Z}_{10}$ van de 10-*adische gehele getallen.*

We hebben

$$\mathbb{N} \subset \mathbb{Z}_{10} \subset \mathbb{Q}_{10}$$

Een 10-adische getal in $\mathbb{Z}_{10}$ is in $\mathbb{N}$ precies dan wanneer $a_n = 0$ voor grote n (d.w.z. vanaf een zeker moment staan naar links toe alleen maar nullen).
Als kind hebben we geleerd hoe optellingen en vermenigvuldigingen uit te voeren. Op een naieve manier gebruiken we deze methode ook voor onze 10-adische getallen. Aan de volgende voorbeelden ziet u direct hoe dit werkt.

```
.... 84271,59
.... 54785,63
___________+
.... 39057,22
```

```
.... 271,59
.... 785,63
___________ ×
..... 81477
.... 62954
... 35795
.. 17272
. 90113
  .
  .
___________ +
... 9,2517
```

Een nauwkeurige definitie van deze optelling en vermenigvuldiging is wat lastiger te geven, maar hij verheldert niets dus laten we dat maar achterwege. Hetzelfde verhaal geldt voor de bewijzen van de volgende regels.

OPTELLING. Stel $x, y, z \in \mathbb{Q}_{10}$. Dan

O.1 $x + y = y + x$

O.2 $(x + y) + z = x + (y + z)$

O.3 $x + 0 = x$
(Hierbij is 0 per definitie het element

$$\ldots 00000, 00000 \ldots \qquad \text{(allemaal nullen)}$$

van $\mathbb{Q}_{10}$.)

VERMENIGVULDIGING. Stel $x, y, z \in \mathbb{Q}_{10}$. Dan

V.1 $xy = yx$

V.2 $(xy)z = x(yz)$

V.3 $x.1 = x$
(Hierbij is 1 per definitie het element

$$\ldots 0001, 0000 \ldots$$

van $\mathbb{Q}_{10}$.)

COMBINATIEVAN BEIDE: Stel $x, y, z \in \mathbb{Q}_{10}$. Dan

D.1 $x(y + z) = xy + xz$.

Door te proberen zie je dat je ook aftrekkingen kunt maken in $\mathbb{Q}_{10}$:

$$\begin{array}{r} \ldots 84271{,}59 \\ \ldots 54785{,}63 \\ \hline \ldots 29485{,}96 \end{array}$$

Maar ook:

$$\begin{array}{r} \ldots 54785{,}63 \\ \ldots 84271{,}59 \\ \hline \ldots 70514{,}04 \end{array}$$

ALGEMEEN:

O.4 Bij iedere $x, y \in \mathbb{Q}_{10}$ is er precies één $z \in \mathbb{Q}_{10}$ met $x + z = y$. Dit getal noemen we $y - x$. Het *tegengestelde* van x is het getal $0 - x$, meestal als $-x$ geschreven.

OPEMERKING. Een verzameling waar een optelling en een vermenigvuldiging is gedefinieerd die aan O.1-O.4 en V.1-V.3 en D.1 voldoen heet een *ring*.
We hebben dus: $\mathbb{Q}_{10}$ is een ring.
("Gewone" voorbeelden: $\mathbb{Z}, \mathbb{Q}, \mathbb{R}$).

1.A OPGAVE. Laat zien dat $\mathbb{Z}_{10}$ ook een ring is.

1.B OPGAVE. Laat zien:

$$\ldots a_2 a_1 a_0 \in \{-1, -2, -3, \ldots\} \iff a_n = 9 \text{ voor grote } n.$$

1.C OPGAVE. Toon aan dat er precies een $x \in \mathbb{Q}_{10}$ is met $3x = 1$.

1.D OPGAVE. Is er (precies één) $x \in \mathbb{Q}_{10}$ met $2x = 1$? Met $4x = 1$? Met $5x = 1$? Met $6x = 1$? Met $7x = 1$? Met $8x = 1$? Met $9x = 1$? Met $10x = 1$?
Hoe loop dit verhaal af?

1.E OPGAVE. Bewijs dat er geen $x \in \mathbb{Z}_{10}$ is met $x^2 + 1 = 0$. Is er wèl een $x \in \mathbb{Q}_{10}$ met $x^2 + 1 = 0$?

1.F OPGAVE. Er bestaan $a, b \in \mathbb{Z}_{10}$ die geen van beide nul zijn terwijl toch $a.b = 0$. Begin bijv. zo:

```
....112
....125
_____×
....560
....24
....2
  .
  .
  .
_____
  000
```

en laat zien dat je a_4 en b_4 kunt verzinnen zo dat er in

```
...a4112
...b4125
_______×
   .
   .
   .
_______+
```

in de uitkomst onderaan de laatste *vier* cijfers nul zijn. Als je deze a_4 en b_4 hebt, hoe moet je dan a_5 en b_5 vinden? Gaat dit verderop altijd goed?

1.G OPGAVE. Heeft de vergelijking $x^2 = 1$, behalve $x = 1$ en $x = -1$ nog andere oplossingen in $\mathbb{Q}_{10}$?

1.H OPGAVE. Toon aan: als $x \in \mathbb{Q}_{10}$, $x^{85} = 0$ dan $x = 0$.
(Hint. Kijk eerst naar de vraag: als $x^2 = 0$ is dan $x = 0$?)

1.I OPGAVE. Iemand oppert het volgende: "Neem ik een 10-adisch getal, bijv.

$$\ldots 3821,716$$

en 'spiegel ik dat om de komma':

$$617,1283\ldots$$

dan krijg ik een ordinaire decimaalontwikkeling van een gewoon reëel getal. Zo kan ik dat doen voor ieder 10-adisch getal. Dus die $\mathbb{Q}_{10}$ is eigenlijk niets nieuws: 10-adische getallen zijn "eigenlijk" reële getallen, maar achterstevoren opgeschreven". Zit daar wat in?

1.J OPGAVE. In $\mathbb{N}$ heb je 't welbekende begrip $>$ (groter dan). Zou je van twee getallen a, b in $\mathbb{Z}_{10}$ kunnen afspreken wanneer je vindt dat $a > b$?

1.K OPGAVE. Je zou kunnen proberen de theorie van de reële getallen en die van $\mathbb{Q}_{10}$ samen te stoppen door één grote verzameling te maken van oneindige rijen

$$\ldots a_2 a_1 a_0, a_{-1} a_{-2} a_{-3} \ldots$$

waarbij $a_i \in \{0, 1, \ldots, 9\}$ maar waarbij verder geen restricties worden opgelegd (geen "staarten" van nullen). Gaat u eens na of dergelijke dingen vermenigvuldigd of opgeteld kunnen worden. □

§2. De n-adische getallen

In §1 zijn we uitgegaan van de schrijfwijze van reële getallen in de decimaalontwikkeling. Grondtal 10 dus. Beschavingen vóór ons hadden andere grondtallen. In het huidige computertijdperk is 't grondtal 2 op de voorgrond gekomen. Er is niets op tegen een ander grondtal dan 10 te nemen, en een verhaal te maken in de stijl van §1 voor dit nieuwe grondtal.

VOORBEELD: In 't zeventallig stelsel beschikken we slechts over de cijfers 0,1,2,3,4,5,6. Wat we vroeger als 7 noteerden wordt nu 10 ($= 0.7^0 + 1.7^1$). Wat vroeger 31 was heet nu 43 (nl. $3.7^0 + 4.7^1$), etc.

2.A OPGAVE. Ga na dat "een zevende" in 't zeventallig stelsel eruit ziet als 0,1. Ga na dat "een zesde" er uitziet als
$0,1111\ldots$ (allemaal énen)

2.B OPGAVE. Maak de volgende sommetjes (*zeventallig stelsel*):

143	143
361	361
___+	___×

We gaan aan de slag. Kies $n \in \{2, 3, 4, \ldots\}$.
DEFINITIE 2.1. Een *n-adisch getal* is een oneindige rij cijfers

$$\ldots a_3 a_2 a_1 a_0, a_{-1} a_{-2} \ldots$$

waarbij $a_{-n} = 0$ voor grote n. Een *cijfer* is een van de getallen $\{0, 1, 2, \ldots, n-1\}$.
De n-adische getallen stoppen we bij elkaar in een verzameling $\mathbb{Q}_n$. Geheel in de stijl van §1 kunnen we optelling en vermenigvuldiging in $\mathbb{Q}_n$ definieren. Die voldoen aan de regels O.1 t/m O.4; V.1 t/m V.3, D.1 (het zou saai worden dit allemaal te gaan controleren!).
Dus: $\mathbb{Q}_n$ is een ring.

2.C OPGAVE. Schrijf de getallen

$$15, -1, -3, \tfrac{1}{2}, \tfrac{1}{3}, \tfrac{1}{4}$$

(oude schrijfwijze)

als elementen

$$\ldots a_2 a_1 a_0$$

van $\mathbb{Q}_5$.

2.D OPGAVE. Hoe schrijf je $\frac{1}{2}$ (oude schrijfwijze) als element van $\mathbb{Q}_n$? (Onderscheid: n even/n oneven).

2.E OPGAVE. Hoeveel oplossingen heeft de vergelijking $x^2 = 1$ in $\mathbb{Q}_7$? Zelfde vraag voor $x^2 = -1$.

§3. DE p-ADISCHE GETALLEN

We hebben gezien in §1 dat $\mathbb{Z}_{10}$ nuldelers had (en heeft) (d.w.z. er zijn a, b geen van beide nul met $ab = 0$). Op soortgelijke manier toon je aan dat $\mathbb{Z}_6, \mathbb{Z}_{28}, \ldots$ nuldelers hebben.
Maar, als n een priemgetal p is lukt dat niet meer:
Als

$$a = \ldots a_2 a_1 a_0$$

$$b = \ldots b_2 b_1 b_0$$

in $\mathbb{Z}_p$ liggen en a_0 en b_0 zijn beide $\neq 0$ dan is

$$a.b = \ldots c_2 c_1 c_0$$

waarbij c_0 = het laatste cijfer van $a_0 b_0$. Daar $a_0 \in \{1, 2, \ldots, p-1\}$, $b_0 \in \{1, 2, \ldots, p-1\}$ niet door p deelbaar zijn is $a_0 b_0$ 't ook niet, dus $c_0 \neq 0$.
Door nu wat met komma's te schuiven kun je algemeen inzien:
Als $a, b \in \mathbb{Q}_p$ (p is een priemgetal) en $a \neq 0$, $b \neq 0$, dan is $ab \neq 0$.
We gaan zelfs bewijzen:

STELLING 3.1. *Zij p een priemgetal. Dan heeft ieder element $a \in \mathbb{Q}_p$, $a \neq 0$ een inverse (d.w.z. er is een $x \in \mathbb{Q}_p$ met $xa = 1$)* Eerst even een hulpstellinkje:

LEMMA 3.2. *Stel p is een priemgetal, en stel t is een van de getallen $1, 2, \ldots, p-1$. Dan is er een $u \in \{1, 2, \ldots, p-1\}$ te vinden zó dat $tu = p$-voud $+1$.*
(Voorbeeld: $p = 1$, $t = 3$. Dan $u = 5$. Algemener: je kunt voor $p = 7$ "een vermenigvuldigingstabel" maken waar in 't schema alleen de rest bij deling door 7 is genoteerd:

	1	2	3	4	5	6
1	1	2	3	4	5	6
2	2	4	6	1	3	5
3	3	6	2	5	1	4
4	4	1	5	2	6	3
5	5	3	1	6	4	2
6	6	5	4	3	2	1

En inderdaad: in elke rij staat precies één 1.)

BEWIJS VAN LEMMA 3.2. Door te tellen: ik maak een afbeelding

$$f : \{1, 2, \ldots, p-1\} \longrightarrow \{1, 2, \ldots, p-1\}$$

als volgt:

$$f(x) := \text{ de rest van } xt \text{ bij deling door } p.$$

Die afbeelding f is "injectief" d.w.z. als $x \neq y$ dan $f(x) \neq f(y)$. Want stel maar $f(x) = f(y)$. Dan (denk even dat $x < y$) heeft yt dezelfde rest als xt bij deling door p. Dus $yt - xt$ is deelbaar door p. Maar t is dat niet. Dus moet $y - x$ deelbaar zijn door p. Maar dat kan ook niet want $y - x \in \{1, 2, \ldots, p-1\}$ zoals je eenvoudig inziet. De "injectiviteit" geeft ons dus: $f(1), f(2), \ldots, f(p-1)$ zijn allemaal verschillend. Maar ze zitten allemaal in $\{1, 2, \ldots, p-1\}$. Dus moet 1 ook voorkomen als f-beeld d.w.z. er is een u met $f(u) = 1$ d.w.z. $uf = p$-voud $+1$.

In 't bewijs kun je ook nog opmaken dat niet alleen 1, maar ook alle anderen in $\{1, 2, \ldots, p-1\}$ moeten voorkomen als beeld onder f. Dus breiden we ons Lemma 3.2 een beetje uit.

LEMMA 3.2^{BIS}. *Stel p is een priemgetal en stel $t \in \{1, 2, \ldots, p-1\}$. Voor ieder geheel getal b is er een $u \in \{0, 1, \ldots, p-1\}$ met $tu - b$ is een p-voud.*

BEWIJS. Is toevallig b een p-voud dan neem ik $u = 0$. Zo niet dan gebruik ik de bovenstaande opmerking.

BEWIJS VAN STELLING 3.1. Door weer wat met komma's te schuiven zie je gemakkelijk in dat 't voldoende is om aan te tonen dat voor 'n element

$$\ldots a_2 a_1 a_0$$

met $a_0 \neq 0$ er een $\ldots x_2 x_1 x_0$ is met

$$(\ldots x_2 x_1 x_0).(\ldots a_2 a_1 a_0) = \ldots 0001.$$

We proberen dus de vermenigvuldiging

$$\begin{array}{r} \ldots a_2 a_1 a_0 \\ \ldots x_2 x_1 x_0 \\ \hline \times \\ \quad . \\ \quad\ . \\ \quad\ \ . \\ \hline \ldots 0\ 0\ 0\ 1 \end{array}$$

kloppend te maken door $x_0, x_1, x_2, \ldots$ successievelijk te kiezen. Om te beginnen moet $x_0 a_0$ rest 1 hebben bij deling door p. Zo'n x_0 kan ik vinden: dat is precies ons lemma 3.2. Nu wil ik vervolgens x_1 kiezen zó dat 't product eindigt op 01. Daartoe heb ik alleen te maken met $a_1 a_0$ en $x_1 x_0$. Dus ik moet x_1 zó kiezen dat

$$(a_0 + a_1 p)(x_0 + x_1 p) \text{ rest 1 heeft bij deling door } p^2$$

d.w.z.

$$a_0 x_0 + p(a_1 x_0 + a_0 x_1) - 1 \text{ deelbaar door } p^2$$

d.w.z. (we weten dat $a_0 x_0 - 1$ deelbaar is door p)

$$\frac{a_o x_0 - 1}{p} + a_1 x_0 + a_0 x_1 \text{ deelbaar door } p$$

oftewel:

$$a_0 x_1 = \text{ iets bekends } + p\text{−voud.}$$

Lemma 3.2^{bis}. levert aan ons 't bestaan van zo'n $x_1 \in \{0, 1, \ldots, p-1\}$.
Zo kun je doorgaan. Als je $x_0, x_1, \ldots, x_n$ al gekozen hebt zó dat 't product eindigt op $\ldots 000 \ldots 01$ (n nullen en 'n 1) dan krijg je voor x_{n+1} weer een voorwaarde van 't type

$$a_0 x_{n+1} = \text{ iets bekends } + p\text{−voud}$$

en hieraan is altijd te voldoen omdat $a_0 \neq 0$.

Een ring waarin ieder element $\neq 0$ een inverse heeft heet 'n lichaam. We hebben dus (p is priemgetal)

STELLING 3.3. $\mathbb{Q}_p$ *is een lichaam.*

$$\mathbb{Q} \subset \mathbb{Q}_p.$$

Voor ieder priemgetal p hebben we zo een "lichaam" $\mathbb{Q}_p$ gebouwd. Binnen $\mathbb{Q}_p$ kun je eigenlijk net zo prettig rekenen als in $\mathbb{R}$ ($+$, $-$, $\times$, $\div$, deling door getallen $\neq 0$).

3.A OPGAVE. (Alweer $x^2 + 1 = 0$!) Heeft de vergelijking $x^2 + 1 = 0$ een oplossing in $\mathbb{Q}_5$?

3.B OPGAVE. Toon aan (p is priemgetal):
Als $a \neq 0$, $a \in \mathbb{Q}_p$ en $xa = 1$, $ya = 1$ voor $x, y \in \mathbb{Q}_p$ dan $x = y$.

3.C OPGAVE. Pak een $n \in \{1, 2, 3, \ldots\}$. Als U $p^n!$ in 't p-tallig stelsel schrijft, hoeveel nullen staan er dan op 't eind? (p is weer priemgetal.
(Hint. Zoek dat eerst uit voor $n = 1$, $n = 2$, $n = 3, \ldots$)

3.D OPGAVE. Wat is de rest van $(p-1)!$ is bij deling door p (p is priemgetal)?

3.E OPGAVE. Wat komt eruit als ik de "optelling" $1 + 5 + 5^2 + 5^3 + 5^4 + \ldots$ maak in $\mathbb{Q}_5$?
Toon aan dat 't antwoord gelijk is aan $-\frac{1}{4}$.

§4. VERGELIJKINGEN OPLOSSEN IN $\mathbb{Q}_p$ (p priemgetal)

VOORBEELD. Los op in $\mathbb{Q}_p$: $x^2 = 1$. We kunnen dat nu zó doen: $x^2 - 1 = 0 \Longrightarrow (x-1)(x+1) = 0 \Longrightarrow$ ($\mathbb{Q}_p$ heeft geen nuldelers) $\Longrightarrow x = 1$ of $x = -1$.

(Vergelijk opgave 1.G.)

We nemen nu eens een wat moeilijker vergelijking:

PROBLEEM. Zoek alle oplossingen van $x^2 = -1$ in $\mathbb{Q}_p$.

4.A OPGAVE. Bewijs: *als* er zo'n oplossing is dan $x \in \mathbb{Z}_p$.

We proberen als voorbeeld $p = 7$. We zoeken dus een $\ldots x_2x_1x_0$ in $\mathbb{Z}_7$ met $(\ldots x_2x_1x_0)^2 = -1 = \ldots 6666$. Dit schrijven we weer als

$$\begin{array}{l}
\ldots x_2x_1x_0 \\
\ldots x_2x_1x_0 \\
\hline
\quad\quad\quad\quad \times \\
\quad\quad\quad\quad\quad \cdot \\
\quad\quad\quad\quad \cdot \\
\quad\quad\quad \cdot \\
\hline
\ldots 6\ 6\ 6
\end{array}$$

Dit geeft ons dus de vraag: zoek $x_0, x_1, x_2, \ldots \in \{0, 1, \ldots, 6\}$ opdat de vermenigvuldiging hierboven kloppend wordt gemaakt.
x_0 moet dus voldoen aan: x_0^2 eindigt op 6 d.w.z. $x_0^2 = 7$voud $+6$.
Eén blik op de diagonaal van de tabel onderaan blz.7 leert ons dat zo'n x_0 niet te vinden is. Dus, het lukt niet:

$$\boxed{\text{Er is geen } x \in \mathbb{Q}_7 \text{ met } x^2 = -1}$$

Niet uit 't veld geslagen proberen we $p = 5$. We krijgen dan: zoek $x_0, x_1, x_2, \ldots \in \{0, 1, \ldots, 4\}$ zó dat

$$
\begin{array}{l}
\ldots\ x_2x_1x_0 \\
\ldots\ x_2x_1x_0 \\
\hline
\qquad\quad . \\
\quad\ . \\
\ \ . \\
\hline
\ldots\ \ 4\ 4\ 4
\end{array}
\times
$$

kloppend wordt. Nu blijkt dat er twee keuzen voor x_0 zijn: $x_0 = 2$ en $x_0 = 3$. Beginnen we met $x_0 = 2$. We zoeken nu x_1 zó dat

$$
\begin{array}{l}
\ldots x_1 2 \\
\ldots x_1 2 \\
\hline
\qquad 4 \\
\ \ . \\
\ \ . \\
\hline
\quad 44
\end{array}
$$

klopt. D.w.z. $4x_1 = 4 + 5$-voud. Hier is maar één mogelijkheid: $x_1 = 1$. Vervolgens:

$$
\begin{array}{l}
\ldots x_2 12 \\
\ldots x_2 12 \\
\hline
\qquad 24 \\
\quad 12 \\
\ \ . \\
\hline
\quad 444
\end{array}
$$

Dan moet $4x_2 + 1 = 5$-voud $+4$ d.w.z. $4x_2 = 3 + 5$-voud. Ook hier maar één oplossing: $x_2 = 2$. Volgende stap:

$$
\begin{array}{l}
\ldots x_3 212 \\
\ldots x_3 212 \\
\hline
\qquad 424 \\
\quad 212 \\
\quad 24 \\
\ \ . \\
\ \ . \\
\hline
\qquad 444
\end{array}
$$

Nu krijgen we: $4x_3 + 2 + 2 + 1 = 5\text{-voud}+4$. Dus $x_3 = 1$.
Volgende stap

$$\begin{array}{r} x_4 1212 \\ x_4 1212 \\ \hline 2424 \\ 1\ 212 \\ 24\ 24 \\ 121\ 2 \\ . \\ . \\ . \\ \hline 4444 \end{array}$$

$4x_4 + 1 + 4 + 1 + 1 = 5\text{-voud} + 4$.

$$\begin{aligned} 4x_4 &= 5\text{−voud} + 2 \\ x_4 &= 3 \end{aligned}$$

en zo ga je verder: steeds krijg je, als je $x_0, \ldots, x_n$ reeds gemaakt hebt voor x_{n+1} een vergelijking van 't type

$$4x_{n+1} = 5\text{−voud} + \text{iets bekends}$$

en dat is steeds op te lossen dank zij Lemma 3.2^{bis}.
Dus:

$$\boxed{\text{Er is wèl een } x \in \mathbb{Q}_5 \text{ met } x^2 = -1}$$

Er zijn zelfs twee oplossingen: de startwaarde $x_0 = 3$ (zie onderaan blz.10) leidt ook tot 'n oplossing.
Nu rijst de vraag: Voor welke priemgetallen p heeft $x^2 + 1 = 0$ een oplossing in Q_p en voor welke niet?
Het antwoord is bijzonder fraai. We zullen bewijzen:

STELLING 4.1.
Als p een 4-voud plus 1 is dan heeft $x^2 + 1 = 0$ (twee) oplossingen in $\mathbb{Q}_p$.
Als p geen 4-voud plus 1 is dan heeft $x^2 + 1 = 0$ geen oplossingen in $\mathbb{Q}_p$.
Dus de stelling zegt:
Wel oplossingen voor $p = 5, 13, 17, 29, 37, 41, \ldots$
Geen oplossingen voor $p = 2, 3, 7, 11, 19, 23, 31, \ldots$
Dit klopt met onze eerdere ervaringen met $p = 7$, $p = 5$.

BEWIJS VAN STELLING 4.1.
Het geval $p = 2$ moet even apart behandeld worden:

4.B OPGAVE. Door in de stijl van blz.9-10 te prutsen met vermenigvuldigingen, toon aan dat $x^2 + 1 = 0$ *geen* oplossingen heeft in $\mathbb{Q}_2$.

In de rest van 't bewijs nemen we aan: p is oneven.

Stel p is een 4-voud plus 1. Ik ga aantonen dat $x^2+1=0$ 'n oplossing heeft in $\mathbb{Q}_p$:

$$\begin{array}{l} \ldots\ x_2 \quad x_1 \quad x_0 \\ \ldots\ x_2 \quad x_1 \quad x_0 \\ \hline \qquad\qquad\quad . \\ \qquad\qquad . \\ \qquad . \\ \hline \ldots\ p-1 p-1 p-1 \end{array}$$

Het cruciale punt is of ik een x_0 kan vinden met $x_0^2 = p$-voud $+(p-1) = p$-voud-1. Die x_0 zal moeten liggen in $\{0,1,\ldots,p-1\}$, maar als ik maar een $x_0 \in \mathbb{Z}$ zou kunnen vinden met $x_0^2 = p$-voud -1 dan ben ik wel klaar want door bij x_0 een p-voud erbij of eraf te doen kan ik in $\{0,1,\ldots,p-1\}$ komen en

$$(x_0 \pm p\text{-voud})^2 = p\text{-voud} - 1$$

blijft overeind.
In opgave 3.D hebt U hopelijk gevonden dat (voor p oneven)

$$1.2.3.\ldots.(p-1) = p\text{-voud} - 1$$

(Hebt U die som niet gemaakt? Dan doe ik 'm nu even: Volgens lemma 3.2 is er bij elke $t \in \{1,2,\ldots,p-1\}$ er (precies) één $u \in \{1,2,\ldots,p-1\}$ te vinden met $tu = p$-voud$+1$. Bijv. voor $p=7$: (zie lijst op blz. 7)

$$t: 1\ 2\ 3\ 4\ 5\ 6$$

$$u: 1\ 4\ 5\ 2\ 3\ 6$$

Dus, voor algemene p:

$$t: 1\ 2\ 3\ \ldots\ p-1$$

$$u: 1\ ?\ ?\ \ldots\ p-1$$

Ik zie nu $1.2.3.\ldots.(p-1)$ zó:

$$1 \times \underline{(2 \times 3 \times \ldots \times (p-2))} \times (p-1)$$

in 't middelste gedeelte heeft ieder getal s een getal t met $st = p$-voud$+1$. Die pak ik in paren bij elkaar:

$$(p-1)! = 1 \times (p\text{-voud}+1)(p\text{-voud}+1)\ldots(p\text{- -voud}+1) \times (p-1).$$

Het product rechts van het $=$-teken heeft $\frac{p-1}{2}$ stuks "(p-voud $+1$)". Vermenigvuldigd levert dit ook nog steeds een p-voud $+1$ op. Dus:

$$(p-1)! = (p\text{-voud}+1)(p-1)$$

hetgeen een p-voud -1 is. Reken maar na.
Dus we hebben

$$1.2.3.\ldots.(p-1) = p\text{-voud min } 1.$$

Splits dit:

$$1.2.3.\ldots.\frac{p-1}{2}\,\underline{\frac{p+1}{2}\ldots(p-1)} = p\text{-voud min } 1.$$

Van ieder van de getallen in 't onderstreepte gedeelte haal ik p af:

$$1.2.\ldots.\frac{p-1}{2}\left(-\frac{p-1}{2}\right)\left(-\frac{p-3}{2}\right)\ldots(-2)(-1) = p\text{-voud min1}.$$

Dus:

$$\left(\frac{p-1}{2}\right)!(-1)^{\frac{p-1}{2}}\left(\frac{p-1}{2}\right)! = p\text{-voud min } 1.$$

Ofwel

$$(-1)^{\frac{p-1}{2}}\left(\left(\frac{p-1}{2}\right)!\right)^2 = p\text{- -voud min } 1.$$

Maar, nu was $p = 4$ -voud $+1$ dus $\frac{p-1}{2}$ is even, dus $(-1)^{\frac{p-1}{2}} = 1$. We krijgen zo:

$$\left(\left(\frac{p-1}{2}\right)!\right)^2 = p\text{-voud min } 1.$$

We hebben een getal in $\mathbb{Z}$ gevonden waarvan 't kwadraat een p-voud -1 is en dus zijn we klaar met dit stuk van het bewijs (zie blz.13): we hebben een x_0 die in de vermenigvuldiging op blz. 13 voldoet.
De rest is nu kinderspel (ga zelf na hoe men nu $x_1, x_2, \ldots$ moet vinden).

Stel p is geen 4-voud plus 1. Ik ga laten zien dat $x^2 + 1 = 0$ geen oplossing heeft in $\mathbb{Q}_p$ door (zie vermenigvuldiging op blz. 13) aan te tonen dat er *geen* $x_0 \in \{1, 2, \ldots, p-1\}$ is met $x_0^2 = p$-voud-1.
Stel er was wel zo'n x_0.
Bekijk van $1^2, 2^2, \ldots, (p-1)^2$ alle resten bij deling door p. Omdat 1^2 en $(p-1)^2$, 2^2 en $(p-2)^2, \ldots$ dezelfde resten opleveren heb ik slechts $\frac{p-1}{2}$ van die resten.
Stel r is zo'n rest: $x^2 = r + p$-voud voor zekere $x \in \{1, \ldots, p-1\}$.
Neem $(x_0 x)^2 = (r + p\text{-voud})\ (p - 1 + p\text{-voud})$
$= (p - r) + p$-voud.
Dus als r als zo'n rest voorkomt dan ook $p - r$. D.w.z. 't aantal resten is even: $\frac{p-1}{2}$ is even. Dus $p - 1 = 4$-voud. Tegenspraak.

Dus: Als p geen 4-voud plus 1 dan heeft $x^2 + 1 = 0$ geen oplossingen in $\mathbb{Q}_p$.

De tussenwaardestelling in MAVO 3

A.J. Goddijn

Inleiding

We moeten het nu gaan hebben over intuïties die gewone mensen over de reële getallen hebben. De organisatie van deze cursus raadt mij aan daar ervaringen met leerlingen bij te gebruiken en verder nog wat wiskunde in het verhaal te stoppen. We beginnen met leerlingen die grafieken moeten tekenen en we eindigen met computers die vloeiende lijnen moeten tekenen; het blijkt allebei op een of ander manier hetzelfde te zijn.

Ik neem daarbij de vrijheid in wat minder strenge stijl over wiskunde te praten. Erger nog, ik zal onaangekondigd dingen beweren die pertinent onjuist zijn maar het wel lijken te zijn.

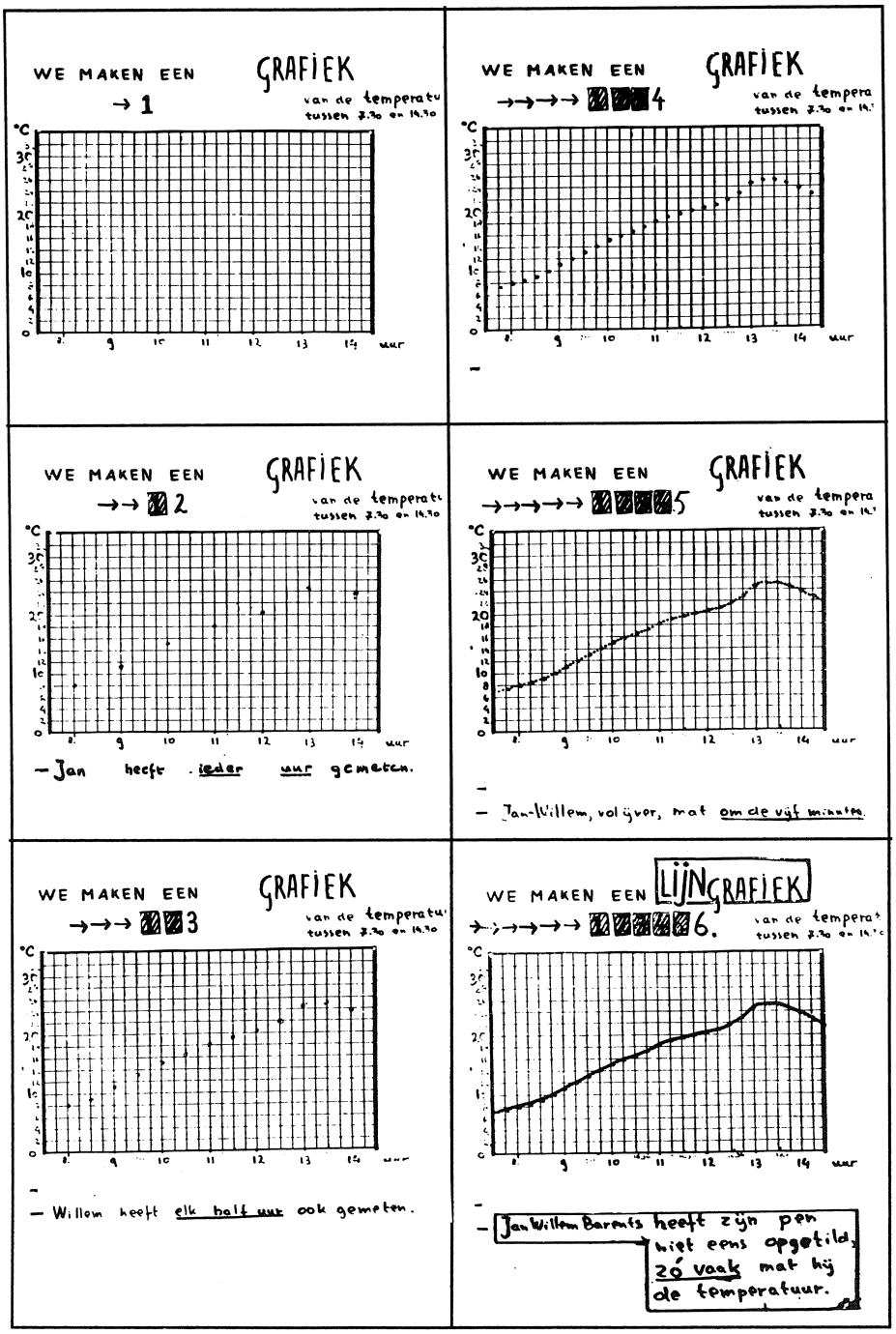
Op naar de reële getallen

Figuur 1 vond ik in 'Klas 1c en de lijngrafieken'. Dat is een beschrijving van de lessen, neerslag van het tussentijdse overleg en sfeertekening rond mijn eerste produkt bij het toenmalige IOWO: Lijngrafieken. Het is november 1978, de beschrijving is gemaakt door George Schoemaker en Nol van 't Riet.

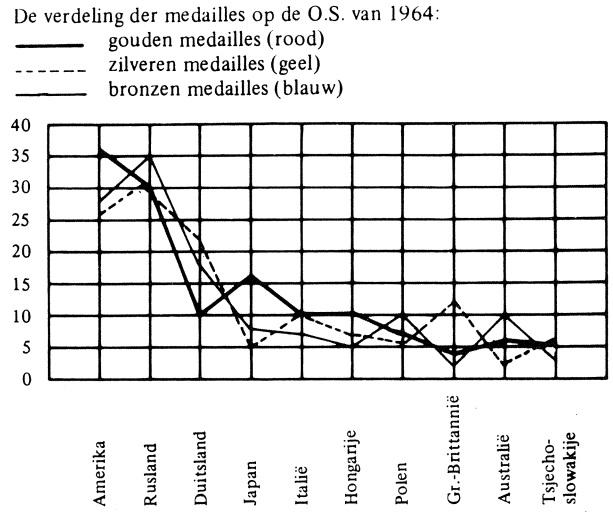
Het pakketje introduceert grafieken met behulp van het temperatuurverloop op een dag. Dat vinden we nu heel gewoon, maar omdat het 1978 was en het eigenlijk over algebra moest gaan was het toch wel bijzonder. Achteraf zie ik dat het pakketje helemaal niet over algebra maar over analyse gaat. Straks meer rond dat verschil.

Op zeker moment, na een hoop activiteiten van de leerlingen is er een show met de overheadprojector. Er wordt iets op de overheadprojector gedaan wat op het bord niet kan en op papier ook moeilijk is te reproduceren: doorzichtige vellen moeten over elkaar heen gelegd worden.

Bij elke nieuwe sheet staan er meer punten in de grafiek, tussen de al aanwezige punten in. Mij leek dat toen belangrijk: de geleidelijke overgang van stippengrafiek naar continue lijn. Want juist dat continue variëren van de variabele, daar moest op voorbereid worden nog zonder de echte algebraletters en ver vóór de merkwaardige produkten de klas in kwamen. En ook moest maar weer eens goed duidelijk worden dat zulke grafieken met lijnen als in figuur 2 absoluut uit den boze waren.



Figuur 1. De 6 sheets bij lijngrafieken, verkleind.



Figuur 2. Medailles in slechte grafiek. Uit een veelgebruikte rekenmethode van de jaren zestig.

In het verslag is te lezen dat de projectorshow prachtig was, maar dat de stap in het zesde plaatje naar de vloeiende lijn waarschijnlijk door weinig leerlingen is opgepakt.

Als er zo'n begripsdrempel opduikt bij leerlingen van pakweg dertien en veertien jaar oud, kun je er vaak wel vanuit gaan dat je met fundamentele wiskundige of filosofische problemen bezig bent. Niet dat je die met de leerlingen op dat moment moet gaan oplossen! In zo'n geval als dit is er niets storends aan de hand wat het werk in de klas betreft; de grafieken worden uiteindelijk heus wel met een lijntje getekend.

