CWI Syllabi

Centrum voor Wiskunde en Informatica (CWI)

P.O. Box 94079, 1090 GB Amsterdam, The Netherlands
Telephone + 31 - 20 592 9333
Telefax + 31 - 20 592 4199
Website http://www.cwi.nl/publications/

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

Actuele wiskunde

25 en 26 augustus in Eindhoven - 1 en 2 september in Amsterdam

Centrum voor Wiskunde en Informatica
CWI SYLLABUS 56

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

Deze cursus is mede mogelijk gemaakt door een subsidie van de Nederlandse Organisatie voor Wetenschappelijk Onderzoek.

Ontwerp omslag: Tobias Baanders naar een illustratie uit de voordracht van Marjolein Dohmen (pagina 61).

ISBN 90 6196 537 3
NUGI-code: 811

Inhoud

Docenten

Prof.dr. J.M. Aarts
Technische Universiteit Delft, Faculteit Elektrotechniek, Wiskunde en Informatica
Van Kinschotstraat 13, 2614 XJ Delft, tel. 015-2126448
j.m.aarts@ewi.tudelft.nl

Prof.ir. B.A.C. Ambrosius
Technische Universiteit Delft, Faculteit Luchtvaart- en Ruimtevaarttechniek
Kluyverweg 1, 2629 HS Delft, tel. 015-2785173
b.a.c.ambrosius@lr.tudelft.nl

Mw. dr.ir. C.M. Dohmen-Janssen
Universiteit Twente, Construerende Technische Wetenschappen
Postbus 217, 7500 AE Enschede, tel. 053 489 4209
c.m.dohmen-janssen@ctw.utwente.nl

Mw. dr. R.M. Elkenbracht-Huizing
ABN AMRO bank N.V. Amsterdam
marije.elkenbracht@nl.abnamro.com

Mw. G. Geuze
Christelijke Scholengemeenschap Walcheren
Elzenlaan 4, 4334 BW Middelburg, tel. 0118-652110
gilliengeuze@xs4all.nl

Prof.dr.ir. J.D. Jansen
Technische Universiteit Delft, Faculteit Civiele Techniek en Geowetenschappen
Mijnbouwstraat 120, 2628 RX Delft, tel. 015-2787838
j.d.jansen@citg.tudelft.nl

Prof.dr.ir. P.J.M. van Oosterom
Technische Universiteit Delft, Faculteit Techniek, Bestuur en Management
Jaffalaan 5, 2628 BX, Delft, tel. 015-2786950
p.v.oosterom@otb.tudelft.nl

Dr. B. de Smit
Universiteit Leiden, Mathematisch Instituut
Postbus 9512, 2300 RA Leiden, 071-5277144
desmit@math.leidenuniv.nl

Prof.dr.ir. H.C.A. van Tilborg
Technische Universiteit Eindhoven, Faculteit Wiskunde en Informatica
Den Dolech 2, 5612 AZ, Eindhoven, tel. 040-2472739
h.c.a.v.tilborg@tue.nl

Vakantiecursus 2006

pp. 1 – 2

Ten geleide

J.M. Aarts
Technische Universiteit Delft
e-mail: j.m.aarts@ewi.tudelft.nl

De eerste vakantiecursus werd gehouden in Amsterdam onder auspiciën van het Mathematisch Centrum (MC) op 29 en 31 oktober 1946. In het programma van 1 en 2 september 2006 in Amsterdam zal plaats worden ingeruimd om even stil te staan bij de 60^{e} verjaardag van het CWI zelf en van de door hem georganiseerde vakantiecursus.

In de laatste tien jaren is de vakantiecursus telkens georganiseerd rond een toepassingsgebied van de wiskunde. Op die manier konden de meest uiteenlopende toepassingsmogelijkheden van de wiskunde voor het voetlicht gebracht worden. Dat bood de leraren aanknopingspunten met de leerstof die op school aan de orde kwam.

Dit jaar is het thema: *Actuele Wiskunde*. Bij zo goed als alle onderwerpen waarover we horen op het nieuws of lezen in de krant speelt de wiskunde een belangrijke rol. In deze cursus willen we dit nog eens benadrukken en er komen een groot aantal actuele onderwerpen aan de orde.

Op beide dagen komt het ABC-vermoeden aan de orde. Het gaat hier om een op het oog eenvoudig probleem dat je kunt uitleggen aan een ieder die weet wat priemgetallen zijn. Het onderzoek aan het ABC-vermoeden is een hot item. Nadat de laatste stelling van Fermat is bewezen, is het ABC-vermoeden de nieuwe heilige graal van de getaltheorie. Dr. B. de Smit zal ons vertellen over de achtergrond en de diepere betekenis van het vermoeden, en ons op de hoogte stellen van de huidige stand van zake. De volgende dag zal Mevrouw G. Geuze ons laten zien hoe we hiermee op school met de leerlingen aan de slag kunnen. Mevrouw Geuze is een van de trekkers van het Kennislink-project Reken mee met ABC. In dit project zal het ABC-vermoeden voor een groot publiek toegankelijk worden gemaakt en zullen scholen kunnen meewerken aan het uitvoeren van berekeningen betreffende het vermoeden.

Tegenwoordig heeft men tal van constructies en financiële producten bedacht die nauw verwant zijn met geld; dat zijn de zogenaamde derivaten. Mevrouw dr. R.M. Elkenbracht-Huizing zal ons inwijden in de geheimen van het moderne sparen. De centrale onderwerpen zijn daarbij de waardebepaling van het derivaat en het beheer van de risico's.

Hoe kan alle informatie, bijvoorbeeld die betreffende banktransacties, op een betrouwbare wijze worden verstuurd? Voor de beveiliging van gegevens zijn er al verschillende crypto-systemen ontwikkeld. Prof.dr.ir. H.C.A. van Tilborg zal een systeem voor cryptografie bespreken dat is gebaseerd op de optelgroep van

de punten op een elliptische kromme. Dit levert een methode die heel geschikt is voor de beveiliging van onder andere organizers en mobieltjes.

Op 2de Kerstdag 2004 werd Zuidoost-Azië getroffen door een aardbeving gevolgd door een tsunami. Mevrouw dr.ir. C.M. Dohmen-Janssen zal de relevante wiskundige beschrijving van de processen die hierbij een rol spelen bespreken en met behulp hiervan een aantal van de optredende verschijnselen uitleggen.

Het toenemende energieverbruik dwingt ons om zuinig om te gaan met de beschikbare (energie)bronnen. Een belangrijk aspect is tegenwoordig het verhogen van de 'winningsfactor'. Prof.dr.ir. J.D. Jansen zal een uiteenzetting geven van het gecombineerde gebruik van modellen en ondergrondse sensoren en kleppen bij het oliereservoir. De wiskunde is nodig voor het aanpassen van de modellen en het optimaal aansturen van de kleppen. (Het zijn dus niet de olievelden die slim zijn, maar...?)

Prof.ir. B.A.C. Ambrosius zal ons inwijden in het zonnezeilen, een techniek waarbij gebruik wordt gemaakt van de stuwende kracht van fotonen. Die kracht is uiterst gering, maar kan bij langdurige blootstelling en goede eigenschappen van het voertuig leiden tot snelheden die voor traditionele voertuigen onbereikbaar zijn.

De klassieke meetkunde is voortgekomen uit de landmeetkunde. Aan de hand van een voorbeeld zal prof.dr.ir. P.J.M. van Oosterom laten zien dat er nog steeds een nauwe relatie bestaat tussen de wiskunde en de systemen die de geografische informatie verwerken welke is verzameld door satellieten en ruimtecapsules.

Gaarne wil ik op deze plaats allen bedanken die dit jaar opnieuw een Vakantiecursus hebben gemaakt. In de eerste plaats natuurlijk de sprekers. In het bijzonder wil ik de Spinozaprijswinnaar 2005, professor Alexander Schrijver bedanken voor zijn originele bijdrage aan het programma in Amsterdam. Het Centrum voor Wiskunde en Informatica te Amsterdam en de Technische Universiteit Eindhoven stelden zaalruimte beschikbaar. De administratieve en praktische organisatie van de cursus was in handen van Wilmy van Ojik, Minnie Middelberg en dr. Miente Bakker.

Allen hartelijk dank!

Vakantiecursus 2006

pp. 3 – 6

Zestig jaar Vacantiecursus

J. van de Craats
Open Universiteit
Universiteit van Amsterdam
e-mail: craats@science.uva.nl

Dit jaar wordt de CWI-Vacantiecursus voor de zestigste maal georganiseerd. Dat mag niet ongemerkt voorbijgaan, temeer daar ook het CWI dit jaar zijn zestigjarig jubileum viert. Het is opgericht in 1946 als Mathematisch Centrum; in datzelfde jaar vond op 29 en 31 oktober de eerste Vacantiecursus voor wiskundeleraren plaats. Dat was het begin van een lange jaarlijkse traditie die nog steeds voortduurt. Alleen in het jaar 1954 is er geen Vacantiecursus gegeven in verband met het *International Congress of Mathematics* dat toen in Amsterdam georganisserd werd. Daarom is de cursus van dit jaar ook echt de zestigste cursus, en niet de eenenzestigste.

Ter gelegenheid van de vijftigste cursus in 1996 heeft prof.dr. A.W. Grootendorst in de Syllabus van dat jaar een overzichtsartikel [1] geschreven vol historische wetenswaardigheden; ik wil zijn werk hier niet overdoen, maar zal me voornamelijk beperken tot de cursussen van de afgelopen tien jaar. Aan zijn artikel ontleen ik slechts een paar gegevens, zoals het feit dat de cursus tot 1965 uitsluitend in Amsterdam werd gegeven, en daarna bijna steeds ook nog een week eerder of later op de Technische Universiteit Eindhoven. Verder dat de cursus in de beginjaren meestal drie dagen besloeg, maar vanaf 1955 twee dagen. Dat 1981 een absoluut topjaar was met 169 deelnemers in Amsterdam en 126 in Eindhoven (het onderwerp was toen "Oriëntatie op de informatica").

In het algemeen schommelde het totale aantal deelnemers meestal tussen de 120 en 150. Ook thans is dat nog steeds het geval. Vermeldenswaard is verder dat het CWI sinds 1972 een aantrekkelijk uitgevoerde Syllabus bij de cursus uitgeeft waarin de teksten van de voordrachten gebundeld zijn. En zeker niet onvermeld mag blijven dat de cursus, waarvan de praktische organisatie nog steeds bij het CWI berust, plaatsvindt onder auspiciën van de Nederlandse Vereniging van Wiskundeleraren. De voorzitter van de NVvW bekleedt ook traditiegetrouw het voorzitterschap van de Voorbereidingscommissie die jaarlijks het onderwerp kiest en sprekers suggereert. Het daadwerkelijk aanzoeken van sprekers en de verdere inhoudelijke coördinatie is daarna in handen van een commissielid, dat daarmee dus ook telkens weer een persoonlijk stempel op de cursus drukt. Jarenlang is dat coördinatorschap in de bekwame handen geweest van Albert Grootendorst. In 1998 werd ik zijn opvolger en na de cursus van vorig jaar heeft Jan Aarts het van me overgenomen.