De dieper liggende problemen lijken hier te zijn:

- de oneindige deelbaarheid van de tijd
- de overgang van eindig veel (maar steeds meer) rationale getallen naar de volle lijn van de reële getallen.

Het eerste lijkt de realiteit te betreffen, maar ook dat weten we niet zo zeker. Kerkvader Augustinus (354-430) beschreef het begrip tijd als puur subjectief. De eeuwigheid van God was juist zijn buiten de tijd staan. Alle momenten in een keer kunnen overzien, zoals wij naar zo'n lijngrafiek als van de zesde sheet in klas 1c kijken. Tijd bestaat niet, alleen wij mensen met onze beperkte vermogens denken in termen van verleden, heden en toekomst.

De problematiek van de oneindige deelbaarheid van de tijd is beroemd geworden door toedoen van Achilles, die een wedstrijd hield met de schildpad. Het

verhaal is bekend: Achilles geeft de schildpad een voorsprong, niet uit grootmoedigheid (lees Homerus over Achilles' karakter) maar uit pure domheid. Immers, als Achilles aankomt op de plek waar de schildpad op het startmoment was, dan is op dat moment de schildpad toch net iets verder. De zaak ligt dus er bij zoals in het begin: weer moet Achilles naar de plek waar de schildpad dan is en, als Achilles daar aangekomen is, ligt de schildpad wéér voor. Dat herhaalt zich oneindig vaak. Ik laat nu even de conclusie weg dat Achilles de schildpad inderdaad niet inhaalt, iedereen weet dat immers.

Voor ons, wiskundige twintigste eeuwers, is het eenvoudig allerlei redeneringen en berekeningen aan te dragen die aantonen dat Achilles de schildpad wél inhaalt. We weten ook wel dat we daarmee niet het vreemde gevoel dat deze paradox van Zeno steeds oproept kunnen weg nemen. Meetkundige reeksen en lineaire vergelijkingen staan machteloos tegenover die oneindige rij achter elkaar liggende punten in de tijd waarop Achilles de schildpad niet heeft ingehaald. Tussen al die punten in zit steeds weer een lapje aaneengesloten tijd, dat houdt nooit op.

Steeds maar méér momenten in eindige tijdsruimte, dat is ook wat vertoond wordt bij de vijf over elkaar liggende puntgrafieken.

Zouden we gered zijn als we in plaats van aan de lopende tijd aan de lijn van de reële getallen gaan denken?

Het klassiek voorbeeld waarmee aangetoond wordt dat er meer moet zijn dan rationale getallen alleen is het zoeken naar de oplossing van de vergelijking $x^2 = 2$.

Met rationale x, dus $x = p/q$ met p en q geheel, lukt dat niet. Na omwerken tot $p^1 = 2q^2$ komen we in moeilijkheden als we naar het aantal factoren 2 links en rechts kijken: link is dat even, rechts oneven. Eenvoudig. Maar toch onbegrijpelijk want je kunt wel 'oneindig dicht' bij die gezochte komen. Wiskundig mooi, maar het helpt de intuïtie niet. Het bewijs, en andere bewijzen van dezelfde bewering, overtuigt alleen het in redeneren gelovend deel van de mens en de rest niet.

Ik zet U nu een paradox voor die wat minder bekend is, maar nog heviger duidelijk maakt dat onze (of laat ik bescheiden wezen: mijn) intuïties over de rationale en reële getallen bepaald niet zijn wat ze moeten wezen. We weten dat de rationale getallen dicht liggen in de reële getallen. Dat is dat 'oneindig dicht' bij die gezochte kunnen liggen. Nu leg ik eens open intervallen om alle rationale getallen. Een open interval, als je daar als getal in zit, dan kun je je nog enigszins roeren. Je hebt een omgevinkje, hoe klein dan ook, om je heen. Zeker zitten er links en rechts van je nog rationale getallen in dat intervalletje om je heen, een klein maar toch niet-nul afstandje van je vandaan. Ook het stukje tussen jouw en dat buur-rationale getal zit helemaal in het intervalletje. Zo'n verzameling van open intervallen om alle rationale getallen overdekt dus uiteraard de hele getallenlijn.

Als U nog geen argwaan heeft moet u het krijgen als we het kiezen van

de open intervallen preciezer maken. We nemen daartoe een aftelling van de rationale getallen, zeg $r_1, r_2, \ldots$. In die rij zitten dus alle rationale getallen, niet naar grootte geordend maar dat hoeft ook niet. Om r_n leg ik een open interval van lengte $^{-n}$. Die intervallen zullen hier en daar wel overlappen, dus de totale maat van het overdekte gebied is zeker kleiner of gelijk aan:

$$\sum_{n=1}^{\infty} 2^{-n} = 1$$

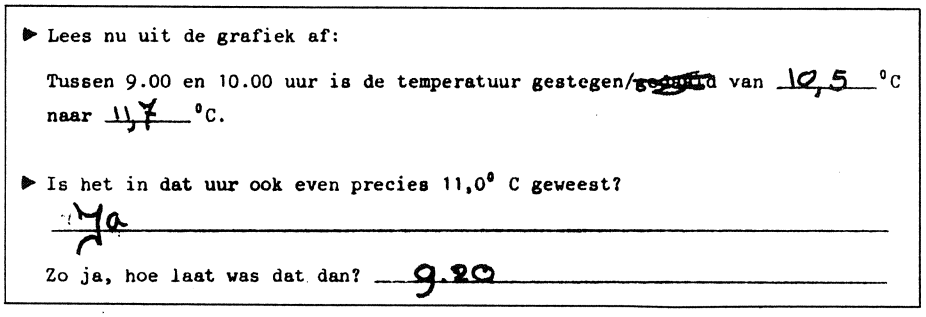
Zo heb ik dan alle reële getallen gevat in een verzameling met maat kleiner of gelijk 1.

Ofwel in de informele redenering, ofwel in het deel met de aftelling en de sommering moet iets misgegaan zijn. Raar is het wel, ook al weet je welke redenering de verkeerde is. Een intuïtieve voorstelling van de rationale getallen in de reële getallen liggend met 'gaten er tussen' zal hier wel de storende factor zijn. De fout zit 'natuurlijk' in de zin hierboven met 'uiteraard' maar dat helpt me eigenlijk weinig.

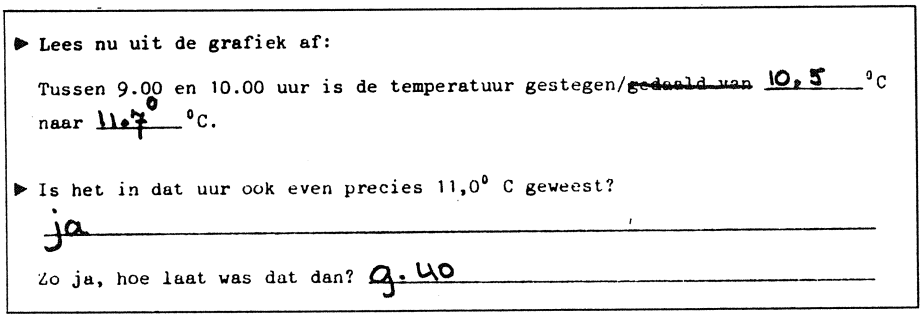
De tussenwaardestelling

Terug naar de lijngrafieken in klas 1c van 1978, in verband met het oplossen van de vergelijking $x^2 = 2$. In het pakket zit een vraag die daar in zekere zin op voorbereid. In figuur 3 zien we wat Minet van der Poel op het werkblad invulde. Als de temperatuur stijgt van $10,5°$ naar $11,7°$, dan wordt de $11°$ gepasseerd. Minet weet ook wanneer.

Andrea van Zijl is het met het passeren eens, zie figuur 4. Minet en Andrea verschillen wel van mening over het juiste tijdstip, maar hun beider 'ja' maakt een solide, zekere indruk.



Figuur 3. Minet van der Poel, 1c, past de tussenwaardestelling toe.



Figuur 4. Andrea van Zijl ook, maar zij vindt een ander tijdstip.

Het is in feite wat we in de analyse de tussenwaardestelling noemen. Als een op een interval $[a, b]$ continue functie f voor a negatief is en voor b positief, dan is er ergens tussen a en b een x met $f(x) = 0$. In de officiële formulering staat iets meer, namelijk dat het geheel zich afspeelt op de reële getallen.

Het bewijs van de stelling loopt als volgt. Op het midden van $[a, b]$ is f positief, negatief of nul. In dat laatste geval hoeven we niets meer te bewijzen. Als $f((a + b)/2)$ positief is gaan we verder met de linkerhelft van ons interval, anders met de rechter. In de gekozen helft geldt nu weer dat f op de uiteinden van het interval links negatief is en rechts positief. Het lijkt dat we zo niet verder komen, maar we doen het nogeens en nogeens. Zo wordt een oneindige rij gesloten intervallen geconstrueerd, waarvan elke volgende een helft is van de vorige. Als een merkwaardig pannennest zitten ze in elkaar.

Nu komt de cruciale stap: er is één getal dat in al in die intervallen ligt en dat is de x die we zoeken. Van dat getal (aangenomen dat het er inderdaad is) is gemakkelijk te bewijzen dat de functiewaarde ervan 0 moet zijn. Zo niet, dan is er namelijk een open intervalletje omheen te vinden waarop f aan één kant van nul blijft (dat is de continuïteit van f). In zo'n open intervalletje ligt een van de intervallen van het pannennest in zijn geheel, omdat de lengte van de intervallen uit het pannennest naar nul gaat. Maar dat gaat fout met de waarden van f op de grenzen van dat interval! De echt cruciale stap is dus die naar het bestaan van het ene getal dat in al die intervallen ligt.
Dat getal is

> *het kleinste getal dat groter of gelijk is aan alle linkergrenzen van de intervallen.*

Daarmee is het probleem vertaald naar een basiseigenschap van de reële getallen, namelijk die van de kleinste bovengrens:

> *als een verzameling reële getallen een bovengrens heeft, dan heeft die verzameling ook een kleinste bovengrens.*

(Een bovengrens van een verzameling getallen is een getal dat groter of gelijk is dan de getallen van de verzameling, maar hoeft zelf niet tot de verzameling te behoren.)

Subtiel, maar Minet en Andrea hadden er blijkbaar intuïtief vertrouwen in.

Iedereen schijnt dat de hebben. Dat is ook de moeilijkheid in de didactiek van dit soort analyse: je staat dingen te bewijzen die zo klaar als een klontje lijken en je maakt ze alleen maar moeilijker.

Maar, merkt nu de oplettende lezer op, ik ben alweer bezig U te bedriegen. Minet en Andrea zijn met temperaturen bezig en niet met reële getallen. Dat is waar, en precies de kern van de vraag naar de intuïtie over de reële getallen. Enerzijds redeneren we met temperatuur alsof het iets glijdends, vloeiends en continuus is. Kortom: alsof het zich laat modelleren naar de reële getallen. Anderzijds gebruiken we beelden (of realiteiten) als temperatuur, tijd en afstand om ons redeneren over de reële getallen te ondersteunen. Het mag niet van de echte wiskunde, maar dat lijken me zondagse praatjes die in de dagelijkse arbeid niet houdbaar zijn.

Het proces in het bewijs van de tussenwaardestelling kunnen we nog van een andere kant bekijken. Neem daarvoor voor het gemak even $a = 0$ en $b = 1$. Bij elke keuze naar het volgende interval noteren we nu of we links of rechts moesten nemen met een 0 of een 1. Een rij nul- len en enen ontstaat, bijvoorbeeld:

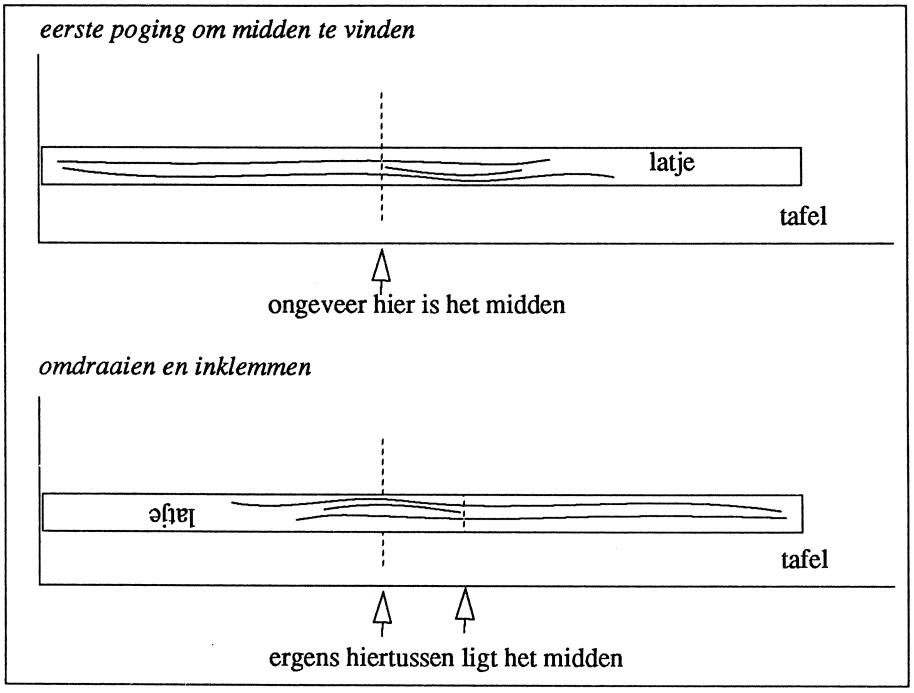
01101001000111.....

Dat is juist de binaire ontwikkeling van het gezochte getal x. Er is ook een opbouw van de reële getallen mogelijk via 'oneindig doorlopende kommagetallen', in basis 10 of 2, dat doet er niet zoveel toe. Weer blijken we met zo'n stelling en zijn bewijs in het hart van de reële getallen te zitten.

In het nieuwe wiskunde programma voor 12-16 komen termen als rationaal, reëel, irrationaal niet meer voor. Merkwaardig genoeg laten we leerlingen wel werken tegen de achtergrond van de intuïties die Minet en Andrea (en de rest van 1c) zo duidelijk hebben. Bij de temperatuurgrafiek was niet precies te bepalen wanneer het nu exact 11 graden was, maar dat is wat in de experimenten van het W12-16 project ook is opgelost: de in het wiskundige bewijs gegeven inklemmethode is daar tot leerstof verheven!

Daartoe nog even verder terug in de tijd.

Toen ik 9 jaar was maakte ik in een bepaalde periode bijna dagelijks een vlieger. Steeds weer andere modellen natuurlijk, maar ook omdat de vliegers snel stuk gingen omdat ik de latjes met de figuurzaag uit hout van sinaasappelkistjes zaagde. Ik piekerde wel over de theoretisch mogelijkheid van asymmetrische vliegers, maar uiteindelijk dorst ik het experiment nooit aan en bepaalde steeds het midden van het gezaagde latje. Ik mat daarvoor niet de lengte op en deelde niet door twee, maar deed het zo: Latje met het uiteinde tegen de tafelrand leggen, ongeveer in het midden een streep over lat én tafel zetten; latje omdraaien. Dan zag ik twee streepje: op de tafel hn op de lat en meestal klopte het niet. Maar ik wist zeker dat de volgende betere poging tussen de streepjes moest liggen. Zie figuur 5.



Figuur 5. Midden van een latje bepalen

Ik itereerde altijd twee keer, maar het inklemproces als proces zat er in principe in.

Deze methode is als voorbeeld opgenomen in het experimentele pakket 'Ergens ertussen', na het volgende spel.

De ene speler heeft een getal onder de 100 in gedachten en de andere moet het raden. Die mag vragen stellen zoals "Is het groter dan 75?" en krijgt alleen ja of nee als antwoord. Hier wordt het een beetje onhandig door de rader gespeeld:

Rader	Weter
groter dan 30	ja
kleiner dan 90	ja
groter dan 25	ja
kleiner dan 89	ja

Verstandige spelers beginnen met "Groter dan 50?" en gaan dan over naar ofwel "Groter dan 25" ofwel "Groter dan 75?". De parallel met het bewijs van de tussenwaardestelling is frappant en in principe gaat het hier ook over een vergelijking. Een orakel spuit functiewaarden uit (een f-waarde of een 'ja' of 'nee') en er is een opzoekproces. In 'Ergens Ertussen' wordt er naar toe gestuurd dat leerlingen zo'n zoekproces georganiseerd opschrijven in tabelvorm

en iets verderop wordt ook gewerkt aan problemen als: *zoek een getal dat samen met 10 keer zijn wortel juist 1000 is.*

Zullen we niet gewoon vierkantsvergelijkingen blijven doen in de MAVO, want daar kan dit toch ook wel mee?

Maar met de inklemmethode kunnen dingen waar je anders niet eens aan mag denken. Een voorbeeld, een beetje boven MAVO-niveau vanwege de subtiele gonio.

Span een touw om de aarde en blijf opletten want dit is niet het bekende touw om de aarde. We knippen het touw open en voegen precies 10 cm in. Nu komt het slap te liggen. Zou je er onder door kunnen kruipen? Hoe hoog moet de (ene!) paal eronder zijn om het weer strak te krijgen?

De vraag is vlug gesteld maar het is moeilijk om ook maar een idee te krijgen van de grootte van die paal!

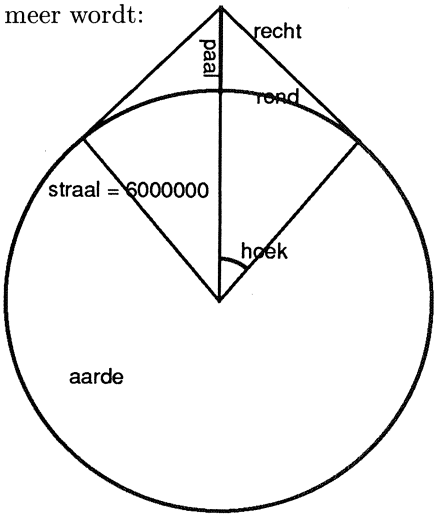
In figuur 6 zijn aarde, touw en paal te zien. Niet op schaal.

Met wat gonio komen we tot de volgende vergelijkingen (alles in meters en radialen uitgedrukt):

$$\begin{aligned} rond &= hoek \times 6000000 \\ recht &= \tan(hoek) \times 6000000 \\ paal &= (1/\cos(hoek) - 1) \times 6000000 \end{aligned}$$

De voorwaarde van de tien centimeter meer wordt:

$$recht = rond + 0.05$$



Figuur 6. Aarde met touw erom. Verleng het touw met 10 cm. Span het met één paal eronder. Hoe hoog is die paal?

Als we *'hoek'* uit die laatste vergelijking kunnen vinden, weten we *'paal'* ook. Eigenlijk is die vergelijking van de vorm:

$$\tan(hoek) = hoek + 0.05/6000000$$

en daar is geen gewone wiskundige methode tegen opgewassen. Je kunt blijven omvormen, maar je houdt goniometrische functies (desnoods vermomd als imaginaire e-macht) en gewone dingen in één vergelijking.

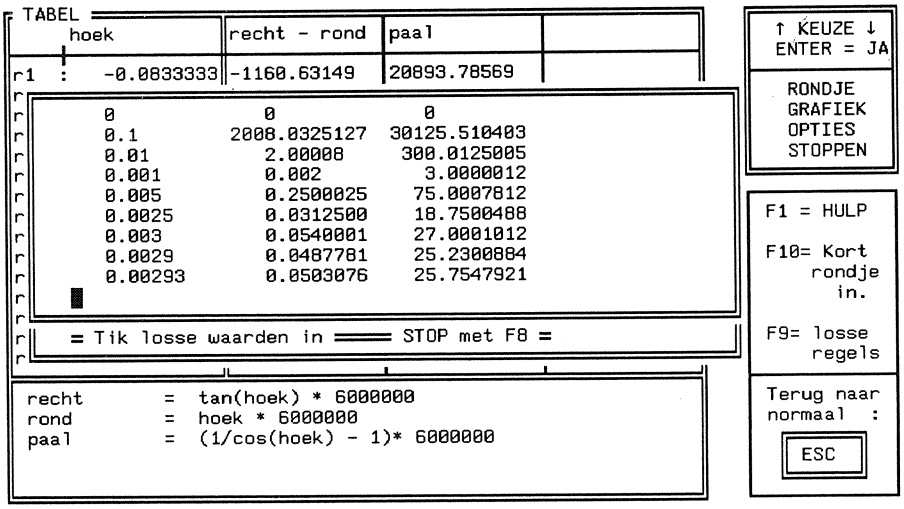
Eigenlijk is het heel vreemd dat we zo weinig meetkunde vraagstukken tegenkomen waar zulke onaangename vergelijkingen uitrollen. Of hebben we na vele eeuwen wiskunde bedrijven, gericht op expliciet kunnen oplossen van vergelijkingen, ons teveel gefixeerd op wat in dat keurslijf past? Dat is een vraag die zo even terzijde hierbij opkomt.

Er zijn overigens veel van die problemen die tot dit soort zogenaamde transcendente vergelijkingen leiden. Zoeken we naar iets waar een recht lijnstuk op een of andere manier van een cirkel afgepeld is, dan is het al gauw raak. Gewoon de vraag naar de omtrek van de cirkel is er al een, maar daar is ook heel wat menselijke energie opgeofferd voor Lindemann in 1882 bewees dat het echt om een transcendent probleem ging.

Tegenwoordig zijn we tevreden met het antwoord 'π', maar dat is vroeger altijd beschouwd als een benaderingsantwoord, waarbij het echte probleem nog opgelost moest worden. Maar terug naar de onbekende paal.

Inklemmen, nu met behulp van een computerprogramma voor het rekenwerk, is het enige wat nog helpt.

In figuur 7 is te zien hoe dat bijvoorbeeld kan gaan.



Figuur 7. recht - rond wordt ingeklemd. Paal wordt 25 meter.

In de tabelkop (bovenaan) staan *hoek, recht - rond en paal.* We proberen waarden voor '*hoek*' in te voeren zodat we *recht - rond* in de buurt van 0.05 krijgen (de helft van de 10 centimeter extra touw). In het middelste kader is dat opzoekproces te volgen. Eerst is 0 en 0.1 ingevoerd. De 0.1 blijkt veel te groot en het oorspronkelijke halveringsproces wordt versneld door ook andere punten dan middens van de intervallen te gebruiken. Met 0.0029 zijn we al heel dichtbij en vol belangstelling lezen we af dat '*paal*' ongeveer 25 meter moet zijn. De hoek zelf speelt niet zo'n grote rol, het is een hulpgrootheid.

Je kunt je nog afvragen waarom we hier van te voren zo weinig schattend-intuïtief een antwoord kunnen voorspellen.

Daarvoor proberen we te vinden hoe de hoogte van de paal van de lengte van het ingezette stuk afhangt. Met ander worden, we proberen de algemene vergelijking

$$\tan(hoek) = hoek + inzet$$

Omdat het om kleine hoeken gaat, is Taylorontwikkeling hier het juiste wapen.

Voor de tangens weten we :

$$\tan(x) \approx x + \frac{1}{3}x^3$$

en daaruit zien we

$$\frac{1}{3}hoek^3 \approx inzet$$

De formule voor 'paal' benaderen we ook met een korte Taylorreeks:

$$paal \approx 6000000\frac{(hoek)^2}{2}$$

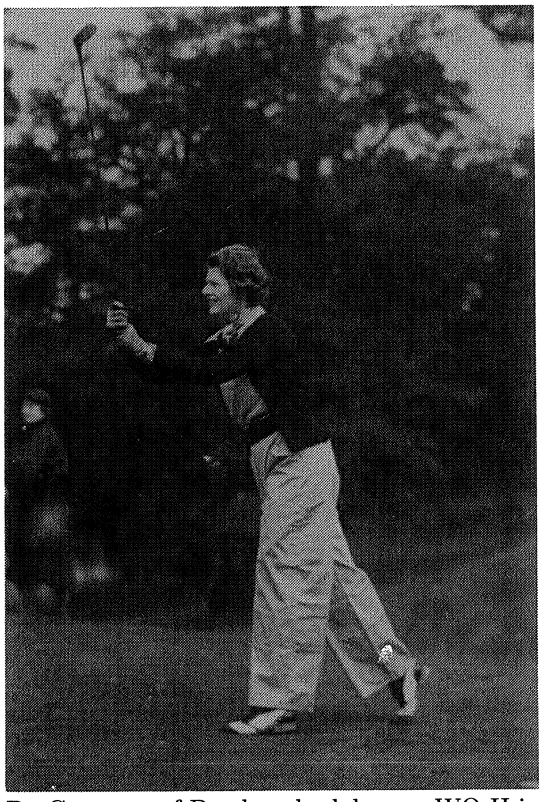
en dan houd je je hart al vast: de paal is ongeveer evenredig met de anderhalve macht van inzet die in ons geval $0.5/6000000$ was. Want als dingen zich dichtbij nul niet een beetje lineair gedragen, schatten we ze in het algemeen slecht in.

Het iteratieve oplossingsproces zelf doet een beetje aan golfspelen denken.

Daar mag je ook tegen het balletje slaan zodat het enigszins in de buurt van de 'hole' komt en daarna mag je nog eens je poging verbeteren. Dat is een aardige sport!

Vergelijk dat nou eens met boogschieten: als de pijl de boog verlaat is het allemaal verder bepaald. Je kunt niets meer. Boogschieten doet denken aan algebra (zoals die tenminste bij vergelijkingen oplossen gebruikt kan worden): je brengt een oplossingsformule in stelling en dan is het in een klap raak. Of je deelt de lengte van het vliegerlatje door twee. Of, als je er niet uit kan komen, je definieert koelweg een mysterieus object dat je $\sqrt{2}$ noemt, en waarvan je niets meer wenst te weten dan dat het een oplossing is van de vergelijking $x^2 = 2$. De basis vormen steeds alleen de rekenbewerkingen en de grootte van zo'n getal(?) speelt voorlopig geen rol, is aanvankelijk niet eens gedefinieerd.

En wat kun je nu feitelijk met de algebraïsche methode aan vergelijkingen oplossen? Oplossen is hier eerst even alleen: via een vast eindig aantal algebraïsche bewerkingen vanuit de coëfficiënten van de vergelijking tot een resultaat komen. Algebraïsche bewerkingen zijn $+, -, *$, / en worteltrekken; wortels ook andere dan 2^e machts, maar wel geheel.



Figuur 8. De Countess of Brecknock vlak voor WO II in een charitatieve match de analyse bedrijvend.

Sinds 4000 jaar weten we hoe dat moet bij tweedegraadsvergelijkingen, sinds ruwweg 400 jaar ook bij derde- en vierdegraadsvergelijkingen. En sinds 172 jaar (Ruffini bewees het onvolledig in 1799, Abel voltooide het bewijs in 1821) jaar weten we dat zo'n algemene oplossing voor vijfde en hogeregraads vergelijkingen niet bestaat. Niet dat die ooit nog eens gevonden zal worden als de wetenschap wat verder voortschrijdt, nee, van het bestaan van zo'n algemeen algoritme is de onmogelijkheid bewezen.

Het geheel is toch wel teleurstellend! Er is wel eens gedroomd (door Descartes o.a.) dat we alle meetkundige problemen via algebra (vergelijkingen dus) definitief konden oplossen. Jammer, of gelukkig maar voor wie van de synthetische meetkunde houdt.

Je zou nog kunnen zeggen: als die wortels mogen, waarom dan niet ook de goniometrische functies en hun inversen in een oplossingsformule toelaten. Okee; maar de 25 meter van de paal van zoëven blijft ook dan buiten schot. Dat komt omdat de tangensfunctie uit te drukken is in complexe e-machten en als we uit de vergelijking

$$\tan(x) = x + c$$

nu een algebraïsch verband tussen x en c konden distilleren,

$$\sum_{i=0}^{N}\sum_{j=0}^{M} a_{ij}x^i c^j = 0$$

dan bestond er ook een algebraïsch verband tussen e^x en x:

$$\sum_{i=0}^{P}\sum_{j=0}^{Q} b_{ij}x^i e^{jx} = 0$$

Daarin zullen de coëfficiënten dan wel complex zijn, maar de term

$$x^P e^{Qx}$$

daarin is voor grote x dermate overheersend dat hij nooit door de lagere orde termen te niet kan worden gedaan. Het zal dus niet lukken.

Vloeiende lijnen schetsen

We gaan nog eens terug naar de lijngrafieken in 1c in november 1978. Natuurlijk moesten er lijnen door de geplotte punten van de grafieken getekende worden. Maar wat voor lijnen? Rechte stukken om de punten netjes te verbinden of vloeiende; we vroegen het de leerlingen zelf.

Het uitgangspunt was de temperatuurgrafiek en daarom kon Miriam zeggen: "Het is een gebogen lijn, 't gaat een beetje in golven, de temperatuur gaat ook niet stotend".

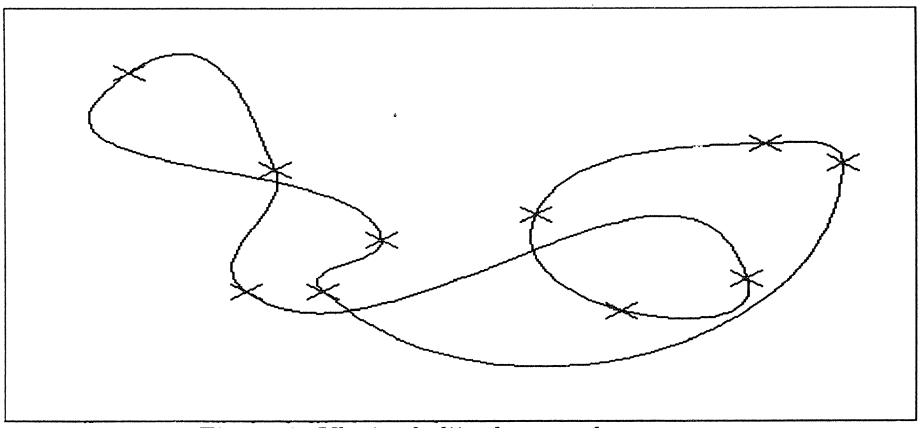
Het schetsen van de vloeiende door gegeven punten is voor ons mensen niet zo moeilijk. Iedereen maakt van

met gemak

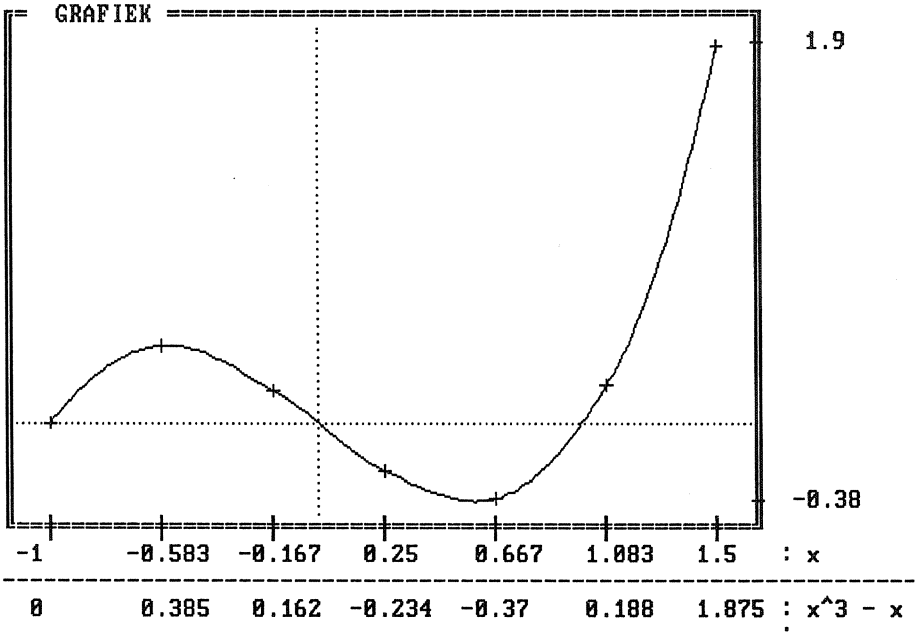
Toch is dat een fantastische prestatie. Uit alle mogelijke curven selecteren we zo uit de hand een heel redelijke. Wat onze hersens in zo'n geval doen is nauwelijks te beschrijven, maar wel te beschrijven is hoe computers zulke dingen doen. Er zijn programma's die ook zo schijnbaar intuïtief vloeiende lijnen kunnen trekken. Figuur 9 is zo gemaakt met een gewone goedkope PC.

De computer zette eerst een stel punten op het scherm (een randomgenerator is hier gebruikt) en tekende daarna de vloeiende gesloten kromme.

Het gebied van de wiskunde waar we nu een klein tipje van de sluier van gaan oplichten, is tamelijk nieuw. Het is alom bekend in de wereld van de informatica en daar moet je dan ook zoeken naar meer details.



Figuur 9. Vloeiende lijn door random punten.



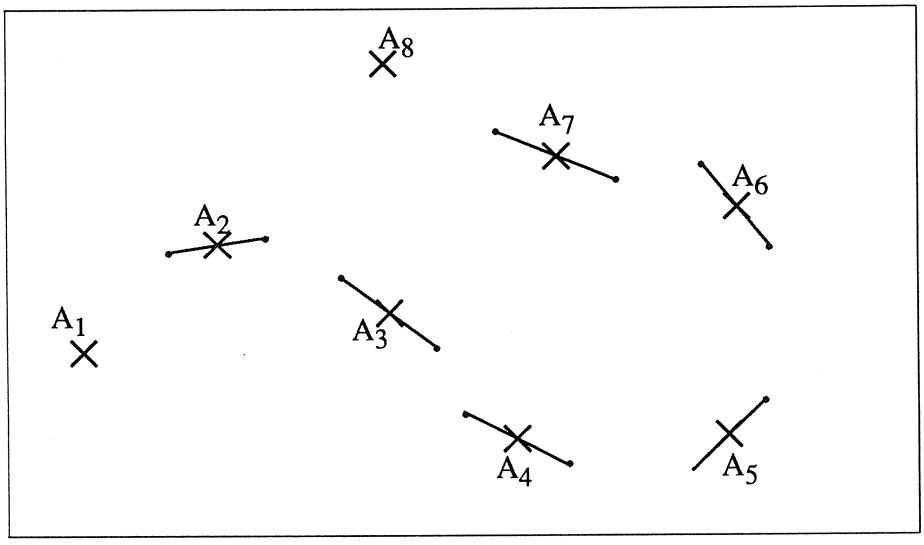
Figuur 10. Derdegraads grafiek, getekend op grond van 7 berekende waarden.

Het TABEL-programma dat in figuur 7 is gebruikt is, werkt samen met een GRAFIEK-programma dat vloeiende grafieken tekent op grond van bijvoorbeeld 7 berekende functiewaarden. In figuur 10 is een derdegraads kromme getekend op grond van de onderaan aangegeven waarden, met zo'n algemeen kromme-schets-algoritme.

Dat is veel sneller dan bijvoorbeeld 200 functiewaarden uitrekenen en het is in dit geval niet te zien. (Als we met dollere functies zoals 1/x gaan werken natuurlijk wel, maar dit programma kan het ook met meer punten als de gebruiker dat wil).

Het algoritme bestaat uit twee fasen. Eerst bepalen we bij de gegeven punten de raakrichtingen van de kromme die we daar willen hebben. Een keuze kan zijn -andere zijn ook mogelijk- : loodrecht op de deellijn van de hoek naar de buurpunten. Op de raaklijnen kiezen we twee hulppunten op gelijke afstanden van het gegeven punt. De afstand is enigszins willekeurig, maar denk aan bijvoorbeeld 1/6 van de afstand van de buurpunten onderling.

In figuur 11 zijn de oorspronkelijke punten, de berekende raakrichtingen en de hulppunten te zien.

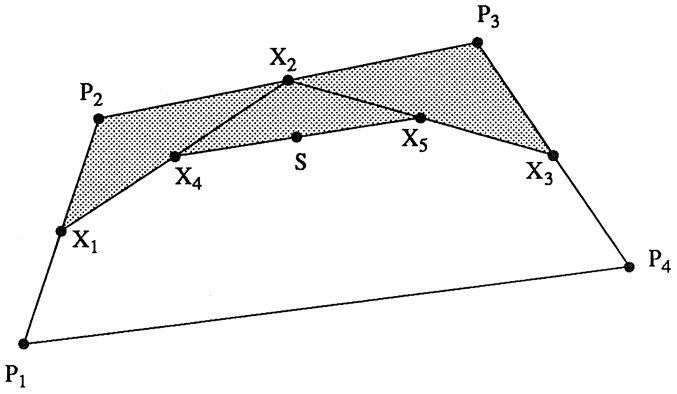


Figuur 11. De oorspronkelijke punten, A_1 t/m A_8, zijn met kruisjes aangegeven; berekende raakrichtingen en hulppunten zijn bepaald.

Nu gaan we stukje voor stukje werken, waarbij we alleen van twee van de oorspronkelijke punten, de gevonden raakrichtingen en twee daarop gekozen hulppunten uitgaan. De uitgangs punten heten P_1 en P_4, de hulppunten op de raaklijnen P_2 en P_3. We willen vanaf P_1 richting P_2 vertrekken en uiteindelijk vanuit de richting P_3 in P_4 aankomen. Het idee is dat we de vierhoek P_1 P_2 P_3 P_4 aan drie zijden bijsnijden, op de manier waarop je met een rechte zaag een cirkel uit een vierkant stuk hout haalt. Voor de hand liggen de snedes die in figuur 12 zijn getekend, waarbij steeds de zaag op de middens van lijnstukken is gezet.

Achtereenvolgens is bepaald:

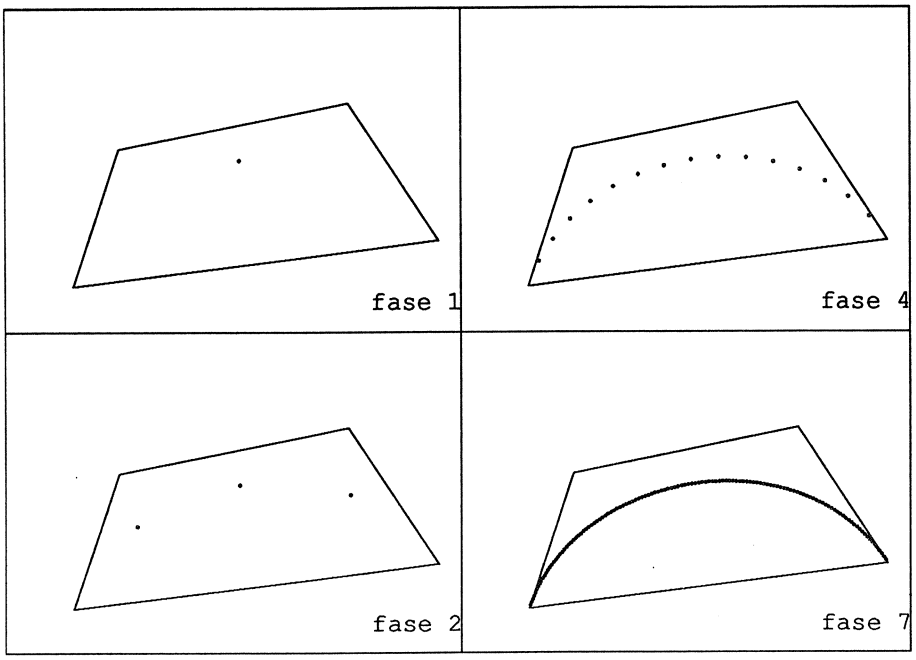
X_1 : midden van P_1 P_2,
X_2 : midden van P_2 P_3,
X_3 : midden van P_3 P_4,
X_4 : midden van X_1 X_2,
X_5 : midden van X_2 X_3,
S : midden van X_4 X_5.



Figuur 12. Een vierhoek afronden. De grijze stukken worden afgezaagd.

Laten we in ieder geval nu maar punt S op onze kromme leggen en de kromme in S aan X_4 X_5 laten raken. De gehoekte lijn van P_1 naar S loopt via X_1 en X_4. Eigenlijk kunnen we op het viertal P_1,X_1, X_4,S het hele proces nog eens toepassen en aan de andere kant op S, X_5, X_3, P_4 ook. Dan vinden we weer punten, in elke helft nu een. X_1 is dus hulppunt om van P_1 naar S te komen. Voor dit deelstuk is de startrichting dus dezelfde als van het oorspronkelijke stuk. Dan volgt de volgende stap: die twee helften worden elk in twee helften verdeeld en vier nieuwe punten ontstaan.

In figuur 12 zien we enkele fasen uit het groeiproces afgebeeld.



Figuur 13. Vier fasen in de opbouw van een kromme.

Dat lijkt toch wel erg veel op de sheets uit 1978!

De krommes die we aan het opbouwen zijn heten Bézier-curves naar Pierre Bézier.

Je zou het proces overzichtelijk zo kunnen noteren:

Zo gaat *Bézier*(p_1, p_2, p_3, p_4):

1. Bereken een aantal middens
   $x1 = midvan(p_1, p_2)$
   $x2 = midvan(p_2, p_3)$
   $x3 = midvan(p_3, p_4)$
   $x4 = midvan(x_1, x_2)$
   $x5 = midvan(x_2, x_3)$
   $s = midvan(x_4, x_5)$
2. Zet een *stip* op s
3. Roep nu twee slaven, die doen:
   Bézier$(p1, x1, x4, s)$ en *Bézier*$(s, x5, x3, p4)$
4. *Klaar.*

Figuur 13 is getekend door een computerprogramma dat als twee druppels water lijkt op wat hier cursief is aangegeven.

Je kunt de beschrijving als definitie van de uiteindelijke kromme opvatten. Als volgt.

De definitie levert in fase 1 eerst een stip, in de volgende slag 2 stippen, daarop 4, enzovoort. Daarbij gaan we ervan uit dat de slaven in de definitie beiden even hard werken en dat ook de door hun geroepen onderslaven dat doen.

Je kunt ook weer denken aan binair bijhouden in welke tak gewerkt wordt. Neem steeds een 0 voor de eerste slaaf en een 1 voor de tweede slaaf. Een rij nullen en enen ontstaat bij het steeds dieper in de recursie gaan, elk punt dat gestipt wordt correspondeert zo met een eindig rijtje; een eindig binaire breuk die tussen 0 en 1 ligt. Je zou je kunnen voorstellen dat de recursie echt oneindig diep gaat, laten we zeggen aftelbaar oneindig. Dan zitten de oneindig lange binaire kommagetallen er ook in en we hopen natuurlijk dat we zo een afbeelding van het interval [0,1] naar een echte kromme te pakken hebben.

Begripsmatig is dat lastig omdat het proces met die werkende slaven als het ware in de tijd wordt voorgesteld. Dat is natuurlijk niet echt de bedoeling. Alles is gewoon door de definitie vastgelegd, ongeacht slaven en praktische moeilijkheden. (Elders in deze vakantiecursus komen nog wel ernstiger vormen van recursie aan de orde om nog meer dan de reële getallen uit het niets te laten ontstaan.)

Hier wordt de recursieve vertakking in de breedte uitgevoerd, zoals een boom groeit. Of zoals een schaker die eerst kijkt welke zetten hij kan doen en dan pas bij elke zet de reactie van de tegenstander bekijkt, en dan pas weer bij al die combinaties kijkt hoe .... enz.

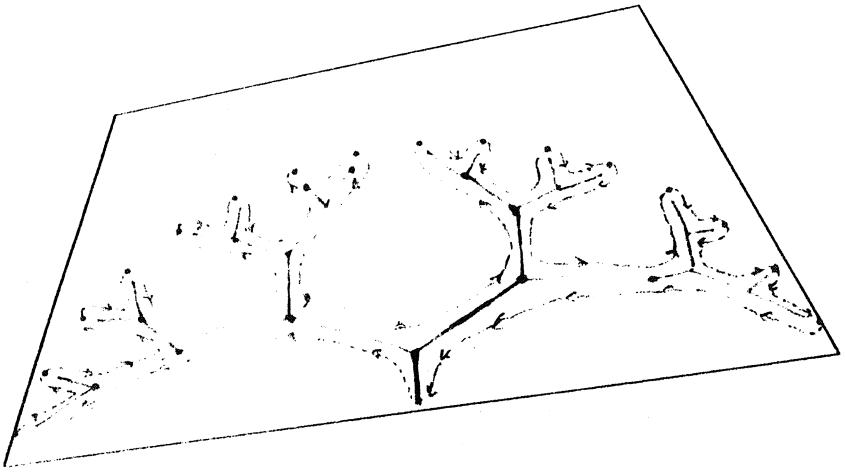
In de PC waar figuur 13 mee getekend is zit maar één slaaf. En dan moet de

eerste zijtak van de recursie klaar zijn voor aan de tweede wordt begonnen. De slaaf is dan ook zijn eigen onderslaaf en als we niet uit kijken komt er alleen een heel diep pad in steeds de linkertak van de recursie. In het werkelijke computerprogramma zit dan ook een stop-voorwaarde. Dat kan zijn: ga niet verder de diepte in dan 6 fasen, of stop als de vierhoek van de vier punten waar mee gewerkt wordt erg smal wordt. In het algemeen zal dan op het moment dat aan de stopvoorwaarde voldaan wordt een recht lijntje getekend worden tussen de hoofdpunten en worden de laatste hulppunten genegeerd.

In figuur 14 is met de stopvoorwaarde 'niet verder dan vier recursieslagen diep' gewerkt en is steed het eerste punt van de argumenten die 'Bézier' meekrijgt gestipt.

Zo gaan we dan in de recursie een stuk links de boom in maar keren steeds weer op de oude takken terug. Merk dat nu de eindig veel punten van links naar rechts netjes in volgorde aan de beurt komen.

Recursie is erg in de mode, vooral in verband met fractals en dan ligt chaos al gauw om de hoek. Hier is dat niet zo, gelukkig maar. Er ontstaat een vloeiende kromme en we kunnen zelfs een parametrisering van de kromme vinden.



Figuur 14. Het zogenaamde backtrackproces getekend bij de vierde fase Bézier benadering.

Hoe stellen we een lijn voor van punt P naar Q?
Dit is de bekende manier:

$$R(t) = tP + (1-t)Q$$

waarbij we P en Q als vectoren opvatten en t van 0 naar 1 loopt.
Je kunt ook noteren:

$$R(t) = \begin{pmatrix} x(t) \\ y(t) \end{pmatrix}$$

In dit geval zijn $x(t)$ en $y(t)$ lineaire functies.

Zoiets proberen we ook bij de vier punten P_1, P_2, P_3 en P_4. Nu hebben we heel wat meer om rekening mee te houden. $x(0), x(1), y(0)$ en $y(1)$ liggen in ieder geval vast, omdat we bij $t = 0$ juist P_1 willen treffen en bij $t = 1$ juist P_4. De afgeleide van $R(t)$ naar t is de snelheidsvector waarmee $R(t)$ van P_1 naar P_4 beweegt. We willen de richtingen en snelheden aan begin en eind vastleggen door

$$R'(0) = c(P2 - P1)$$

en

$$R'(1) = c(P4 - P3).$$

Dat legt bij gegeven c nog eens 4 voorwaarden op aan x en y en we moeten nu wel x en y van de derde graad in t gaan nemen, anders lukt het niet. De geschikte waarde van c vinden we zo dadelijk nog wel.
We zoeken zoiets als:

$$R(t) = f_1(t)P_1 + f_2(t)P_2 + f_3(t)P_3 + f_4(t)P_4$$

Als we $R(0) = P_1$ willen hebben, gaat dat het gemakkelijkst door te zorgen dat f_1 geen factor t bevat en f_2, f_3 en f_4 wel.

We willen ook $R(1) = P_4$ en dat valt te regelen door f_4 geen factor $(1 - t)$ te geven en de andere drie wel. Laten we eens een gooi doen

$$R(t) = (1 - t)^3P_1 + c_2(1 - t)^2tP_2 + c_3(1 - t)t^2P_3 + t^3P_4$$

Als we gaan differentiëren komen we er snel achter dat we $c_1 = c_2 = 3$ moeten nemen. De c in $R'(0) = c(P_2 - P_1)$ is dan ook 3. Resultaat:

$$R(t) = (1 - t)^3P_1 + 3(1 - t)^2tP_2 + 3(1 - t)t^2P_3 + t^3P_4$$

Dat $R(t)$ precies de gewenste eigenschappen heeft is door rekenen en differentiëren snel na te gaan. Als t dicht bij 0 ligt staat R nog het meest onder invloed van P_1 en P_2, als t dichtbij 1 komt, juist meer onder die van P_3 en P_4.

Merk nu dat we een prachtige analogie hebben met het rechtlijnige geval. Daar was R een gewogen gemiddelde van P en Q. Nu is R een gewogen gemiddelde van vier punten, omdat de vier functies f_1, f_2, f_3 en f_4 samen juist identiek gelijk aan 1 zijn. R is een mengsel van de vier punten en in de literatuur heten de functies f_1, f_2, f_3 en f_4 de '*blending functions*'.

Er volgt nog uit: de door R voorgestelde kromme blijft binnen het convex omhulsel van P_1, P_2, P_3 en P_4.

Maar is deze $R(t)$ precies de kromme die ons Bézier-algoritme opleverde? Het is eenvoudig na te gaan dat $R(t)$ voor $t = 0.5$ door het bovengeconstrueerde punt S gaat en dat ook het raken aan de lijn X_4X_5 in orde is. De linkerhelft van onze kromme kan ook door

$$Rl(t) = R(t/2)tin[0, 1]$$

worden voorgesteld en de rechterhelft door

$$Rr(t) = R(1/2 + t/2).$$

$Rl(t)$ heeft nu in P_1 juist de halve snelheid en dat klopt uitstekend met het starten in de richting X_1 die door de linkertak van de recursie wordt gekist, want $(X_1 - P_1) = 1/2(P_2 - P_1)$

De snelheidsaansluitingen in S en P_4 kloppen ook, en dus blijkt de gekozen $R(t)$ prachtig aan de recursieve structuur te voldoen.

In de praktijk blijkt de methode met de middens sneller te werken dan het berekenen via de derdegraadspolynomen. Wie op laag niveau in de computer duikt kan dan ook nog uitbuiten dat er alleen opgeteld wordt en gedeeld wordt door twee en dat zijn nu net heel simpele activiteiten voor de hardware.

De gekozen hulppunten ter weerszijden van de uitgangspunten legden we op gelijke afstanden van de oorspronkelijke uitgangspunten.

Dat is van belang als we de verschillende Bézierkromme's die van A_1 naar A_2 naar A_3 (zie figuur 11) leiden aan elkaar plakken. Want die afstand bepaalt de grootte van de afgeleide van $R(t)$. En op deze manier komt er geen plotselinge snelheidsverandering tot stand als we van de het ene deelgebied over zo'n punt naar het andere deelgebied gaan, wat in wiskundige termen betekent: de totale kromme als geheel is tweemaal differentieerbaar.

Of om het in de terminologie van Miriam uit 1c te zeggen: 'dan gaat het niet zo stotend'.

De continuum-hypothese

J.M. Aarts

In de artikelen [1, 2] heeft Cantor de aanzet gegeven tot de verzamelingenleer. Hij liet met behulp van het begrip kardinaalgetal zien hoe oneindige verzamelingen in "omvang" kunnen verschillen. Wij behandelen hiervan enkele onderwerpen in Paragraaf 1. De kardinaalgetallen kunnen geordend worden naar "grootte". Dit wordt behandeld in Paragraaf 2. In Paragraaf 3 komt het eigenlijke onderwerp van deze voordracht aan de orde, namelijk de Continuumhyphothese, een fundamenteel probleem over de ordening van de kardinaalgetallen. Tenslotte bespreken we in Paragraaf 4 de Souslin-hypothese. Het gaat hierbij om een karakterisering van de reële getallen onder de lineair geordende verzamelingen. Dit onderdeel sluit aan bij hetgeen door professor Grootendorst is behandeld.

1. Kardinaalgetallen

Indien men de omvang van twee eindige verzamelingen wil vergelijken, dan kan men van beide verzamelingen het aantal elementen tellen en op grond van de uitkomst vaststellen welke verzameling de grootste is, dan wel dat de verzamelingen even groot zijn. In feite wordt hierbij elk van de beide verzamelingen vergeleken met een deelverzameling van de natuurlijke getallen $\mathbb{N}$. Cantor heeft er op gewezen dat men óók de verzamelingen direct met elkaar kan vergelijken en aldus op zinvolle wijze kan komen tot de vorming van een getalbegrip voor oneindige verzamelingen.

Definitie 1.1. De verzamelingen X en Y heten *gelijkmachtig* indien er een bijectie $f: X \to Y$ bestaat. Als X en Y gelijkmachtig zijn dan schrijven we $X \sim Y$.

De relatie $\sim$ is een equivalentierelatie. Er is immers voldaan aan de volgende eigenschappen. Voor alle verzamelingen X, Y en Z geldt er:

(reflexiviteit) $X \sim X$,

(symmetrie) als $X \sim Y$, dan $Y \sim X$,

(transitiviteit) als $X \sim Y$ en $Y \sim Z$, dan $X \sim Z$.

De equivalentieklassen met betrekking tot de relatie $\sim$ worden nu de (kardinaal)getallen door middel van de volgende definitie.

Definitie 1.2. Het *kardinaalgetal* van de verzameling X is de klasse van alle verzamelingen Y zó dat $X \sim Y$. We duiden het kardinaalgetal van X aan met $|X|$.

Blijkbaar is $X \sim Y$ dan en slechts dan indien $|X| = |Y|$. We brengen de notatie van de kardinaalgetallen in overeenstemming met die van de gewone getallen door de eerste twee afspraken uit de volgende lijst:

1. $|\emptyset| = 0$.
2. $|\{0, 1, \ldots, n-1\}| = n$, voor iedere n in $\mathbb{N} \setminus \{0\}$.
3. $|\mathbb{N}| = \aleph_0$ (spreek uit: aleph nul).
4. $|\mathbb{R}| = \aleph$.

De kardinaalgetallen 0, 1, 2, ... heten *eindige* kardinaalgetallen. De kardinaalgetallen die niet eindig zijn heten *oneindig*. Een verzameling is *eindig* als haar kardinaalgetal eindig is. Een verzameling die niet eindig is, heet *oneindig*. Een verzameling waarvan het kardinaalgetaal eindig is of gelijk is aan $\aleph_0$ heet *aftelbaar*. We noemen een verzameling die niet aftelbaar is *overaftelbaar*.

We beginnen met te onderzoeken welke verzamelingen zoal aftelbaar zijn.

VOORBEELD 1.3. De verzameling van de gehele getallen $\mathbb{Z}$ is aftelbaar. Men kan bijvoorbeeld een bijectie $f: \mathbb{Z} \to \mathbb{N}$ definiëren door

$$f(n) = \begin{cases} 2n, & \text{voor } n \geq 0, \\ -(2n+1), & \text{voor } n < 0. \end{cases}$$

De volgende stelling is bijzonder nuttig.

STELLING 1.4. Een deelverzameling van een aftelbare verzameling is aftelbaar.

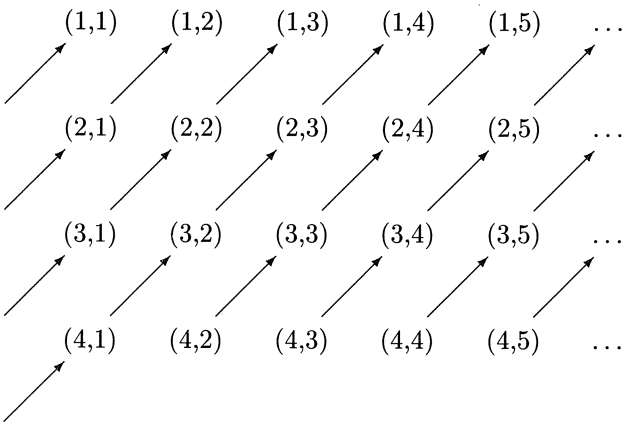
BEWIJS. Laat Y een deelverzameling zijn van de aftelbare verzameling X. Zonder beperking der algemeenheid mogen we aannemen dat X en Y oneindig zijn. We maken gebruik van de volgende eigenschap van $\mathbb{N}$, die met volledige inductie bewezen kan worden:

iedere niet-lege deelverzameling van $\mathbb{N}$ heeft een kleinste element.

Omdat X aftelbaar is, bestaat er een bijectie $f: X \to \mathbb{N}$. We definiëren $g: \mathbb{N} \to Y$ inductief als volgt.

1. $g(0) = f^{-1}\big(\min f[B]\big)$,
2. $g(n) = f^{-1}\big(\min(f[B] \setminus \{g(0), \ldots, g(n-1)\})\big)$.

VOORBEELD 1.5. $|\mathbb{Q}| = \aleph_0$.
Gelet op hetgeen we in Voorbeeld 1.3 gedaan hebben is het voldoende om te bewijzen dat de verzameling $\mathbb{Q}^+$ van de positieve rationale getallen aftelbaar is. We beschouwen de verzameling van alle breuken $\frac{t}{n}$ met $t > 0$ en $n > 0$ en merken op dat $\mathbb{Q}^+$ opgevat kan worden als een deelverzameling van deze verzameling, namelijk de deelverzameling van alle breuken $\frac{t}{n}$ met $\text{ggd}(t, n) = 1$.



FIGUUR 1. De eerste diagonaalmethode

In plaats van $\frac{t}{n}$ schrijven we omwille van de eenvoud van de notatie (n,t) en we ordenen deze paren in een matrix-schema. De verzameling van al deze paren noteren we met P.

We tellen nu P volgens de pijlen. Deze telmethode heet de *eerste diagonaalmethode* of ook wel de *diagonaalmethode van Cauchy.* Zo krijgen we een bijectie $f:\mathbb{N} \to P$. Voor de inverse functie f^{-1} is er de formule

$$f^{-1}(n,t) = \frac{(t+n-1)(t+n-2)}{2} + t.$$

Op deze wijze is P aftelbaar. Dat $\mathbb{Q}^+$ het ook is volgt uit Stelling 1.4.

Er zijn verschillende stellingen die met de methode uit het laatste voorbeeld bewezen kunnen worden. We laten de bewijzen aan de lezer over.

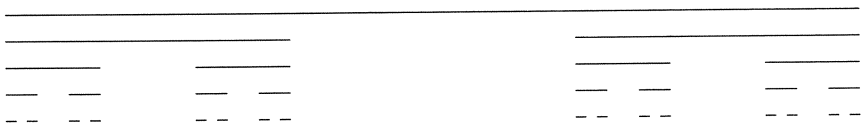
STELLING 1.6. *Het Cartesisch product van twee aftelbare verzamelingen is aftelbaar.*

STELLING 1.7. *De vereniging van een aftelbare collectie van aftelbare verzamelingen is aftelbaar.*

Zoveel voor dit moment over aftelbare verzamelingen. In [3] voerde Cantor een bijzondere verzameling in, die nu naar hem genoemd wordt.

VOORBEELD 1.8. Het Cantordiscontinuum C

Hieronder staan de eerste vier stappen van de constructie van het Cantordiscontinuum. Het eenheidsinterval wordt in drie intervallen van gelijke lengte verdeeld die alleen eindpunten gemeenschappelijk hebben. Het middelste open interval wordt weggelaten. Dit proces wordt herhaald bij ieder van de overgebleven intervallen. Elk interval wordt weer in drie intervallen van gelijke lengte verdeeld en het middelste open interval wordt weggelaten. Wat uiteindelijk



FIGUUR 2. Vier stappen op de weg naar het Cantordiscontinuum

overblijft is het Cantordiscontinuum C. Met behulp van reeksen kan men C als volgt beschrijven.

$$C = \{\, x : x = \sum_{i=1}^{\infty} x_i 3^{-i},\ x_i = 0 \text{ of } x_i = 2 \,\}.$$

We merken op dat de voorstelling van x uit C met behulp van zo'n reeks uniek is; verschillende rijen $x_1, x_2, x_3, \ldots$ geven verschillende elementen van C. Dat x geschreven wordt als $x = \sum_{i=1}^{\infty} x_i 3^{-i}$, waarin $x_i = 0$ of $x_i = 2$, betekent eigenlijk niets anders dan dat x in het drietalling stelsel geschreven wordt als $x = 0, x_1\, x_2\, x_3\, \ldots$, waarbij alleen 0-en en 2-en gebruikt worden. Deze schrijfwijze is uniek.

We zijn nu toegekomen aan een uiterst belangrijk resultaat.

STELLING 1.9. *C en $\mathbb{R}$ zijn overaftelbaar. In het bijzonder, $\aleph_0 \neq \aleph$.*

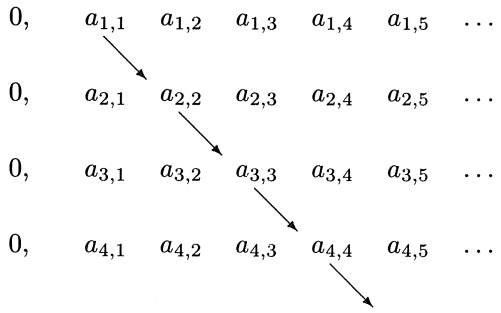
BEWIJS. In verband met Stelling 1.4 is het voldoende om te laten zien dat C overaftelbaar is. We bewijzen dit als volgt. Zij $g\colon \mathbb{N} \to C$ een willekeurige afbeelding. We zullen aantonen dat g niet surjectief is. In het bijzonder volgt hieruit dat er geen bijectie van $\mathbb{N}$ naar C kan bestaan. Voor ieder natuurlijk getal n schrijven we $g(n) = \sum_{k=1}^{\infty} a_{n,k} 3^{-k}$, waarin $a_{n,k} = 0$ of $a_{n,k} = 2$. We definiëren $b_n = 2 - a_{n,n}$. Dan kan voor geen enkel natuurlijk getal m gelden dat $g(m) = \sum_{k=1}^{\infty} b_k 3^{-k}$. Dit is zo omdat $b_m \neq a_{m,m}$. De hier beschreven methode staat bekend als de *tweede diagonaalmethode* of ook wel de *diagonaalmethode van Cantor*. De ontwikkelingen van $g(n)$ in het drietallig stelsel worden onder elkaar opgeschreven. Langs de hoofddiagonaal (aangegeven door de pijlen) gaande, maken we de ontwikkeling van een getal in het drietallig stelsel dat van iedere $g(n)$ op de diagonaalplaats verschilt.

Cantor gebruikte de laatste stelling om een elegant bewijs te maken voor het bestaan van transcendente getallen. Hierover gaat de volgende stelling. Dedekind was over Cantor's resultaat zó enthousiast dat hij zijn eerdere, nogal gereserveerde, houding ten opzichte van de verzamelingenleer moest herzien.

Wij brengen in herinnering dat een reeël getal β *algebraïsch* genoemd wordt indien β nulpunt is van een polynoom

$$p(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0 \tag{1}$$

waarbij $a_n \neq 0$ en alle coëfficiënten a_i tot $\mathbb{Q}$ behoren. Zonder beperking der algemeenheid mogen we aannemen dat alle coëfficiënten in formule (1) geheel



FIGUUR 3. De tweede diagonaalmethode

zijn en dat $a_n > 0$. Een reëel getal dat niet algebraïsch is heet *transcendent* (zo'n getal overstijgt de grenzen van de Algebra). In 1844 had Liouville als eerste het bestaan van transcendente getallen aangetoond. In 1873 bewees Hermite dat e transcendent is en in 1882 bewees Lindemann de transcendentie van π. In 1874 gaf Cantor in [2] op de volgende wijze een bijzonder eenvoudig bewijs van het bestaan van transcendente getallen.

STELLING 1.10. *De verzameling van alle algebraïsche getallen is aftelbaar.*

Deze stelling in combinatie met Stelling 1.9 laat zien dat er transcendente getallen bestaan. Sterker nog, met Stelling 1.7 volgt dat de verzameling van alle transcendente getallen overaftelbaar is.

BEWIJS. We beschouwen alle polynomen $p(x)$ van de vorm (1), waarbij alle coëfficiënten gehele getallen zijn (en $a_n > 0$). De *hoogte* van het polynoom $p(x)$ uit 1 is het getal

$$h = n + a_n + |a_{n-1}| + \cdots + |a_1| + |a_0|.$$

Het is duidelijk dat er bij gegeven hoogte maar eindig veel polynomen zijn met die hoogte en, bijgevolg, er ook maar eindig veel algebraïsche getallen zijn die nulpunt zijn van een polynoom met die hoogte. Met Stelling 1.7 volgt nu dat de verzameling van alle algebraïsche getallen aftelbaar is.

2. ORDENING

De ordening van de kardinaalgetallen is op de volgende wijze gedefinieerd.

DEFINITIE 2.1. Laat $|X|$ en $|Y|$ kardinaalgetallen zijn. Dan is per definitie

1. $|X| \leq |Y|$ als er een injectie $f: X \to Y$ bestaat,
2. $|X| < |Y|$ als $|X| \leq |Y|$ en $|X| \neq |Y|$,
3. $|X| \geq |Y|$ als $|Y| \leq |X|$,

4. $|X| > |Y|$ als $|Y| < |X|$.

Zo is bijvoorbeeld $|\mathbb{N}| < |C| \leq |\mathbb{R}|$ en $\aleph_0 < \aleph$. Voor alle n uit $\mathbb{N}$ geldt $0 \leq n < \aleph_0$. Men gaat eenvoudig na dat de Definitie 2.1 onafhankelijk is van de keuze van X en Y, d.w.z., als $|X| = |X'|$ en ook $|Y| = |Y'|$ dan is $|X| < |Y|$ dan en slechts dan als $|X'| < |Y'|$.

Welke eigenschappen heeft deze ordening nu? We brengen de definitie van lineaire ordening in herinnering.

DEFINITIE 2.2. Een *lineaire ordening* op een verzameling X is een binaire relatie $\leq$ zódat voor alle x, y en z in X geldt

1. $x \leq x$,
2. als $x \leq y$ en $y \leq z$, dan $x \leq z$,
3. $x \leq y$ of $y \leq x$,
4. als $x \leq y$ en $y \leq x$, dan $x = y$.

We schrijven $x < y$ als $x \leq y$ en $x \neq y$. Een verzameling met een lineaire ordening heet een *lineair geordende verzameling.*

De reële rechte $\mathbb{R}$ met de natuurlijke ordening $\leq$ is een linair geordende verzameling. Ook het Cantordiscontinuum C en de rationale getallen $\mathbb{Q}$ zijn met de ordening, die ze van $\mathbb{R}$ erven, lineair geordende verzamelingen.

Het is duidelijk dat de ordening van de kardinaalgetallen de eigenschappen (1) en (2) heeft. De klasse van kardinaalgetallen heeft ook eigenschap (3); twee kardinaalgetallen zijn altijd vergelijkbaar. Het bewijs van deze eigenschap is verre van eenvoudig. Omdat in de situaties waarin we deze eigenschap nodig hebben, de juistheid ervan direct duidelijk is, zullen we verder niet op het bewijs ingaan. Eigenschap (4) wordt bewezen in de volgende stelling van Cantor en Bernstein.

STELLING 2.3. *Als er injecties $f\colon A \to B$ en $g\colon B \to A$ bestaan, dan bestaat er een bijectie $h\colon A \to B$.*

BEWIJS. Eerst merken we op dat $g(B) \subset A$ en dat $g \circ f$ een injectie van A in $g(B)$ is. Als we nu een bijectie h van A naar $g(B)$ kunnen maken, dan is de samengestelde afbeelding $g^{-1} \circ h$ de gezochte bijectie. We mogen dus aannemen dat $B \subset A$.

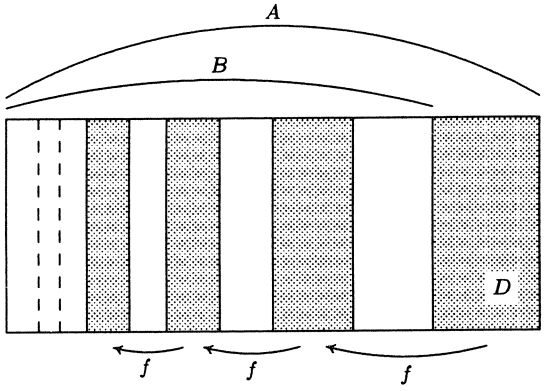
Zij $D = A \setminus B$ en $E = D \cup f[D] \cup f\big[f[D]\big] \cup \cdots$. Merk op dat $f[E] = E \cap B$. Definieer nu

$$h(x) = \begin{cases} f(x), & \text{als } x \in E, \\ x, & \text{als } x \in B \setminus E. \end{cases}$$

Ga na dat h inderdaad een bijectie is.

De zojuist bewezen eigenschap is erg handig voor het maken van berekeningen. We geven hiervan een voorbeeld.

VOORBEELD 2.4. $|C \times C| = |C| = |\mathbb{R}| = |\mathbb{R} \times \mathbb{R}|$.



FIGUUR 4. De Stelling van Cantor-Bernstein

De gelijkheden volgen met Stelling 2.3 uit het bestaan van de volgende rij van injecties

$$C \times C \xrightarrow{f} C \xrightarrow{i} \mathbb{R} \xrightarrow{j} \mathbb{R} \times \mathbb{R} \xrightarrow{g} C \times C.$$

Voor i en j kunnen we de natuurlijke inbeddingen nemen. De afbeelding f wordt als volgt gedefinieerd. In Voorbeeld 1.8 hebben we gezien dat ieder punt x van C op unieke wijze geschreven kan worden als $x = \sum_{i=1}^{\infty} x_i 3^{-i}$, waarin $x_i = 0$ of $x_i = 2$. Definieer

$$f(x) = f\big(\textstyle\sum_{i=1}^{\infty} x_i 3^{-i}\big) = \big(\sum_{i=1}^{\infty} x_{2i} 3^{-i}, \sum_{i=1}^{\infty} x_{2i-1} 3^{-i}\big).$$

Men gaat gemakkelijk na dat f een bijectie is. Voor de definitie van g merken we eerst op dat er een surjectie p van C op $\mathbb{R}$ bestaat. Deze wordt verkregen als de compositie $p_1 \circ p_2 \circ p_3$ van de afbeeldingen p_1, p_2 en p_3. De afbeelding p_3 beeld C af op het interval $[0,1]$ en is gedefinieerd door

$$p\big(\textstyle\sum_{i=1}^{\infty} x_i 3^{-i}\big) = \sum_{i=1}^{\infty} x_i 2^{-i-1}.$$

De afbeelding p_2 is een surjectie van $[0,1]$ op $(0,1)$ en is bijvoorbeeld gedefinieerd door $p_2(x) = x$ voor $x \neq 0$ en $x \neq 1$, $p_2(0) = p_2(1) = \frac{1}{2}$. De afbeelding p_1 tenslotte is een surjectie van $(0,1)$ op $\mathbb{R}$. Hiervoor kan men bijvoorbeeld nemen het voorschrift $p_1(x) = \tan \pi(x - \frac{1}{2})$. Met behulp van de surjectie p definieert men een injectie h van $\mathbb{R}$ in C door $h(x)$ gelijk te stellen aan een element van $p^{-1}\big[\{x\}\big]$. De afbeelding $g\colon \mathbb{R} \times \mathbb{R} \to C \times C$ is nu gedefinieerd door $g(x,y) = \big(h(x), h(y)\big)$.

Tot nu toe hebben we twee oneindige kardinaalgetallen gezien, namelijk $\aleph_0$ en $\aleph$. Zijn er nog andere? Cantor liet zien hoe bij ieder oneindig kardinaalgetal een groter kardinaalgetal gevonden kan worden. In de volgende stelling is $\mathcal{P}(X)$ de machtsverzameling (Potenzmenge) van X, dat is de familie van alle deelverzamelingen van X.

STELLING 2.5. *Voor iedere verzameling X geldt $|X| < |\mathcal{P}(X)|$.*

BEWIJS. Het bewijs van de stelling lijkt veel op dat van Stelling 1.9; men zou kunnen zeggen dat het bewijs gebruik maakt van een abstracte diagonaalmethode. Het is duidelijk dat $|X| \leq |\mathcal{P}(X)|$; definieer $f: X \to \mathcal{P}(X)$ door $f(x) = \{x\}$. We bewijzen uit het ongerijmde dat $|X| \neq |\mathcal{P}(X)|$. Neem eens aan dat er een bijective afbeelding $f: X \to \mathcal{P}(X)$ zou bestaan. Definieer Z door $Z = \{x \in X : x \notin f(x)\}$. Omdat Z een deelverzameling is van X en omdat f surjectief is, bestaat er een z in X zó dat $f(z) = Z$. Maar dan geldt: $z \in Z$ dan en slechts dan als $z \notin Z$. Dit is een tegenspraak.