Aan het overzichtsartikel [1] van Grootendorst is ook een lijst van onderwerpen uit de periode 1946-1966 toegevoegd. Historisch geïnteresseerden wil

1990	Getaltheorie
1991	Meetkundige structuren
1992	Systeemtheorie
1993	Het reële getal
1994	Computeralgebra
1995	Kegelsneden en kwadratische vormen
1996	Chaos
1997	Rekenen op het toeval
1998	Meetkunde: oud en nieuw
1999	Onbewezen vermoedens
2000	Is wiskunde nog wel mensenwerk?
2001	Experimentele wiskunde
2002	Wiskunde en gezondheid
2003	Wiskunde in het dagelijks leven
2004	Structuur in schoonheid
2005	De schijf van vijf
2006	Actuele wiskunde

Tabel 1. *Titels van de CWI-vacantiecursussen in de periode 1990-2006*

ik daarnaar verwijzen. In Tabel 1 vermeld ik slechts de titels van de meer recente Vacantiecursussen. Uit de titellijsten valt op te maken dat de cursussen de eerste vijftig jaren bijna allemaal gegroepeerd waren rond één duidelijk afgebakend, meestal wiskundig onderwerp. Dat onderwerp werd dan van tal van kanten door voornamelijk universitaire wiskundigen belicht. De laatste tien jaar is er in dit opzicht sprake van een trendbreuk. Nu markeert de titel van de cursus veelal een thema, een soort kapstok waaraan een heel scala van onderwerpen opgehangen wordt. Ook valt te signaleren dat vaak sprekers worden aangezocht die vanuit een wiskundige achtergrond kunnen vertellen over onverwachte technologische, natuurwetenschappelijke, of zelfs medische of forensische toepassingen. Op die manier dragen de vacantiecursussen in belangrijke mate bij aan de bewustwording bij wiskundeleraren van de enorme rijkdom aan toepassingsmogelijkheden en beroepsperspectieven die de moderne wiskunde biedt. Daarnaast geven de lezingen en de syllabusteksten dikwijls aanknopingspunten voor leraren voor lesmateriaal of onderwerpen voor praktische opdrachten of profiewerkstukken van leerlingen. Ook niet onvermeld mag blijven dat veel syllabusteksten een tweede leven, en dus een verdiende, ruimere verspreiding hebben gekregen in de vorm van een artikel in het Nieuw Archief voor Wiskunde (het tijdschrift van het Koninklijk Wiskundig Genootschap), het blad Euclides van de Nederlandse Vereniging van Wiskundeleraren of de Nieuwe Wiskrant. Bij alle aandacht die er de laatste jaren voor toepassingen van de wiskunde in de meest uiteenlopende disciplines is, wordt in de vacantiecursussen ook de schoonheid van de zuivere wiskunde niet vergeten. De cursus van 2004, met als titel "Structuur in schoonheid", is daar een mooi voorbeeld van. En ook dit jaar is er weer een menu samengesteld dat voor een heel breed lerarenpubliek aantrekkelijk is, vol onverwachte toepassingen en prachtige, uit-

dagende zuivere wiskunde. Moge de CWI-Vacantiecursus daarom nog een lang en succesvol leven beschoren zijn!
Verwijzing:

1. A.W. Grootendorst: *Inleiding bij de CWI-Vacantiecursus 1996*, in: *Vacantiecursus 1996 – Chaos*, CWI-Syllabus 41, Amsterdam, 1996, ISBN 90-6196-464-4, pp. 1-7

Appendix. Vacantiecursussen sinds 1946

1946	Wiskunde / Didactiek van de wiskunde
1947	Topologie
1948	Cursus had betrekking op: grondslagen - problemen (i.s.m. Nederlandse Vereniging voor Logica)
1949	De Groepentheorie
1950	De waarschijnlijkheidsrekening, haar grondslagen en haar toepassingen
1951	De wiskunde in haar onderscheidene toepassingen ’
1952	De mechanica
1953	Diverse onderwerpen, die tezamen een groot gebied van de zuivere en toegepaste wiskunde besloegen
1954	Geen vakantiecursus i.v.m. Internationaal Mathematisch Congres
1955	Op verzoek van vereniging WIMECOS gewijd aan het ontwerpleerplan voor wiskunde 1954 bij het M.O.
1956	De wetenschappelijke grondslagen der Elementaire Wiskunde
1957	Historische en methodische aspecten van de Meetkunde
1958	De Algebra met haar Historische en Methodische aspecten
1959	De Vectoren
1960	Het wiskunde-onderwijs in het V.H.M.O. van morgen
1961	De moderne Algebra
1962	Rondom de vernieuwing van het wiskunde-onderwijs bij het V.H.M.O.
1963	Topologie
1964	Toegepaste analyse
1965	Getallentheorie
1966	Zadelpuntsmethoden
1967	Besliskunde
1968	De geschiedenis van de Wiskunde tot omstreeks 1900
1969	Waarschijnlijkheidsrekening en Statistiek
1970	Computer en Onderwijs
1971	De ontwikkeling van de Wiskunde in de afgelopen 25 jaar

zie volgende pagina

1972	Grafentheorie en haar toepassingen
1973	Abstracte informatica
1974	Algebraïsche vergelijkingen
1975	Discrete Wiskunde
1976	Functionaalanalyse
1977	Mathematische Logica (algorithmen en hun beperkingen)
1978	Meetkunde, van kunst tot kunde, vroeger en nu
1979	Nieuwe toepassingsgebieden van de wiskunde (econometrie, sociale wetenschappen, biomathematica en linguistiek)
1980	Vertellingen over tellingen
1981	Oriëntatie op informatica
1982	Wiskunde in het vrije veld; Golfverschijnselen Cryptografie
1983	Complexe getallen
1984	Hewet-plus wiskunde
1985	Variatierekening
1986	Matrices
1987	De personal computer en de Wiskunde op school
1988	Differentie- en differentiaalvergelijkingen
1989	Wiskunde in de Gouden Eeuw
1990	Getaltheorie
1991	Meetkundige Structuren
1992	Systeemtheorie
1993	Het reële getal
1994	Computer-Algebra
1995	Kegelsneden en kwadratische vormen
1996	Chaos
1997	Rekenen op het toeval
1998	Meetkunde oud en nieuw
1999	Onbewezen vermoedens
2000	Is wiskunde nog wel mensenwerk?
2001	Experimentele wiskunde
2002	Wiskunde en gezondheid
2003	Wiskunde in het dagelijks leven
2004	Structuur in schoonheid
2005	De schijf van vijf – meetkunde, algebra, analyse, discrete wiskunde, stochastiek
2006	Actuele wiskunde

Vakantiecursus 2006

pp. 7 – 20

Elliptische krommen in de cryptografie

H.C.A. van Tilborg
Technische Universiteit Eindhoven
Faculteit Wiskunde en Informatica
e-mail: H.C.A.v.Tilborg@tue.nl

1. Inleiding

Machtsverheffen, bijvoorbeeld de berekening van 3^i, kan behoorlijk efficiënt uitgevoerd worden. Je kunt dit nagaan door langzaam de exponent "af te pellen": even exponenten kun je door 2 delen, bij oneven exponenten moet je eerst een factor 3 afzonderen. Bijvoorbeeld voor de berekening van 3^{2185} begint dit proces als volgt:

$$3^{2185} = 3^{2184}3 = \left(3^{1092}\right)^2 3 = \left(\left(3^{546}\right)^2\right)^2 3 = \left(\left(\left(3^{273}\right)^2\right)^2\right)^2 3 = \ldots$$

Als je zo doorgaat kom je uit op

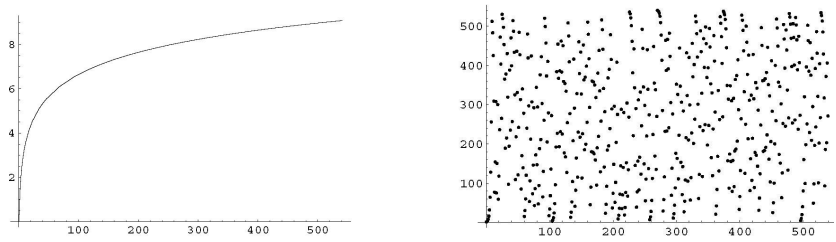
$$3^{2185} = \left(\left(\left(\left(\left(\left(\left(\left(\left(\left((3)^2\right)^2\right)^2\right)^{\mathbf{2}}\right)3\right)^2\right)^2\right)^2\right)^{\mathbf{2}}\right)3\right)^2\Bigg)^2\Bigg)^{\mathbf{2}}\Bigg)3.$$

Dit is meteen in te zien door gebruik te maken van de binaire schrijfwijze 100010001001 van de exponent 2185 (een 0 betekent kwadrateren, een 1 betekent kwadrateren gevolgd door een vermenigvuldiging met 3. Ook *modulair* machtsverheffen, bijvoorbeeld de berekening van $3^{2185} \pmod{3583}$, kan op dezelfde efficiënte manier: na elke kwadratering of vermenigvuldiging met 3 reduceer je antwoord modulo 3583 en dan ga je weer verder. (Het antwoord is 220.)

Om i op te lossen uit $3^i \equiv 2217 \pmod{3583}$, d.w.z. om ${}^3\log 2217$ in $\mathbb{Z}_{3583}$ te bepalen, is geen methode bekend met een vergelijkbare lage complexiteit. Deze constatering ligt ten grondslag aan een modern public key cryptosysteem: het zgn. discrete logaritme systeem (zie [2], maar ook verderop).

De reden waarom het zo gemakkelijk is om een gewone logaritme te bepalen, maar niet een modulaire/discrete logaritme wordt misschien een beetje duidelijk door de twee funktieverlopen in Figuur 1 te vergelijken.

Volledigheidshalve merken we hier op dat het oplossen van $g^{2185} \equiv 2217 \pmod{3583}$ wel simpel is omdat 3583 een priemgetal is. Volgens een bekende



Figuur 1. *De funktie* $^2\log x$ *over* $\mathbb{R}$ *en de discrete logaritme funktie* $^2\log x$ *in* $\mathbb{Z}_{541}$.

stelling van Fermat geldt namelijk dat $g^{3582} \equiv 1 (\mathrm{mod}\ 3583)$. Als we nu beide kanten verheffen tot de macht 1741 (de multiplicatieve inverse van 2185 modulo 3582), dan vinden we de oplossing $g \equiv 2217^{1741} \equiv 68 (\mathrm{mod}\ 3583)$. Als de modulus echter het product van twee onbekende priemgetallen is dan wordt het vinden van de oplossing in zijn algemeenheid voor grote moduli rekentechnisch ondoenlijk, simpelweg omdat het factoriseren van grote getallen ondoenlijk is. Het bekende RSA cryptosysteem maakt hiervan gebruik (zie [10]).