In plaats van $|\mathcal{P}(X)|$ schrijft men ook wel $2^{|X|}$. In het volgende voorbeeld berekenen we $2^{\aleph_0}$.

VOORBEELD 2.6. $2^{\aleph_0} = \aleph$.
Voor een deelverzameling A van $\mathbb{N}$ is de karakteristieke functie χ_A gedefinieerd door

$$\chi_A(x) = \begin{cases} 1, & \text{als } x \in A, \\ 0, & \text{als } x \notin A. \end{cases}$$

Het is niet moeilijk om na te gaan dat de afbeelding $f: \mathcal{P}(X) \to C$ gedefinieerd door $f(A) = \sum_{i=1}^{\infty} 2\,(\chi_A(i-1))3^{-i}$ een bijectie is. Hieruit volgt de juistheid van de formule.

3. DE CONTINUUM-HYPOTHESE

Bij de opbouw van de verzamelingenleer stuitte Cantor in 1884 op het *continuumprobleem*:

> als X een oneindige deelverzameling van $\mathbb{R}$ is, is dan noodzakelijkerwijs $|X| = \aleph_0$ of $|X| = \aleph$? Of bestaat er een deelverzameling X van $\mathbb{R}$ zó dat $\aleph_0 < |X| < \aleph$?

Aan dit op het oog nogal eenvoudige probleem is door velen hard gewerkt. Het is het eerste van de drieëntwintig problemen die Hilbert [11] in 1900 aan het Internationaal Mathematisch Congres voorlegde. Cantor en velen met hem geloofden dat er géén kardinaalgetal bestond dat in de ordening van de kardinaalgetallen tussen $\aleph_0$ en $\aleph$ ligt. Dit vermoeden is de continuum-hypothese.

CH Als $X \subset \mathbb{R}$ en X is oneindig, dan is $|X| = \aleph_0$ of $|X| = \aleph$.

We kunnen de continuum-hypothese ook anders formuleren. Daartoe maken we gebruik van een eigenschap van de ordening van de kardinaalgetallen die erg veel lijkt op een eerder genoemde eigenschap van de natuurlijke getallen:

> iedere niet-lege verzameling van kardinaalgetallen heeft een kleinste element.

Met deze eigenschap vinden we een kleinste kardinaalgetal groter dan $\aleph_0$. Dat getal wordt aangeduid met $\aleph_1$. Met behulp van het resultaat uit Voorbeeld 2.6 vinden we de volgende formulering van de continuum-hypothese, die equivalent is met de oorspronkelijke.

CH $2^{\aleph_0} = \aleph_1$.

De oplossing van het continuum-probleem ligt niet in de verzamelingenleer, maar in de mathematische logica.

Door de axiomatische opbouw is de verzamelingenleer uitgegroeid tot een volwaardige wiskundige discipline. Een vaak gebruikt axiomasysteem is dat van Zermelo en Fraenkel waaraan het keuzeaxioma nog is toegevoegd. Wij zullen dat systeem aanduiden met **ZFC**. Men noemt een axiomasysteem *consistent* indien in de daaruit opgebouwde theorie geen tegenspraken, bijvoorbeeld $0 = 1$, voorkomen. Zou men nu kunnen bewijzen dat **ZFC** consistent is ? Gödel [9] heeft in 1930 laten zien dat de consistentie van **ZFC** niet bewezen kan worden. Hij toonde tevens aan dat het tegendeel ook niet bewezen kan worden.

In 1938 bewees Gödel[10] dat het systeem **ZFC** + **CH** *relatief consistent* is: als uit **ZFC** + **CH** een tegenspraak kan worden afgeleid, dan kan die tegenspraak al uit **ZFC** worden afgeleid. Uitgaande van de consistentie van **ZFC** kunnen we de volgende conclusie trekken. Omdat $\neg$ **CH** + **CH** een contradictie is, kan $\neg$ **CH** niet bewezen kan worden in het axiomasysteem **ZFC**. M.a.w., $\neg$ **CH** is onafhankelijk van **ZFC**. (Met $\neg$ **CH** wordt de ontkenning van **CH** aangeduid.)

Het continuum-probleem werd uiteindelijk opgelost door Cohen [5, 6]. Hij bewees dat ook **ZFC** $+\neg$ **CH** relatief consistent is. Op dezelfde wijze als boven volgt hieruit dat **CH** niet bewezen kan worden in **ZFC**, m.a.w., **CH** is onafhankelijk van **ZFC**.

Een interessante uiteenzetting over de continuum-hypothese met veel aandacht voor historische details is te vinden in het eerste deel van het boek [8] van Van Dalen en Monna.

4. De Souslin-hypothese

De reële rechte $\mathbb{R}$ heeft een zeer rijke struktuur. Het belangrijkste element daarvan is wellicht de lineaire ordening. Dat is ook de reden dat men door abstractie deze struktuur geïsoleerd heeft. Dit heeft geleid tot de studie van lineair geordende verzamelingen. De Souslin-hypothese [14] hangt samen met pogingen om $\mathbb{R}$ te karakteriseren binnen de klasse van de lineair geordende verzamelingen. Om de hypothese te formuleren moeten we eerst enige kennis over lineair geordende verzamelingen ophalen.

Behalve de snede (van Dedekind), die in de voordracht van professor Grootendorst besproken is, moeten we nu ook nog andere ontbindingen van een lineair geordende verzameling beschouwen.

DEFINITIE 4.1. Een *ontbinding* van een lineair geordende verzameling X is een paar $L : R$ van deelverzamelingen van X zó dat geldt:

1. $L \neq \emptyset$ en $R \neq \emptyset$,
2. $L \cup R = X$ en $L \cap R = \emptyset$,
3. als $x \in L$ en $y \in R$, dan $x < y$.

We onderscheiden drie soorten van ontbindingen $L : R$, namelijk de

sprong - L heeft een grootste element en R heeft een kleinste element,

leemte - L heeft geen grootste element en R heeft geen kleinste element,

snede - óf L heeft een grootste element óf R heeft een kleinste element.

Het Cantordiscontinuum heeft sprongen en sneden, maar geen leemten. De verzameling $\mathbb{Q}$ heeft geen sprongen, maar wel leemten en ook (rationale) sneden. In de voordracht van professor Grootendorst werden de leemten irrationale sneden genoemd. De reële rechte heeft geen sprongen en leemten, maar iedere ontbinding is een snede: men zou kunnen zeggen dat alle leemten van $\mathbb{Q}$ zijn opgevuld.

We geven nu een karakterisering van de lineair geordende verzameling $\mathbb{R}$. We zeggen dat de lineair geordende verzamelingen X en Y met de ordeningen $\leq_X$ respectievelijk $\leq_Y$ *isomorf* zijn indien er een *orde-behoudende* bijectie $f: X \to Y$ bestaat, d.w.z., voor alle x_1 en x_2 in X geldt: $x_1 \leq_X x_2$ dan en slechts dan als $f(x_1) \leq_Y f(x_2)$. Zo is bijvoorbeeld het interval $(-\frac{\pi}{2}, \frac{\pi}{2})$ isomorf met $\mathbb{R}$ via de orde-behoudende bijectie tan.

STELLING 4.2. *Zij X een lineair geordende verzameling met ordening $\leq_X$. Dan is X isomorf met $\mathbb{R}$ dan en slechts dan als X aan de volgende drie voorwaarden voldoet:*

1. *X heeft geen kleinste en ook geen grootste element,*
2. *iedere ontbinding van X is een snede,*
3. *er is een aftelbare deelverzameling D van X zó dat voor alle x en y in X met $x <_X y$ er een d in D is met $x <_X d <_X y$.*

BEWIJS. We geven een schets van het bewijs. Het is duidelijk dat R aan de drie genoemde voorwaarden voldoet (Neem $D = \mathbb{Q}$). Dus als X isomorf is met $\mathbb{R}$, dan voldoet X aan de drie voorwaarden.

Laat omgekeerd gegeven zijn dat X aan de drie voorwaarden voldoet. Het is eenvoudig in te zien dat D aftelbaar oneindig is en de volgende eigenschap heeft: voor alle p en r in D met $p <_X r$ is er een q in D met $p <_X q <_X r$.

We beginnen nu met de definitie van een orde-behoudende bijectie f van $\mathbb{Q}$ naar D. De ordening van $\mathbb{Q}$ geven we aan met $<$. Omdat $\mathbb{Q}$ en D aftelbaar oneindig zijn, kunnen we deze verzamelingen indiceren met $\mathbb{N}$: $\mathbb{Q} = \{ q_n : n \in \mathbb{N} \}$ en $D = \{ d_n : n \in \mathbb{N} \}$. We nemen, zoals te doen gebruikelijk is, aan dat bij verschillende indices verschillende elementen horen. Om te bereiken dat f een bijectie wordt, definiëren we f en f^{-1} om en om. Om te beginnen: $f(q_0) = d_0$ en $f^{-1}(d_0) = q_0$. Zoek nu in $\mathbb{Q}$ het element q_{n_1} met de kleinste index zó dat geldt: $q_0 < q_{n_1}$ dan en slechts dan als $d_0 < d_1$. Definieer $f^{-1}(d_1) = q_{n_1}$ en $f(q_{n_1}) = d_1$. Voor de volgende stap in de definitie van de bijectie zoeken we in $\mathbb{Q} \setminus \{ q_0, q_{n_1} \}$ het element q_{n_2} met de kleinste index. We bepalen nu in D het

element d_{n_2} met de kleinste index dat in de ordening $<_X$ ten opzichte van d_0 en d_1 dezelfde positie heeft als q_{n_2} ten opzichte van q_0 en q_{n_1}. We definiëren $f(q_{n_2}) = d_{n_2}$ en $f^{-1}(d_{n_2}) = q_{n_2}$. Zo voortgaande met inductie krijgen we een bijectie $f\colon \mathbb{Q} \to D$ die orde-behoudend is. Deze afbeelding kunnen we voort zetten tot een bijectie $\tilde{f}\colon \mathbb{R} \to X$ op de volgende wijze. Als α een irrationaal getal is, dan bepaalt α een ontbinding van $\mathbb{Q}$, zeg $L : R$. Omdat f ordebehoudend is, is $f(L) : f(R)$ een ontbinding van D. Dit bepaalt een ontbinding $L^* : R^*$ van X via

$$\begin{aligned} L^* &= \{\, z \in X : \text{ er is een } x \in f(L) \text{ zó dat } z \leq_X x \,\}, \quad \text{en} \\ R^* &= \{\, z \in X : \text{ voor alle } x \in f(L) \text{ zó dat } z \not\leq_X x \,\}. \end{aligned}$$

Op grond van de voorwaarde 2 is $L^* : R^*$ een snede. Het hierdoor bepaalde element (het grootste element van L^* of het kleinste element van R^*) stellen we gelijk aan $\tilde{f}(\alpha)$. Daarmee is de isomorfie tussen $\mathbb{R}$ en X vastgelegd.

Om de Souslin-hypothese te formuleren hebben we nog een begrip nodig. Als X een geordende verzameling is met ordening $\leq$ dan definieren we een *open interval* op de gebruikelijke wijze: voor a, b in X met $a < b$, is $(a, b) = \{\, x \in X : a < x < b \,\}$.

DEFINITIE 4.3. Een lineair geordende verzameling heeft de eigenschap c.c.c. (*countable chain condition*) indien iedere collectie van paarsgewijs disjuncte open intervallen aftelbaar is.

VOORBEELD 4.4. $\mathbb{R}$ heeft de eigenschap c.c.c..
Dat is niet zo moeilijk in te zien. Zij $\{\, U_\lambda : \lambda \in \Lambda \}$ een collectie van paarsgewijs disjuncte intervallen. In iedere U_λ kiezen we een rationaal getal q_λ. Dan is door $f(\lambda) = p_\lambda$, $\lambda \in \Lambda$, een injectie f van Λ in $\mathbb{Q}$ gedefinieerd. Dus is Λ aftelbaar.
De Souslin-hypothese luidt nu als volgt

> **SH** Zij X een lineair geordende verzameling. Dan is X isomorf met $\mathbb{R}$ dan en slechts dan indien X aan de volgende drie voorwaarden voldoet
>
> 1. X heeft geen kleinste en ook geen grootste element,
> 2. iedere ontbinding van X is een snede,
> 3. X heeft de eigenschap c.c.c..

Met het Voorbeeld 4.4 zien we dat $\mathbb{R}$ aan de drie genoemde voorwaarden voldoet.

Nadat door Cohen het raadsel van de Continuum-hypothese ontsluierd was, zijn topologen naarstig gaan onderzoeken of ook hun problemen met de methoden van de mathematische logica opgelost konden worden. De Souslin-hypothese is één van de problemen waar dit tot succes heeft geleid: **SH** is onafhankelijk van **ZFC**. In het bewijs spelen zowel **CH** als $\neg$ **CH** een rol. We gaan hier verder niet op in. Er is een zeer lezenswaardig artikel van Rudin [13] over de Souslin-hypothese. Een gedegen uiteenzetting over onafhankelijkheidsbewijzen is te vinden in het boek van Kunen [12].

REFERENTIES

1. G. CANTOR, Über die Ausdehnung eines Satzes aus der Theorie der trigoniometrischen Reihen, *Mathematische Annalen* **5**, 1872, 123–132
2. G. CANTOR, Über eine Eigenschaft des Inbegriffs aller reellen algebraischen Zahlen, *Crelles Journal für Mathematik* **77**, 1874, 258–262
3. G. CANTOR, Über unendliche lineare Punktmannigfaltigkeiten, Nr.5, *Mathematische Annalen* **21**, 1883, 545–586
4. G. CANTOR, *Gesammelte Abhandlungen mathematischen und philosophischen Inhalts*, bezorgd door E. Zermelo, (Berlin 1932), (Herdruk Springer 1990)
5. P.J. COHEN, The independence of the continuum hypothesis I, *Proceedings of the National Academy of Sciences USA* **50**, 1963, 1143–1148
6. P.J. COHEN, The independence of the continuum hypothesis II, *Proceedings of the National Academy of Sciences USA* **51**, 1964, 105–110
7. P.J. COHEN, *Set Theory and the Continuum Hypothesis* (Benjamin, New York 1966)
8. D. VAN DALEN en A.F. MONNA, *Sets and Integration, An Outline of the Development* (Wolters-Noordhoff, Groningen 1972)
9. K. GÖDEL, Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme, I, *Monatshefte für Mathematik und Physik* **38**, 1931, 173–198
10. K. GÖDEL, Consistency proof for the Generalized Continuum Hyphotesis, *Proceedings of the National Academy of Sciences U.S.A.* **25**, 1939, 220–224
11. D. HILBERT, Mathematische Probleme, Vortrag, gehalten auf dem internationalen Mathematiker-Kongress zu Paris 1900, *Archiv der Mathematik und Physik*, derde serie **1**, 1901, 44–63, 213–237
12. K. KUNEN, *Set Theory, An Introduction to Independence Proofs* (North-Holland, Amsterdam 1973)
13. M.E. RUDIN, Souslin's conjecture, *American Mathematical Monthly* **76**, 1969, 1113–1119
14. M. SOUSLIN, Problème 3, *Fundamenta Mathematicae* **1**, 1920, 223

Een intuïtionistische kijk op het reële getal

A.S. Troelstra

Worden wiskundige objecten ontdekt of bedacht (gemaakt)? Wie neigt tot de eerste mogelijkheid (en dat doen waarschijnlijk de meeste wiskundigen) is een "wiskundig platonist": wiskundige objecten bestaan als abstracta in een rijk van ideeën, en beweringen over dergelijke objecten zijn waar dan wel onwaar — het is aan ons om te ontdekken wat het geval is.

Maar wellicht helt U over naar de tweede mogelijkheid: wiskundige objecten worden *bedacht*, d.w.z. het zijn constructies van de menselijke geest, en ze hebben geen existentie buiten het menselijke kenvermogen. Als U er zo over denkt, voelt U zich waarschijnlijk eerder thuis bij het *intuïtionisme* van L.E.J. Brouwer (1881–1966). In de visie van Brouwer bestaat de hele wiskunde uit gedachtenconstructies, en heeft het geen zin om te vragen of een bewering A over dergelijke constructies waar is, onafhankelijk van onze kennis omtrent A. Consequent doordenken en toepassen van deze opvatting leidt tot een herziening van tal van bekende resultaten uit de klassieke (d.w.z. de gebruikelijke of "gewone") wiskunde; we zullen dat laten zien aan de hand van de theorie van het reële getal. Kennisnemen van de constructieve benadering kan daarbij ons inzicht in de klassieke theorie verdiepen.

Het constructieve uitgangspunt brengt met zich mee dat alle wiskundige objecten waarover we spreken in zekere (geïdealiseerde) zin *geconstrueerd* moeten kunnen worden. Wat we daarmee bedoelen is onproblematisch en intuïtief duidelijk bij de natuurlijke getallen. De volgende definitie beschrijft dan ook, intuïtionistisch gezien, *geen* natuurlijk getal:

> $p = 1$ als het vermoeden van Fermat waar is, en $p = 2$ als het vermoeden onwaar is.

Evenzo zijn rationale getallen en gehele getallen onproblematisch; om een rationaal getal aan te geven, moeten we teller en noemer precies kunnen berekenen. Reële getallen daarentegen worden vastgelegd door hun rationale benaderingen: een reëel getal is intuïtionistisch goed gedefiniëerd, als we weten hoe we rationale benaderingen voor het getal kunnen vinden met elke gewenste graad van nauwkeurigheid.

Ook onze vertrouwde redeneerprincipes (d.w.z. de logica) ontkomen niet aan een intuïtionistische revisie. Terwille van een compactere schrijfwijze zal ik in het volgende veelvuldig bepaalde afkortingen voor logische uitdrukkingen hanteren, met name

$\wedge$	voor "en"
$\vee$	voor "of"
$\rightarrow$	voor "impliceert"
$\neg$	voor "niet"
$\forall x$	voor " voor alle x"
$\forall x \in Y$	voor "voor alle x in Y"
$\exists x$	voor "er bestaat een x zodat ...", etc.

Het is niet de bedoeling dat we formele logica gaan doen, en de uitdrukkingen in woorden zullen steeds als synoniemen van de logische symbolen behandeld worden.

Bij het wiskundig platonisme zijn beweringen A altijd waar of onwaar (i.e. $\neg A$ waar), ook al weten *wij* niet welk van de twee het geval is. Compact uitgedrukt:

$$A \vee \neg A.$$

Maar intuïtionistisch beschouwd kan "A is waar" alleen maar betekenen dat we een *bewijs* voor A hebben, en moet "A is niet waar" (oftewel "$\neg A$ is waar") betekenen dat we uit de onderstelling dat A waar is een tegenspraak kunnen afleiden (bij "tegenspraak" dan wel "contradictie" kunt U denken aan een vast erkend onware uitspraak zoals bijv. 0=1). Derhalve kunnen we bij deze interpretatie ook niet $A \vee \neg A$ als algemeen geldig principe accepteren, want dat zou betekenen dat we elke uitspraak konden bewijzen dan wel weerleggen, d.w.z. niets minder dan de oplosbaarheid van elk wiskundig probleem!

Maar: het feit dat we $A \vee \neg A$ niet accepteren als (constructief) principe, betekent nog niet dat voor een *gegeven* A dit principe tegenstrijdig is! Immers, "niet (A of $\neg A$)" betekent dat zowel "niet A" als "niet $\neg A$" zou moeten gelden — en dat is een regelrechte tegenspraak. Kortom, er is een groot verschil tussen "(nog) niet bewijsbaar" en "tegenstrijdig".

Laten we dit met een voorbeeld concreet maken.

n is een *rijtjesgetal* als $A(n)$ geldt, waarbij $A(n) \equiv$ "n is het nummer van de laatste decimaal achter de komma van het eerste rijtje van tien opeenvolgende zevens in de decimaalbreuk voor π", meer aanschouwlijk

$$3,14159\ldots\underset{\uparrow}{7}777777777\ldots$$
$$n$$

□

De "kennissituatie" is hier als volgt:

1. voor elke n kunnen we nagaan of $A(n)$ dan wel $\neg A(n)$ geldt (we hoeven alleen maar de decimalen van π voldoende ver uit te rekenen);

2. we weten echter niet of er wel zo'n rijtje is. (Althans ik weet het niet; misschien heeft inmiddels iemand wel zo'n rijtje gevonden — er zijn al zoveel decimalen van π uitgerekend — maar je kan "tien" natuurlijk door "honderd" of "duizend" vervangen.)

In compacte symboliek:

$$\forall n(A(n) \vee \neg A(n)), \qquad ?\exists n A(n) \vee \neg A(n)?$$

Beschouw nu een oneindige decimale breuk

$$a = 0, d_1 d_2 d_3 \ldots$$

gedefiniëerd door

$d_i = 3$ als voor alle $n \leq i$ $\neg A(n)$
$d_i = 0$ als voor zekere $n \leq i$ $A(n)$

of meer aanschouwlijk

$$\begin{array}{rl} & \text{rijtjesgetal} \\ & \downarrow \\ \pi = & 3,14159\ldots 7777777777\ldots \\ a = & 0,33333\ldots 3333333330000\ldots \end{array}$$

Als er geen n is zodat $A(n)$, dan zijn alle d_i gelijk aan 3, en de breuk is te schrijven als $\frac{1}{3}$; als daarentegen $A(n)$ geldt voor zekere n, dan is de breuk onvereenvoudigbaar te schrijven als $3 \cdot \sum_{i=0}^{n_1} /10^{i-n} = 3 \cdot 10^{-n} \sum_{i=0}^{n_1} 10^i$. In het ene geval heeft de breuk dus kleine teller en noemer, in het andere geval zeer grote teller en noemer. Merk op dat a als reëel getal goed gedefiniëerd is: we kunnen a met elke gewenste graad van nauwkeurigheid benaderen door rationale getallen. a kan onmogelijk niet rationaal zijn (d.w.z. $\neg\neg$(a is rationaal)), maar daarmee kunnen we nog niet bewijzen dat a rationaal is, want dat vraagt *expliciet* berekenen van teller en noemer van de breuk, en daarvoor zouden moeten weten of π een rijtjesgetal bevat. Dit voorbeeld laat zien dat het principe van het bewijs uit het ongerijmde "als $\neg A$ tegenstrijdig is, dan A" (of symbolisch $\neg\neg A \rightarrow A$) intuïtionistisch ook niet algemeen geldig is.
Samenvatting. Intuïtionistisch dienen we te onderscheiden tussen

- A is onbewezen,
- A geldt, oftewel A is waar, d.w.z. we hebben een bewijs voor A,
- $\neg A$ geldt, oftewel A is onwaar of tegenstrijdig, d.w.z. we hebben een methode om uit een willekeurig bewijs van A een tegenspraak af te leiden,
- $\neg\neg A \equiv$ geldt, oftewel $\neg A$ is tegenstrijdig.

Indien we (bijv. blijkens voorbeelden geconstrueerd uit onopgeloste problemen, zgn. *zwakke tegenvoorbeelden*) tot op *heden* A niet als waar kunnen accepteren, betekent dat nog niet dat we $\neg A$ kunnen aantonen (we kunnen altijd later nog een bewijs van A vinden). Voorts leert nadere overweging leert dat de beweringen $\neg A$ en $\neg\neg\neg A$ wel gelijkwaardig zijn (het nadenken over de betekenis van $\neg\neg\neg A$ pleegt aanvankelijk lichte duizelingen te veroorzaken).

Ik geef hier het argument: enerzijds $B \rightarrow \neg\neg B$ voor alle B (ga na), dus ook $(\neg A) \rightarrow \neg\neg(\neg A)$. Stel $\neg\neg\neg A$. Als nu A zou gelden, dan ook $\neg\neg A$ (wegens het

voorgaande); maar $\neg(\neg\neg A)$ en $\neg\neg A$ spreken elkaar tegen, dus hebben we $\neg A$ afgeleid uit $\neg\neg\neg A$. Geef een reëel getal b aan d.m.v. een oneindige decimale breuk, waarvoor het onbekend is of $b = 0$ dan wel $b > 0$.

Tot dusverre hebben we diverse reële getallen geïntroduceerd d.m.v. oneindige decimale breuken. Als we een methode hebben om de decimalen te berekenen zo ver als we maar willen, dan zijn dergelijke getallen constructief goed gedefiniëerd: de eindige beginstukken van de oneindige breuk geven immers rationale benaderingen met elke gewenste graad van nauwkeurigheid.

Maar is deze methode ook algemeen genoeg? Krijgen we zo alle reële getallen die we bij een (constructieve, intuïtionistische) opbouw van de wiskunde nodig hebben? Het antwoord luidt ontkennend, zoals we zullen zien.

Laten we de verzameling reële getallen die gedefiniëerd kunnen worden door oneindig voortlopende decimale breuken $\mathbb{R}^{\mathrm{dec}}$ noemen. Vergelijk nu $b \in \mathbb{R}^{\mathrm{dec}}$ gegeven door $b = 0,4\not3 = 0,43333\ldots$ met $b' \equiv 0, b'_1 b'_2 b'_3 \ldots$, waarbij

$$b'_i = \begin{cases} 3 \text{ als } i \text{ geen rijtjesgetal is,} \\ 4 \text{ als } i \text{ een rijtjesgetal is} \end{cases}$$

Ook $b' \in \mathbb{R}^{\mathrm{dec}}$, krachtens zijn definitie. Maar hoe staat het met $b - b'$? Als er geen rijtjesgetal bestaat, dan geldt $b - b' = 0,1$; als er wel een rijtjesgetal bestaat, zeg i, dan is $b - b' = 0, c_1 c_2 c_3 \ldots$ met

$$\begin{aligned} c_1 &= 0 \\ c_k &= 9 \text{ voor } 1 < k \leq i \\ c_k &= 0 \text{ voor } i < k \end{aligned}$$

We kunnen de eerste decimaal achter de komma van $b - b'$ dus alleen uitrekenen als we weten of er een rijtjesgetal is. Dit weten we niet, en dus kunnen we ook niet aantonen dat $\mathbb{R}^{\mathrm{dec}}$ gesloten is onder aftrekken.

De moeilijkheid is kennelijk, dat een minieme verandering in een getal b', door een wijziging in een decimaal met een zeer hoog rangnummer, kan resulteren in een verandering in de eerste decimaal van het verschil. Elementen van $\mathbb{R}^{\mathrm{dec}}$ die willekeurig dicht bij elkaar liggen op de getallenlijn kunnen toch een verschillende eerste decimaal hebben (men drukt dit ook wel zo uit: de waarde van de eerste decimaal is niet *continu* afhankelijk van het getal zelf).

Aangezien $\mathbb{R}^{\mathrm{dec}}$ niet alle gewenste eigenschappen voor een wiskundig bruikbare verzameling der reële getallen bezit, en met name dus geen goede arithmetiek heeft, bespreken we nog twee andere zeer bekende methoden voor het invoeren van de reële getallen. De eerste is de methode der

Fundamentaalrijen (van rationale getallen)

Een fundamentaalrij (f.r.) is niets anders dan een Cauchy-rij van rationale getallen, d.w.z. een rij die aan het convergentiecriterium van Cauchy voldoet. In het vervolg zullen de letters r, r', s steeds voor rationale getallen gebruikt worden.

Een rij $\langle r_n \rangle_n$ heet een *fundamentaalrij*, als bij elke rationale $\epsilon > 0$ er een n is zodat $|r_{n+m} - r_{n+m'}| < \epsilon$ voor alle m, m', of meer compact

$$\forall \epsilon > 0 \exists n \forall m, m' (|r_{n+m} - r_{n+m'}| < \epsilon)$$

Twee fundamentaalrijen $\langle r_n\rangle_n, \langle r'_n\rangle_n$ heten equivalent (notatie $\sim$) als

$$\forall\epsilon > 0 \exists n \forall m(|r_{n+m} - r'_{n+m}| < \epsilon).$$

De *Cauchy-reële getallen*, $\mathbb{R}^c$, zijn nu de equivalentieklassen van de f.r. t.o.v. deze equivalentierelatie. □

Nu is het niet moeilijk som en product te definiëren voor f.r.:

$$\langle r_n\rangle_n + \langle r'_n\rangle_n := \langle r_n + r'_n\rangle_n$$
$$\langle r_n\rangle_n \cdot \langle r'_n\rangle_n := \langle r_n \cdot r'_n\rangle_n$$

Aangezien deze operaties de equivalentie tussen f.r. respecteren, d.w.z.

$$\langle r_n\rangle_n \sim \langle r'_n\rangle_n \text{ en } \langle s_n\rangle_n \sim \langle s'_n\rangle_n \rightarrow \langle r + s_n\rangle_n \sim \langle r' + s'_n\rangle_n$$

en analoog voor het product, worden hiermee ook goede operaties op de reële getallen zelf gedefiniëerd. We kunnen zonder dubbelzinnigheid de som van twee reële getallen definiëren door middel van de som van twee f.r. die deze getallen representeren; het resultaat is niet van de keuze van de representanten afhankelijk. Zo kunnen we volgens de gebruikelijke, van de gewone wiskunde welbekende wijze, de rekenkunde van de reële getallen ontwikkelen.

Het idee achter de constructie van $\mathbb{R}^c$ is de Cauchy-completering van de rationale getallen: niet alle Cauchy-rijen van rationale getallen (d.w.z. f.r.) hebben een limiet in de rationale getallen zelf; we bedden nu de rationale getallen in in een grotere structuur $\mathbb{R}^c$ waarin alle f.r. als het ware per definitie een limiet hebben. Daarbij hebben we geluk: net als in de gebruikelijke wiskunde, geldt ook intuitionistisch, dat we dit procedé niet hoeven te herhalen: de Cauchy-rijen van elementen uit $\mathbb{R}^c$ hebben zelf weer een limiet in $\mathbb{R}^c$; dit wordt *Cauchy-volledigheid* genoemd. We schetsen een bewijs.

We definiëren eerst de afstand tussen twee elementen x, x' van $\mathbb{R}^c$:

$$|x - x'| := \langle |r - r'|_n\rangle_n/\sim, \text{ waarbij } \langle r_n\rangle_n \in x, \langle r'_n\rangle_n \in x'.$$

Eenvoudige eigenschappen zoals de driehoeks-ongelijkheid $|x - x'| \leq |x - x''| + |x'' - x|$ kunnen zonder veel moeite bewezen worden. De Cauchy-volledigheid van $\mathbb{R}^c$ zegt: als $\langle x_n\rangle_n$ een Cauchy-rij in $\mathbb{R}^c$ is, d.w.z. als

$$\forall k, m, m'(|x_{\alpha k+m} - x_{\alpha k+m'}| < 2^{-k})$$

dan heeft $\langle x_n\rangle_n$ een limiet in $\mathbb{R}^c$. (Zonder beperking mogen we α strikt monotoon onderstellen). Hoe vinden we die limiet? Aangezien x_n door een fundamentaalrij gegeven wordt, kunnen we altijd een voldoende snel convergerende rij $\langle r_{n,k}\rangle_k$ vinden zodanig dat

$$|r_{n,k} - x_n| < 2^{-k} \quad \text{voor alle } k.$$

Definiëer nu

$$r'_i := r_{\alpha(i),i},$$

dan is $\langle r'_n\rangle_n/\sim$ de gezochte limiet. Immers

$$|r'_i - r'_{i+j}| = |r_{\alpha i,i} - r_{\alpha(i+j),i+j}| \leq$$
$$|r_{\alpha i,i} - x_{\alpha i}| + |x_{\alpha i} - x_{\alpha(i+j)}| + |x_{\alpha(i+j)} - r_{\alpha(i+j),i+j}| \leq$$
$$2^{-i} + 2^{-i} + 2^{-i} < 2^{-i+2}.$$

Maar ook $\mathbb{R}^c$ heeft intuïtionistisch nog niet alle fraaie eigenschappen die we van de reële getallen uit de gewone wiskunde kennen. Het volgende voorbeeld laat zien dat de klassieke stelling: "Elke begrensde, monotoon niet-dalende rij van rationale getallen heeft een limiet" voor geen enkele intuïtionistische notie van reële getallen op kan gaan. (In de gewone wiskunde is dit een speciaal geval van de stelling van Bolzano-Weierstrass: elke oneindige begrensde verzameling van reële getallen heeft een verdichtingspunt.) We definiëren een rij $\langle r_n \rangle_n$ door

$$\begin{cases} r_n = 1 - 2^n \text{ als er geen rijtjesgetal} \leq n \text{ is,} \\ r_n = 4 - 2^n \text{ als er een rijtjesgetal} \leq n \text{ is.} \end{cases}$$

Deze rij is begrensd (nl. door 4) en monotoon stijgend. Maar we kunen geen limiet berekenen, want daartoe zouden we een benadering moeten kunnen geven van de limiet met elke gewenste graad van nauwkeurigheid, d.w.z. we zouden moeten kunnen zeggen of de limiet < 3 dan wel > 2 is, en dat kunnen we niet zolang we niet weten of er een rijtjesgetal bestaat.

Anders gezegd, de klassieke stelling van de kleinste bovengrens: elke begrensde verzameling van reële getallen heeft een kleinste bovengrens (waarvan de voorgaande bewerring een speciaal geval is) is niet intuïtionistisch geldig.

In de klassieke theorie vormen de reële getallen niet alleen de Cauchy-completering, maar ook de *orde-completering* van de rationale getallen, d.w.z. $\mathbb{R}$ is de kleinste verzameling waarin de rationale getallen met behoud van ordening ingebed kunnen worden en waarin elke begrensde verzameling van rationale getallen (en in feite elke begrensde verzameling van reële getallen) een kleinste bovengrens heeft. Zoals het voorbeeld van de begrensde monotone rij getallen zonder limiet laat zien, kunnen de reële getallen in de intuïtionistische wiskunde nooit de orde-completering van de rationale getallen vormen. Verderop zullen we zien dat er wel een intuïtionistische orde-completering van de rationale getallen bestaat, maar die verschilt sterk van $\mathbb{R}^c$.

Laat $\langle r_n \rangle_n \in x$, $\langle r'_n \rangle_n \in y$, $x, y \in \mathbb{R}^c$. Definieer

$$x < y := \exists k, n \forall m (r'_{n+m} - r_{n+m} > 2^{-k}) \ (x \textit{ kleiner} \text{ dan } y),$$
$$x \leq y := \neg(y < x) \ (x \textit{ niet groter} \text{ dan } y),$$
$$x \sharp y := (x < y) \vee (y < x) \ (x \textit{ apart} \text{ of } \textit{verwijderd} \text{ van } y).$$

$\sharp$ wordt *apartheids-* of *verwijderingsrelatie* genoemd. □

We merken op dat in het algemeen $\neg\neg(x < y) \equiv \neg(y \leq x)$ iets zwakkers uitdrukt dan $x < y$. Als $x < y$, dan wil dat zeggen dat x een *positieve* afstand beneden y ligt. Evenzo drukt $x \sharp y$ uit dat er een positieve afstand tussen x en y bestaat, terwijl $x \neq y$ alleen maar zegt dat x en y niet gelijk kunnen zijn. Anderzijds is het niet zo eenvoudig een (zwak) tegenvoorbeeld tegen

$$x \neq y \rightarrow x \sharp y$$

te geven. Zo wordt een principe dat hiermee equivalent is door de Russische constructivisten uit de school van A.A. Markov (het "Markov-principe") wel geaccepteerd, terwijl de geldigheid van $x \neq y \to x \sharp y$ door Brouwer expliciet verworpen werd. Hier wil ik er niet verder op ingaan; voor ons is het voldoende te beseffen dat $x \sharp y$ op positieve wijze het verschillen van x en y uitdrukt. Het is niet moeilijk in te zien dat

$$\begin{aligned} \neg x \sharp y &\leftrightarrow x = y, \\ x \sharp y &\to y \sharp x, \\ x \sharp y &\to x \sharp z \vee z \sharp y. \end{aligned}$$

Een gevolg van de eerste eigenschap is

$$\neg\neg\neg x \sharp y \leftrightarrow \neg\neg x = y;$$

derhalve, aangezien algemeen $\neg\neg\neg A \leftrightarrow \neg A$,

$$x = y \leftrightarrow \neg x \sharp y \leftrightarrow \neg\neg\neg x \sharp y \leftrightarrow \neg\neg x = y.$$

(Gelijkheid tussen reële getallen is dus *stabiel*, d.w.z. $\neg\neg x = y \to x = y$.)

In $\mathbb{R}^c$ kunnen we gemakkelijk een reëel getal a aangeven waarvoor $a \leq 0 \vee 0 \leq a$ onbeslist is. Laat $\langle s_n \rangle_n$ een fundamentaalrij zijn gegeven door

$$s_n = \begin{cases} 0 \text{ als er geen rijtjesgetal } \leq n \text{ is} \\ 2^{-k} \text{ als er een even rijtjesgetal } k \leq n \text{ is} \\ --2^{-k} \text{ als er een oneven rijtjesgetal } k \leq n \text{ is} \end{cases}$$

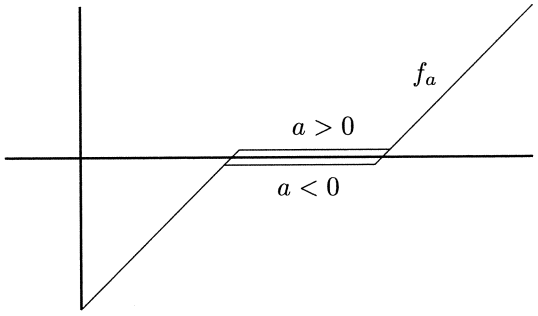
Aangezien we niet weten of er een rijtjesgetal is, en evenmin van te voren kunnen zeggen of een rijtjesgetal, als het gevonden zou worden, even dan wel oneven zou zijn, kunnen we voor het reële getal a beschreven door $\langle s_n \rangle_n$ niet zeggen of $a \leq 0$ dan wel $a \geq 0$ (in het eerste, resp. tweede geval moeten we het voorkomen van een oneven, resp. even rijtjesgetal a priori kunnen uitsluiten).

Het bestaan van zo'n getal heeft weer consequenties, bijv. het niet geldig zijn van de tussenwaardestelling.

In zijn eenvoudigste vorm luidt deze: laat f een functie van het gesloten interval $[b, c]$ naar $\mathbb{R}$ zijn, met $f(b) < 0, f(c) > 0$, dan is er een reëel getal x, $b < x < c$, zodat $f(x) = 0$. (We schrijven $\mathbb{R}$ voor de "gewone" reële getallen uit de klassieke theorie.) Intuïtionistisch is de tussenwaardestelling niet waar voor $\mathbb{R}^c$.

Bekijk daartoe de functie f_a, afhangend van een parameter $a \in \mathbb{R}^c$, gegeven door

$$f_a(x) = \begin{cases} x - 2 \text{ als } x - 2 \leq a, \\ a \text{ als } x - 2 \leq a \leq x - 4, \\ x - 4 \text{ als } a \leq x - 4. \end{cases}$$



Als we niet weten of $a \leq 0$, dan wel $a \geq 0$, dan kunnen we niet uitrekenen waar f_a een nulpunt heeft. Een willekeurig *kleine* verandering van een waarde van $a < 0$ naar een waarde > 0 doet het nulpunt springen van 4 naar 2; we kunnen daardoor het nulpunt niet met elke gewenste graad van nauwkeurigheid benaderen, en voor een constructief reëel getal is dat vereist.

Dit betekent echter niet dat we de tussenwaardestelling nu maar geheel moeten afschrijven als onbruikbaar binnen de context van de constructieve wiskunde. Op verschillende manieren kunnen we bruikbare zwakkere versies krijgen: door verzwakking van de conclusie dan wel versterking van de vooronderstellingen. Als voorbeeld van de eerste mogelijkheid kan de volgende stelling dienen:

Laat $f : [0,1] \longrightarrow \mathbb{R}$ een continue functie zijn, $f(0) < 0, f(1) > 0$. Dan geldt $\forall \epsilon > 0 \exists x \in [0,1](|f(x)| < \epsilon)$.
Het is niet moeilijk hiervoor een constructief bewijs te geven. Als tweede mogelijkheid geven we Lat $f : [b,c] \longrightarrow \mathbb{R}$ een continue functie zijn, $b < c$, $f(b) < 0, f(c) > 0$, en voor alle x, y met $b \leq x < y \leq c$ geldt $\exists z \in (x,y)(f(z) \sharp 0)$. Dan heeft f een nulpunt in $[b,c]$.

Deze stelling kan met een variant van een welbekend bewijs bewezen worden: beschouw $[b + \frac{1}{3}(c-b), b + \frac{2}{3}(c-b)]$, dan is er een $z \in [b + \frac{1}{3}(c-b), b + \frac{2}{3}(c-b)]$, en $f(z) > 0 \vee f(z) < 0$. In het eerste geval gaan we door met het interval $[b,z]$ i.p.v. $[b,c]$, in het tweede geval met $[z,c]$ i.p.v. $[b,c]$. De intervallen worden snel kleiner (bij elke stap met minstens een factor $\frac{2}{3}$), en convergeren naar een punt a waarvoor $f(a) = 0$.

Deze tweede stelling is bijvoorbeeld van toepassing op polynomen van oneven graad (het is niet triviaal aan te tonen dat die aan de voorwaarde van de stelling voldoen!), en levert dan een bewijs van de hoofdstelling van de algebra voor het reëel-afgesloten zijn van $\mathbb{R}^c$: elke polynoom van oneven graad met reële coëfficienten heeft een nulpunt in $\mathbb{R}^c$.

Het bestaan van een getal a waarvoor de juiste ligging t.o.v. 0 niet bekend is, leidt er ook toe dat een functie als: $f(x) = 0$ voor $x < 0$, $f(x) = 1$ voor $x \geq 0$ niet als voorbeeld van een discontinue functie kan dienen, althans niet als voorbeeld van een *totale* discontinue functie. Immers, voor het twijfelgetal a kunnen we niet bepalen of $f(a) > 0$ dan wel $f(a) < 1$. Dit maakt de schijnbaar paradoxale stelling van Brouwer: "alle functies van $\mathbb{R}$ naar $\mathbb{R}$ zijn continu" enigszins begrijpelijk. Brouwer maakt deze stelling aannemelijk door een ruime interpretatie van reëel getal (d.m.v. "keuzerijen"); vgl. ook het model

voor de intuïtionistische reële getallen aan het eind van deze syllabus, waar de verzameling van de reële getallen ook "groter" is dan de gebruikelijke $\mathbb{R}$.

Reële getallen als sneden van Dedekind

Er zijn verschillende, equivalente, manieren om de sneden van Dedekind te definiëren: soms symmetrisch, als een paar (S, S') van deelverzamelingen van $\mathbb{Q}$, die aan zekere voorwaarden voldoen, dan weer asymmetrisch, als "linkssneden" dan wel "rechtssneden". Om technische redenen kiezen wij hier voor linkssneden (rechtssneden zijn spiegelbeeldig) i.p.v. een symmetrische versie. We brengen eerst de bekende definitie in herinnering.

Een *linkssnede* S is een verzameling rationale getallen die voldoet aan

(a) $\exists r(r \in S)$, $\exists r(r \notin S)$ (S is *begrensd*).

(b) $\forall r, r'(r \in S \text{ en } r' < r \rightarrow r' \in S)$ (S is *monotoon*). In woorden: met elk rationaal getal in S zit elk kleiner rationaal getal ook in S.

(c) $\forall r{\in}S \exists r'{\in}S(r' > r)$ (S is *open*). In woorden: als $r \in S$, dan is een open intervalletje rond r geheel in S bevat.

□

$\emptyset$ en $\mathbb{Q}$ zijn zelf geen sneden, want niet begrensd. $\{r : r < 0\} \cup \{r' : 1 \leq r' < 2\}$ is wel begrensd en open, maar niet monotoon; en $\{r : r \leq 2\}$ is monotoon en begrensd, maar niet open.

De conditie van openheid houdt in dat een snede zijn "grenspunt" nooit kan bevatten. De Dedekind reële getallen $\mathbb{R}^{\mathrm{d}}$ kunnen we nu definiëren als de collectie van sneden, geordend door:

$$S \leq S' := S \subset S'.$$

Het is duidelijk dat de rationale getallen met behoud van hun ordening ingebed liggen in $\mathbb{R}^{\mathrm{d}}$; voor de inbedding neemt men eenvoudig de toevoeging $r \mapsto \{r' : r' < r\}$. Is nu de bovenstaande definitie ook een goede constructieve definitie? Nee, zoals uit het volgende voorbeeld blijkt:

$$S := \{r : r < 1\} \cup \{r : r < 2 \text{ en } F\}$$

waarbij F een of ander onopgelost wiskundig probleem voorstelt. S is inderdaad monotoon, begrensd en open, maar we kunnen de "grens" voorgesteld door S niet nauwkeurig benaderen, d.w.z. we kunnen geen twee rationale getallen r_1 en r_2 vinden zodat $r_2 - r_1 < 1$ en $r_1 \in S$ en $r_2 \notin S$; daarvoor zouden we over de waarheid van F moeten kunnen beslissen. De simpelste oplossing lijkt de volgende: intuïtionistisch eisen we van sneden dat ze behalve aan monotonie, openheid en begrensdheid ook nog voldoen aan

(d') $\forall r(r \in S \vee r \notin S)$ (sterke gelocaliseerdheid).

(In de gewone theorie is deze eis triviaal vervuld!) Dan kunnen we inderdaad de grens van S benaderen. Neem bijv. $r \in S$, $r' \notin S$, kies N zo groot dat $(r' - r)N^{-1} < 2^{-k}$, vorm een verdeling van het interval $[r, r']$ met

$$r_0 = r,\ r_N = r',\ r_i = r + i(r' - r)N^{-1}.$$

We bepalen de kleinste i waarvoor $r_i \in S$, $r_{i+1} \notin S$, dan $|r_{i+1} - r_i| < 2^{-k}$.

De voorwaarde van sterke gelocaliseerdheid is echter *te sterk*: als we niet weten of het getal $a \in \mathbb{R}^c$ groter dan wel gelijk aan nul is, dan definiëert

$$(*) \qquad S \equiv \{r : r < a\}$$

een open, begrensde en monotone verzameling waarvan de grens a met elke gewenste graad van nauwkeurigheid benaderd kan worden, maar die niet sterk gelocaliseerd is, aangezien we niet kunnen zeggen of 0 al dan niet in S zit.

Daarom vervangen we "sterke gelocaliseerdheid" door

(d) $\forall r, s(r < s \to r \in S \vee s \notin S)$ (*gelocaliseerdheid*).

Dit is sterk genoeg om de grens van een snede te kunnen benaderen, en zwak genoeg om ook de S van $(*)$ als snede te accepteren.

Opgemerkt dient te worden dat de voorwaarde van monotonie uit die van gelocaliseerdheid volgt: als $r \in S$ en $r' < r$, dan met gelocaliseerdheid $r \in S$ of $r' \notin S$; het tweede is echter uitgesloten. In feite is een gelocaliseerde snede zelfs sterk monotoon:

(b′) $\forall r, s(r < s \wedge \neg\neg s \in S \to r \in S)$ (*sterke monotonie*)

wordt met hetzelfde argument aangetoond.

Om in te zien dat (d) volstaat om de grens van een snede S te benaderen, geven we het volgende argument. Laat weer $r \in S$, $r' \notin S$, en stel $(r - r')N^{-1} < 2^{-k-1}$, r_i als te voren. Bij elk paar r_i, r_{i+1} geldt hetzij $r_i \in S$, dan wel $r_{i+1} \notin S$ (deze mogelijkheden hoeven elkaar niet uit te sluiten). Laat $\phi : \{0, \ldots, N-1\} \longrightarrow \{0, 1\}$ een functie zijn zodanig dat

$$\phi i = 0 \to r_i \in S, \quad \phi i = 1 \to r_{i+1} \notin S$$

Bepaal dan de kleinste i zodat hetzij $i = 0$ en $\phi 0 = 1$, hetzij $\phi i = 0$, $\phi(i+1) = 1$, dan wel $i = N - 1$ en $\phi i = 0$. Dan geldt respectievelijk $r_0 \in S \wedge r_1 \notin S$, $r_i \in S \wedge r_{i+2} \notin S$ $(0 < i < N - 1)$, $r_{N-1} \in S \wedge r_N \notin S$, en daarmee hebben we de gewenste benadering van de grens van S gevonden.

In feite is de voorwaarde van gelocaliseerdheid equivalent met de volgende, meer aanschouwlijke, maar minder elegante voorwaarde:

$$\forall k \exists r \in S \exists r' \notin S(r - r < 2^{-k}).$$

Het is niet moeilijk in te zien dat ook constructief $\mathbb{R}^d$ en $\mathbb{R}^c$ isomorf zijn. Laat ϕ de functie zijn die een $x \in \mathbb{R}^c$ gegeven door een fundamentaalrij $\langle r_n \rangle_n$ afbeeldt op de snede $\{r : \exists k \forall n > k(r_n > r)\}$. Dan beeldt ϕ $\mathbb{R}^c$ in $\mathbb{R}^d$ af met behoud van de ordening. Omgekeerd, laat S een snede zijn; wegens de gelocaliseerdheid is er een $\alpha : \mathbb{Q}^2 \longrightarrow \{0, 1\}$ zodat

$$\forall r, r' \in \mathbb{Q}(r < r' \to (\alpha(r, r') = 0 \wedge r \in S) \vee (\alpha(r, r') = 1 \wedge r' \notin S))$$

Beeld dan S af op

$$\psi(S) := \{r : \alpha(r) = 0\}.$$

OPMERKING. Het bestaan van α op grond van gelocaliseerdheid doet in feite een beroep op een aftelbaar keuzeaxioma van de volgende vorm: als er bij elke n een m te vinden is zodat $A(n, m)$, dan is er een functie α zodat voor alle n $A(n, \alpha n)$. We zullen dit als onproblematisch beschouwen.

Laat zien dat het beeld onder ψ niet van de precieze keuze van α afhangt, en dat ψ de inverse is van ϕ.

Laten we nu de voorwaarde van gelocaliseerdheid vallen, dan rijst de vraag: wat voor structuur vormt de collectie van sneden die open, monotoon en begrensd zijn? Kennelijk iets veelomvattenders dan $\mathbb{R}^{\mathrm{d}}$. Het antwoord luidt: de sneden die alleen aan (a)–(c) voldoen vormen *bijna* de orde completering van $\mathbb{Q}$, d.w.z. de kleinste structuur waarin $\mathbb{Q}$ met behoud van ordening ingebed kan worden zodanig dat elke begrensde verzameling van rationale getallen een kleinste bovengrens en grootste benedengrens in de structuur heeft. Daarbij wordt de ordening op deze collectie net zo gedefiniëerd als te voren voor $\mathbb{R}^{\mathrm{d}}$. Het is *bijna* de orde completering, maar niet helemaal: om een precieze representatie te krijgen zouden we nog een equivalentierelatie moeten invoeren op de sneden die aan (a)–(c) voldoen. Met name willen we bijv. de volgende twee sneden identificeren

$$S = \{r : r < 1\} \quad \text{en} \quad S' = \{r : r < 0\} \cup \{r : r < 1 \wedge F\}$$

waarbij F een zodanige wiskundige bewering is dat we geen bewijs hebben voor F, maar wel voor $\neg\neg F$. Immers, de rationale getallen s met $0 \leq s < 1$ kunnen *onmogelijk niet* in S' zitten, terwijl ze wel in S zitten; het ligt voor de hand dat we S en S' willen identificeren. Inplaats van een wat moeizame hantering van equivalentieklassen kunnen we ook een unieke representant in elke klasse kiezen; dit kan door de voorwaarde van monotonie te verscherpen tot sterke monotonie. Dus:

De verzameling *begrensd-uitgebreide reële getallen* $\mathbb{R}^{\mathrm{be}}$ bestaat uit de sneden die open, sterk monotoon en begrensd zijn. Dergelijke sneden noemen we *zwakke sneden.* De niet-stricte ordening $\leq$ op $\mathbb{R}^{\mathrm{be}}$ wordt gegeven door $S \leq S' := S \subset S'$. □

(Opmerking: in het bovenstaande voorbeeld is S' niet sterk monotoon en dus geen zwakke snede.)

Het is nu niet moeilijk in te zien dat elke begrensde verzameling $X \subset \mathbb{Q}$ een kleinste bovengrens heeft; neem daarvoor

$$\{r : \exists r'(r < r' \wedge \neg\neg r' \in X\}$$

(De simpelste oplossing lijkt op het eerste gezicht $\{r : \exists r' \in X(r < r')\}$. Echter, deze verzameling is wel open en monotoon, maar we kunnen geen *sterke* monotonie aantonen.)

Algemener: als $\mathcal{X} \subset \mathbb{R}^{\mathrm{be}}$, $\mathcal{X}$ is bewoond, d.w.z. bevat minstens een element, en naar boven begrensd, d.w.z. er is een $r \in \mathbb{Q}$ zodat $r \notin S$ voor alle $s \in \mathcal{X}$, dan heeft $\mathcal{X}$ een kleinste bovengrens in $\mathbb{R}^{\mathrm{be}}$. In eerste benadering is dit

$$\bigcup \mathcal{X} = \{r' : \exists S \in \mathcal{X}(r' \in S)\}$$

maar om de sterke monotonie te garanderen nemen we

$$\{s : \exists r' > s \neg\neg \exists S \in \mathcal{X}(r' \in S)\}$$

($\neg\neg$ vanwege de sterke monotonie, $\exists r > s$ om daarna de openheid weer te herstellen).

Met wat meer moeite dan voor $\mathbb{R}^{\mathrm{c}}$ kunnen we ook de rekenkunde voor $\mathbb{R}^{\mathrm{be}}$ doorvoeren als voor de gewone reële getallen. Aangezien, zoals men gemakkelijk in zal zien, Cauchy-volledigheid een gevolg is van orde-volledigheid, lijkt het erop dat $\mathbb{R}^{\mathrm{be}}$ alle mogelijke fraaie afsluitingseigenschappen heeft. Toch vormt $\mathbb{R}^{\mathrm{be}}$ geen goede kandidaat voor de verzameling der constructieve reële getallen: het is in het algemeen niet mogelijk de "plaats van de rand van de snede" met elke gewenste graad van nauwkeurigheid te benaderen. De positie van een zwakke snede is in het algemeen in hoge mate onbepaald, het enige wat altijd bekend is, is een onder- en een bovengrens.

Tot dusverre hebben we alleen de niet-strikte ordening $\leq$ op $\mathbb{R}^{\mathrm{d}}$ en $\mathbb{R}^{\mathrm{be}}$ gedefiniëerd. Een voor de hand liggende definitie van de strikte ordening $<$ op $\mathbb{R}^{\mathrm{d}}$ en $\mathbb{R}^{\mathrm{be}}$ is

$$S < S' := \exists r > 0(S + r \subset S')$$

waarbij $S + r := \{r' + r : r \in S\}$. In het geval van $\mathbb{R}^{\mathrm{d}}$ is een equivalente, meer elegante definitie mogelijk:

$$S < S' := \exists r(r \in S' \wedge r \notin S).$$

Bewijs dat voor $\mathbb{R}^{\mathrm{d}}$ de twee definities van $<$ gelijkwaardig zijn.

Het verschil tussen $\mathbb{R}^{\mathrm{d}}$ en $\mathbb{R}^{\mathrm{be}}$

Tot besluit zullen we een op het eerste gezicht paradoxaal gevolg van het verschil tussen $\mathbb{R}^{\mathrm{be}}$ en $\mathbb{R}^{\mathrm{d}}$ bespreken. Zoals we uit het voorgaande hebben gezien, kunnen we, als we $\vee, \neg$ op intuïtionistische wijze lezen, noch $P \vee \neg P$, noch de zwakkere versie $\neg P \vee \neg\neg P$ accepteren. "Niet accepteren" betekent dat we, met een onopgelost probleem als voorbeeld voor P, geen bewijs voor $P \vee \neg P$, dan wel $\neg P \vee \neg\neg P$ hebben (de voorbeelden van bekende onopgeloste problemen demonstreren niet alleen $?P \vee \neg P?$, maar ook $?\neg P \vee \neg\neg P?$). Maar dat betekent niet dat we voor een dergelijke P zouden hebben aangetoond $\neg(P \vee \neg P)$ dan wel $\neg(\neg P \vee \neg\neg P)$; dat is zelfs tegenstrijdig. Het probleem P kan immers altijd later in positieve dan wel negatieve zin opgelost worden. Anderzijds, aangezien er een in principe onbeperkte voorraad wiskundig onopgeloste problemen bestaat, lijkt het intuitionistisch niet vergezocht te geloven dat ze *onmogelijk* allemaal opgelost kunnen worden. Kortom we zijn geneigd te geloven in het principe van de *ontestbaarheid*

$$\neg\forall P(\neg P \vee \neg\neg P)$$

als een plausibel intuïtionistisch principe. (Brouwer heeft in zijn theorie van de keuzerijen, waarvan de bespreking hier ons te ver zou voeren, meer concrete redenen geleverd om beweringen veel sterker dan ontestbaarheid te accepteren.)

Nu wordt het interessant: tot dusverre konden we met de zwakke tegenvoorbeelden alleen maar *minder* bewijzen dan in de gebruikelijke theorie, maar als we het principe van de ontestbaarheid accepteren dan hebben we een principe omhelsd dat afwijkt van de klassieke theorie en kunnen we dus ook resultaten verwachten die niet in de klassieke theorie passen. We geven een voorbeeld.

Testbaarheid, d.w.z. $\forall P(\neg P \vee \neg\neg P)$, is equivalent met

1. $\mathbb{R}^{\mathrm{be}} = \mathbb{R}^{\mathrm{d}}$, en met
2. Er bestaat een niet-constante functie $f : \mathbb{R}^{\mathrm{be}} \longrightarrow \mathbb{R}^{\mathrm{d}}$, d.w.z. er zijn $x_1, x_2 \in \mathbb{R}^{\mathrm{be}}$ zodat $f(x_1) \sharp f(x_2)$.

Als we ontestbaarheid accepteren, zijn alle functies van $\mathbb{R}^{\mathrm{be}}$ naar $\mathbb{R}^{\mathrm{d}}$ constant.

Intuïtief kan men zich de betekenis hiervan zo voorstellen: de objecten in $\mathbb{R}^{\mathrm{be}}$ zijn "diffuus", niet scherp gelocaliseerd; de objecten in $\mathbb{R}^{\mathrm{d}}$ daarentegen zijn scherp gelocaliseerd. De enige soort functies die bij diffuse objecten steeds precieze data kunnen produceren zijn triviale functies, nl. constante functies.

BEWIJS. (a) Stel $\forall P(\neg P \vee \neg\neg P)$ en laat $X \in \mathbb{R}^{\mathrm{be}}$. We moeten aantonen dat X ook gelocaliseerd is, d.w.z.

$$r < s \to r \in X \vee s \notin X.$$

Laat $r < s$; we hebben $s \notin X \vee \neg\neg s \in X$; in het tweede geval wegens strikte monotonie $r \in X$. Dus $\mathbb{R}^{\mathrm{be}} = \mathbb{R}^{\mathrm{d}}$.

(b) Stel nu $\mathbb{R}^{\mathrm{be}} = \mathbb{R}^{\mathrm{d}}$, dan kunnen we de identiteitsfunctie $f(x) = x$ als voorbeeld van een niet-constante functie nemen.

(c) Laat tenslotte f een niet-constante functie zijn met $x_1, x_2 \in \mathbb{R}^{\mathrm{be}}$, $f(x_1) \sharp f(x_2)$. We definiëren een z_P afhankelijk van een willekeurige propositie (=wiskundige bewering) P zodanig dat

$$z_P = \begin{cases} x_1 \text{ als } P \\ x_2 \text{ als } \neg P. \end{cases}$$

Daartoe kunnen we bijvoorbeeld nemen

$$z_P := \{r : \exists r'(r < r' \wedge \neg\neg((r' \in x_1 \wedge P) \vee (r' \in x_2 \wedge \neg P))\}.$$

(In eerste instantie wordt de gezochte z_P benaderd door $\{r' : (r' \in x_1 \wedge P) \vee (r' \in x_2 \wedge \neg P)\}$. Maar dan moeten we $\neg\neg$ invoegen om sterke monotonie te garanderen; en dan moeten we weer $\exists r'(r < r' \ldots$ voorschakelen om de openheid te herstellen.) Dan geldt

$$f(z_P) \sharp f(x_1) \vee f(z_P) \sharp f(x_2),$$

dus $z_P \neq x_1 \vee z_P \neq x_2$, d.w.z. $\neg P \vee \neg\neg P$. $\square$.

Een klassiek model voor $\mathbb{R}^{\mathrm{d}}$ Aangezien het, zeker aanvankelijk, niet eenvoudig is consequent intuïtionistisch te redeneren, kunnen we ons afvragen of we niet een soort model van de intuïtionistische reële getallen kunnen maken dat vanuit het gewone klassieke standpunt begrijpelijk is, en toch bepaalde eigenaardigheden van de intuïtionistische reële getallen demonstreert.

Aangezien "intuïtionistisch waar" zich anders gedraagt dan het gebruikelijke waarheidsbegrip, ligt het voor de hand in zo'n model het begrip "waarheid" te wijzigen. Waar we in de klassieke theorie maar twee *waarheidswaarden*, waar en onwaar kennen, hebben we hier behoefte aan "tussentoestanden" tussen waar en onwaar.

Dit kan bereikt worden in een zogenaamd topologisch model. De waarheidswaarden zijn de *open* verzamelingen van een topologische ruimte T; *onwaar* correspondeert met de lege verzameling, *waar* met de hele ruimte. Twee disjucte open verzamelingen representeren twe onvergelijkbare waarheidswaarden, die noch helemaal waar, noch helemaal onwaar zijn.

Laten we als speciaal geval $T = \mathbb{R}$ uitwerken. Nu is $X \subset \mathbb{R}$ open als voor alle $x \in X$ een ϵ-omgeving van x geheel bevat is in X. We gaan $\mathbb{R}^{\mathrm{d}}$ nu interpreteren als de verzameling $\mathcal{R}$ van *continue* functies van $\mathbb{R}$ naar $\mathbb{R}$. $\mathbb{N}$ en $\mathbb{Q}$ worden geïnterpreteerd door de constante functies $\mathbb{R} \longrightarrow \mathbb{N}$, $\mathbb{R} \longrightarrow \mathbb{Q}$.

Voor uitspraken A is $[\![A]\!]$ de waarheidswaarde van de uitspraak in dit model; $[\![A]\!]$ is een open verzameling in $\mathbb{R}$. Basisuitspraken zijn van de vorm $f = g, f < g, f \leq g$, waarbij $f, g, h \in \mathcal{R}$ (d.w.z. f, g, h zijn "reële getallen" in het model). We leggen vast

$$[\![f = g]\!] := \mathrm{Int}\{t : f(t) = g(t)\},$$
$$[\![f < g]\!] := \{t : f(t) < g(t)\},$$
$$[\![f \leq g]\!] := \mathrm{Int}\{t : f(t) \leq g(t)\}.$$

$\mathrm{Int}(X)$, het *inwendige* van X is de grootste open verzameling die in X bevat is, d.w.z. voor $\mathbb{R}$

$$\mathrm{Int}(X) := \{x : \exists \epsilon > 0 (U_\epsilon(x) \subset X)\}.$$

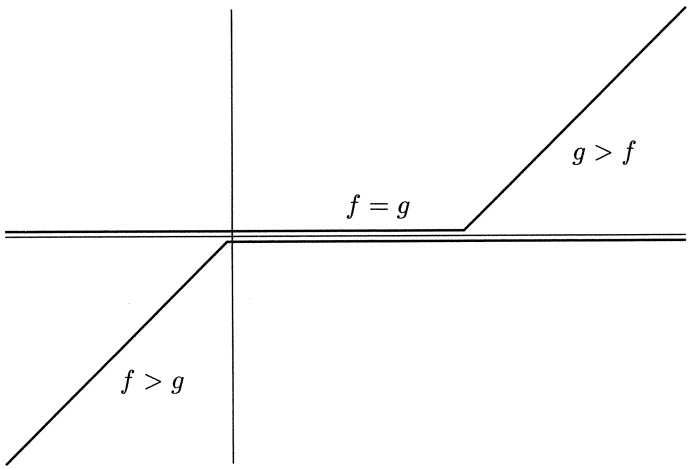
Voor logisch samengestelde uitspraken berekenen we de waarheidswaarden volgens

$$\begin{array}{lcl}
[\![A \wedge B]\!] & := & [\![A]\!] \cap [\![B]\!], \\
[\![A \vee B]\!] & := & [\![A]\!] \cup [\![B]\!], \\
[\![A \to B]\!] & := & \mathrm{Int}([\![B]\!] \cup (\mathbb{N} \setminus [\![A]\!]), \\
[\![\neg A]\!] & := & \mathrm{Int}(\mathbb{R} \setminus [\![A]\!]), \\
[\![\forall x \in \mathbb{R} A(x)]\!] & := & \mathrm{Int}(\bigcap_{f \in \mathcal{R}} [\![A(f)]\!]), \\
[\![\exists x \in \mathbb{R} A(x)]\!] & := & \bigcup_{f \in \mathcal{R}} [\![A(f)]\!]).
\end{array}$$

Nu zien we gemakkelijk in dat $[\![f = g \vee \neg f = g]\!]$ niet altijd waar is in het model: neem

$$f(x) = x \text{ voor } x \leq 0, \;\; 0 \text{ voor } 0 \leq x$$
$$g(x) = 0 \text{ voor } x \leq 0, \;\; x - 1 \text{ voor } 1 \leq x.$$

In het onderstaande plaatje hebben we de horizontale delen van de functies op kleine afstand van de x-as getekend, hoewel ze er in feite langs lopen.



Dan $[\![f = g]\!] = (0,1)$, $[\![f \neq g]\!] = \mathrm{Int}(\mathbb{R} \setminus [\![f = g]\!]) = (-\infty, 0) \cup (1, \infty)$. Derhalve $[\![f = g \vee f \neq g]\!] = (-\infty, 0) \cup (0,1) \cup (1, \infty)$, maar dit is $\neq \mathbb{R}$. Ga zelf na dat wel altijd

$$[\![f < g \to f < h \vee h < g]\!] = \mathbb{R}.$$

Geef een voorbeeld dat laat zien dat $f \leq g \vee g \leq f$ niet algemeen waar is. Laat zien dat $[\![\forall f, g(f = g \vee f \neq g]\!] = \emptyset$.

Het is niet moeilijk in te zien dat het principe $x \neq y \to x \sharp y$ in dit model *niet* geldig is. Immers, vergelijk f (als boven) met h gedefiniëerd door

$$h(x) = 0 \text{ als } x \leq 0, \quad h(x) = x \text{ als } x \geq 0.$$

Dan $[\![f = h]\!] = \mathrm{Int}\{t : f(t) = g(t)\} = \mathrm{Int}\{0\} = \emptyset$, en dus $[\![f \neq h]\!] = \mathbb{R}$. Anderzijds, $[\![f \sharp h]\!] = [\![f < h \vee h < f]\!] = [\![f < h]\!] \cup [\![h < f]\!] = (-\infty, 0) \cup (0, \infty)$ en daaruit volgt weer $[\![f \neq h \to f \sharp h]\!] = (-\infty, 0) \cup (0, \infty) \neq \mathbb{R}$.

Dit soort topologische modellen kan men variëren door de topologische ruimte te variëren. In het beschreven model is $\mathbb{R}^c \neq \mathbb{R}^d$ (en het aftelbare keuzeaxioma, nodig om de isomorfie van $\mathbb{R}^c$ en $\mathbb{R}^d$ aan te tonen geldt dus niet), maar over andere T (bijv. het zgn. Cantor-discontinuum) geldt wel $\mathbb{R}^c = \mathbb{R}^d$.

De keuze van de interpretatie van $\mathbb{N}, \mathbb{Q}, \mathbb{R}^d$ in het model lijkt ad hoc, komt a.h.w. uit de lucht vallen. Door het begrip topologisch model nader uit te werken tot "model van schoven over een topologische ruimte" kan men echter laten zien dat de keuze van de interpretatie van $\mathbb{N}, \mathbb{Q}, \mathbb{R}^d, \mathbb{R}^c$ etc. onderling samenhangt en a.h.w. dwingend voorgeschreven wordt door de keuze van de onderliggende topologische ruimte.

Literatuur Voor de elementaire theorie van het reële getal, zie hoofdstuk 5 uit deel 1 van: A.S. Troelstra, D. van Dalen, *Constructivism in Mathematics*, North-Holland Publ. Co. Amsterdam 1988. Daar vindt men ook verdere literatuur. Topologische modellen worden behandeld in hoofdstuk 13–14 van deel 2 van hetzelfde werk. Merk op dat in zekere zin de "gewone" reële getallen in dit model bevat zijn in de vorm van constante functies van $\mathbb{R}$ naar $\mathbb{R}$.

Bekende reële getallen

F. van der Blij

Paragraaf 1: Inleiding

De meeste reële getallen zijn onbekend. Wanneer U zegt: maar ik ken ze toch allemaal, het zijn de oneindig-voortlopende decimale breuken, spreek ik U vandaag tegen, want U kent

$$0.37284192...$$

niet als U mij niet vertelt hoe het verder gaat.

Natuurlijk kent U wel een aantal reële getallen, zoals alle rationale getallen en alle algebraïsche getallen, dat zijn wortels van veelterm-vergelijkingen met gehele rationale coëfficiënten. En met bekende bewerkingen en functies, zoals de logaritme, de exponentiële functie, goniometrische functies enzovoorts kunt U weer nieuwe getallen construeren. Maar dit alles vormt toch maar een aftelbare deelverzameling van de verzameling van alle reële getallen. En het helpt niet als u de getallen π, e en alle samenstellingen hiervan nog extra toevoegt.

We willen ons in deze voordracht bezighouden met de echt bekende reële getallen.

Hoe kunnen die dan bekend zijn? Ik noemde reeds de rationale getallen.

Sommige andere reële getallen zijn goed gedefinieerd, bijvoorbeeld het positieve getal waarvan het kwadraat gelijk aan 2 is. Of het getal π, de verhouding van omtrek en middellijn van de cirkel, of het getal e.

Toch willen we vaak het positieve getal waarvan het kwadraat 2 is beter leren kennen. We schrijven dan bijvoorbeeld

$$\sqrt{2} = 1.41421356...$$

maar wat moet er op de puntjes komen te staan?

Extra vervelend is nog dat ik alleen maar te weten schijn te kunnen komen, welk cijfer op de duizendste plaats achter de komma staat als ik alle 999 voorgangers heb bepaald. Nu is

$$3 : 37 = \ 0.081081081...$$

maar omdat het drietal 081 steeds herhaald wordt, kan ik zonder moeite vertellen welk cijfer op de duizendste plaats achter de komma komt. Dit is een heel andere situatie dan bij wortel 2. U weet en kunt eenvoudig bewijzen dat alleen bij rationale getallen periodieke decimale breuken te voorschijn komen. Nu is zo'n decimale ontwikkeling eigenlijk een representatie als som van een reeks of wat op hetzelfde neerkomt als limiet van een rij. Van de reeks

$$\sqrt{2} = 1 + \sum_{n=1}^{\infty} \frac{c_n}{10^n}$$

weten we van de gehele getallen c_n, die we tussen 0 en 9 kiezen, helaas weinig te zeggen. Maar we weten wel

$$\sqrt{2} = \lim_{n\to\infty} a_n,$$

waarbij

$$a_{n+1} = \frac{1}{2}(a_n + \frac{2}{a_n}); \quad a_1 = 1.$$

Even een stukje opschrijven:

$$a_1 = 1,\ a_2 = \frac{3}{2},\ a_3 = \frac{17}{12},\ a_4 = \frac{577}{408},\ a_5 = \frac{665857}{470832}, \cdots$$

Helpt dit om $\sqrt{2}$ beter te leren kennen? Om de duizendste term te kennen, moeten we toch alle voorgaanden eerst uit rekenen.

Is de voorstelling als

$$\sqrt{2} = \frac{7}{5}\left[1 + \frac{1}{2}\cdot\frac{1}{49} - \frac{1}{8}\cdot\frac{1}{49^2} + \frac{1}{16}\cdot\frac{1}{49^3} - \cdots\right] = \frac{7}{5}\sum_{n=0}^{\infty}\binom{1/2}{n}49^{-n}$$

beter? Hier hebben we direct een uitdrukking voor de n-de term gegeven, in tegenstelling tot de boven gegeven recurrente betrekking.