Verderop zullen we het Discrete Logaritme probleem toelichten. In het volgende hoofdstuk zullen we uitleggen hoe dit probleem gebruikt kan worden door twee personen om over een open communicatielijn (radio, telefoon, internet; allemaal met mogelijke afluisteraars) toch tot overeenstemming te komen over een gemeenschappelijke geheime sleutel. Deze methode heet de Diffie-Hellman sleuteluitwisseling. Deze gemeenschappelijke sleutel kan daarna door hen gebruikt worden als geheime sleutel voor een "conventioneel" (symmetrisch) cryptosysteem. We zullen in hetzelfde hoofdstuk ook een idee geven van de verschillende technieken die bekend zijn om discrete logaritmes te nemen. Relevant hierbij is wat de rekencomplexiteit is van deze technieken. Dit is van belang bij de keuze van het lichaam $\mathbb{Z}_p$. Door p voldoende groot te kiezen kun je er voor zorgen dat ook de snelste methodes om discrete logaritmen te nemen nog geen bedreiging vormen voor de veiligheid van de Diffie-Hellman sleuteluitwisseling.

In 1985 realiseerden Victor Miller (IBM, [8]) en Neil Koblitz (University of Washington) zich dat de optelgroep die bij elliptische krommen hoort op een soortgelijke manier gebruikt kan worden om te komen tot een "openbare" sleuteluitwisseling. In hetzelfde jaar richt Scott Vanstone (zie Figuur 2) samen met andere onderzoekers van de University of Waterloo in Ontario, Canada, een bedrijf op dat ze Mobius noemen, maar dat later omgedoopt wordt tot Certicom, met als doel dit idee in een commercieel product om te zetten. Momenteel werken ongeveer driehonderd mensen bij Certicom. Blijkbaar realiseerden de mensen rond Vanstone zich dat de uitvinding van Miller en Koblitz niet zomaar een interessante wiskundige generalisatie inhield van het oorspronkelijke idee van Diffie en Hellman, maar dat sommige aspecten ervan zouden kunnen leiden tot een veelbelovend praktisch (lees: commercieel) alternatief. We zullen hun ideeën preciezer uitleggen in hoofdstuk 3. De essentie van hun inzicht is dat

Een van de oprichters van Certicom, Dr. Vanstone is ook Professor of Mathematics and Computer Science aan de University of Waterloo. Dr. Vanstone wijdt veel van zijn onderzoekstijd aan een efficiënte implementatie van het elliptische kromme cryptosysteem (ECC) met als doel het realiseren van informatiebeveiligingsdiensten in handcomputers, smart cards, draadloze communicatieonderdelen en geïntegreerde circuits.
Dr. Vanstone heeft meer dan 150 wetenschappelijke artikelen en verschillende boeken over onderwerpen als cryptografie, coderingstheorie, eindige lichamen, eindige meetkundes en combinatorische designs op zijn naam staan. Hij is mede-auteur van het Handbook of Applied Cryptography. Hij is Fellow van de Royal Society of Canada, Academy of Sciences.

***Figuur 2**. Dr. Scott A. Vanstone, oprichter van Certicom.*

de enige methode met een subexponentiële complexiteit (zowel in rekentijd als in benodigd geheugen) om een gewone discrete logaritme te nemen niet langer opgaat in de context van elliptische krommen. Dat betekent dat alleen methodes overblijven met een exponentiële complexiteit. Deze eigenschap maakt het mogelijk te werken met veel kleinere priemgetallen, m.a.w. hetzelfde niveau van veiligheid kan bereikt worden met veel handzamere parameters.

2. Het discrete logaritme sleuteluitwisselingssysteem

Twee mensen, die we Alice en Bob noemen willen vertrouwelijke informatie uitwisselen door gebruik te maken van een verbinding die niet beveiligd is. Ze maken hiervoor gebruik van een zgn. *symmetrisch* cryptosysteem, ook wel "secret key" systeem genoemd. In zo'n systeem moeten ze alle twee beschikken over dezelfde geheime *sleutel*. In de Enigma, door de Duitsers gebruikt in WW-II, bestond de sleutel uit de keuze en beginposities van drie rotoren [5]. In moderne secret key systemen is de sleutel een rijtje bestaande uit 128, 192, 256 of zelfs meer bits. Het probleem is dat Alice en Bob elkaar nooit eerder ontmoet hebben om een gemeenschappelijke sleutel af te spreken. Ze gaan hiervoor het Diffie-Hellman protocol voor een sleuteluitwisseling gebruiken. We merken nogmaals op dat Alice en Bob alleen maar informatie kunnen uitwisselen over een publiek kanaal, d.w.z. dat afluisteraars in principe alles kunnen afluisteren.

Alice en Bob kiezen een heel groot priemgetal p en een element g in $\mathbb{Z}_p$ dat een voldoende grote vermenigvuldigingsorde heeft, denk aan 128 bits lang. We duiden deze orde met q aan; dus q is de kleinste positieve exponent van g die 1 oplevert modulo p. In het vervolg zullen we aannemen dat deze orde zelf een priemgetal is (merk op dat q een deler is van $p-1$). Alice kan de parameters p en g vrijelijk kiezen en aan Bob vertellen, maar deze parameters mogen ook tot de standaard van het communicatiesysteem behoren. Hoe dan ook, we zullen aannemen dat ook de afluisteraar ze kent.

Vervolgens kiezen Alice en Bob allebei een willekeurige exponent kleiner dan q, zeg s_A en s_B. Ze berekenen vervolgens $k_A = g^{s_A}$, resp. $k_B = g^{s_B}$ in

$\mathbb{Z}_p$ en wisselen vervolgens deze waarden uit over het publieke communicatie kanaal. Wel houden zij hun exponent s_A resp. s_B geheim.

De gemeenschappelijke, geheime waarde (sleutel) van Alice en Bob is nu

$$g^{s_A s_B} \text{ in } \mathbb{Z}_p.$$

Alice vindt die uit de berekening $(k_B)^{s_A}$ en Bob uit $(k_A)^{s_B}$.

VOORBEELD 1 *Neem* $p = 93179$. *Dan heeft* $g = 3$ *multiplicatieve orde* $q = 46589$, *zoals men met niet te veel moeite kan nagaan. Stel Alice kiest de geheime sleutel* $s_A = 38009$ *en Bob kiest* $s_B = 62952$. *Alice berekent haar openbare sleutel* $k_A \equiv 3^{38009} \equiv 84555 \pmod{93179}$ *en Bob berekent zijn openbare sleutel* $k_B \equiv 3^{62952} \equiv 51910 \pmod{93179}$.

Dan is de gemeenschappelijke sleutel door Alice uit te rekenen met $(k_B)^{s_A} \equiv 51910^{38009}$ *en door Bob met* $(k_A)^{s_B} \equiv 84555^{62952}$. *Beiden vinden* 17905 *waarvan de binaire representatie* $1, 0, 0, 0, 1, 0, 1, 1, 1, 1, 1, 0, 0, 0, 1$ *gebruikt kan worden als sleutel van een symmetrisch systeem.*

Er is een praktische complicatie met dit systeem. Hoe kan Alice er zeker van zijn dat s_B inderdaad Bob's openbare waarde is? Er zijn hierop verschillende antwoorden. Je zou aan een soort openbaar boek kunnen denken (zoals een telefoongids) waarin van iedereen de correcte openbare waarde staat, maar praktisch is deze oplossing niet. Het is handiger dat de een of andere betrouwbare instantie elke openbare waarde van een "handtekening" voorziet.

Het moge duidelijk zijn dat het bovenstaande sleuteluitwisselingsysteem gebroken is zo gauw een afluisteraar in staat is om uit de openbare waarde k_A de geheime waarde s_A uit te rekenen (of s_B uit k_B). Dit dwingt ons om wat langer stil te staan bij de verschillende manieren om

$$g^s \equiv k \pmod{p}, \tag{1}$$

op te lossen, waarbij g, k, and p bekend zijn en s berekend moet worden. Hoe (1) op te lossen staat bekend als het *discrete logaritme probleem.* Om meer te weten te komen over de technieken die we verderop beschrijven, verwijzen we de lezer naar [7], dat een uitstekend handboek is, of naar [13], een handboek met interactieve cd-rom.

EXHAUSTIVE KEY SEARCH

Je kunt s met *brute kracht* vinden door $s = 0, 1, 2, \ldots$ in te vullen totdat er de oplossing gevonden is. Als alternatief kun je eerst de waarden $g^0, g^1, g^2, \ldots$ in een tabel opslaan als voorberekening en dan de plaats van k opzoeken zodra deze gegeven is. Hoe dan ook, de complexiteit van deze methode is p. Als p een getal van t bits is, dan kunnen we ook zeggen dat de complexiteit gegeven wordt door 2^t, m.a.w. de complexiteit groeit exponentieel in t.

BABY-STEP GIANT-STEP

Een aardige manier om bovenstaande rekentijd en benodigde geheugenruimte beter in balans te brengen is de zgn. *baby-step giant-step* methode. Neem aan

9	(4,6)	(5,5)	(1,7)	(7,5)	(0,0)	(5,3)	(1,8)	(3,5)	(8,4)	(7,9)
8	(1,8)	(2,2)	(6,5)	(9,9)	(9,6)	(2,0)	(4,4)	(6,1)	(2,3)	(9,1)
7	(7,9)	(1,0)	(2,4)	(2,3)	(4,8)	(0,3)	(3,7)	(8,2)	(4,2)	(6,5)
6	(9,9)	(0,6)	(1,5)	(8,5)	(7,3)	(8,4)	(9,5)	(4,3)	(3,8)	(1,3)
5	(4,2)	(9,9)	(7,8)	(4,8)	(3,1)	(5,1)	(9,7)	(1,9)	(3,7)	(6,9)
4	(1,2)	(0,4)	(7,6)	(6,4)	(0,5)	(1,6)	(8,7)	(5,7)	(1,7)	(5,8)
3	(0,2)	(5,0)	(0,9)	(4,9)	(2,6)	(3,3)	(7,3)	(5,0)	(8,7)	(6,9)
2	(3,9)	(1,8)	(5,5)	(4,7)	(0,1)	(4,7)	(0,1)	(4,8)	(3,7)	(0,6)
1	(1,1)	(2,2)	(7,9)	(8,5)	(5,1)	(4,6)	(8,7)	(5,2)	(1,6)	(5,3)
0	(3,1)	(3,2)	(0,6)	(4,3)	(0,2)	(1,7)	(5,4)	(7,9)	(0,1)	(9,5)
	0	1	2	3	4	5	6	7	8	9