Er is nog een heel andere manier om wortel 2 te leren kennen:

$$\sqrt{2} = 1 + \frac{1}{x}.$$

Duidelijk is dat $x > 1$, zelfs dat

$$x = 2 + \frac{1}{y}, \quad y > 1$$

en

$$y = 2 + \frac{1}{z}, \quad z > 1.$$

Maar als we x, y, z met de zakrekenmachine berekenen zien we iets bijzonders. Voor zover we kunnen zien lijkt te gelden

$$x = y = z.$$

Zou dat echt waar zijn? Als we schrijven

$$\sqrt{2} = 1 + \frac{1}{x}$$

vinden we

$$x = 1 + \sqrt{2}.$$

We schrijven dit een beetje anders:

$$\sqrt{2} - 1 = \frac{1}{2 + (\sqrt{2} - 1)} = \frac{1}{2 + \dfrac{1}{2 + (\sqrt{2} - 1)}} =$$

en dus geldt

$$\sqrt{2} - 1 = \cfrac{1}{2 + \cfrac{1}{2 + \cfrac{1}{2 + \cdots}}}$$

en het is duidelijk wat op de stippeltjes staat.

We kiezen een wat eenvoudiger notatie voor deze kettingbreuk:

$$\sqrt{2} - 1 = [2, 2, 2, \ldots]$$

In de volgende paragraaf geven we een overzicht van de zaken die we in het vervolg van kettingbreuken zullen gebruiken.

Laten we eens een ander bekend reëel getal bezien, namelijk e. Natuurlijk moeten we ons eerst herinneren welk reëel getal dit is. Ik noem een paar mogelijkheden:

$$e = \lim_{n \to \infty} \left(1 + \frac{1}{n}\right)^n$$

of

$$e = \sum_{n=0}^{\infty} \frac{1}{n!}$$

of e is het getal met de eigenschap dat de afgeleide van e^x gelijk is aan e^x. Maar dan moeten we wel bewijzen dat zo'n getal bestaat!

Natuurlijk kunnen we uit de boven gegeven definities afleiden

$$e = 2.7182818284590\ldots$$

Maar wat staat er op de puntjes? Zou een kettingbreuk hier uitkomst geven? Gaan we met de rekenmachine een stukje van de kettingbreuk voor e bepalen, dan vinden we:

$$e - 2 = [1, 2, 1, 1, 4, 1, 1, 6, 1, 1, 8, 1 \ldots]$$

Gaat het op de stipjes echt zo verder, dat wil zeggen de opeenvolgende even getallen, ieder voorafgegaan en gevolgd door het cijfer 1? Inderdaad kan men bewijzen dat de kettingbreuk voor $e - 2$ deze mooie regelmatige, voorspelbare vorm heeft. In paragraaf 3 komen we hier op terug.

Proberen we hetzelfde met π. We vinden voor de kettingbreuk

$$\pi - 3 = [7, 15, 1, 292, 1, \ldots]$$

en zien geen enkele regelmaat. Jammer, maar ik kan er niets aan doen. Wanneer we een generalisatie van de kettingbreuken invoeren, is wel een regelmatige voorstelling van π te vinden. In de eerste aantekening aan het einde van ons betoog zullen we hier verder op in gaan. Maar is

$$\pi = 4 \sum_{n=0}^{\infty} \frac{(-1)^n}{2n+1}$$

of ook

$$\frac{\pi}{2} = \frac{2.2}{1.3} \cdot \frac{4.4}{3.5} \cdot \frac{6.6}{5.7} \cdot \frac{8.8}{7.9} \cdots \frac{2n.2n}{(2n-1) \cdot (2n+1)} \cdots$$

eigenlijk ook niet heel regelmatig?

We kunnen vragen naar manieren om voor bekende reële getallen rationale benaderingen te vinden. Door partiële sommen van de bovenvermelde reeksen te berekenen, vinden we rationale benaderingen, maar deze zijn niet altijd even mooi. De klassieke benadering voor π, namelijk $\frac{22}{7}$, die we van oudsher kennen, is zo slecht nog niet.

Als we stukjes kettingbreuk als benadering voor π gebruiken, vinden we achtereenvolgens:

$$3\frac{1}{7}, 3\frac{15}{106}, 3\frac{16}{113}, 3\frac{4687}{33102}, 3\frac{4703}{33215}, \cdots$$

Het optreden van het "grote" getal 292 in de kettingbreukontwikkeling van π is de "oorzaak" van het bestaan van de eenvoudige, zeer goede benadering van π.

In het algemeen is het een aardige vraag op welke manieren reële getallen alzo voorgesteld kunnen worden. In het boekje *Irrationalzahlen* van O. Perron zijn verschillende manieren aangegeven.

Wij merken hier op dat de machtreeks

$$\sum c_n x^n, \quad x \textit{ rationaal}$$

in het geval dat de rij c_n periodiek is, een rationaal getal voorstelt; c_n veronderstellen we rationaal. Dit is ook het geval als c_n een polynoom in n is (coëfficiënten rationaal). In het geval dat c_n een gebroken rationale functie van n is, komen er niet-rationale antwoorden. De gevallen

$$\sum \frac{x^n}{n}$$

en

$$\sum \frac{x^n}{2n+1}$$

laten zien dat nu logaritmische en goniometrische irrationaliteiten gaan optreden.

De bekende machtreeks

$$\sum \frac{x^n}{n!}$$

geeft weer andere irrationaliteiten.

PARAGRAAF 2: KETTINGBREUKEN

In de vorige paragraaf kwamen we kettingbreuken tegen. We willen nu een wat formelere introductie van deze objecten geven en enkele nuttige eigenschappen afleiden.

We definiëren het symbool

$$[a_1, a_2, a_3, \ldots, a_n; x]$$

door volledige inductie naar het aantal voorkomende variabelen:

$$[a_1; x] = \frac{1}{a_1 + x}$$

en

$$[a_1, a_2, \ldots, a_n; x] = [a_1, a_2, \ldots, a_{n-1}; \frac{1}{a_n + x}].$$

We kiezen de getallen a_k uit de positieve gehele getallen en het getal x zal een niet negatief reëel getal zijn. In veel toepassingen zal bovendien gelden

$$0 \leq x < 1.$$

De kettingbreuk

$$[a_1, a_2, a_3, \ldots, a_n; 0],$$

die we veelal zullen voorstellen als

$$[a_1, a_2, a_3, \ldots, a_n],$$

is te schrijven als het quotiënt van twee veeltermen in de variabelen a_k. We vinden door rechttoe rechtaan te rekenen:

$$[a_1] = \frac{1}{a_1};$$

$$[a_1, a_2] = \frac{a_2}{a_1 a_2 + 1};$$

$$[a_1, a_2, a_3] = \frac{a_2 a_3 + 1}{(a_1 a_2 + 1)a_3 + a_1}.$$

Maar hoe gaat het verder?

STELLING

$$[a_1, a_2, \ldots, a_n; x] = \frac{T_n + xT_{n-1}}{N_n + xN_{n-1}}$$

waarin T_k *en* N_k *polynomen in* $a_1, a_2, \ldots, a_k$ *zijn;* $T_{-1} = 1$, $N_{-1} = 0$, $T_0 = 0$, $N_0 = 1$; *die voor* $k \geq 1$ *gedefinieerd worden:*

$$T_k = a_k T_{k-1} + T_{k-2},$$

$$N_k = a_k N_{k-1} + N_{k-2}.$$

BEWIJS
We bewijzen met volledige induktie.

$$[a_1; x] = \frac{1}{a_1 + x} = \frac{T_1 + xT_0}{N_1 + xN_0}.$$

Stel nu

$$[a_1, a_2, \ldots, a_{k-1}; y] = \frac{T_{k-1} + yT_{k-2}}{N_{k-1} + yN_{k-2}}$$

dan

$$[a_1, a_2, \ldots, a_{k-1}, a_k; x] = [a_1, a_2, \ldots, a_{k-1}; \frac{1}{a_k + x}] =$$

$$\frac{T_{k-1} + \frac{1}{a_k+x}T_{k-2}}{N_{k-1} + \frac{1}{a_k+x}N_{k-2}} = \frac{(a_k T_{k-1} + T_{k-2}) + xT_{k-1}}{(a_k N_{k-1} + N_{k-2}) + xN_{k-1}} = \frac{T_k + xT_{k-1}}{N_k + xN_{k-1}}.$$

Het onderstaande schema werkt handig om bij gegeven getallen a_k de getallen T_k en N_k te bepalen:

k	-1	0	1	2	3	
a_k			a_1	a_2	a_3	
T_k	1	0	1	a_2	$a_2a_3 + 1$	
N_k	0	1	a_1	$a_1a_2 + 1$	$(a_1a_2 + 1)a_3 + a_1$	

STELLING

$$[a_1, a_2, \ldots, a_n; x]$$

ligt altijd tussen $\frac{T_n}{N_n}$ *en* $\frac{T_{n-1}}{N_{n-1}}$, *terwijl* $\frac{T_n}{N_n} - \frac{T_{n-1}}{N_{n-1}} = \frac{(-1)^{n-1}}{N_n \cdot N_{n-1}}$.

BEWIJS
We berekenen

$$\frac{T_n + xT_{n-1}}{N_n + xN_{n-1}} - \frac{T_n}{N_n} = \frac{(-1)^n x}{N_n \cdot (N_n + xN_{n-1})},$$

$$\frac{T_n + xT_{n-1}}{N_n + xN_{n-1}} - \frac{T_{n-1}}{N_{n-1}} = \frac{(-1)^{n-1}}{N_{n-1} \cdot (N_n + xN_{n-1})}.$$

Substitutie van $x = 0$ geeft de te bewijzen formule.

We geven nu een reëel irrationaal getal t tussen 0 en 1, en willen dit als een kettingbreuk voorstellen. We schrijven:

$$\frac{1}{t} = a_1 + s$$

met a_1 is een natuurlijk getal en $0 < s < 1$. We zien dat

$$a_1 = [\frac{1}{t}],$$

waarbij $[p]$ het grootste gehele getal kleiner of gelijk aan p is. Via een eenvoudige rekenprocedure met de zakrekenmachine is voor iedere (niet te grote) n een voorstelling

$$t = [a_1, a_2, a_3, \ldots, a_n; t_n]$$

te vinden. We voeren immers t in, berekenen $\frac{1}{t}$, noteren a_1 als het gehele deel van $\frac{1}{t}$, trekken dit van $\frac{1}{t}$ af. Nu hebben we weer een getal tussen 0 en 1, bepalen weer de inverse; a_2 is het gehele deel, dat we noteren en weer aftrekken enzovoort, enzovoorts. We noemen

$$[a_1, a_2, a_3, \ldots, a_n] = \frac{T_n}{N_n}$$

benaderende breuken voor t.

Passen we dit procedé toe op een rationaal getal, dan breekt het af en we vinden voor het rationale getal een eindige kettingbreuk. We zien dat uit bovenstaande stelling volgt

$$\left| t - \frac{T_{n-1}}{N_{n-1}} \right| \leq \frac{1}{N_{n-1}^2}$$

zodat we een goede schatting voor de gemaakte fout bij de benadering hebben.

Even een tussenopmerking: men kan zelfs bij ieder reëel, niet rationaal, getal t tussen 0 en 1 oneindig veel breuken

$$\frac{g}{h}$$

vinden zodat

$$\left| t - \frac{g}{h} \right| < \frac{1}{h^2.\sqrt{5}}.$$

Beter kan het niet. Als $c > \sqrt{5}$ zijn er bijvoorbeeld maar eindig veel onvereenvoudigbare breuken $\frac{g}{h}$ zodat

$$\left| \frac{-1 + \sqrt{5}}{2} - \frac{g}{h} \right| < \frac{1}{ch^2}.$$

We geven hier geen bewijs van. In de aantekening nummer twee aan het einde van dit opstel gaan we hier verder op in.

STELLING
Laat t een irrationaal reëel getal tussen 0 en 1 zijn. Stel

$$t = [a_1, a_2, \ldots, a_n; x] \text{ met } 0 < x < 1$$

dan geldt

$$\lim_{n\to\infty} [a_1, a_2, \ldots, a_n] = t.$$

Het bewijs is eenvoudig te leveren met behulp van de boven aangegeven schattingen, omdat $a_n > 0$ geldt immers dat N_n tot oneindig nadert. We schrijven voortaan

$$[a_1, a_2, a_3, \ldots, a_n, \ldots]$$

als afkorting voor

$$\lim_{n\to\infty} [a_1, a_2, a_3, \ldots, a_n].$$

We beschouwen nu zuiver periodieke kettingbreuken:

$$[a_1, a_2, \ldots, a_k, a_1, a_2, \ldots, a_k, a_1, a_2, a_3, \ldots]$$

Hiervoor geldt dus

$$x = [a_1, a_2, \ldots, a_k; x]$$

en dus

$$x = \frac{T_k + xT_{k-1}}{N_k + xN_{k-1}}.$$

We zien dat x de positieve wortel van de vierkantsvergelijking

$$N_{k-1}x^2 + (N_k - T_{k-1})x - T_k = 0$$

is.

Hebben we een niet-zuiver periodieke kettingbreuk

$$y = [b_1, b_2, \ldots, b_m, a_1, a_2, \ldots, a_k, a_1, a_2, \ldots, a_k, a_1, a_2, \ldots]$$

dan geldt

$$y = [b_1, b_2, \ldots, b_m; x] = \frac{T_m + xT_{m-1}}{N_m + xN_{m-1}},$$

waarin x een getal voorstelt met een zuiver-periodieke kettingbreuk. Omdat x de wortel is van een vierkantsvergelijking met gehele coëfficiënten geldt dit evenzo voor y.

We zien dat periodieke kettingbreuken corresponderen met wortels van vierkantsvergelijkingen met gehele coëfficiënten.

Dit is een mooi analogon met de klassieke stelling dat rationale getallen corresponderen met periodieke decimale breuken. Helaas ken ik geen generalisatie in die zin dat er een representatie van reële getallen is, waarbij algebraïsche getallen corresponderen met periodieke representaties!

PARAGRAAF 3: EEN KETTINGBREUK VOOR e

Uit de elementaire analyse kennen we het grondtal van de natuurlijke logaritmen. We weten dat

$$e = \sum_{n=0}^{\infty} \frac{1}{n!} = \lim_{n\to\infty} (1 + \frac{1}{n})^n = 2.7182818284590\ldots$$

Met deze voorstelling konden we een stukje van de kettingbreuk van e bepalen:

$$e - 2 = [1, 2, 1, 1, 4, 1, 1, 6, 1, 1, \ldots].$$

We vermoeden als regelmaat: de opeenvolgende even getallen ieder voorafgegaan en gevolgd door het getal 1. Euler bewees dat de kettingbreuk voor $e - 2$ inderdaad deze regelmatige, maar niet-periodieke vorm heeft. Hij merkte ook nog op dat

$$[1, 3, 5, 7, \ldots] = \frac{e^2 + 1}{e^2 - 1}.$$

Dit is een speciaal geval van een algemene formule:

$$[k, 3k, 5k, 7k, \ldots] = \frac{e^{\frac{2}{k}} + 1}{e^{\frac{2}{k}} - 1}.$$

Het bewijs van de algemene formule is in principe niet moeilijk, maar wel een beetje gecompliceerd. We zullen ons beperken tot een schets van het bewijs van de kettingbreuk voor $e - 2$.

Merkwaardig is dat een formeel bewijs eenvoudig is via een listige substitutie, maar de terecht gestelde didaktische vraag is hoe de te substitueren formule in vredesnaam gevonden is. We zullen in onze derde aantekening aan het eind hier iets meer over zeggen. We proberen nu de benaderende breuken van een kettingbreuk

$$[1, 2, 1, 1, 4, 1, 1, 6, 1, \ldots]$$

te bepalen.

k	-1	0	1	2	3	4	5	6	7	8	9
a_k			1	2	1	1	4	1	1	6	1
T_k	1	0	1	2	3	5	23	28	51	334	385
N_k	0	1	1	3	4	7	32	39	71	465	536

We voeren nu nieuwe variabelen in:

$$T_{3k} = A_k,\ T_{3k+1} = B_k,\ T_{3k+2} = C_k$$

$$N_{3k} = M_k,\ N_{3k+1} = Q_k,\ N_{3k+2} = R_k.$$

Dan geldt:

$$B_k = 2(2k+1)\,B_{k-1} + B_{k-2},$$

$$Q_k = 2(2k+1)\,Q_{k-1} + Q_{k-2}.$$

Wanneer de getallen B_k en Q_k gevonden zijn, kunnen de overige eenvoudig gevonden worden uit:

$$A_k = \frac{1}{2}(B_k + B_{k-1});\ C_k = (A_{k+1} - B_k) = (B_{k+1} - A_{k+1}),$$

$$M_k = \frac{1}{2}(Q_k + Q_{k-1});\ R_k = (M_{k+1} - Q_k) = (Q_{k+1} - M_{k+1}).$$

PROBLEEM
Zoek de algemene oplossing van de lineaire differentievergelijking van de tweede orde met niet-constante coëfficiënten

$$X_k = 2(2k+1)\,X_{k-1} + X_{k-2},$$

bepaal daarna de speciale oplossingen B_k en Q_k met beginwaarden

k	-1	0	1	2	3	4
$2(2k+1)$			6	10	14	18
B_k	-1	1	5	51	719	12993
Q_k	1	1	7	71	1001	18089

In de literatuur vinden we tot onze verbazing een expliciete oplossing van de differentievergelijking:

$$X_k = 2(2k+1)X_{k-1} + X_{k-2}.$$

HULPSTELLING

$$P_k = (-1)^k \sum_{\nu=0}^{\infty} \frac{(k+\nu+1)!}{\nu!(2k+2\nu+3)!} \left(\frac{1}{2}\right)^{2\nu+1};\ k \geq -1$$

voldoet aan

$$X_k = 2(2k+1)\,X_{k-1} + X_{k-2}.$$

Het bewijs is direct te leveren door eenvoudige substitutie. De beginwaarden van deze oplossing zijn:

k	-1	0	1	$\alpha = \frac{1}{2}e^{-1/2}$
P_k	$\alpha - \beta$	$3\alpha - \beta$	$19\alpha - 7\beta$	$\beta = \frac{1}{2}e^{1/2}$

Wanneer we van de betrokken differentievergelijking twee lineair onafhankelijke oplossingen kennen, is door lineaire combinatie de algemene oplossing, met willekeurig voorgeschreven beginwaarden te vinden. We proberen nu in het algemeen bij een differentievergelijking

$$X_k = c_k\, X_{k-1} + X_{k-2}$$

waarvan we één oplossing

$$P_k$$

kennen, een tweede te vinden. We proberen een oplossing van de vorm

$$X_k = P_k \cdot Y_k.$$

Substitutie geeft

$$P_k Y_k = c_k P_{k-1} Y_{k-1} + P_{k-2} Y_{k-2}$$

hetgeen we kunnen omschrijven als

$$P_k \cdot (Y_k - Y_{k-1}) = -P_{k-2} \cdot (Y_{k-1} - Y_{k-2}),$$

$$Y_k - Y_{k-1} = -\frac{P_{k-2}}{P_k}(Y_{k-1} - Y_{k-2}),$$

en we zien dat

$$P_k Y_k = P_k \sum_{\nu=1}^{k} (-1)^{\nu+1} \frac{P_0 P_1}{P_\nu P_{\nu-1}},\ k \geq 0$$

een tweede oplossing is, met $Y_0 = 0$ en $Y_1 = 1$. Opdat deze procedure werkt, is nodig dat voor de gegeven oplossing geldt:

$$P_k \neq 0.$$

De tweede oplossing heeft als beginwaarden:

k	-1	0	1	2	3
c_k			c_1	c_2	c_3
$P_k Y_k$	P_1	0	P_1	$c_2 P_1$	$(1 + c_2 c_3) \cdot P_1$

Uit de nu bekende gegevens:

k	-1	0	1	2
$2(2k+1)$			6	10
P_k	$\alpha-\beta$	$3\alpha-\beta$	$19\alpha-7\beta$	$193\alpha-71\beta$
P_kY_k	$19\alpha-7\beta$	0	$19\alpha-7\beta$	$190\alpha-70\beta$
B_k	-1	1	5	51
Q_k	1	1	7	71

vinden we

$$B_k = \frac{1}{3\alpha-\beta}\left[P_k + \frac{-4\alpha+2\beta}{19\alpha-7\beta}P_kY_k\right],$$

$$Q_k = \frac{1}{3\alpha-\beta}\left[P_k + \frac{2\alpha}{19\alpha-7\beta}P_kY_k\right].$$

We onderzoeken nu eerst het gedrag voor k naar oneindig van Y_k. We schrijven daartoe:

$$Y_{2k} = \sum_{\nu=1}^{2k}(-1)^{\nu+1}\frac{P_0P_1}{P_\nu P_{\nu-1}} = P_0P_1\left\{\frac{P_2-P_0}{P_0P_1P_2} + \frac{P_4-P_2}{P_2P_3P_4} + \frac{P_6-P_4}{P_4P_5P_6} + \cdots\right\}.$$

We vinden de volgende schatting:

$$P_{2l} = \sum_{\nu=0}^{\infty}\frac{(2l+2\nu+1)!}{\nu!(4l+2\nu+3)!}\frac{1}{2^{2\nu+1}} \leq \sum_{\nu=0}^{\infty}\frac{1}{\nu!} = e.$$

$$P_0 = 3\alpha-\beta > 10^{-2} \qquad P_1 = 19\alpha-7\beta < -10^{-3}$$

en daaruit volgt:

$$Y_{2k} < -10^{-5}\left\{\frac{6}{P_0P_2} + \frac{10}{P_2P_4} + \frac{14}{P_4P_6} + \cdots\right\} < -10^{-6}\{6+10+14+\cdots\}.$$

Maken we hiervan gebruik dan vinden we

$$\lim_{k\to\infty} Y_{2k} = -\infty$$

en daaruit volgt direct:

$$\lim_{k\to\infty}\frac{B_{2k}}{Q_{2k}} = \lim_{k\to\infty}\frac{19\alpha-7\beta+(-4\alpha+2\beta)Y_{2k}}{19\alpha-7\beta+2\alpha Y_{2k}} = \frac{-4\alpha+2\beta}{2\alpha} = e-2$$

waarmee bewezen is dat:

$$e-2 = [1,2,1,1,4,1,1,6,1,1,8,1,\ldots].$$

Paragraaf 4: Aantekeningen

Eerste aantekening
Door wat manipulaties met de productvoorstelling van $\frac{\pi}{2}$ (van Wallis) vond W. Brouncker in 1656 de volgende "Kettingbreuk":

$$\frac{\pi}{4} = \cfrac{1}{1+\cfrac{1^2}{2+\cfrac{3^2}{2+\cfrac{5^2}{2+\cfrac{7^2}{2+\cfrac{}{\cdots}}}}}}$$

Dit is geen kettingbreuk van het type dat we bestudeerden in paragraaf 2. We kunnen een aantal afkappingen berekenen.

k	-1	0	1	2	3	4	5	6
T_k	1	0	1	2	13	76	789	7334
N_k	0	1	1	3	15	105	945	10095

Met enig manipuleren is deze kettingbreuk om te zetten in een kettingbreuk:

$$\frac{\pi}{4} = [a_1, a_2, a_3, a_4, \ldots]$$

met niet gehele getallen a_k. We geven een rijtje van de waarden van deze nieuwe coëfficiënten

$$a_1 = 2, \quad a_2 = 2 \cdot \frac{1^2}{3^2}, \quad a_3 = 2 \cdot \frac{1^2 \cdot 5^2}{3^2 \cdot 7^2}$$

$$a_4 = 2 \cdot \frac{3^2 \cdot 7^2}{1^2 \cdot 5^2 \cdot 9^2}, \quad a_5 = 2 \cdot \frac{1^2 \cdot 5^2 \cdot 9^2}{3^2 \cdot 7^2 \cdot 11^2}$$

Tweede aantekening

Er is een opmerkelijk verschil tussen de mogelijkheden voor rationale benadering van algebraïsche getallen (wortels van veelterm-vergelijkingen met gehele coëfficiënten) enerzijds en niet-algebraïsche, transcendente getallen anderzijds.

Laat α een positief algebraïsch getal zijn dat wortel is van een irreducibele veelterm van de vergelijking van de graad n, $n > 1$. Een klassiek voorbeeld (Liouville) is de stelling dat er voor ieder positief algebraïsch getal α van de graad n en voor iedere positieve δ maar eindig veel onderling ondeelbare paren van natuurlijke getallen g en h zijn zodat

$$\left|\alpha - \frac{g}{h}\right| \leq \frac{1}{h^{n+\delta}}.$$

Deze stelling is eenvoudig te bewijzen. Stel α is een wortel van

$$f(x) = a_n x^n + a_{n-1}x^{n-1} + a_{n-2}x^{n-2} + \cdots + a_0 = 0; \ (a_k \in \mathbb{Z}).$$

Uit

$$\left|\alpha - \frac{g}{h}\right| \leq \frac{1}{h^{n+\delta}}$$

volgt voor h voldoende groot:

$$\frac{g}{h} < \alpha + 1$$

en hieruit volgt dat er een van g en h onafhankelijk getal M is zodat voor iedere η tussen α en $\frac{g}{h}$

$$|f'(\eta)| \leq n|a_n|\eta^{n-1} + (n-1)|a_{n-1}|\eta^{n-2} + \cdots \leq M.$$

Voor h voldoende groot geldt dus

$$\left|\frac{f(\alpha) - f(\frac{g}{h})}{\alpha - \frac{g}{h}}\right| = |f'(\eta)| \leq M < h^{\delta}.$$

Nu is

$$\left|f(\frac{g}{h})\right| = \left|\frac{a_n g^n + a_{n-1}g^{n-1}h + \cdots}{h^n}\right| \geq \frac{1}{h^n}.$$

Dus geldt voor alle breuken $\frac{g}{h}$ met h voldoende groot dat

$$\left|\alpha - \frac{g}{h}\right| > \left|\frac{f(\frac{g}{h})}{h^{\delta}}\right| \geq \frac{1}{h^{n+\delta}}.$$

Later zijn belangrijke verscherpingen van deze stelling gevonden.

Roth bewees in 1955 dat voor een niet-rationaal algebraïsch getal α en een willekeurig positief getal δ geldt dat er slechts eindig veel breuken $\frac{g}{h}$ zijn met

$$\left|\alpha - \frac{g}{h}\right| < \frac{1}{h^{2+\delta}}.$$

Met de opmerking van Liouville lukt het al vele transcendente getallen te construeren. Dit kan bijvoorbeeld door de constructie van kettingbreuken waarvan de definiërende getallen a_n voldoende snel groeien.

Derde aantekening

Hoe komt men op de gedachte dat

$$P_k = (-1)^k \sum_{\nu=0}^{\infty} \frac{(k+\nu+1)!}{\nu!(2k+2\nu+3)!}\left(\frac{1}{2}\right)^{2\nu+1}, \quad k \geq -1$$

een oplossing van de differentievergelijking

$$X_k = 2(2k+1)X_{k-1} + X_{k-2}$$

is?

Of hoe tot het vermoeden dat

$$(-1)^k \sum_{\nu=0}^{\infty} \frac{2^k(k+\nu+1)!}{\nu!(2k+2\nu+3)!}x^{k+2\nu}$$

een oplossing van

$$X_k = \frac{2k+1}{x}\,X_{k-1} + X_{k-2}$$

is?

Enkele mogelijkheden. Een toegepast wiskundige die aan problemen met cylindersymmetrie gewerkt heeft of een astronoom die zich expliciet met de planetenbeweging heeft bezig gehouden, kent de Besselfuncties. Deze zijn bijvoorbeeld gedefinieerd door:

$$J_k(z) = \frac{1}{\pi} \int_0^{\pi} \cos[z \sin \varphi - k\varphi] d\varphi.$$

Een hieruit direct af te leiden bekende formule is:

$$J_{k+1}(z) + J_{k-1}(z) = \frac{2k}{z} J_k(z).$$

En deze relatie lijkt wel erg veel op de op te lossen differentievergelijking. Met een klein beetje handigheid is een relatie tussen een oplossing van de gestelde differentievergelijking en de Besselfuncties te vinden. De machtreeksontwikkelingen van de Besselfuncties voeren dan tot het beoogde resultaat.

Werkers op het gebied van de combinatoriek of van de getaltheorie weten het algemene idee om de elementen van een rij onbekende getallen op te vatten als coëfficiënten van een (formele) machtreeks. Men definieert dan bijvoorbeeld:

$$\tilde{\varphi}(t) = \sum_{k=0}^{\infty} X_k t^k.$$

Dan is

$$\tilde{\varphi'}(t) = \sum_{k=0}^{\infty} k X_k t^{k-1},$$

en gezien het feit dat in de differentievergelijking

$$(2k+1)X_{k-1}$$

voorkomt, voeren we liever in

$$\varphi(t) = \sum_{k=0}^{\infty} X_{k-1} t^{2k+1}.$$

Dan geldt

$$\varphi'(t) = \sum_{k=0}^{\infty} (2k+1) X_{k-1} t^{2k}$$

en de differentievergelijking voert tot de differentiaalvergelijking (van de eerste orde en de eerste graad):

$$\frac{1}{4}(t^2 + \frac{1}{t^2})'\varphi(t) + \varphi'(t) = \frac{1}{2}(X_{-1} - X_0) - \frac{1}{2t^2} X_{-1}.$$

Deze is met elementaire methoden uit de leer van de differentiaalvergelijkingen op te lossen bijvoorbeeld via een integrerende factor. Dan komt een factor

$$e^{\frac{1}{4}(t^2+\frac{1}{t^2})}$$

te voorschijn en de te bepalen primitieve vertoont overeenkomst met formules over Besselfuncties, dus.... Zouden er nog andere en betere methoden zijn om op het geniale idee te komen? We laten deze vraag nu nog onbeantwoord.

Vierde aantekening
Het klassieke standaardwerk over kettingbreuken is:
O. Perron: Die Lehre von den Kettenbrüchen Band I/II Stuttgart 1954/1957, 3te Auflage.
Maar in het eenvoudige boekje:
O. Perron: Irrationalzahlen, Berlin, 1960, vierte Auflage is al heel wat te vinden over ons onderwerp.
Over de approximatiestellingen is een klassiek boek:
S. Lang: Diophantine Geometry, New York, 1962.
Voor oudere resultaten verwijs ik naar:
J.F. Koksma: Diophantische Approximationen, Berlin, 1936.
Formules over Besselfuncties vindt men in ieder boek over "Advanced Calculus".
Over algemene zaken is veel te vinden in:
H.-D. Ebbinghaus, et al. Zahlen, Berlin 1983.

OPGAVEN VOOR HET OPGAVENUURTJE

ZAKREKENMACHINE MEEBRENGEN !!

Het is niet verboden er thuis al mee te beginnen, mocht U ze allemaal af hebben dan kunt U tijdens het middagoefenuurtje nog wel extra aandacht aan opgave 11 geven.

1) Bepaal natuurlijke getallen t en n met $0 < t < n < 10000$, zodat $t/n = 0.772181325\ldots$

2) Zoudt U een voor leerlingen begrijpelijke methode weten om natuurlijke getallen t en n te vinden met $0 < t < n < 100$, zodat $t/n = 0.39784\ldots$

3) $1/17$ heeft een periodieke decimale breuk. Wat is de lengte van de periode en vindt met zo min mogelijk moeite alle cijfers van de periode op een eenvoudige zakrekenmachine.
Kunt U in het algemeen iets zeggen over de lengte van de periode van decimale ontwikkeling van een gewone breuk? En over methoden om de decimalen van de ontwikkeling met een zakrekenmachine te vinden?

4) Geef een voor leerlingen begrijpelijke methode om effectief een aantal decimalen van de ontwikkeling van π te bepalen. Probeer ook benaderingen met wortels af te leiden, zoals bijvoorbeeld

$$6\sqrt{2-\sqrt{3}}.$$

5) Zoals bekend zijn 3 1/7 en 355/113 goede benaderingen voor π. Zoek goede benaderingen voor:

$$\pi^2, e^2, e^{\pi}.$$

6) Bepaal de lengte van de periode van de kettingbreuk voor:

$$\sqrt{n},\ n \in \mathbb{N},\ 2 \leq n \leq 27.$$

(U mag het werk met uw buurpersoon verdelen.) Heeft U een vermoeden van een stelling? Zo ja, kunt U deze bewijzen?

7) Bereken voor verschillende waarden van het natuurlijke getal n met de zakrekenmachine
$x = 1/n$
$y = 1/x$
$z = y - n$
en verklaar het antwoord. Doe hetzelfde met
$x = sqr(n)$
$y = x * x$
$z = y - n$

Een antwoord als "onnauwkeurigheid" of "afronden" is niet voldoende, U moet met het antwoord voorspellingen kunnen doen en deze via experimenten kunnen controleren.

8) Wat voor getallen worden voorgesteld door

$$\sum \frac{c_n}{n!} x^n; \quad x \textit{ geheel of } x \textit{ rationaal}$$

waarbij c_n een polynoom met gehele coëfficiënten in n is.

9) Zet de zakrekenmachine op grad. Begin met 0 en toets een aantal malen achter elkaar de toets cos in. Verklaar het resultaat.

10) Op sommige rekenmachines lukt het om natuurlijke getallen a en b te vinden zodat $sqr(a) * sqr(b)$ - $sqr(b) * sqr(a)$ ongelijk aan nul is. Ik weet geen verklaring. U wel?

11) In the Mathematical Intelligencer, 15, (1993), p. 52 komt een artikel "The Patron Saint of Mathematics" van R.J. Duffin voor, dat mij inspireerde tot de volgende opgave:
Definieer
$B(t) = t - n$ als $n \leq t < n + 1$; n geheel
$G(t) = m$ als $m - 1 < t \leq m$; m geheel
en bepaal
$G(7B(n\pi))$
voor $\leq n \leq 80$.
Bepaal ook
$1 + [7n\pi - 7[n\pi]]$; $[t]$ stelt het grootste gehele getal kleiner of gelijk t voor.
Verklaar eventuele bijzonderheden.

12) Bedenk zelf nog enkele zinvolle opgaven en probeer deze op te lossen.

On numbers and games

Chapter 0: All Numbers Great and Small

J.H. Conway

Whatever is not formidden, is permitted.
J.C.F. von Schiller, Wallensteins Lager

This book is in two = {zero, one | } parts. In this zeroth part, our topic is the notion of *number*. As examples we have the finite numbers 0,1,2,..., -1,$\frac{1}{2}$, $\sqrt{2}, \pi, \ldots$; infinite numbers such as ω (the first infinite ordinal); and also infinitesimal number such as $1/\omega$. If we were to adopt the axiom of choice, then the infinite cardinal numvers like $\aleph_0$ could be identified with the least corresponding ordinal numbers, so that we can regard these too as part of our system (althoug the arithmetic is different).

In the system we shall describe, every number has its own unique name and properties and many remarkable numbers, such as

$$\sqrt[3]{(\omega + 1)} - \frac{\pi}{\omega}$$

appear. But the 'number' $i = \sqrt{-1}$ will not arise in the same way (though we add it in Chapter 4), since there is no property enjoyed by i which is not shared by $-i$. In fact we reply to questions about 'the square root of -1' by simply asking exactly which square root of -1 is meant?

Let us see how those who were good at constructing numbers have approached this problem in the past.

Dedekind (and before him the author–thought to be Eudoxus–of the fifth book of Euclid) constructed the real numbers from the rationals. His method was to divide the rationals into two sets L and R in such a way that no number of L was greater than any number of R, and use this 'section' to define a new number $L|R$ in the case that neither L nor R had an extremal point.

His method produces a logically sound collection of real numbers (if we ignore some objections on the grounds of ineffectivity, etc.), but has been critised on several counts. Perhaps the most important is that the rationals are supposed to have been already constructed in some other way, and yet are 'reconstructed' as certain real numbers. The distinction between the 'old' and 'new' rational seems artificial but essential.

Cantor constructed the infinite ordinal numbers. Supposing the integers 1,2,3,... given, he observed that their *order-type* ω was a new (and infinite) number greater than all of them. Then the order-type of $\{1, 2, 3, \ldots, \omega\}$ is a still greater number $\omega + 1$, and so on, and on, and on. The similar objections to Cantor's procedure have already been met by von Neumann, who observes

that it is unnecessary to suppose 1,2,3,... given, and that it is natural to start at 0 rather than 1. He also takes each ordinal as the *set* (rather than the order-type) of all previous ones. Thus for von Neumann, 0 is the empty set. 1 the set $\{0\}$, 2 the set $\{0,1\},\ldots,\omega$ the set $\{0,1,2,\ldots\}$, and so on.

In this chapter we intend to show that these two methods are in reality part of a simpler and more general method by which we can construct a very large Class **No** of numbers, including at the same time the real numbers and the ordinal numbers, and others such as those menthoned above. Because of the generality of this Class, we shall simply describe its members as *numbers*, without adding any restricting adjective. It will turn out that **No** is a Field (i.e. a field whose domain is a proper Class)– in general we shall capitalise the initial letter of any 'big' concept, on the grounds that proper Classes, like proper names, deserve capital letters. So, for instance, the word *Group* will mean any group whose fomain is a proper class.

CONSTRUCTION
If L, R are any two sets of numbers, and no member of L is $\geq$ any member of R, then there is a number $\{L|R\}$. All numbers are constructed in this way.

CONVENTION
If $x = \{L|R\}$ we write x^L for the typical member of L, and x^R for the typical member of R. For x itself we then write $\{x^L|x^R\}$.

$x = \{a,b,c,\ldots|d,e,f,\ldots\}$ means that $x = \{L|R\}$, where $a,b,c,\ldots$ are the typical members of L, and $d,e,f,\ldots$ the typical members of R.

DEFINITIONS

Definition of $x \geq y, x \geq y$.
We say $x \geq y$ iff (no $x^R \leq y$ and $x \leq$ no y^L), and $x \leq y$ iff $y \leq x$.

We write $x \not\leq y$ to mean that $x \leq y$ does not hold.

Definition of $x = y, x > y, x < y$.
$x = y$ iff ($x \geq y$ and $y \geq x$). $x > y$ iff ($x \geq y$ and $y \not\geq x$).
$x < y$ iff $y > x$.

Definition of $x + y$.
$x + y = \{x^L + y, x + y^L | x^R + y, x + y^R\}$.

Definition of $-x$.
$-x = \{-x^r | - x^L\}$.

Definition of xy.
$xy = \{x^L y + xy^L - x^L y^L, x^R y + xy^R - x^R y^R|$
$|x^L y + xy^R - x^L y^R, x^R y + xy^L - x^R y^L\}$.

It is remarkable that these few lines already define a real-closed field with a

very rich structure.

We now comment on the definitions. A most important comment whose logical effects will be discussed later is that *the notion of equality is a defined relation.* Thus apparently different definitions will produce the same number, and we must distinguish between the $form\{L|R\}$ of a number and the number itself.

All the definitions are inductive, so that to decide, for instance, whether $x \geq y$ we must consider a number of similar questions about the pairs x^R, y and x, y^L, but these problems are all simpler than the given one. It is perhaps not quite so obvious that the inductions require no basis, since ultimately we are reduced to problems about members of the empty set.

In general when we wish to establish a proposition $P(x)$ for all numbers x, we will prove it inductively by deducing $P(x)$ from the truth of all the propositions $P(x^L)$ and $P(x^R)$. We regard the phrase 'all numbers are constructed in this way' as justifying the legitimacy of this procedure. When proving propositions $P(x, y)$ involving two variables we may use *double induction*, $P(x, y)$ from the truth of all propositions of the form $P(x^L, y)$, $P(x^R, y)$, $P(x, y^L)$, $P(x, y^R)$ (and, if necessary, $P(x^L, y^L)$, $P(x^L, y^R)$, $P(x^R, y^R)$). Such multiple inductions can be justified in the usual way in terms of repeated single inductions.

We shall allow ourselves to use certain expressions $\{L|R\}$ which are not numbers, since they do not satisfy the condition that no member of L shall be $\geq$ any member of R. In general we may write down any expression $\{L|R\}$ and even discuss inequalities between such expressions before establishing that they are members, but if we wish such an expression to represent a number we must establish the condition on L and R. In the more general theory developed in the next part of the book, we show that when the condition on L and R is omitted we obtain the more general notion of a *game.*

Our next comments concern the motives for these particular definitions. Now it is our intention that each new number x shall lie between the numbers x^L (to the left) and x^R (to the right), and that $\geq, +, -, .$, etc., shall have their usual properties. So that if (say) $y \geq$ some x^R we would not have $x \geq y$, for then $x \geq x^R$. Similarly, we could not allow $x \geq y$ if $x \leq$ some y^L. So we define $x \geq y$ in all other cases. (This conforms with our motto, and helps to ensure that numbers are totally ordered.)

The spirit of the definitions is to ask what we know already (i.e. by the answers to *simpler* questions) about the object being defined, and to make the answers part of our definition. Thus if addition is to have nice properties and if x is between x^L and x^R, and y between y^L and y^R, then we know 'already' that $x+y$ must lie between both x^L+y and $x+y^L$ (on the left) and x^R+y and $x+y^R$ (on the right), which yields the definition of $x+y$. Similarly $-x$ will lie between $-x^R$ (on the left) and $-x^L$ (on the right), which suffice to define $-x$.

It is not nearly so easy to find exactly what we 'already' know about xy. It might seem, for instance, that we know that xy lies between x^Ly and xy^L (on the left) and x^Ry and xy^R (on the right), which would yield the definition

$$xy = \{x^Ly, xy^L | x^Ry, xy^R\}.$$

But this fails in two ways. Firstly, what we 'knew' here is sometimes false (consider negative numbers), and secondly, even when it is true it need not be the strongest information we 'already' know. In fact, of course, this defines the same function as $x + y$.

It takes a great deal of thought to find the correct definition, which comes from the observation that (for instance) from $x - x^L > 0$ and $y - y^L > 0$ we can deduce $(x - x^L)(y - y^L) > 0$, so that we must have $xy > x^L y + xy^L - x^L y^L$. Since all the products here are simpler ones, and since we regard addition and subtraction as already already defined, we can regard this inequality as already known when we come to define xy, and the other inequalities in the definition are similar. [Note that for positive numbers x and y the inequality $xy > x^L y + xy^L - x^L y^L$ is stronger than both inequaliteis $xy > x^L y$, $xy > xy^L$.]

We can summarise our comments by saying that the definitions of the various operations and relations are just the simplest possible definitions which are consistent with their intended prperties. In the next chapter, we shall verify that these intended properties really hold of all numbers, but in the rest of this chapter we shall simply explore the system in a more informal way. To simplify the notation, when L is the set $\{a, b, c, \ldots\}$ and R the set $\{\ldots, x, y, z\}$, we shall simply write $\{a, b, c, \ldots | \ldots, x, y, z\}$ for $\{L|R\}$.

Examples of numbers, and some of their properties

The number 0

According to the construction, every number has the form $\{L|R\}$, where L and R are two sets of earlier constructed numbers. So how can the system possibly get 'off the ground', since initially there will be no earlier constructed numbers?

The answer, of course, is that even before we have any numbers, we have a certain *set* of numbers, namely *the empty set* $\emptyset$! So the earliest constructed number can only be $\{L|R\}$ with both $L = R = \emptyset$, or in the simplified notation, the number $\{|\}$. This number we call 0.

Is 0 a number? Yes, since we cannot have any inequality of the form $0^L \geq 0^R$, there is neither a 0^L nor a 0^R!

Is $0 \geq 0$? Yes, for we can have no inequality of the form $0^R \leq 0$ or $0 \leq 0^L$. So by the definition, and happily, we have $0 = 0$. We also see from the definitions that $-0 = 0 + 0 = 0$, since there is no number of any of the forms -0^R, -0^L, $0^L + 0, 0 + 0^L$, $0^R + 0.0 + 0^R$. A slightly more complicated observation of the same type is that $x0 = 0$, since in every one of the terms defining xy there is a mention of y^L or y^R, so that when $y = 0$ no term is needed and the expression for xy reduces to $\{|\} = 0$. So the number 0 has at least some of the properties we know and love.

The numbers 1 an -1

We can now use the sets $\{\}$ and $\{0\}$ for L and R, obtaining hopefully the numbers $\{|\}$, $\{0|\}$, $\{|0\}$, $\{0|0\}$. But since we have already proved that $0 \geq 0$, $\{0|0\}$ is *not* a number, and we have only two new cases, which we call $1 = \{0|\}$ and $-1 = \{|0\}$. Note that -1 is indeed a case of the definition $-x$.

Is $0 \geq 1$? This will be true unless there is 0^R with $0^R \leq 1$ (there isn't) or 1^L with $0 \leq 1^L$ (there is , namely $1^L = 0$). So we do *not* have $0 \geq 1$.

Is $1 \geq 0$? This is true unless there is 1^R with '...' or 0^L with '...' (whatever '...', there plainly can't be). So we have $1 \geq 0$, and so $1 > 0$.

By symmetry, we have $-1 < 0$, and so if inequalities 'behave', then we should have $-1 < 1$. We check this:

Is $-1 \geq 1$? This will happen unless there is $(-1)^R \leq 1$ or ... (there is, namely $(-1)^R = 0$). So we do not have $-1 \geq 1$.

Is $1 \geq -1$? This will happen unles there is 1^R with ... or $(-1)^L$ with ... (there isn't). So $1 > -1$, as we hoped.

We can generalise a part of this last argument. If there is no x^R and no y^L, then $x \geq y$ holds vacuously.

We forgot to check that $1 \geq 1$. Why not do this yourself?

The numbers 2, $\frac{1}{2}$, and their negatives

We now have three numbers $-1 < 0 < 1$, and so a whole battery of particular sets

$$\{\}, \{-1\}, \{0\}, \{1\}, \{-1, 0\}, \{-1, 1\}, \{0, 1\}, \{-1, 0, 1\}$$

to use for L and R. But the condition that no member of L should be $\geq$ any member of R restricts us to the possibilities

$$\{|R\}, \{L|\}, \{-1|0\}, \{-1|0, 1\}.\{-1|1\}, \{0|1\}, \{-1, 0|1\}.$$

If our hopes are fulfilled, we should have $\{1|\} > 0$ and $0 < \{0|1\} < 1$. So we anticipate their probable values, and define $\{1|\} = 2$, $\{0|1\} = \frac{1}{2}$. We then have, of course, $\{|-1\} = -2$, and $\{-1|0\} = -\frac{1}{2}$, from the definition of negation.

Before we justify these names, let us ask about some of the other possibilities. For example, what about the number $x = \{0, 1|\}$? This x is presumably restricted by the conditions $0 < x$, $1 < x$. But since $0 < 1$, if inequalities behave (and we shall suppose from now on that they do), the condition $1 < x$ already implies $0 < x$, so in some sense the entry 0 isn't telling us anything. We can therefore hope that $x = \{0, 1|\} = \{1|\} = 2$. We test this, supposing $2 > 1 > 0$.

Is $x \geq 2$? This is so unless there is $x^R \leq 2$ (no) or $x \leq$ some 2^L (no, because the only 2^L is 1, and we believe $x > 1$). So we think that $x \geq 2$.

Is $2 \geq x$? Yes, unles some 2^R... (no) or $2 \leq$ some x^L (no, since the only x^L are 1 and 0). So indeed $x = 2$, if all our expectations are fulfilled.

In a similar way, we should expect all the equalities in the table:

$$\begin{aligned}
-2 &= \{|-1\} = \{|-1.0\} = \{|-1, 1\} = \{|-1, 0, 1\} \\
-1 &= \{|0\} = \{|0, 1\} \\
-\tfrac{1}{2} &= \{-1|0\} = \{-1|0, 1\} \\
0 &= \{|\} = \{-1|\} = \{|1\} = \{-1|1\} \\
\tfrac{1}{2} &= \{0|1\} = \{-1.0|1\} \\
1 &= \{0|\} = \{-1, 0|\} \\
2 &= \{1|\} = \{0, 1|\} = \{-1, 1|\} = \{-1, 0, 1|\}.
\end{aligned}$$

Clearly we need some way of automating our expectations. Let us ask when the number $X = \{y, x^L | x^R\}$ obtained by adding a new entry y to the left of x is still equal to x.

If $X \geq x$? Yes, unless some $X^R \leq x$ (no, since every X^R is an x^R) or $X \leq$ some x^L (no, since every x^L is an X^L).

Is $x \geq X$? Yes, unless some $x^R \leq X$ (no, since every x^R is and X^R) or $x \leq$ some X^L (and so $x \leq y$, since every other X^L is an x^L). We conclude that provided $y \not\geq x$, we can add y to the left of x in this way without affecting x. This justifies all the equalities in the table. (We allow also, of course, y to be inserted on the right if $y \not\leq x$.)

[In the case $\{-1|1\}$ we need to use the process twice. Thus since $-1 \not\geq 0 = \{|\}$, we have $0 = \{-1|\}$. Then since $1 \not\leq 0 = \{-1|\}$, we have $0 = \{-1|1\}$.]

It is not hard to check theinequalities

$$-2 < -1 < -\frac{1}{2} < 0 < \frac{1}{2} < 1 < 2,$$

which shows that at least these numbers have the right order properties. What else do we require to justify their names?

According to the definition

$$1 + 1 = \{0 + 1, 1 + 0\},$$

since 0 is the only 1^L, and there is no 1^R. So provided $0 + 1$ and $1 + 0$ behave as expected, we have $1 + 1 = 2$, as we might hope, But provided $x^L + 0 = x^L$ and $x^R + 0 = x^R$, we have

$$x + 0 = \{x^L + 0 | x^R + 0\} = \{x^L | x^R\} = x,$$

and similarly $0 + x = x$. Since we already know $0 + 0 = 0$, this shows that $1 + 0 = 0 + 1 = 1$, as we wanted for the proof of $1 + 1 = 2$, but in fact it gives us an inductive prrof that $x + 0 = 0 + x = x$ for all x.

It is much harder to show that $\frac{1}{2} + \frac{1}{2} = 1$, justifying the name of $\frac{1}{2}$. From the definition (supposing) that $x + y = y + x$ for all x, y, which is quite easy to prove inductively) we see that

$$\frac{1}{2} + \frac{1}{2} = \{\frac{1}{2} | 1\frac{1}{2}\},$$

where we are using $1\frac{1}{2}$ as a temporary name for $1 + \frac{1}{2}$.

Is $\frac{1}{2} + \frac{1}{2} \geq 1$? Yes, unless $1\frac{1}{2} \leq 1$ or $\frac{1}{2} + \frac{1}{2} \leq 0$. Oh my, we have to do these first. Let's get on with it.

Is $1 \geq 1\frac{1}{2}$? Yes, unless (empty) or $1 \leq$ some $1\frac{1}{2}^L$ is $1 + 0 = 1$, so $1 \not\geq 1\frac{1}{2}$.

Is $0 \geq \frac{1}{2} + \frac{1}{2}$? Yes, unless (empty) or $0 \leq$ some $(\frac{1}{2} + \frac{1}{2})^L$. But since $0 \leq \frac{1}{2} + 0$, we have $0 \not\geq \frac{1}{2} + \frac{1}{2}$. So (at last) $\frac{1}{2} + \frac{1}{2} \geq^L$.

Now is the time to leave the question

"*is* $1 \geq \frac{1}{2} + \frac{1}{2}$?"

to the reader. He should conclude that indeed $\frac{1}{2} + \frac{1}{2} = 1$.

In most of our examples x^L and x^R have been fairly close to each other, so that there was an obvious candidate for $\{x^L|x^R\}$. When they are far apart, there will be many simple numbers in between–which one of these will $\{x^L|x^R\}$ be? We consider $x = \{-1|2\}$.

Is $x \geq 0$? yes, unless $2 \leq 0$ (false) or $x \leq$ some 0^L (false). So in this case we have $x \geq 0$.

Is $0 \geq x$? Yes, unless some $0^R \leq x$ (false) or $0 \leq -1$ (false). So in fact $x = 0$.

More generally, the argument proves that if every $x^L < 0$ and every $x^R > 0$, then $x = 0$, so for instance $\{-1, -\frac{1}{2}|2, 3\} = 0$.

But when we have defined $2\frac{1}{2}$ and 17 we shall have to decide about $\{2\frac{1}{2}|17\}$. A first guess might be their mean, $9\frac{3}{4}$, but since we have just seen that the mean rule does not always hold, this seems unlikely. A clue is given in the form of the preceding argument–since we must ask the quenstions 'is $x = y$?' for the various possible y in order of simplicity, the answer should be the *simplest* y that is not prohibited. This rule will be established in Chapter 2, and it implies, for instance, that $\{2\frac{1}{2}|17\} = 3$, and $\{\frac{1}{4}|1\} = \frac{1}{2}$.

The numbers $\frac{1}{4}, \frac{3}{4}, 1\frac{1}{2}, 3$, *and so on*

Once we have settled all the trivialities like $x \geq x$ for all x (which we have begun to take for granted), we can proceed a little faster. For instance, if L and R are sets of numbers chosen from those we already have, then since we suspect these numbers are totally ordered, in any expression $x = \{x^L|x^R\}$ we need only consider the greatest x^L (if any) and the least x^R (ditto). This gives us for the next 'day' only the numbers

$$0 < \{0|\frac{1}{2}\} < \frac{1}{2} < \{\frac{1}{2}|1\} < 1 < \{1|2\} < 2 < \{2|\}$$

and their negatives. What are the proper names for these numbers? We suspect that $\{2|\} = 3$, and indeed we can verify that

$$1 + 1 + 1 = \{0 + 1 + 1, 1 + 0 + 1, 1 + 1 + 0|\} = \{2|\}.$$

The equation $\{1|2\} = 1\frac{1}{2}$ is almost as easy to guess and verify. So we shall make $1\frac{1}{2}$ a permanent name for this number.

The two likely guesses for $\{0|\frac{1}{2}\}$ are $\frac{1}{3}$ and $\frac{1}{4}$. If anything, the first might seem the better guess, since otherwise it's hard to see what $\frac{1}{2}$ will be. But in fact it turns out that $\{0|\frac{1}{2}\}$ is $\frac{1}{4}$–at least we can verify that twice this number is $\frac{1}{2}$. In a similar way, $\{\frac{1}{2}|1\}$ turns out to be $\frac{3}{4}$ rather that $\frac{2}{3}$.

It is now easy to quess the pattern for the numbers which take only finitely much work to define. Let us imagine the numbers created on successive 'days', in such a way that on day number n we create all numbers $x = \{L|R\}$ for which every member of each of the two sets L, R has already been constructed. We number the day on which 0 was created with the number 0 itself, so that our creation story begins (or began)? on *the zeroth day.*

Then the numbers seem to form a tree, as shown in Fig. 0. Each node of the tree has two 'childeren', namely the first later numbers born just to the

left and right of it. We guess that on the n'th day the extreme numbers to be born are n and $-n$, and that each other number is the arithmetic mean of the numbers to the left and right of it. Happily, of course, this turns out to be the case. Supposing all this, we know all numbers born on finite days.

The numbers born on day ω

Of course the process doesn't stop with these numbers. The next day we call day ω. Let's consider some of the numbers born then. The largest number is the number ω itself, defined as $\{0,1,2,\ldots|\}$. Of course, ω has many other forms, for instance $\omega = \{1,2,4,8,16,\ldots|\}$, or even $\omega = \{\text{all numbers } (m/2^n)|\}$. But since the collection of numbers to the left of ω has no largest member in these expressions, we cannot simply eliminate all but one of the numbers appearing on the left.

Of course the most negative number born on day ω will be

$$-\omega = \{|0,-1,-2,-3,\ldots\}.$$

The smallest positive number born on this day is the number $\{0|1,\frac{1}{2},\frac{1}{4},\frac{1}{8},\ldots\}$, which turns out to be $1/\omega$, surplrisingly and fortunately.

But besides these strange new numbers, some quite ordinary numbers are born at the same time. For instance, we have

$$\frac{1}{4} < \frac{1}{4}+\frac{1}{16} < \frac{1}{4}+\frac{1}{16}+\frac{1}{64} < \ldots < \frac{1}{3} < \ldots < \frac{1}{2}-\frac{1}{8} < \frac{1}{2},$$

so we might expect the number

$$\{\frac{1}{4},\frac{1}{4}+\frac{1}{16},\frac{1}{4}+\frac{1}{16}+\frac{1}{64},\ldots|\frac{1}{2},\frac{1}{2}-\frac{1}{8},\ldots\} = x, \text{ say}$$

to be $\frac{1}{3}$, and behold, it can in fact be proved that $x+x+x=1$! In a similar way, all of the real numbers defined by Dedekind, including in particular all the remaining rational numbers can be defined as 'Dedekind sections' of the dyadic rational numbers (by which we mean the numbers of the form $m/2^n, m$ and n integers), rather than as sections of *all* rationals. So $\sqrt{2}, e$, and π are all born on day ω.

It is rather nice that our definition of equality ensures automatically that the number (for example)

$$\{\text{ dyadic rationals } < \frac{3}{8}|\text{ dyadic rationals } > \frac{3}{8}\}$$

turns out to be the same as the number $\frac{3}{8} = \{\frac{1}{4}|\frac{1}{2}\}$, so that the dyadic rationals 'recreated' on day ω are 'the same' as those created before.

It is also rather nice that Cantor's ordinal numbers (as modified by von Neumann) fit smoothly into our system. Thus we have

$$0 = \{|\},\ 1 = \{0|\},\ 2 = \{0,1|\},\ldots,\ \omega = \{0,1,2,3,\ldots|\},$$

$$\alpha = \{\beta < \alpha|\},$$

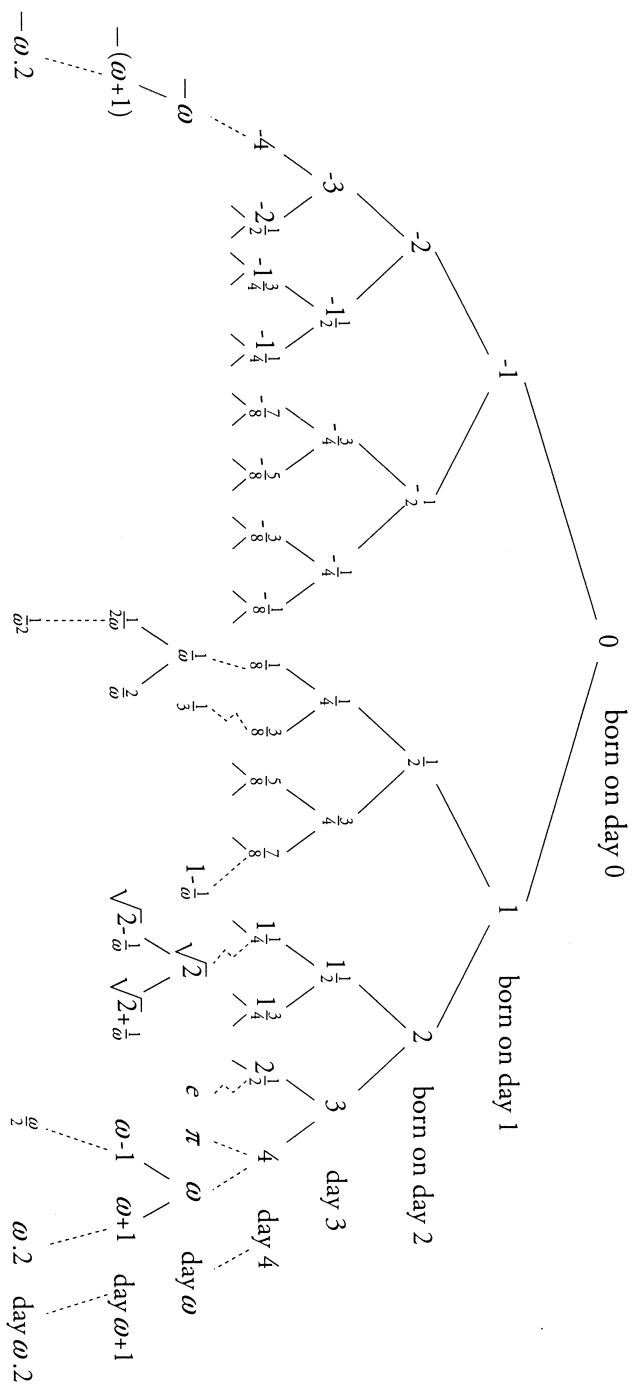

0
born on day 0
-1
1
born on day 1
-2
-1/2
1/2
2
born on day 2
-3
-1 1/2
-1/4
1/4
1 1/2
3
day 3
-4
-2 1/2
-1 3/4
-1 1/4
-7/8
-3/4
-5/8
-3/8
-1/4
-1/8
1/8
3/8
5/8
3/4
7/8
1 1/4
1 3/4
2 1/2
4
day 4
−ω
−(ω+1)
−ω.2
1/2ω
1/ω²
1/ω
2/ω
1/3
1-1/ω
√2-1/ω
√2
√2+1/ω
e
π
ω
day ω
ω/2
ω-1
ω+1
day ω+1
ω.2
day ω.2

Fig. 0

where von Neumann has

$$0 = \{\},\ 1 = \{0\},\ 2 = \{0,1\},\ldots,\ \omega = \{0,1,2,\ldots\},\ \alpha = \{\beta < \alpha\},\ldots .$$

In orther words, the ordinal numbers are those we obtain by requiring always that the set R be empty. We may say that Cantor was only interested in moving ever rightwards, wheras Dedekind stopped to fill in the graps, so that R was *always* empty for Cantor, *never* empty for Dedekind. It is remarkable that by dropping these restrictions we obtain a theory which is both more general and more easy to work with. (Compare the theory developed in the next chapter with the classical foundation for the real numbers, in which we must first construct or postulate the 'natural numbers', then rationals as equivalence classes of ordered pairs, then reals as sections of rationals, with negative numbers being introduced at some stage in the process.)

Some more numbers

After ω, the number $\{0,1,2,3,\ldots,\omega|\} = \omega + 1$ need come as no surprise, but perhaps the number $\{0,1,2,3,\ldots|\omega\}$ will. This number, call it x, should satisfy $n < x < \omega$ for all finite integers n, in other words, x should be an infinite number less than the 'least' infinite number ω. Adding 1 to x, we find the number

$$\{1,2,3,\ldots,x|\omega+1\} = y, \text{ say.}$$

Here, since $x < \omega$, and $\omega + 1 \not\leq \omega$, we see that $y = \omega$, for the new entries x on the left and $\omega + 1$ on the right have made no difference. So $x + 1 = \omega$, $x = \omega - 1$.

Check that we get the same result on subtractic 1 from ω.

In a similar way, we find successively that

$$\omega - 2 = \{0,1,2,3,\ldots|\omega,\omega-1\},\ldots,$$

$$\omega - n = \{0,1,2,3,\ldots|\omega,\omega-1,\omega-2,\ldots,\omega-(n-1)\}.$$

Plainly the text number to consider is

$$z = \{0,1,2,3,\ldots|\omega,\omega-1,\omega-2,\ldots\} = \{n|\omega-n\}, \text{ say.}$$

It should not take the reader too long to verify that $z = \omega/2$. When he has done this, and defined $\omega/4$, $\omega/8,\ldots$ as well, he should be in a position to define $\omega/3$ (for instance), and to verify our assertion that

$$\{0,1,2,3,\ldots|\omega,\omega/2,\omega/4,\omega/8,\ldots\}$$

is a square root of ω.

Other easy exercises are

$$\left\{0\,\middle|\,\frac{1}{\omega}\right\} = \frac{1}{2\omega}\,,\quad \left\{\frac{1}{\omega}\,\middle|\,1,\frac{1}{2},\frac{1}{4},\ldots\right\} = \frac{2}{\omega}\,,\quad \left\{0\,\middle|\,\frac{1}{\omega},\frac{1}{2\omega},\frac{1}{4\omega},\ldots\right\} = \frac{1}{\omega^2}\,,$$

and so on.

If the reader prefers to try his hand at 'constructing' new numbers rather than examining values of those given here, let him try to find definitions for $\sqrt[3]{\omega}, \omega^{1/\omega}, \omega - \pi, (\omega + 1)^{-1}$, $\sqrt{(\omega - 1)}$, and to show, making any reasonable assumptions, that they have the properties we should expect.

In the next chapter, we shall prove that the Class of all numbers really is a Field, making no use of any of the supposed 'facts' from this chapter. It will be some time before we see so many particular numbers mentioned again. In the third chapter, we shall produce a 'canonical form' for numbers, and learn how to manipulate them a little more freely, and in the process will see exactly how general our class of numbers turns out to be.

Chapter 1: The Class No is a Field

Ah! why, ye Gods, should tw and tw make four?
Alexander Pope, 'The Dunciad'

Preliminary comments

There are two problems which arise in the precise treatment which need special comment. The first is that it is necessary to have an expression $\{L|R\}$ *existing* even before we have proved that it is a number. The second concerns the fact that equality is a defined relation, which must initially be distinguished from identity.

Games

The construction for numbers generalises immediately to the following construction for what we call *games*.

Construction

If L and R are any two sets of games, then there is a game $\{L|R\}$. All games are constructed in this way.

Although games are properly the subject of thw first part of this book (where the name will be justified), it is logically necessary to introduce them before numbers. Order-relations and arithmetic operations on games are defined by the same definitions as for numbers. The most important distinction between numbers and general games is that numbers are totally ordered, but games are not–there exist games x and y for which we have neither of $x \geq y, y \geq x$.

To show that a game $x = \{x^L|x^R\}$ is a number, we must show *firstly* that all of the games x^L, x^R are numbers, and *secondly*, that there is no inequality of the form $x^L \geq x^R$.

Identity and equality

We shall call games x and y *identical* ($x \equiv y$) if their left and right sets are identical–that is, if every x^L is identical to some y^L, every x^R identical to some y^R, and vice versa. Recall that x and y are defined to be *equal* ($x = y$) if and only if we have both $x \geq y$ and $y \geq x$. The distinction causes no great problems until we come to multiplication, where the trouble is that there can exist equal games x and y for which xz and yz are unequal. But all goes well as long as we restrict ourselves to the multiplication of numbers.

Finally, we note that almost all our proofs are inductive, so that, for instance, in proving something about the pair (x, y) we can suppose that thing already known about all pairs (x^L, y), (x^R, y), (x, y^L), (x, y^R). After a time we feel free to suppress all references to these inductive hypotheses. We remind the reader

again that since ultimately we are reduced to questions about members of the empty set, no one of our inductions will require a 'basis'. The games x^L, x^R will be called the Left, Right *options* of x.

PROPERTIES OF ORDER AND EQUALITY
Recall that $x > y$ if we havw no inequality of form $x^R \leq y$ or $x \leq y^L$.

THEOREM 0. *For all games x we have*

(i) $x \not\geq x^R$,

(ii) $x^L \not\geq x$,

(iii) $x \geq x$,

(iv) $x = x$.

PROOF.

(i) Taking y as x^R in the definition of $\geq$, and using the inductively true relation $x^R \leq x^R$, we seen that we cannot have $x \geq y$.

(ii) Is similar.

(iii) Taking y as x, we now know that we have no $x^R \leq y$ and $x \leq$ no y^L, whence $x \geq y$.

(iv) From $x \geq x$ and $x \leq x$, we deduce $x = x$. □

THEOREM 1. *If $x \geq y$ and $y \geq z$, then $x \geq z$.*

PROOF. Since $x \geq y$, we cannot have $x^R \leq y$, and so by induction we cannot have $x^R \leq z$. Similarly we cannot have $x \leq z^L$, and so we must have $x \geq z$. □

SUMMARY. We now know that $\geq$ is a partial order relation on games, and that $=$ has the right properties (for instance $x = y$ and $x < z$ imply $y < z$).

THEOREM 2. *For any number x we have $x^L < x < x^R$ for all x^L, x^R. Also, for any two numbers x and y we must have $x \leq y$ or $x \geq y$.* □

PROOF.

(i) Since we know $x \not\geq x^R$, it suffices to prove $x^R \geq x$. This will be true unless some $x^{RR} \leq x$ or $x^R \leq$ some x^L. But the former inductively implies $x^R < x^R \leq x$, a contradiction, and the latter is prohibited by the difinition of number.

(ii) The inequality $x \not\geq y$ implies either some $x^R \leq y$ or $x \leq$ some y^L, whence either $x < x^R \leq y$ or $x \leq y^L, y$. □

SUMMARY. Numbers are totally ordered.

PROPERTIES OF ADDITION

DEFINITION. $0 = \{|\}$.
We recall that $x + y = \{x^L + y, x + y^L | x^R + y, x + y^R\}$.

THEOREM 3. *For all x, y, z we have*

$$x + 0 \equiv x, \quad x + y \equiv y + x, \quad (x + y) + z \equiv x + (y + z).$$

PROOF.

$$\begin{aligned}
x + 0 &\equiv \{x^L + 0 | x^R + 0\} \equiv \{x^L | x^R\} \equiv x \\
x + y &\equiv \{x^L + y, x + y^L | x^R + y, x + y^R\} \equiv \\
&\equiv \{y + x^L, y^L = x | y + x^R, y^R + x\} \equiv y + x. \\
(x + y) + z &\equiv \{(x + y)^L + z, (x + y) + z^L | \ldots\} \equiv \\
&\equiv \{(x^L + y) + z, (x + y^L) + z, (x + y) + z^L | \ldots\} \equiv \\
&\equiv \{x^L + (y + z), z+)y^L + z), x + (y + z^L) | \ldots\} \equiv \\
&\equiv \ldots \equiv x + (y + z).
\end{aligned}$$

□

In each case the middle identity follows from the inductive hypothesis. Proofs like these we call *1-line proofs* even when as here the 'line' is too long for our page. We shall meet still longer 1-line proofs later on, but they do not get harder–one simply transforms the left-hand side through the definitions and inductive hypotheses until one gets the right hand side.

SUMMARY. Addition is a commutative Semigroup operation with 0 as zero, even when we demand identity rather than equality.

PROPERTIES OF NEGATION
Recall the definition $-x = \{-x^R | - x^L\}$.

THEOREM 4.

(i) $-(x + y) \equiv -x + -y$

(ii) $-(-x) \equiv x$

(iii) $x + -x = 0$

PROOF. (i) and (ii) have easy 1-line proofs. Note that (iii) is an equality rather than an identity. If, say, $x + -x \not\geq 0$, we should have some $(x + -x)^R \leq 0$, that is, $x^R + -x \leq 0$ or $x + -x^L \leq 0$. But these are false, since we have by induction $x^R + -x^R \geq 0$, $x^L + -x^L \leq 0$. □

SUMMARY. With equality rather than identity, addition is a commutative Group operation, with 0 for zero, and $-x$ for the negative of x. All this is true for general games.

PROPERTIES OF ADDITION AND ORDER

THEOREM 5. *We have $y \geq z$ iff $x + y \geq x + z$.*

PROOF. If $x + y \geq x + z$, We cannot have

$$x + y^R \leq x + z \text{ or } x + y \leq x + z^L,$$

and so by induction we cannot have $y^R \leq x + z$ or $y \leq z^L$, so that $y \geq z$.

Now supposing $x + y \not\geq x + z$, we must have one of

$$x^R + y \leq x + z, \quad x + y^R \leq x + z, \quad x + y \leq x^L + z, \quad x + y \leq x + z^L,$$

and if we further suppose $y \geq z$, we deduce one of

$$x^R + y \leq x + y, \quad x + y^R \leq x + y, \quad x + z \leq x^L + z, \quad x + z \leq x + z^L,$$

all of which imply contradictions by cancellation. □

Theorem 5 implies in particular that we have $y = z$ iff $x+y = x+z$, justifying replacement by equals in addition.

THEOREM 6.

(i) *0 is a number;*

(ii) *If x is a number, so is $-x$;*

(iii) *If x and y are numbers, so is $x + y$.*

PROOF.

(i) We cannot have $0^L \geq 0^R$, since there exists neither a 0^L nor a 0^R.

(ii) From $x^L < x < x^R$ and x^L, x^R numbers, we inductively deduce $-x^R < -x < -x^L$ and $-x^R, -x^L$ numbers.

(iii) We deduce inductively that each of

$$x^L + y, x + y^L < x + y < \text{ each of } x^R + y, x + y^R,$$

all of $x^L + y$, etc., being numbers. □

SUMMARY. Numbers form a totally ordered Group under addition.

PROPERTIES OF MULTIPLICATION

DEFINITION. $1 = \{0|\}$

We recall the definition of multiplication

$$xy = \{x^L y + xy^L - x^L y^L, \quad x^R y + xy^R - x^R y^R|$$

$$|x^L y + xy^R - x^L y^R, x^R y + xy^L - x^R y^L\}.$$

THEOREM 7. *For all x, y, z we have the identities*

$$x0 \equiv 0,\ x1 \equiv x,\ xy \equiv yx,\ (-x)y \equiv -xy,$$

and the equalities

$$(x+y)z = xz + yz, \quad (xy)z = x(yz).$$

PROOF. The identies have easy 1-line proofs. The equalities also have 1-line proofs, as follows:

$$\begin{aligned}(x+y)z &\equiv \{(x+y)^L z + (x+y)z^L - (x+y)^L z^L, \ldots | \ldots\} \equiv \\ &\equiv \{(x^L + y)z + (x+y)x^L - (x^L + y)z^L, \\ &\quad (x+y^L)z + (x+y)z^L - (x+y^L)z^L - (x+y^L)z^L, \ldots | \ldots\} = \\ &= \{(x^L z + xz^L - x^L z^L)yz,\ xz + (y^L z + yz^L - y^L z^L), \ldots | \ldots\} \\ &\equiv xz + yz. \end{aligned}$$ □

[This fails to yield an identity since the law $x + -x = 0$ is invoked.]
The central expression for xyz has expressions like

$$x^L yz + xy^L z + xyz^L - x^L y^L z - x^L yz^L + z^L y^L z^L$$

(with perhaps some even number of x^L, y^L, z^L replaced by x^R, y^R, z^R) on the left, and four similar expressions (with an odd number of such replacements) on the right.