Figuur 3. *Een random afbeelding f van $\{0,1,\ldots,9\}^2$ naar zichzelf.*

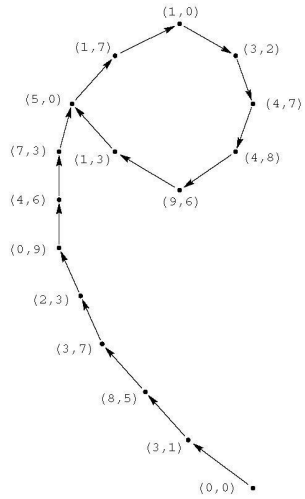
dat er voldoende geheugen beschikbaar is om m waarden in $\mathbb{Z}_p$ op te slaan. Bereken dan de eerste m machten van g, dus $g^0, g^1, g^2, \ldots, g^{m-1}$, sorteer ze om ze gemakkelijker te kunnen opzoeken en plaats ze in een tabel. (Merk op dat de exponenten met 1 omhoog gaan: de "baby-steps"). Vervolgens kijk je of k in de tabel staat, zo niet, kijk dan of k/g^m in de tabel staat, zo niet kijk voor k/g^{2m}, etcetera (de "giant-steps"). Zo gauw je de waarde k/g^{jm} in de tabel tegenkomt, zeg als g^i met $0 \leq i \leq m-1$, kun je de onbekende exponent uitrekenen uit $s - jm = i$ (dus $s = jm + i$). Merk op dat de grootte van de reuzenstappen precies de lengte van de tabel is. Je kunt er dus niet in een keer overheen stappen. De rekencomplexiteit van deze methode is p/m. Het product van rekentijd en benodigde geheugenruimte is dus nog steeds $p \approx 2^t$, maar de onderlinge balans kun je nu zelf kiezen.

Pollard's methodes

Twee verwante technieken om de oplossing s van (1) te bepalen zijn de Pollard-ρ en de Pollard-λ methodes [9]. Wij zullen alleen de eerste methode uitleggen. We nemen nog steeds aan dat g een (onder)groep genereert van de grootte q met q priem. We beginnen met de volgende observatie over random functies.

Laat f een willekeurige (random) afbeelding zijn van een eindige verzameling A naar zichzelf. Pak een willekeurige startwaarde a_0 in A en definieer de rij $\{a_i\}_{i\geq 0}$ recursief door $a_{i+1} = f(a_i)$. De rij $\{a_i\}_{i\geq 0}$ zal op den duur in een loop geraken, simpelweg omdat A eindig is. Zo gauw $a_i = a_j$ geldt ook $a_{i+1} = a_{j+1}$ etc. en zit je in een loop. De te verwachten lengte van deze loop en het te verwachten aantal stappen voordat je aan de loop begint zijn beide gegeven door $\sqrt{\pi|A|/8}$ (zie [3]). De situatie is grafisch weergegeven door de letter ρ in Figuur 4, waarbij we $A = \{0,1,\ldots,9\} \times \{0,1,\ldots,9\}$ hebben genomen en waarbij $f(i,j)$ is gegeven door de waarde op plaats (i,j) in de rechthoek in Figuur 3.

Terugkerend naar het probleem van het vinden van een oplossing van (1) is het idee om een geschikte, recurrente betrekking op $\mathbb{Z}_p$ te definiëren. Zo



Figuur 4. *De wandeling met een ρ figuur in $\{0,1,\ldots,9\}^2$ die start in $(0,0)$.*

gauw je twee keer dezelfde waarde tegenkomt is er een goede kans dat je het probleem opgelost hebt. Het blijkt handig om ook nog twee hulpgrootheden bij te houden. We definiëren de afbeelding $F : \mathbb{Z}_p \times \mathbb{Z}_q \times \mathbb{Z}_q \to \mathbb{Z}_p \times \mathbb{Z}_q \times \mathbb{Z}_q$ door

$$F(x,u,v) = \begin{cases} (x^2, 2u, 2v), & \text{als } x \equiv 0 \pmod 3, \\ (kx, u, v+1), & \text{als } x \equiv 1 \pmod 3, \\ (gx, u+1, v), & \text{als } x \equiv 2 \pmod 3. \end{cases}$$

De rij $\{(x_i, u_i, v_i)\}_{i\geq 0}$ bestaande uit drietallen is recursief gedefinieerd door $(x_0, u_0, v_0) = (1, 0, 0)$ and $(x_{i+1}, u_{i+1}, v_{i+1}) = F(x_i, u_i, v_i)$. Met inductie is gemakkelijk na te gaan dat $x_i = g^{u_i} k^{v_i}$ voor alle $i \geq 0$. Bijvoorbeeld, als $x \equiv 0 \pmod 3$, dan geldt dat $g^{u_{i+1}} k^{v_{i+1}} = g^{2u_i} k^{2v_i} = (g^{u_i} k^{v_i})^2 = (x_i)^2 = x_{i+1}$.

We moeten nu indices $0 \leq i < j$ vinden met $x_i = x_j$ maar we willen niet alle tussenliggende waarden $x_0, x_1, \ldots$ in het geheugen moeten opbergen. Daartoe gaan we alleen de eerste coördinaten van de twee deelrijen (x_i, u_i, v_i) en (x_{2i}, u_{2i}, v_{2i}) for $i > 0$ vergelijken. Als $x_i \neq x_{2i}$, berekenen we simpelweg $(x_{i+1}, u_{i+1}, v_{i+1}) = F(x_i, u_i, v_i)$ en $(x_{2i+2}, u_{2i+2}, v_{2i+2}) = F(F(x_{2i}, u_{2i}, v_{2i}))$ en vergelijken dan de eerste coördinaten opnieuw.

Als $x_i = x_{2i},\ i > 0$, dan geldt $g^{u_i} k^{v_i} = g^{u_{2i}} k^{v_{2i}}$, en dus $g^{u_i} g^{s.v_i} = g^{u_{2i}} g^{s.v_{2i}}$. Hiermee kan bijna altijd de onbekende exponent s uitgerekend worden: $s = (u_{2i} - u_i)/(v_i - v_{2i}) \pmod{p-1}$. (In het geval dat $\gcd(v_i - v_{2i}, p-1) \neq 1$, kan men proberen $kg^l \equiv g^{s+l} \pmod p$ op te lossen met Pollard's methode voor $l = 1, 2, \ldots$ totdat een oplossing gevonden is.)

Vanwege de ρ vorm van de $\{x_i\}_{i\geq 0}$-wandeling weten we dat de verwachte rekentijd van deze methode ongeveer $\sqrt{\pi p/2} \approx 2^{t/2}$ is, terwijl er maar een een paar geheugenplaatsen nodig zijn.

Als voorbeeld gaan we $121^s \equiv 3435 \pmod{4679}$ oplossen. We merken op dat 121 multiplicatieve orde $q = 2339$ heeft. Beginnend met $(x_0, u_0, v_0) = (1, 0, 0)$ vinden we dat $x_{76} = x_{152}$ (en $a_{76} = 84, b_{76} = 2191, a_{152} = 286, b_{152} = 915$). Hiermee vinden we dan $s \equiv (286 - 84)/(2191 - 915) \equiv 1111 \pmod{2339}$.

INDEX-CALCULUS

We volstaan met een typisch voorbeeld van deze methode. Stel dat we

$$11^k \equiv 3333 \pmod{4679}.$$

willen oplossen. We beginnen met een voorberekening die onafhankelijk is van het rechterlid 3333. We nemen de verzameling $\{2, 3, 5, 7, 11\}$, zijnde de lijst van de eerste vijf priemgetallen. Deze verzameling wordt *factor base* genoemd. We gaan eerst het logaritme probleem oplossen voor alle (!) elementen in de factor base. Met andere woorden, we gaan oplossen

$$\begin{aligned} 11^{k_1} &\equiv 2 \pmod{4679}, \\ 11^{k_2} &\equiv 3 \pmod{4679}, \\ 11^{k_3} &\equiv 5 \pmod{4679}, \\ 11^{k_4} &\equiv 7 \pmod{4679}, \\ 11^{k_5} &\equiv 11 \pmod{4679}. \end{aligned}$$

Het lijkt wel of we het probleem erger gemaakt hebben in plaats van beter, maar dat blijkt niet zo te zijn. Kies een willekeurige exponent r, bereken $11^r \pmod{4679}$, en kijk of het antwoord volledig gefactoriseerd kan worden met de getallen in de factor base. Bij voorbeeld, met $r = 12208$ krijgen we $11^{2208} \equiv 4182 \pmod{4679}$ maar $4182 = 2^1 \times 3^1 \times 697$ en 697 kan niet verder ontbonden worden over $\{2, 3, 5, 7, 11\}$. Merk op dat we het rechterlid niet hoeven te factoriseren; we hoeven alleen maar de elementen in de factor base eruit te delen en dan te kijken of er nog iets overblijft of niet. Het moge duidelijk zijn dat hoe groter de factor base is hoe groter de kans is dat het rechterlid volledig ontbonden kan worden in factoren uit de factor base. De prijs die hier echter voor betaald wordt is het groter aantal k_i's dat bepaald moet worden.

Zo gauw het rechterlid volledig ontbonden kan worden over de factor base, krijgen we een lineair verband tussen de onbekende k_i's. Bijvoorbeeld

$$11^{1006} \equiv 315 = 3^2.5.7 \pmod{4679}$$

geeft de relatie

$$1006 \equiv 2m_2 + m_3 + m_4 \pmod{4678}.$$

Zo gauw je genoeg van dit soort lineaire betrekkingen hebt verzameld kun je de onbekende k_i's zo uitrekenen. Bijvoorbeeld vind je uit

$$\begin{aligned} 11^{104} &\equiv 1280 = 2^8.5 \pmod{4679}, \\ 11^{208} &\equiv 750 = 2.3.5^3 \pmod{4679}, \\ 11^{1006} &\equiv 315 = 3^2.5.7 \pmod{4679}, \\ 11^{2303} &\equiv 198 = 2.3^2.11 \pmod{4679}, \\ 11^{3506} &\equiv 4050 = 2.3^4.5^2 \pmod{4679}, \end{aligned}$$

	time	memory
Exhaustive key search	2^t	1
Baby-step Giant-step	$2^t/m$	m
Pollard	$2^{t/2}$	1
Index calculus	$e^{1.923\ t^{1/3}(\ln t)^{2/3}}$	$e^{1.923\ t^{1/3}(\ln t)^{2/3}/2}$

Figuur 5. *De complexiteit van de verschillende methodes om discrete logaritmen te nemen voor* $p \approx 2^t$.

de lineaire betrekkingen

$$\begin{array}{lll} 104 & \equiv 8k_1 + k_3, & (\text{mod } 4678), \\ 208 & \equiv k_1 + k_2 + 3k_3 & (\text{mod } 4678), \\ 1006 & \equiv 2k_2 + k_3 + k_4 & (\text{mod } 4678), \\ 2303 & \equiv k_1 + 2k_2 + k_5 & (\text{mod } 4678), \\ 3506 & \equiv k_1 + 4k_2 + 2k_3 & (\text{mod } 4678). \end{array}$$

De oplossing hiervan is $k_1 \equiv 352$, $k_2 \equiv 3314$, $k_3 \equiv 1966$, $k_4 \equiv 1768$, en $k_5 \equiv 1$, allemaal modulo 4678.

We zijn nu klaar om het oorspronkelijke probleem op te lossen: $11^k \equiv 3333$ (mod 4679). Kies een willekeurige exponent r en kijk of $3333 \times 11^r (\text{mod } 4679)$ volledig ontbonden kan worden over de getallen in de factor base. Na een paar pogingen vinden we

$$3333 \times 11^{573} \equiv 540 = 2^2.3^3.5 \ (\text{mod } 4679).$$

Hieruit halen we

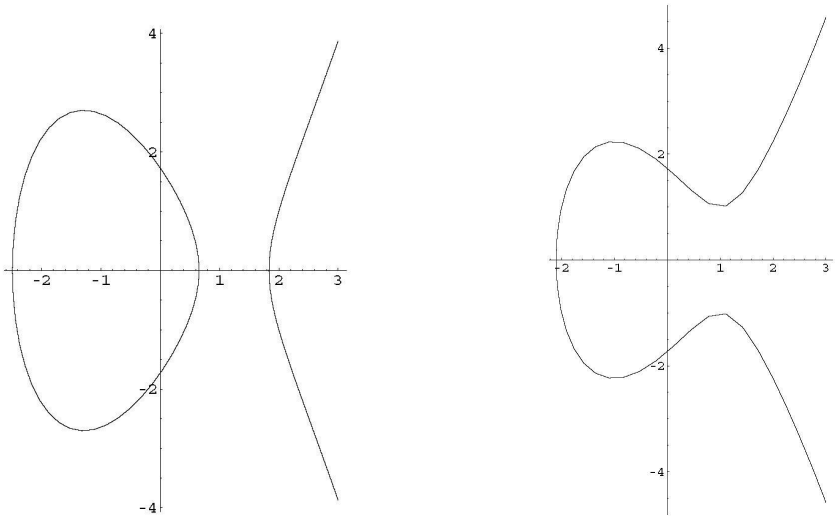
$$k + 573 \equiv 2k_1 + 3k_2 + k_3 \ (\text{mod } 4678).$$

Aangezien de k_i's bekend zijn, is het nu gemakkelijk om k te bepalen. Men vindt zo de oplossing $k \equiv 2\times 352 + 3\times 3314 + 1\times 1966 - 573 \equiv 2683 \,(\text{mod } 4678)$. En inderdaad, $11^{2683} \equiv 3333 \,(\text{mod } 4679)$.

De benodigde rekentijd van deze methode is $e^{1.923\ t^{1/3}(\ln t)^{2/3}}$ (see [4]). Aan geheugenruimte is ongeveer de wortel van dit getal nodig. Hiermee is de index calculus methode (met zijn variaties) de enige manier om de discrete logaritme te bepalen met een complexiteit die subexponentieel groeit (zie ook Figuur 5).

3. Elliptische krommen cryptosystemen

Alhoewel het heel natuurlijk is om de vermenigvuldigingsgroep van een eindig lichaam te gebruiken om de sleuteluitwisselingsmethode van Diffie en Hellman op te zetten, kun je misschien net zo goed een andere voldoende grote, cyclische groep gebruiken. Victor Miller en Neil Koblitz stelden in 1985 een variatie voor die gebruik maakt van de optelgroep die bij een elliptische kromme (EC) hoort. In deze uitleg beperken we ons tot krommen over $\mathbb{Z}_p$ met p priem.



Figuur 6. *De elliptische krommen* $y^2 = x^3 - 5x + 3$ *en* $y^2 = x^3 - 3x + 3$ *over* $\mathbb{R}$.

DEFINITIE 2 *Neem* $a, b, c \in \mathbb{Z}_p$. *De elliptische kromme* $\mathcal{E}$ *over* $\mathbb{Z}_p$ *bestaat uit de punten* (x, y) *die voldoen aan*

$$y^2 = x^3 + ax^2 + bx + c, \tag{2}$$

tezamen met een zgn. punt in het oneindige dat we met $\mathcal{O}$ *aanduiden.*

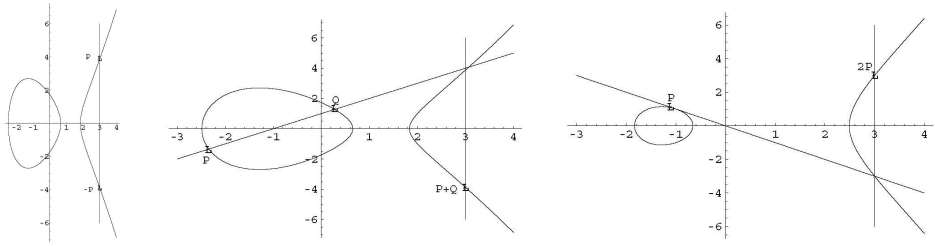
Om enig gevoel te krijgen zijn twee elliptische krommen over $\mathbb{R}$ weergegeven in Figuur 6. Het punt $\mathcal{O}$ kan het beste gevisualiseerd worden als het snijpunt in het oneindige van alle verticale lijnen. Als het rechterlid van (2) positief is voor een reëele waarde van x dan zijn er twee punten op $\mathcal{E}$ met die x coördinaat, eentje met een positieve y-coördinaat en het spiegelbeeld ervan (dus $-y$). Als het rechterlid nul is, is er maar één oplossing namelijk $y = 0$ en als het negatief is, zijn er geen oplossingen.

Over $\mathbb{Z}_p$ is de situatie precies hetzelfde: twee oplossingen als $x^3+ax+bx+c$ het kwadraat is van een niet-nul element in $\mathbb{Z}_p$, één oplossing als het rechterlid nul is en geen oplossing in het overgebleven geval. Er bestaat geen formule voor het precieze aantal punten op een EC-kromme over $\mathbb{Z}_p$. In [12] kan men de volgende schatting vinden:

STELLING 3 (HASSE) *Het aantal punten* $\mathcal{N}$ *op een elliptische kromme over* $\mathbb{Z}_p$ *voldoet aan:*

$$|N - (p + 1)| \leq 2\sqrt{p}.$$

Er bestaan wel algoritmen om het aantal punten op een elliptische kromme te bepalen (zie [11]) maar die hebben maar een beperkte snelheid en verder onderzoek zal zeker nodig blijven.



Figuur 7. *Optelling over elliptische krommen.*

Een nuttige eigenschap van elliptische krommen is dat elke lijn door twee punten op de kromme $\mathcal{E}$ deze ook nog zal snijden in een derde punt (voor een verticale lijn speelt het punt $\mathcal{O}$ die rol). Hetzelfde geldt voor een (eventueel dubbele) raaklijn: hij snijdt de kromme in een derde punt. De reden hiervoor is simpel. Laat $y = mx + n$ de vergelijking zijn van de lijn l door (x_1, y_1) en (x_2, y_2), beide op $\mathcal{E}$ gelegen. Invulling van $y = mx + n$ in (2) geeft de derdegraads vergelijking $(mx + n)^2 = x^3 + ax^2 + bx + c$. Aangezien x_1 en x_2 twee oplossingen zijn, kun je $x^3 + ax^2 + bx + c - (mx + n)^2$ delen door $(x - x_1)$ en $(x - x_2)$. Er blijft dan een factor $(x - x_3)$ over. Het punt (x_3, y_3), met $y_3 = mx_3 + n$, is het derde snijpunt van l met $\mathcal{E}$. Gevallen van een raaklijn kunnen op een soortgelijke manier afgehandeld worden.

We kunnen nu een groepsoperatie definiëren.

DEFINITIE 4 *Optelling op $\mathcal{E}$ volgt de volgende regels:*

1. $\mathcal{O} + \mathcal{P} = \mathcal{P} + \mathcal{O} = \mathcal{P}$,
2. *als* $\mathcal{P} = (x, y)$ *dan* $-\mathcal{P} = (x, -y)$,
3. $\mathcal{P} + \mathcal{Q} = -\mathcal{R}$, *waarbij* $\mathcal{R}$ *het derde snijpunt is van de lijn door* $\mathcal{P}$ *and* $\mathcal{Q}$ *met* $\mathcal{E}$.

In Figuur 7 zijn typische situaties weergegeven over $\mathbb{R}$. De lezer kan uit de afleiding hierboven opmaken dat optellingen heel gemakkelijk uit te voeren zijn. In het algemene geval is $-(x_1 + x_2 + x_3)$ gelijk aan de coefficiënt van x^2 in de derde-graads vergelijking hierboven.

VOORBEELD 5 *Beschouw de elliptische kromme $\mathcal{E}$ gegeven door $y^2 = x^3 + 100x^2 + 10x + 1$ over $\mathbb{Z}_{863}$ en de punten $\mathcal{P} = (121, 517)$ en $\mathcal{Q} = (211, 667)$ op $\mathcal{E}$. De lijn door $\mathcal{P}$ en $\mathcal{Q}$ heeft vergelijking $y = 577x + 603$ over $\mathbb{Z}_{863}$, zoals gemakkelijk na te gaan is.*

Oplossen van $-(121 + 211 + x_3) \equiv -(x_1 + x_2 + x_3) \equiv a - m^2 \equiv 100 - 577^2 \pmod{863}$ *geeft* $x_3 = 242$. *Uit* $y_3 = 577x_3 + 603$ *verkrijgen we* $y_3 = 431$. *Uit de definitie van optelling volgt dat* $\mathcal{P} + \mathcal{Q} = (242, -431) = (242, 432)$.

Aangezien optellen over $\mathcal{E}$ gemakkelijk gaat, is het ook simpel om scalaire veelvouden van een punt, zeg $s.\mathcal{P}$ te berekenen. Net zoals bij machtsverheffen

systeem parameters	elliptische kromme $\mathcal{E}$ over $\mathbb{Z}_p$ punt $\mathcal{P}$ op $\mathcal{E}$ met hoge orde v
geheime sleutel van gebruiker U openbare sleutel van gebruiker U	$s_U, 0 < s_U < v$, het punt $\mathcal{K}_U = s_U\mathcal{P}$
gemeenschappelijke sleutel van Alice en Bob	het punt $\mathcal{S}_{A,B} = s_A s_B \mathcal{P}$
Alice berekent Bob berekent	$\mathcal{S}_{A,B} = s_A\mathcal{K}_B$ $\mathcal{S}_{A,B} = s_B\mathcal{K}_A$

***Figuur 8**. Het Diffie-Hellman sleuteluitwisselingssysteem over elliptische krommen.*

helpt hier ook de binaire schrijfwijze van de constante s. Bijvoorbeeld, als $s = 2185$ met binaire schrijfwijze 100010001001, dan geldt:

$$2185\mathcal{P} = 2(2(2(2(2(2(2(2(2\mathcal{P}))) + \mathcal{P})))) + \mathcal{P}))) + \mathcal{P}.$$

Het tegenovergestelde probleem van s te bepalen uit $s.\mathcal{P} = \mathcal{Q}$, waarbij $\mathcal{P}$ en $\mathcal{Q}$ op $\mathcal{E}$ liggen, is, net als bij het discrete logaritme probleem, in het algemeen ondoenlijk, als de orde van $\mathcal{P}$ groot genoeg is.