NOTE. We now have the more illuminating form

$$\begin{aligned}&\{xy - (x - x^L)(y - y^L), \quad xy - (x^R - x)(y^R - y)| \\ &\quad |xy + (x - x^L)(y^R - y), \quad xy + (x^R - x)(y - y^L)\}\end{aligned}$$

for the product xy.

THEOREM 8.

(i) *If x and y are numbers, so is xy*

(ii) *If $x_1 = x_2$, then $x_1 y = x_2 y$*

(iii) *If $x_1 \le x_2$, and $y_1 \le y_2$, then $x_1 y_2 + x_2 y_1 \le x_1 y_1 + x_2 y_2$, the conclusion being strict if both the premises are.*

PROOF. We shall refer to the inequality of (iii) as $P(x_1, x_2 : y_1, y_2)$. Note that if $x_1 \le x_2 \le x_3$, then we can deduce $P(x_1, x_3 : y_1, y_2)$ from the inequalities $P(x_1, x_2 : y_1, y_2)$ and $P(x_2, x_3 : y_1, y_2)$ by adding these and cancelling common terms from the two sides.

Now to prove (i), we observer first that inductively, all options of xy are numbers, so that we have only to prove a number of inequalities like

$$x^{L_1} y + xy^L - x^{L_1} y^L < x^{L_2} y + xy^R - x^{L_2} y^R.$$

But if $x^{L_1} \leq x^{L_2}$ we have

$$x^{L_1}y + xy^L - x^{L_1}y^L + xy^L - x^{L_2}y^L < x^{L_2}y + xy^R = x^{L_2}y^R$$

(these two inequalities reducing respectively to $P(x^{L_1}, x : y^L, y)$ and $P(x^{L_2}, x : y^L, y^R)$), while if $x^{L_2} \leq x^{L_1}$ we have instead

$$x^{L_1}y + xy^L - x^{L_1}y^L < x^{L_1}y + xy^R - x^{L_1}y^R \leq x^{L-2}y + xy^R - x^{L_2}y^R,$$

(these being $P(x^{L_1}, x : y^L, y^R)$ and $P(x^{L_2}, x^{L_1} : y, y^R)$).

Now to prove (ii). This implication follows immediately from the fact that every Left option of either is strictly less than the other, and every Right option strictly greater, the relevant inequalities all being easy.

If $x_1 = x_2$ or $y_1 = y_2$ we can use (ii) to show that the terms on Left of (iii) are equal to those on the Right.

So we need only consider the case $x_1 < X_2$, $y_1 < Y_2$. Since $X_2 < Y_2$, we have either $x_1 < x_1^R \leq x_2$ or $x_1 \leq x_2^L < x_2$, say the former. But then $P(x_1, x_2 : y_1, y_2)$ can be deduced from $P(x_1, x_1^R; y_1, y_2)$ and $P(x_1^R, x_2 : y_1, y_2)$, of which the latter is strictly simpler than the original. A similar argument now reduces our problem to proving strict inequalities of the four forms

$$P(x_L, x : y^L, y),\ P(x^L y : y, y^R),\ P(x, x^R : y^L, y),\ \text{and}\ P(x, x^R : y, y^R)$$

which merely assert that xy has right order relations with its options. □

THEOREM 9. *If x and y are positive numbers, so is xy.*

PROOF This follows from $P(0, x : 0, y)$. □

SUMMARY. Numbers form a totally ordered Ring. Note that in view of Theorem 8 and the distributive law, we can assert, for example, that $x \geq 0$, $y \geq z$ together imply $xy \geq xz$, and that if $x \neq 0$, we can deduce $y = z$ from $xy = xz$.

PROPERTIES OF DIVISION

We have just shown that if there is any number y such that $xy = t$, then y is uniquely determined by x and t provided that $x \neq 0$. We must now show how to produce such a y. It suffices to show that for positive x there is a number y such that $xy = 1$. We first put x into a sort of standard form.

LEMMA. *Each positive x has a form in which 0 is one of the x^L, and every other x^L is positive.*

PROOF. Let y be obtained from x by inserting 0 as a new Left option, deleting all negative Left options. Then it is easy to check that y is a number, and that $y = x$.

We write $x = \{0, x^L | x^R\}$ in this section, and prove that y is a number and that $xy = 1$. □

Now we shall define a number y, explain the definition, and proof that t is a number and that $xy = 1$.

DEFINITION

$$y = \left\{0, \frac{1+(x^R-x)y^L}{x^R}, \frac{1+(x^L-x)y^R}{x^L} \,\middle|\, \frac{1+(x^L-x)y^L}{x^L}, \frac{1+(x^R-x)y^R}{x^R}\right\}$$

Note that expressions involving y^L and y^R appear in the definition of y. It is this that requires us to 'explain' the definition. The explanation is that we regard these parts of the definition as defining new options for y in terms of old ones. So even the definition of this y is an inductive one.† [This is in addition to the 'other' induction by which we suppose that inverses for the x^L and x^R have already been found.]

THEOREM 10. *We have*

(i) $xy^L < 1 < xy^R$ for all y^L, y^R.

(ii) y *is a number.*

(iii) $(xy)^L < 1 < (xy)^R$ *for all* $(xy)^L, (xy)^R$.

(iv) $xy = 1$.

PROOF We observe that the options of y are defined by formulae of the form

$$y'' = \frac{1+(x'-x)y'}{x'}$$

where y' is an 'earlier' option of y, and x' some non-zero option of x. This formula can be written

$$1 - xy'' = (1 - xy')\frac{x'-x}{x'}$$

which shows that y'' satisfies (i) if y' does. Plainly 0 does. Part (ii) now follows, since we cannot have any inequality $y^L \geq y^R$. The typical form of an option of xy is $x'y + xy' - x'y'$, which can be written as $1 + x'(y - y'')$ with the above definition of y'', and this sffices to prove (iii). For (iv), we observe first that $z = xy$ has a left option 0 (take $x^L = y^L = 0$), and that (iii) asserts that $z^L < 1 < z^R$ for all z^L, z^R. Then

$z \geq 1$, since no $z^R \leq 1$, and $z \leq$ no 1^L (since some $z^L = 0$), and also
$1 \geq z$, since no $1^R \leq z$, and $1 \leq$ no z^L,

so that indeed $z = 1$. □

†To see how the definition works, take $x = \{0, 2|\} = 3$. Then there is no x^R and the only x^L is 2, so $x^L - x = -1$ and the formula for y becomes $y = \{0, \frac{1}{2}(1-y^R)|\frac{1}{2}(1-y^L)\}$. The initial value $y^L = 0$ gives us $\frac{1}{2}(1-0) = \frac{1}{2}$ for a new y^R, whence $\frac{1}{2}(1-\frac{1}{2}) = \frac{1}{4}$ as a y^L, then $\frac{1}{2}(1-\frac{1}{4}) = \frac{3}{8}$ for a y^R, and so on, yieling $y = \{0, \frac{1}{4}, \frac{5}{16}, \ldots | \frac{1}{2}, \frac{3}{8}, \ldots\}$, which certainly looks like $\frac{1}{3}$.

SUMMARY. The Class **No** of all numbers forms a totally ordered Field.

Clive Bach has found a similar definition for the square root of a nonnegative number x. He defines

$$\sqrt{x} = y = \left\{ \sqrt{x^L}, \frac{x + y^L y^R}{y^L + y^R} \,\middle|\, \sqrt{x^R} \frac{x +^L y^{L*}}{y^L + y^{L*}}, \frac{x + y^R y^{R*}}{y^R + y^{R*}} \right\}$$

where x^L and x^R are non-negative options of x, and y^L, y^{L*}, y^R, y^{R*} are options of y chosen so that no one of the three denonminators is zero. We shall leave to the reader the easy inductive proof that this is correct.

Chapter 2: The Real and Ordinal Numbers

Don't let us make imaginary evils, when you know we have so many real ones to encounter.

Oliver Goldsmith, 'The Good-Natured Man'

The following theorem gives us a very easy way of evaluating particular numbers. We call it *the simplicity theorem.*

THEOREM 11. *Suppose for* $x = \{x^L|x^R\}$ *that some number* z *satisfies* $x^L \not\geq z \not\geq x^R$ *for all* x^L, x^R, *but that no option of* z *satisfies the same condition. Then* $x = z$.
[Note: this holds even when x is only given to be a game.]

PROOF. We have

$$x \geq z \text{ unless some } x^R \leq z \text{ (no!) or } x \leq \text{ some } z^L.$$

But from $x \leq z^L$, we can deduce $x^L \not\geq x \leq z^L < z \not\geq x^R$ for all x^L, x^R, from which we have $x^L \not\geq z^L \not\geq x^R$ contradicting the supposition about z. So $x \geq z$, similarly $z \geq x$, and so $x = z$. □

The main assertion of the theorem is that when x is given as a number, it is always the *simplest* number lying between the x^L and the x^R, where *simplest* means *earliest created.* [For if z is this simplest number, the simpler numbers z^L, z^R cannot satisfy the same condition.] But the exact version presented above has several advantages, since it holds when x is given as a game not necessarily known to equal a number, and it is perhaps not quite obvious exactly what is meant by 'the simplest number such that...'. In the applications below, there is never any problem.

THEOREM 12. *If* x *is a rational number whose denominator devides* 2^n, *then* $x = \{x - (1/2^n)|x + (1/2^n)\}$.

PROOF. For $n = 0$ the theorem holds, since it asserts that x is the simplest number between $x-1$ and $x+1$, whereas we know that in fact it is, if positive, the simplest number greater than $x - 1$, if negative the simplest number less than $x + 1$, and if zero the simplest number of all. [These statements follow from the usual definition of integers as sums of 1 or -1.]

For $n > 0$, we double $z = \{x-(1/2^n)|x+(1/2^n)\}$ to see that z is the simplest number between these limits, so that $2z = 2x, z = x$. □

Theorem 12 justifies all the assertions of Chapter 0 about numbers born on finite days. Every such number is a *dyadic rational* number, that is, a rational

number of the form $m/2^n$. Of course, we can speak of 'the' rational number p/q without ambiguity, since we have shown that **No** is a totally ordered Field, and therefore contains a uniquely defined image of each rational number, supposed defined in any of the usual ways.

CONTAINMENT OF THE REAL NUMBERS

DEFINITION. x is a *real number* if and only if $-n < x < n$ for some integer n, and

$$x = \{x - 1, x - \frac{1}{2}, x - \frac{1}{3}, \ldots | x + \frac{1}{2}, x + \frac{1}{3}, \ldots\},$$

or in short, $x = \{x - (1/n) | x + (1/n)\}_{n>0}$. [It is to be understood that n ranges over the positive integers.]

THEOREM 13.

(i) *Dyadic rationals are real numbers.*

(ii) $-x, x + y$, *and* xy *are real if* x *and* y *are.*

(iii) *Each real number has a unique expression in the form* $\{L|R\}$, *where* L *and* R *are non-empty sets of rationals,* L *has no greatest,* R *no least, and there is at most one rational in neither* L *nor* R. *Also,* $y' < y \in L$ *implies* $y' \in L$, $z' > z \in R$ *implis* $z' \in R$.

(iv) *Each section* $\{L|R\}$ *as in (ii) equals a unique real number.*

PROOF. (i) follows from Theorem 11 and 12. (ii) follows from the formulae defining the operations (it might be helpful to use the version of the product formula in the note before Theorem 8). As for (iii), for any real number x, let $L =$ the set of rationals less than $x, R =$ the set of rationals greater than x. Then L and R are non-empty by the condition $-n < x < n$ for some n. Also every member of L is less than $x - (1/n)$ for some n, and so we can add $1/2^n$ and still be less than x. This shows that L has no greatest, and similarly R no least member. A rational in neither L nor R must equal x, so at most one is in neither. Since the expression is obviously unique, this proves (iii). As for (iv), note that $\{L|R\}$ is certainly *some* number, x, say, and that easily $-n < x < n$ for some integer n. So we need only show

$$x = \left\{x - \frac{1}{n} \middle| x + \frac{1}{n}\right\}_{n>o}.$$

But since L has no greatest, for any $y \in L$ we have $y + (1/n) \in L$ for all sufficiently large n. This shows that for sufficiently large n there is a member of L greater than $x - (1/n)$ and similarly a member of R less than $x + (1/n)$, which suffices. □

NOTE. We could obviously replace rationals throughout by dyadic rationals in (iii) and (iv). On doing so, we deduce that every real number not a dyadic rational is born on day ω, as asserted in Chapter 0.

SUMMARY. The real numbers as defined here behave exactly like the real numbers defined in any of the more usual ways. So we shall use the name $\mathbb{R}$ for the set of all real numbers.

THE LOGICAL THEORY OF REAL NUMBERS

We have here regarded the ordinary real numbers and their theory as known, so that Theorem 13 merely serves to identify 'our' real numbers with the familiar ones. But of course one could use our ideas to give a new logical foundation for the real numbers. We digress to discuss the usual classical treatments and the advantages and disadvantages of the possible new approach.

Figure 1 shows the lattice of inclusions between the sets $\mathbb{Z}, \mathbb{Q}, \mathbb{R}$ of integers, rationals, and reals, and the corresponding sets $\mathbb{Z}^+, \mathbb{Q}^+ \mathbb{R}^+$ of positive

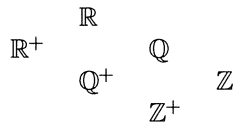


FIGURE 1.

elements. [It does not matter very much whether we add here the number 0 or not.] We shall suppose $\mathbb{Z}^+$ and its properties already known. Then one sees at once that there are several possible paths through the lattice from $\mathbb{Z}^+$ to $\mathbb{R}$. Some experience in teaching convinces one that there is a unique best possible path, which is *not* one that seems natural at first sight!

For $\mathbb{X} = \mathbb{Z}$ or $\mathbb{Q}$ or $\mathbb{R}$ we can proceed from $\mathbb{X}^+$ to $\mathbb{X}$ by introducing ordered pairs (a, b) meaning $a - b$, and the equivalence relation $(a, b) \sim (c, d)$ iff $a + d = b + c$. [The alternative of adding new elements 0 and $-x (x \in \mathbb{X}^+)$ leads to too much case-splitting.]

Similarly we can proceed from $\mathbb{Z}$ to $\mathbb{Q}$ or $\mathbb{Z}^+$ to $\mathbb{Q}^+$ by introducing ordered pairs (a, b) meaning a/b and the equivalence relation $(a, b) \sim (c, b)$ iff $ad = bc$.

We proceed from $\mathbb{Q}$ to $\mathbb{R}$ or $\mathbb{Q}^+$ to $\mathbb{R}^+$ by the method of Dedekind sections, or that of Cauchy sequences.

In practice the main problem is to avoid tedious discussions. [Nobody can seriously pretend that he has ever discussed even eight cases in such a theorem– yet I have seen a presentation in which one theorem actually had 64 cases!] Now if we define $\mathbb{R}$ in terms of Dedekind sections in $\mathbb{Q}$, then there are at least four cases in the definition of the product xy according to the sign of x and y. [And zero often requires special treatment!] This entails eight cases in the associative law $(xy)z = x(yz)$ and strictly more in the distributive law $(x + y)z = xz + yz$ (since we must consider the sign of $x + y$). Of course an elegant treatment will manage to discuss several cases at once, but one has to work very hard to find such a treatment.

This discussion convinces *me* that if one is to use Dedekind sections then the best treatment does not use the branch of our lattice fram $\mathbb{Q}$ to $\mathbb{R}$, and so must be the unique shortest path passing through $\mathbb{R}^+$. This seems surprising, since

the algebraic theory (introduction of negatives and inverses) should naturally be logically prior to the analytic (limits, etc.).

[The reader should be cautioned about difficulties in regarding the construction of the reals as a particular case of the completion of a metric space. If we take this line, we plainly must not start by defining a metric space as one with a real-valued metric! So initially we must allow only rational values for the metric. But then we are faced with the problem that the metric on the completion must be allowed to have arbitrary real values! Of course, the problem here is not actually insoluble, the answer being that the completion of a space whose metric takes values in a field $\mathbb{F}$ is one whose metric takes values in the completion of $\mathbb{F}$. But there are still sufficient problems in making this approach coherent to make one feel that it is simpler to first produce $\mathbb{R}$ from $\mathbb{Q}$, and later repeat the argument when one comes to complete an arbitrary metric space, and of course this distroys the economy of the approach. My own feeling is that in any case the apparatus of Cauchy sequences is logically too complicated for the simple passage from $\mathbb{Q}$ to $\mathbb{R}$–one should surely wait until one has the real numbers before doing a piece of analysis!]

This dicussion should convince the reader that the construction of the real numbers by any of the standard methods is really quite complicated. Of course the main advantage of an approach like that of the present work is that there is just one kind of number, so that one does not spend large amounts of time proving the associative law in several different guises. I think that this makes it the simplest so far, from a purely logical point of view.

Nevertheless there are certain disadvantages. One that can be dealt with quickly is that it is quite tricky to make the process *stop* after constructing the reals! We can cure this by adding to the construction the proviso that if L is non-empty but with no greatest member, than R is non-empty with no least member, and vice versa. This happily restricts us exactly to the reals.

The remaining disadvantages are that the dyadic rationals receive a curiously special treatment, and that the inductive definitions are of an unusual character. From a purely logical point of view these are unimportant quibbles (we discuss the induction problems later in more detail), but they would predispose me against teaching this to undergraduates as 'the' theory of real numbers.

There is another way out. If we adopt a classical approach as far as the rationals $\mathbb{Q}$, and then define the reals as sections of $\mathbb{Q}$ with the definitions of addition and multiplication given in this book, then all the formal laws have 1-line proofs and there is no case-splitting. The definition of multiplication seems complicated, but is fairly easy to motivate. Altogether, this seems the easiest possible approach.

[Perhaps I may add some comments about the multiplication definition. In fact the whole theory was developed even as far as a version of the canonical form theorem of Chapter 3 before any general notion of product appeared, and at first the product was defined in terms of canonical forms. Only several weeks' hard thought, sustained by the conviction that there *must* be a 'genetic' definition, finally led to the 'correct' formula. The genetic definition of $1/x$ at the end of Chapter 1 only appeared a year later.]

CONTAINMENT OF THE ORDINAL NUMBERS

DEFINITION. α is an *ordinal number* if α has an expression of the form $\alpha = \{L|\}$.
[*Note that α is then automatically a number.*]

THEOREM 14. *For any x, the class of all ordinal numbers $\not\geq x$ is a set (i.e. not a proper Class).*

PROOF. Since there is no α^R, the condition $\alpha \not\geq x$ implies $\alpha \leq$ some x^L, and so $\alpha < x^L$ or $\alpha = x^L$. Since the collection of $\alpha <$ any particular x^L is a set by induction, α belongs to a union of a set of sets, and so to a certain set. □

THEOREM 15. *For each ordinal α, we have $\alpha = \{$ordinals $\beta < \alpha|\}$, In any non-empty Class C of ordinals there is a least. For any set S of ordinals there is an ordinal α greater than every member of S.*

PROOF. The first part is immediate from the simplicity theorem and the facy that the collection of $\beta < \alpha$ is a set. For the second part, we observe that the collection L of all β less than all $\alpha \in C$ is a set, for since C is non-empty L is included in the set of all $\beta <$ *some* $\alpha \in C$. Then defining $\delta = \{L|\}$, we find that for all $\alpha \in C$ we have $\alpha \geq \delta$, since there is no α^R, and we never have $\alpha \leq \delta^L$. Then if $\alpha > \delta$ for all $\alpha \in C$, we gat $\delta \in L$, so $\delta < \delta$, a contradiction, and so δ must be equal to some member of C. Finnaly, the ordinal $\{S|\}$ is greater than every member of S. □

SUMMARY. We have proved enough to show that there is a one-to-one order-preserving correspondence between the ordinal numbers as defined here and as defined in any of the more usual ways. So we shall use **On** for the Class of all ordinal numbers.

NOTE. We have regarded the ordinal numbers and their properties as known, so that Theorem 15 merely identifies 'our' ordinal numbers with the familiar ones. Naturally it would be possible to develop the logical theory of ordinals directly from our approach. But the standard set theory of Zermelo and Fraenkel does not seem to be the right vehicle in which to develop such a suggestion, since obviously it should be modified so as to allow two notions of membership (Left and Right) first. There is no logical problem, but we prefer to pospone the discussion till later

The reader should be aware that the operations $\alpha+\beta$ and $\alpha\beta$ as defined here are not the usual *ordinal* operations, but rather the *maximal* sum and product (sometimes called the *natural* sum and product) which can be obtained by treating the Cantor Normal form like a polynomial. [The *maximal* sum $\alpha + \beta$ is the largest order-type of any well-ordered ste $A \cup \beta$ for which A and B have the respective order-types α and β. The *ordinal* sum is the order-type of such a union in which A precedes B. There are similar definitions of the two product notions.]

We consider a generalization of the Cantor Normal form in Chapter 3, and in the first part of the book we shall define an operation $G : H$ (for all games G, H) which will generalise the notion of ordinal sum.

CWI SYLLABI

1 Vacantiecursus 1984: *Hewet - plus wiskunde.* 1984.
2 E.M. de Jager, H.G.J. Pijls (eds.). *Proceedings Seminar 1981-1982. Mathematical structures in field theories.* 1984.
3 W.C.M. Kallenberg, et al. *Testing statistical hypotheses: worked solutions.* 1984.
4 J.G. Verwer (ed.). *Colloquium topics in applied numerical analysis, volume 1.* 1984.
5 J.G. Verwer (ed.). *Colloquium topics in applied numerical analysis, volume 2.* 1984.
6 P.J.M. Bongaarts, J.N. Buur, E.A. de Kerf, R. Martini, H.G.J. Pijls, J.W. de Roever. *Proceedings Seminar 1982-1983. Mathematical structures in field theories.* 1985.
7 Vacantiecurus 1985: *Variatierekening.* 1985.
8 G.M. Tuynman. *Proceedings Seminar 1983-1985. Mathematical structures in field theories, Vol.1 Geometric quantization.* 1985.
9 J. van Leeuwen, J.K. Lenstra (eds.). *Parallel computers and computations.* 1985.
10 Vacantiecursus 1986: *Matrices.* 1986.
11 P.W.H. Lemmens. *Discrete wiskunde: tellen, grafen, spelen en codes.* 1986.
12 J. van de Lune. *An introduction to Tauberian theory: from Tauber to Wiener.* 1986.
13 G.M. Tuynman, M.J. Bergvelt, A.P.E. ten Kroode. *Proceedings Seminar 1983-1985. Mathematical structures in field theories, Vol.2.* 1987.
14 Vacantiecursus 1987: *De personal computer en de wiskunde op school.* 1987.
15 Vacantiecursus 1983: *Complexe getallen.* 1987.
16 P.J.M. Bongaarts, E.A. de Kerf, P.H.M. Kersten. *Proceedings Seminar 1984-1986. Mathematical structures in field theories, Vol.1.* 1988.
17 F. den Hollander, H. Maassen (eds.). *Mark Kac seminar on probability and physics. Syllabus 1985-1987.* 1988.
18 Vacantiecursus 1988. *Differentierekening.* 1988.
19 R. de Bruin, C.G. van der Laan, J.R. Luyten, H.F. Vogt. *Publiceren met LATEX.* 1988.
20 R. van der Horst, R.D. Gill (eds.). *STATAL: statistical procedures in Algol 60, part 1.* 1988.
21 R. van der Horst, R.D. Gill (eds.). *STATAL: statistical procedures in Algol 60, part 2.* 1988.
22 R. van der Horst, R.D. Gill (eds.). *STATAL: statistical procedures in Algol 60, part 3.* 1988.
23 J. van Mill, G.Y. Nieuwland (eds.). *Proceedings van het symposium wiskunde en de computer.* 1989.
24 P.W.H. Lemmens (red.). *Bewijzen in de wiskunde.* 1989.
25 Vacantiecursus 1989: *Wiskunde in de Gouden Eeuw.* 1989.
26 G.G.A. Bäuerle et al. *Proceedings Seminar 1986-1987. Mathematical structures in field theories.* 1990.
27 Vacantiecursus 1990: *Getallentheorie en haar toepassingen.* 1990.
28 Vacantiecursus 1991: *Meetkundige structuren.* 1991.
29 A.G. van Asch, F. van der Blij. *Hoeken en hun Maat.* 1992.
30 M.J. Bergvelt, A.P.E. ten Kroode. *Proceedings seminar 1986-1987. Lectures on Kac-Moody algebras.* 1992.
31 Vacantiecursus 1992: *Systeemtheorie.* 1992.
32 F. den Hollander, H. Maassen (eds.). *Mark Kac seminar on probability and physics. Syllabus 1987-1992.* 1992.
33 P.W.H. Lemmens (ed.). *Meetkunde van kunst tot kunde, vroeger en nu.* 1993.
34 J.H. Kruizinga. *Toegepaste wiskunde op een PC.* 1992.
35 Vacantiecursus 1993: *Het reële getal.* 1993.

MC SYLLABI

1.1 F. Göbel, J. van de Lune. *Leergang besliskunde, deel 1: wiskundige basiskennis.* 1965.
1.2 J. Hemelrijk, J. Kriens. *Leergang besliskunde, deel 2: kansberekening.* 1965.
1.3 J. Hemelrijk, J. kriens. *Leergang besliskunde, deel 3: statistiek* 1966
1.4 G. de Leve, W. Molenaar. *Leergang besliskunde, deel 4: Markovketens en wachttijden.* 1966.
1.5 J. Kriens, G. de Leve. *Leergang besliskunde, deel 5: inleiding tot de mathematische besliskunde.* 1966.
1.6a B. Dorhout, J. Kriens. *Leergang besliskunde, deel 6a: wiskundige programmering 1.* 1968.
1.6b B. Dorhout, J. kriens, J.Th. van Lieshout. *Leergang besliskunde, deel 6b: wiskundige programmering 2.* 1977. 1.7a G. de Leve. *Leergang besliskunde, deel 7a: dynamische programmering 1.* 1968.
1.7b G. de Leve, H.C. Tijms. *Leergang besliskunde, deel 7b: dynamische programmering 2.* 1970.
1.7c G. de Leve, H.C. Tijms. *Leergang besliskunde, deel 7c: dynamische programmering 3.* 1971.
1.8 J. Kriens, F. Göbel, W. Molenaar. *Leergang besliskunde, deel 8: minimaxmethode, netwerkplanning, simulatie.* 1968.
2.1 G.J.R. Förch, P.J. van der Houwen, R.P. van de Riet. *Colloquium stabiliteit van differentieschema's, deel 1.* 1967.
2.2. L. Dekker, T.J. dekker, P.J. van der Houwen, M.N. Spijker. *Colloquium stabiliteit van differentieschema's, deel 2.* 1968.
3.1. H.A. Lauwerier. *Randwaardeproblemen, deel 1.* 1967
3.2. H.A. Lauwerier. *Randwaardeproblemen, deel 2.* 1968
3.3. H.A. Lauwerier. *Randwaardeproblemen, deel 3.* 1968
4 H.A. Lauwerier. *Representaties van groepen.* 1968
5 J.H. van Lint, J.J. Seidel, P.C. Baayen. *Colloquium discrete wiskunde.* 1968.
6. K.K. Koksma. Cursus ALGOL 60. 1969.
7.1 *Colloquium moderne rekenmachine, deel 1.* 1969.
7.2 *Colloquium moderne rekenmachine, deel 2.* 1969.
8 H. Bavinck, J. Grasman. *Relaxatietrillingen.* 1969.
9.1 T.M.T. Coolen, G.J.R. Förch, E.M. de Jager, H.G.J. Pijls. *Colloquium elliptische differentiaalvergelijkingen, deel 1.* 1970
9.2 W.P. van den Brink, T.M.T. Coolen, B. Dijkhuis, P.P.N. de Groen, P.J. van der Houwen, E.M. de Jager, N.M. Temme, R.J. de Vogelaere. *Colloquium elliptische differentiaalvergelijkingen, deel 2.* 1970
10 J. Fabius, W.R. van Zwet. *Grondbegrippen van de waarschijnlijkheidsrekening.* 1970
11 H. Bart, M.A. Kaashoek, H.G.J. Pijls, W.J. de Schipper, J. de Vries. *Colloquium halfalgebra's en positieve operatoren.* 1971.
12 T.J. Dekker. *Numerieke algebra.* 1971
13 F.E.J. Kruseman Aretz. *Programmeren voor rekenautomaten; de MC ALGOL 60 vertaler voor de EL X8.* 1971
14 H. Bavinck, W. Gautschi, G.M. Willems. *Colloquium approximatietheorie.* 1971
15.1 T.J. Dekker, P.W. Hemker, P.J. van der Houwen *Colloquium stijve differentiaalvergelijkingen, deel 1.* 1972.
15.2 P.A. Beentjes, K. Dekker, P.W. Hemker, S.P.N. van Kampen, G.M. Willems. *Colloquium stijve differentiaalvergelijkingen, deel 2.* 1973.
15.3 P.A. Beentjes. K. Dekker, P.W. Hemker, M. van Veldhuizen *Colloquium stijve differentiaalvergelijkingen, deel 3.* 1975.
16.1 L. Geurts. *Cursus programmeren, deel 1: de elementen van het programmeren.* 1973
16.2 L. Geurts. *Cursus programmeren, deel 2: de programmeertaal ALGOL 60.* 1973
17.1 P.S. Stobbe. *Lineaire algebra, deel 1.* 1973.
17.2 P.S. Stobbe. *Lineaire algebra, deel 2.* 1973.
17.3 N.M. Temme. *Lineaire algebra, deel 3.* 1976.
18 F. van der Blij, H. Freudenthal, J.J. de Songh, J.J. Seidel, A. van Wijngaarden. *Een kwart eeuw wiskunde 1946-1971, syllabus van de vakantiecursus 1971.* 1973.
19 A. Hordijk, R. Potharst, J.Th. Runnenburg. *Optimaal stoppen van Markovketens.* 1973.
20 T.M.T. Coolen, P.W. Hemker, P.J. van der Houwen, E. Slagt. *ALGOL 60 procedures voor begin- en randwaardeproblemen.* 1976.
21 J.W. de Bakker (red.). *Colloquium programmacorrectheid.* 1975.
22 R. Helmers, J. Oosterhoff, F.H. Ruymgaart, M.C.A. van Zuylen. *Asymptotische methoden in de toetsingstheorie; toepassingen van naburigheid.* 1976.
23.1 J.W. de Roever (red.). *Colloquium onderwerpen uit de biomathematica, deel 1.* 1976.
23.2 J.W. de Roever (red.). *Colloquium onderwerpen uit de biomathematica, deel 2.* 1977.
24.1 P.J. van der Houwen. *Numerieke integratie van differentiaalvergelijkingen, deel 1: eenstapsmethoden.* 1975.
25 *Colloquium structuur van programmeertalen.* 1976.
26.1 N.M. Temme (ed.). *Nonlinear analysis, volume 1.* 1976.
26.2 N.M. Temme (ed.). *Nonlinear analysis, volume 2.* 1976.
27 M. Bakker, P.W. Hemker, P.J. van der Houwen, S.J. Polak, M. van Veldhuizen. *Colloquium discretiseringsmethoden.* 1976.
28 O. Diekmann, N.M. Temme (eds.). *Nonlinear diffusion problems.* 1976.
29.1 J.C. Bus (red.). *Colloquium numerieke programmatuur, deel 1A, deel 1B.* 1976.
29.2 H.J.J. te Riele (red.). *Colloquium numerieke programmatuur, deel 2.* 1977.
30 J. Heering, P. Klint (red.). *Colloquium programmeeromgevingen.* 1983.
31 J.H. van Lint (red.). *Inleiding in de coderingstheorie.* 1976.
32 L. Geurts (red.). *Colloquium bedrijfssystemen.* 1976.
33 P.J. van der Houwen. *Berekening van waterstanden in zeeën en rivieren.* 1977.
34 J. Hemelrijk. *Oriënterende cursus mathematische statistiek.* 1977.
35 P.J.W. ten Hagen (red.). *Colloquium computer graphics.* 1978.
36 J.M. Aarts, J. de Vries. *Colloquium topologische dynamische systemen.* 1977.
37 J.C. van Vliet (red.). *Colloquium capita datastructuren.* 1978.
38.1 T.H. Koornwinder (ed.). *Representations of locally compact groups with applications, part I.* 1979.
38.2 T.H. Koornwinder (ed.). *Representations of locally compact groups with applications, part II.* 1979.
39 O.J. Vrieze, G.L. Wanrooy. *Colloquium stochastische spelen.* 1978.
40 J. van Tiel. *Convexe analyse.* 1979.
41 H.J.J. te Riele (ed.). *Colloquium numerical treatment of integral equations.* 1979.
42 J.C. van Vliet (red.). *Colloquium capita implementatie van programmeertalen.* 1980.
43 A.M. Cohen, H.A. Wilbrink. *Eindige groepen (een inleidende cursus).* 1980.

44 J.G. Verwer (ed.). *Colloquium numerical solution of partial differential equations*. 1980.
45 P. Klint (red.). *Colloquium hogere programmeertalen en computerarchitectuur*. 1980.
46.1 P.M.G. Apers (red.). *Colloquium databankorganisatie, deel 1*. 1981.
46.2 P.M.G. Apers (red.). *Colloquium databankorganisatie, deel 2*. 1981.
47.1 P.W. Hemker (ed.). *NUMAL, numerical procedures in ALGOL 60: general information and indices*. 1981.
47.2 P.W. Hemker (ed.). *NUMAL, numerical procedures in ALGOL 60, vol. 1: elementary procedures; vol. 2: algebraic evaluations*. 1981.
47.3 P.W. Hemker (ed.). *NUMAL, numerical procedures in ALGOL 60, vol. 3A: linear algebra, part I*. 1981.
47.4 P.W. Hemker (ed.). *NUMAL, numerical procedures in ALGOL 60, vol. 3B: linear algebra, part II*. 1981.
47.5 P.W. Hemker (ed.). *NUMAL, numerical procedures in ALGOL 60, vol. 4: analytical evaluations; vol. 5A: analytical problems, part I*. 1981.
47.6 P.W. Hemker (ed.). *NUMAL, numerical procedures in ALGOL 60, vol. 5B: analytical problems, part II*. 1981.
47.7 P.W. Hemker (ed.). *NUMAL, numerical procedures in ALGOL 60, vol. 6: special functions and constants; vol. 7: interpolation and approximation*. 1981.
48.1 P.M.B. Vitányi, J. van Leeuwen, P. van Emde Boas (red.). *Colloquium complexiteit en algoritmen, deel 1*. 1982.
48.2 P.M.B. Vitányi, J. van Leeuwen, P. van Emde Boas (red.). *Colloquium complexiteit en algoritmen, deel 2*. 1982.
49 T.H. Koornwinder (ed.). *The structure of real semisimple Lie groups*. 1982.
50 H. Nijmeijer. *Inleiding systeemtheorie*. 1982.
51 P.J. Hoogendoorn (red.). *Cursus cryptografie*. 1983

Sprekers

Prof.dr. J.M. Aarts
(TU Delft) Van Kinschotstraat 13; 2614 XJ Delft; 015-126448

Prof.dr. H.P. Barendregt
(KU Nijmegen) Rijksstraatweg 72; 6573 AN Beek-Ubbergen; 08895-43576

Prof.dr. F. van der Blij
(RU Utrecht) Ruysdaellaan 6; 3723 CC Bilthoven; 030-283168

Drs. A.J. Goddijn
(Freudenthal Instituut Utrecht) Tiberdreef 4; 3561 GG Utrecht; 030-611611

Prof.dr. A.W. Grootendorst
(TU Delft) Aardbeistraat 11; 2564 TM Den Haag; 070-3232936.

Dr. W.H. Schikhof
(KU Nijmegen) Weezenhof 36-07; 6536 HC Nijmegen; 080-443085

Prof.dr. A.S. Troelstra
(Univ. van Amsterdam) P. de Hooghlaan 4; 1399 GA Muiderberg; 02942-61964