Het punt $\mathcal{P} = (121, 517)$ in Voorbeeld 5 blijkt orde 432 te hebben (dus $432P = \mathcal{O}$). Het is vrij gemakkelijk om nu $300\mathcal{P}$ uit te rekenen, maar als gevraagd wordt om s te bepalen uit $s.\mathcal{P} = (101, 496)$ dan zit er niet veel meer op dan $s = 1, 2, 3, \ldots$ uit te proberen.

Het is nu vrij gemakkelijk om een Diffie-Hellman-achtige sleuteluitwisselingsmethode voor elliptische krommen op te zetten (en ook om alle latere variaties na te bootsen). Begin met een elliptische kromme $\mathcal{E}$ (niet elke kromme is geschikt, net zoals niet elk priemgetal geschikt is in het oorspronkelijke systeem) en kies een punt $\mathcal{P}$ op $\mathcal{E}$ met een voldoende grote orde (we noemen de orde v).

Alice en Bob (en elke andere deelnemer) kiezen ieder een willekeurige coefficiënt, zeg s_A resp. s_B, beide kleiner dan v. Alice en Bob berekenen de scalaire veelvouden $\mathcal{K}_A = s_A\mathcal{P}$ resp. $\mathcal{K}_B = s_B\mathcal{P}$ en maken die bekend of wisselen die uit op een openbare manier. Beide kunnen nu het punt

$$\mathcal{S}_{A,B} = s_A s_B \mathcal{P} = s_A\mathcal{K}_B = s_B\mathcal{K}_A$$

uitrekenen. Inderdaad kan Alice met de publieke waarde $\mathcal{K}_B$ van Bob en haar eigen geheim s_A de waarde $s_A\mathcal{K}_B = \mathcal{S}_{A,B}$ uitrekenen. Voor Bob geldt een vergelijkbare situatie.

Aangezien de x-coördinaat van $\mathcal{S}_{A,B}$ de y-coördinaat uniek bepaalt op het teken na (wat overeenkomt met een bit), hoeven Alice en Bob alleen maar de x-coördinaat te gebruiken.

In Figuur 8 is het Diffie-Hellman sleuteluitwisselingsschema over elliptische krommen schematisch weergegeven. Merk op dat de machtsverheffing in het oorspronkelijke systeem hier vervangen is door scalaire vermenigvuldiging, net zoals een vermenigvuldiging hier vervangen is door een optelling.

# digits in p	Pollard	index calculus
100	$1.13\ 10^{15}$	$6.79\ 10^{15}$
150	$3.78\ 10^{22}$	$1.03\ 10^{19}$
200	$1.27\ 10^{30}$	$4.05\ 10^{21}$
250	$4.25\ 10^{37}$	$6.87\ 10^{23}$
300	$1.43\ 10^{45}$	$6.49\ 10^{25}$

Figuur 9. *Pollard's methode vergeleken met de "index calculus"-methode.*

De vraag is nu natuurlijk waarom het Diffie-Hellman sleuteluitwisselingssysteem zo interessant is in de context van elliptische krommen? Misschien dat sommige sceptici zich in het begin weleens afgevraagd hebben of dit niet gewoon een generalisatie om de generalisatie was, of dat het alleen maar om het verkrijgen van eigen patenten ging. Beide aannames bleken onterecht. Er is echt veel meer aan de hand. Om dat te begrijpen moeten we nagaan of de verschillende manieren om de discrete logaritme te kunnen nemen ook aangepast kunnen worden in de context van elliptische krommen.

Het is gemakkelijk in te zien dat "exhaustive search" en "baby-step giant-step" zonder problemen toepasbaar blijven. Men kan bijvoorbeeld een tabel maken van $\mathcal{O}, \mathcal{P}, 2\mathcal{P}, \ldots, 9\mathcal{P}$ en dan kijken of $\mathcal{Q}$ in de tabel staat. Als dat niet is, kan men kijken of $\mathcal{Q} - 10\mathcal{P}$ in de tabel staat, of anders $\mathcal{Q} - 20\mathcal{P}$, etc.

Ook "Pollard's methode" kan op een natuurlijke manier overgedragen worden (zie bijv. [1]). Maar de "index calculus" methode heeft elke poging tot generalisatie naar de elliptische krommen context weerstaan. De reden lijkt te liggen in het feit dat niet op een natuurlijke en bruikbare manier een factor base gedefinieerd kan worden. Zoiets als priemgetallen in de gehele getallen en onontbindbare veeltermen in de verzameling van alle veeltermen lijkt niet te bestaan.

Bovenstaande constatering is heel belangrijk en is de verklaring van alles. Om de gewone Diffie-Hellman sleuteluitwisseling over $\mathbb{Z}_p$ veilig uit te kunnen voeren moet p zo groot gekozen worden dat ook de "index-calculus"-methode met zijn subexponentiële complexiteit ondoenlijk is. Maar met de Diffie-Hellman sleuteluitwisseling over elliptische krommen hoeft men geen rekening te houden met index-calculus-achtige methoden! Dat betekent dat de ontwerper alleen naar Pollard's methode hoeft te kijken en die heeft een exponentieel groeiende rekentijd. In Figuur 9 worden deze complexiteiten vergeleken. Men ziet meteen dat voor eenzelfde beveiligingsniveau veel kleinere priemgetallen nodig zijn als elliptische krommen gebruikt worden dan als men werkt in de vermenigvuldingsgroep van $\mathbb{Z}_p$. Helemaal eerlijk is deze vergelijking niet. Het werken over elliptische krommen is een beetje bewerkelijker (het scheelt een constante factor).

In de normale Diffie-Hellman sleuteluitwisseling zijn sommige parameter keuzes niet veilig, bijvoorbeeld als het priemgetal p de eigenschap heeft dat $p - 1$ alleen maar kleine priemdelers heeft. Voor EC-cryptosystemen geldt iets soortgelijks bij de keuze van de specifieke kromme. De sterkste aanval

is de zgn. MOV-aanval [6]. Deze herleidt het EC-logaritme probleem tot het gewone logaritme probleem over een eindig lichaam. En dat houdt in dat index calculus-achtige methodes met hun subexponentiële rekentijd wel degelijk gebruikt kunnen worden! De MOV aanval werkt voor zgn. singuliere en supersinguliere elliptische krommen. Gelukkig zijn er heel eenvoudige manieren om krommen te testen op deze eigenschappen.

Er is een andere zorg die overblijft: elliptische krommen zijn in het verleden niet in dezelfde mate bestudeerd als het klassieke discrete logaritme probleem over $\mathbb{Z}_p$ en nog veel minder als het ontbinden van een getal in zijn factoren. Nieuwe inzichten zouden dus EC-cryptosystemen alsnog waardeloos kunnen maken voor concrete toepassingen. Met het verstrijken van de tijd wordt deze zorg steeds minder.

Nog een laatste algemene opmerking: men kan natuurlijk ook andere groepsstructuren bestuderen om te komen tot Diffie-Hellman-achtige sleuteluitwisselingssystemen. Tot op heden hebben echter alle voorstellen in die richting problemen met een efficiënte implementatie.

4. Toepassingen

Zoals eerder al uitgelegd werd, groeit de veiligheid van EC-cryptosystemen exponentiëel in de lengte (in bits gemeten) van de gebruikte parameters, terwijl dit voor alternatieve public key cryptosystemen, zoals RSA, slechts subexponentiëel is. Om een getallenvoorbeeld te geven; een 256 bits EC-cryptosysteem dient vergeleken te worden met een 3072 bits RSA modulus! De rekenkracht om deze systemen te gebruiken is hetzelfde; het groeit als t^3, waarbij t het aantal bits in de sleutel is.

Bij de implementatie van cryptosystemen heeft men vaak te maken met praktische beperkingen, zoals benodigde geheugenruimte, rekentijd of stroomgebruik (denk aan smartcards). Een snel groeiend gebied is draadloze communicatie. Het is in dit soort toepassingen dat EC-cryptosystemen zo gunstig afsteken bij andere systemen. Zij kunnen op smartcards geïmplementeerd worden die niet beschikken over een wiskundige coprocessor. Door de veel kortere sleutel is veel minder energie nodig. Ook als straks "handhelds" draadloos gaan communiceren hebben EC-cryptosystemen voordelen die andere systemen niet kennen: sleutels aanmaken kost minder tijd, het begroetingsprotocol kost minder tijd, etc. Om die redenen zullen we in de toekomst meer en meer apparaten zien die gebruik maken van EC-cryptosystemen.

Op dit moment zijn er reeds vele internationale standaarden met betrekking tot EC-cryptosystemen: ISO, ANSI, IEEE en SECG.

Literatuur

1. I. Blake, G. Seroussi, and N. Smart, *Elliptic curves in cryptography,* London Math. Soc. Lect. Notes Ser. **265**, Cambridge University Press, 1999.
2. W. Diffie and M.E. Hellman, *New directions in cryptography,* IEEE Trans. Inf. Theory, IT-22, pp. 644–654, Nov. 1976.
3. P. Flajolet and A.M. Odlyzko, *Random mapping statistics*, in Advances in

Cryptography: Proc. of Eurocrypt '89, J.-J. Quisquater and J. Vandewalle, Eds., Lecture Notes in Computer Science 434, Springer Verlag, Berlin etc., pp. 329–354, 1990.
4. D.M. Gordon, *Discrete logarithms in $GF(p)$ using the number field sieve,* SIAM Journal on Discrete Mathematics, **6**, pp. 124-138, 1993.
5. A.G. Konheim, *Cryptography, a primer,* John Wiley & Sons, New York, etc., 1981.
6. A. Menezes, T. Okamoto, and S. Vanstone, *Reducing elliptic curve logarithms to logarithms in a finite field,* IEEE Transactions on Information Theory, IT-39, pp. 1639–1646, 1993.
7. Menezes, A.J., P.C. van Oorschot, and S.A. Vanstone, *Handbook of Applied Cryptography,* CRC Press, Boca Raton, etc. 1997.
8. S. Miller, *Use of elliptic curves in cryptography*, abstract in Advances in Cryptography: Proc. of Crypto '85, H.C. Williams, Ed., Lecture Notes in Computer Science 218, Springer Verlag, Berlin etc., p. 417, 1986.
9. J.M. Pollard, *Monte Carlo methods for index computation (mod p)*, Math. Comp. **32**, pp. 918–924, 1978.
10. R.L. Rivest, A. Shamir and L. Adleman, *A method for obtaining digital signatures and public key cryptosystems,* Comm. ACM, Vol. **21**, pp. 120–126, Febr. 1978.
11. R. Schoof, *Counting points on elliptic curves over finite fields*, Journal de Théorie des Nombres de Bordeaux, 7, pp. 219–254, 1995.
12. J.H. Silverman, *The Arithmetic of Elliptic Curves,* Springer Verlag, Berlin, etc., 1986.
13. Henk C.A. van Tilborg, *Fundamentals of cryptology; A professional reference and interactive tutorial,* Kluwer Academic Publishers, Boston etc., 2000.

Vakantiecursus 2006

pp. 21 – 31

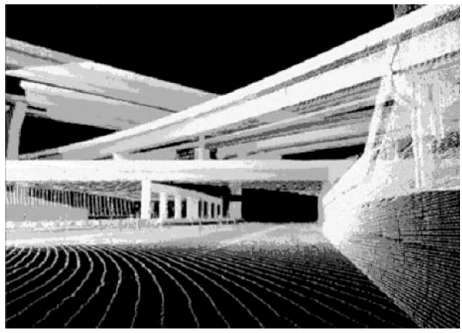
Wiskunde van GIS – Het Poincaré formalisme van de simpliciale homologie voor 3D volume modellen

P.J.M. van Oosterom
F. Penninga
Technische Universiteit Delft
e-mail: p.v.oosterom@otb.tudelft.nl

In dit artikel zal nader worden ingegaan op het opslaan (modelleren) van 3D topografie gebaseerd op een wiskundige theorie. Topografische objecten zoals gebouwen, infrastructuur en kunstwerken worden steeds complexer door toenemend meervoudig ruimtegebruik. Vergrote bewustwording van het belang van duurzame (stedelijke) ontwikkelingen verhoogt de behoefte aan 3D planning- en analysemogelijkheden. Als gevolg hiervan moeten topografische producten worden uitgebreid naar de derde dimensie. In dit artikel wordt een nieuw 3D topologisch datamodel gepresenteerd, gebaseerd op het Poincaré formalisme van de simpliciale homologie. De interne structuur bestaat uit een netwerk van simplexen (knopen, zijden, driehoeken en tetraheders), die goed gedefinieerd zijn, en zeer geschikt zijn voor het consistent bijhouden van de 3D data. Complete 3D objecten bestaan uit een verzameling van deze simplexen.

1. Introductie

De meeste topografische producten representeren de werkelijkheid via een afbeelding in twee dimensies. Echter de topografische objecten (fysieke objecten, zoals gebouwen, wegen, tunnels, viaducten, etc.) worden steeds complexer door het toenemende meervoudige ruimtegebruik. Dit vraagt om 3D data modellen ter ondersteuning van 3D planning en analyse, dit in tegenstelling tot de eerdere 3D GIS ontwikkelingen die vooral op de visualisatie gericht waren. Een kenmerk van topografische producten is hun brede variëteit aan toepassingen, met als gevolg dat het niet mogelijk is om de productdefinitie voor één specifieke toepassing te optimaliseren. Dankzij de continue ontwikkelingen op het gebied van de sensor technologie (Vosselman 2005) komen er meer en meer 3D gegevens beschikbaar. Bovendien neemt de nauwkeurigheid, puntdichtheid en daarmee ook de gegevenshoeveelheden toe. Een voorbeeld van terrestisch laserscannen is in Figuur 1 te zien.



Figuur 1. *Terrestisch laser scannen geeft inzicht in complexe constructies.*

Er is een groot aantal onderzoeken verricht naar 3D data modellering. Veel van deze onderzoeken zijn samengevat en vergeleken in (Zlatanova 2002). Het uitbreiden van topografische datamodellen naar de derde dimensie is vooral relevant voor grootschalige topografie. Het waarborgen van de integriteit en het zorgen voor goede performance zijn belangrijke eisen. Vandaar dit onderzoek om de 3D datastructuur te implementeren in een ruimtelijke database gebaseerd op een 'Tetraheder Netwerk' (TEN). Een TEN heeft namelijk een aantal goede eigenschappen, zoals: eenduidige definitie van driehoeken (per definitie vlak), aanwezigheid van topologische relaties, goede onderhoudbaarheid, visualisatie via driehoeken, en flexibiliteit om meer complexe objecten te vormen. Het model is gebaseerd het Poincaré formalisme van de simpliciale homologie (Poincaré 1895) en heeft daarmee een solide fundament.

2. 3D Topografisch model in een TEN datastructuur

Als we topografie zien als de verzameling van fysieke objecten, kunnen er twee opmerkingen worden gemaakt voor wat betreft 3D topografisch modelleren:

1. De fysieke objecten hebben per definitie een inhoud (volume). Er bestaan in de werkelijkheid geen echte punt-, lijn- of vlakobjecten; er bestaan slechts punt-, lijn-, of vlakrepresentaties op een gegeven generalisatieniveau. Welke weergave gebruikt zou moeten worden, is aangegeven in het Digitaal Cartografisch Model (DCM), maar niet in het Digitaal Landschaps Model (DLM), welke de 3D topografie bevat.
2. De fysieke werkelijkheid kan beschouwd worden als een volume partitie: een verzameling van niet-overlappende volume objecten, die tezamen de te modelleren ruimte geheel vullen. Als gevolg hiervan zijn objecten als 'aarde' en 'lucht' expliciet onderdeel van de fysieke werkelijkheid en dus van het model.

De topografische gegevenverzameling bestaat dus uit volumeobjecten. Toch kunnen in sommige gevallen ook vlakobjecten nuttig zijn, aangezien deze belangrijke overgangen markeren tussen twee volumeobjecten. Vlakobjecten kunnen hun eigen attributen hebben, zoals oppervlaktemateriaal en kleur, maar

ze kunnen niet bestaan zonder de aanwezigheid van deze volumeobjecten. Een vlakobject kan worden gezien als de eerste afgeleide van een volumeobject (en dit zou herhaald kunnen worden voor lijn- en puntobjecten). In het UML klassediagram (zie Figuur 3) zijn de vlakobjecten dan ook als associatieklassen gemodelleerd.

De keuze om expliciet 'lucht' en 'aarde' objecten mee te nemen – eigenlijk de 'lege' ruimte tussen de fysieke objecten – is mede ingegeven door het feit dat deze 'lege' ruimte vaak het onderwerp van analyse is. In geval van modelleren van luchtverontreiniging, dijkdoorbraak of overstroming, is de gebruiker geïnteresseerd wat er met deze 'lege' ruimte gebeurd.

2.1. Poincaré formalisme van de simpliciale homologie

De TEN is de 3D tegenhanger van het welbekende TIN (getrianguleerd onregelmatig netwerk). Behalve knopen, zijden en driehoeken, bestaat een TEN uit tetraheders voor het representeren van volumeobjecten. Knopen, zijden, driehoeken en tetraheders zijn allen simplexen; d.w.z. de meest eenvoudige primitieven in de gegeven dimensie. Het modelleren van 3D objecten d.m.v. simplexen is beschreven door Carlson (1987). Het gebruik van simplexen heeft een aantal voordelen:

1. Simplexen zijn goed gedefinieerd: een kD simplex wordt begrensd door $k+1$ simplexen van de dimensie $k-1$ (Egenhofer et al.1989). Bijvoorbeeld: een 2D simplex (driehoek) wordt begrensd door 3 1D simplexen (zijden).
2. Vlakheid van elke driehoek, aangezien 3 punten liggen per definitie in een vlak.
3. Elke simplex is convex, ongeacht de dimensie.

Een direct gevolg van het goed gedefinieerde karakter van simplexen en daarmee een TEN, is de beschikbaarheid van 3D topologische relaties. Daar waar in 2D (TIN) veel belangrijke topologische relaties gerelateerd zijn aan de zijde (b.v. een zijde heeft een driehoek links en een driehoek rechts), zijn in 3D veel van de belangrijke relaties te vinden op het niveau van de driehoek. Elke driehoek is onderdeel van de begrenzing van twee tetraheders. Links en rechts zijn betekenisloos in 3D, maar via de ordening van de zijden binnen een driehoek kan de richting van de normaalvector van de driehoek bepaald worden. Hiermee kan dus een tetraheder in positieve en negatieve richting worden aangegeven. De n-dimensionele simplex wordt gedefinieerd door $n+1$ knopen en dit wordt als volgt genoteerd $S_n =< x_0, \ldots, x_n >$.
De eerste vier simplexen zijn dus respectievelijk van 0D naar 3D: $S_0 = < x_0 >$, $S_1 =< x_0, x_1 >, S_2 =< x_0, x_1, x_2 >$, and $S_3 = < x_0, x_1, x_2, x_3 >$. De $(n+1)$ knopen geven $(n+1)!$ permutaties van deze knopen, of te wel resp. 1, 2, 6 en 24 opties voor 0D, 1D, 2D en 3D simplexen. Voor S_1 zijn de twee permutaties $< x_0, x_1 >$ and $< x_1, x_0 >$, waarvan de eerste (van begin naar einde) positief $(+)$ wordt genoemd en de tweede negatief $(-)$. De 2D simplex heeft zes permutaties $S_2 :< x_0, x_1, x_2 >, < x_1, x_2, x_0 >, < x_2, x_0, x_1 >, < x_2, x_1, x_0 >, < x_0, x_2, x_1 >$ en $< x_1, x_0, x_2 >$. De eerste drie hebben de tegenovergestelde oriëntatie aan de tweede set van drie, dus er kan gesteld worden dat

$< x_0, x_1, x_2 > = - < x_2, x_1, x_0 >$. De positieve oriëntatie is tegen de klok indraaiend (+) en de negatieve oriëntatie is met de klok meedraaiend (−). Voor de 3D simplex $S_3 = < x_0, x_1, x_2, x_3 >$ zijn er 24 verschillende permutaties, waarvan er 12 gerelateerd zijn aan de positief georiënteerde tetraheder (+, alle normaalvectoren wijzen naar buiten) en de overige 12 betreffen de negatief georiënteerde tetraheder (−, alle normaalvectoren wijzen naar binnen). Aangezien er dus verschillende equivalente notaties zijn, is het handig om een afspraak te maken over de voorkeursnotatie; b.v. de combinatie met positieve oriëntatie met knopen met de laagste id's als eerste. Volgens het Poincaré formalisme van de simpliciale homologie, bestaat de grens van een simplex uit de volgende som van $(n-1)$ dimensionale simplexen (weglaten van de i^{de} knoop en het om en om afwisselen van de + en − tekens):

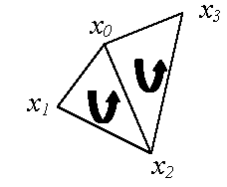
$$\partial S_n = \sum_{i=0}^{n} (-1)^i < x_0, \ldots, \hat{x}_i \ldots, x_n >$$

Dus de grens van $\partial S_1 =< x_0, x_1 >$ is $< x_1 > - < x_0 >$ en de grens van $\partial S_1^{neg} =< x_1, x_0 >$ zou zijn $< x_0 > - < x_1 >$. De grens van $\partial S_2 = < x_0, x_1, x_2 >$ is $< x_1, x_2 > - < x_0, x_2 > + < x_0, x_1 >$. Op soortgelijke manier kunnen de grenzen van de andere 5 permutaties van S_2 gegeven worden. Tenslotte de grens $\partial S_3 =< x_0, x_1, x_2, x_3 >$ is $< x_1, x_2, x_3 > - < x_0, x_2, x_3 > + < x_0, x_1, x_3 > - < x_0, x_1, x_2 >$ (en soortgelijk voor de 23 andere permutaties). Kijkend naar de grenzen van de grenzen van een tetraheder, dat wil zeggen de grenzen van de driehoeken (de zijden dus), blijkt dat elke zijde exact één keer in de positieve richting en één keer in de negatieve richting voorkomt binnen de tetraheder. Een ander interessant resultaat van het Poincaré formalisme van de simpliciale homologie is het aantal lager dimensionale simplexen die voorkomen als (in)directe grens van een gegeven simplex:

$$S_n \text{ heeft } \binom{n+1}{p+1} \text{ grens simplexen van dimensie } p \text{ met } (0 \leq p < n)$$

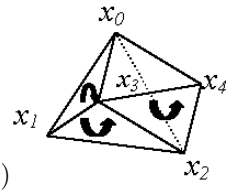
Dus S_2 (driehoek) bestaat uit 3 0D simplexen (knopen) en 3 1D simplexen (zijden). De simplex S3 heeft respectievelijk 4, 6 en 4 0D, 1D and 2D grenssimplexen. Wanneer buursimplexen van gelijke dimensie worden samengevoegd, dan wordt hun gemeenschappelijke grens verwijderd zoals aangegeven in Figuur 2a. Neem bijvoorbeeld de buurdriehoeken $< x_0, x_1, x_2 >$ en $< x_0, x_2, x_3 >$ dan resulteert het samenvoegen van al hun grenzen (zijden) in: $(< x_1, x_2 > - < x_0, x_2 > + < x_0, x_1 >) + (< x_2, x_3 > - < x_0, x_3 > + < x_0, x_2 >) = < x_1, x_2 > + < x_0, x_1 > + < x_2, x_3 > - < x_0, x_3 > = < x_1, x_2 > + < x_0, x_1 > + < x_2, x_3 > + < x_3, x_0 >$. Merk dus op dat de gemeenschappelijke grens $< x_0, x_2 >$ is verdwenen. Op een zelfde wijze resulteert het samenvoegen van twee buurtetraheders $< x_0, x_1, x_2, x_3 >$ en $< x_0, x_2, x_4, x_3 >$ en het optellen van hun grenzen (driehoeken) in $< x_1, x_2, x_3 > + < x_0, x_1, x_3 > + < x_2, x_1, x_0 > + < x_2, x_4, x_3 > + < x_3, x_4, x_0 > + < x_4, x_2, x_0 >$. Wanneer ook hier weer naar de zijden wordt gekeken, dan valt het op dat elke zijde één keer in positieve en één keer in negatieve richting

wordt gebruikt. De samengevoegde buur n-simplexen worden een 'simpliciaal complex' of n-cell genoemd. Het is goed mogelijk om een topologische structuur te bouwen voor een verzameling verbonden n-simplexen (inclusief al hun lager dimensionale grenzen: $0, \ldots, n-1$ simplexen), welke het n-dimensionale domein geheel opdeelt. In 3D wordt dit het tetrahedron netwerk (TEN) genoemd. Binnen zo'n netwerk is het niet alleen interessant om naar de grens van een simplex te kijken, maar ook naar de co-grens, oftewel van welke hoger dimensionale simplexen de gegeven simplex zelf een grens is. Zo bestaat bijvoorbeeld de grens van een driehoek uit drie zijden en de co-grens van de driehoek uit twee tetraheders. Op soortgelijke manier bestaat de grens van een zijde uit twee knopen en de co-grens uit twee of meer driehoeken.



(a)

$S_{21} =< x_0, x_1, x_2 >$ en $S_{22} =< x_0, x_2, x_3 >$
$C_2 =< x_1, x_2 > - < x_0, x_2 > + < x_0, x_1 >$
$+ < x_2, x_3 > - < x_0, x_3 > + < x_0, x_2 >$
$=< x_1, x_2 > + < x_0, x_1 > + < x_2, x_3 > + < x_3, x_0 >$



(b)

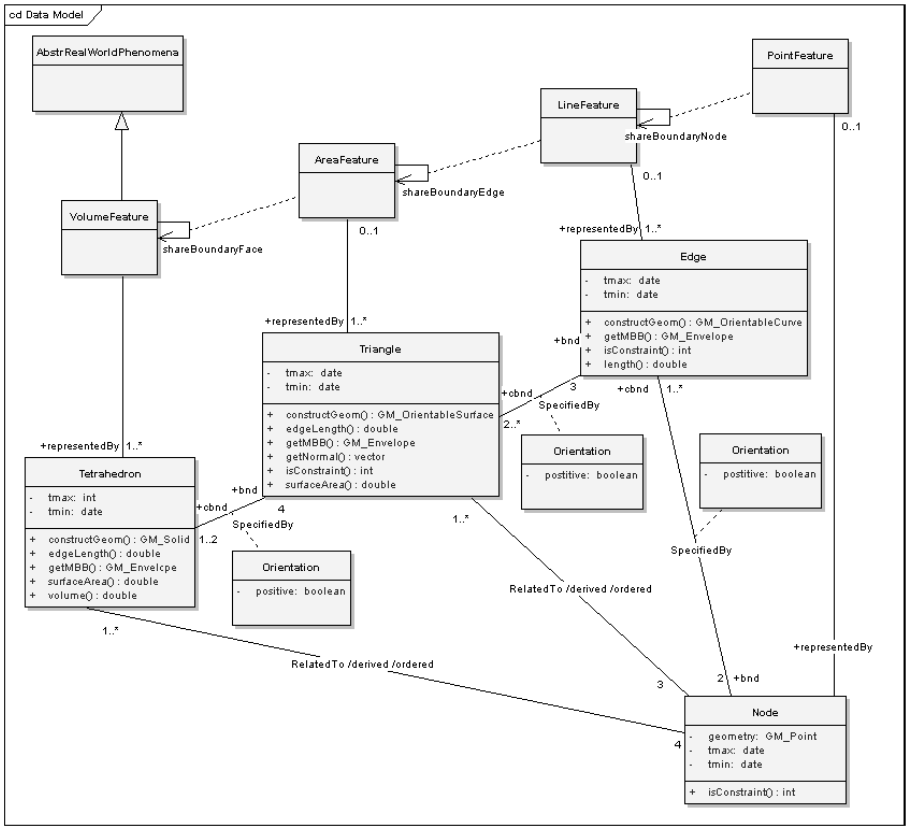
$S_{31} =< x_0, x_1, x_2, x_3 >$ en $S_{32} =< x_0, x_2 x_4, x_3 >$
$C_3 =< x_1, x_2, x_3 > - < x_0, x_2, x_3 > + < x_0, x_1, x_3 >$
$- < x_0, x_1, x_2 > + < x_2, x_4, x_3 > - < x_0, x_4, x_3 >$
$+ < x_0, x_2, x_3 > + < x_0, x_1, x_3 > - < x_0, x_1, x_2 >$
$=< x_1, x_2, x_3 > + < x_0, x_1, x_3 > - < x_0, x_1, x_2 >$
$+ < x_2, x_4, x_3 > - < x_0, x_4, x_3 > + < x_4, x_2, x_0 >$

Figuur 2. *Samengevoegde simplex buren vormen een complex in (a) 2D en (b) 3D.*

2.2. 3D Topografie model gebaseerd op een TEN

Redelijk dicht bij dit model komen de implementaties in Panda (Egenhofer et al. 1989) en Oracle Spatial, beide echter gaan maar tot twee dimensies en zijn gebaseerd op complexen (samengevoegde buur simplexen van zelfde dimensie, of te wel de n-cellen) en niet op de simplexen zelf.

Er zijn drie verschillende conceptuele TEN modellen ontwikkeld. Hoewel deze alle drie ongeveer een gelijke bedoeling hebben zijn de verschillen toch opvallend. Daar een conceptueel model de start van een implementatie vormt zijn deze verschillen wel belangrijk. Het eerste model is een min of meer intuïtieve modellering van een TEN; zie Figuur 3: de primitieven zijn gericht (positief) en de grens/co-grens associaties tussen knoop en zijde, zijde en driehoek en driehoek en tetraheder zijn van een teken voor voorzien via de associatie klasse 'Orientation'. Een efficiënte implementatie zal waarschijnlijk geen expliciete extra klassen gebruiken, maar de verwijzingen van teken $(+/-)$ voorzien. De associatie tussen tetraheder (of driehoek) en knoop kan worden afgeleid (en geeft de juiste ordening van de knopen).



Figuur 3. *Het eerste TEN model (d.w.z. model 1)*

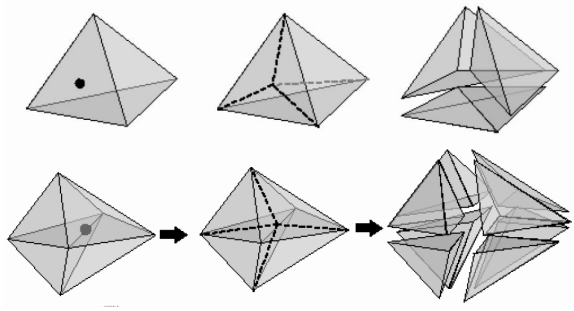
Het tweede model (appendix A) is een variant op het eerste model, maar in plaats van associatie klassen, zijn er nu extra klassen geïntroduceerd: de ongerichte versie van hun gerichte tegenhangers. Net als bij het eerste model kunnen nu ook weer de associaties tussen tetraheder (of driehoek) en knoop worden afgeleid. Het derde model is direct gebaseerd op het Poincaré formalisme van de simpliciale homologie: directe associaties tussen knoop en alle drie de andere primitieven (zijde, driehoek en tetraheder). De ordening van de knopen is hierbij belangrijk, want deze definieert tevens de oriëntatie. De andere associaties (tussen tetraheder en driehoek en tussen driehoek en zijde) kunnen worden afgeleid inclusief de richting/oriëntatie. De belangrijkste verschillen tussen de 3 modellen zijn: (1) welke associaties worden expliciet opgeslagen en welke worden afgeleid en (2) in geval van georiënteerde associaties (verwijzingen), worden deze gemodelleerd via een associatie klasse of via extra ongerichte primitieven? Het derde model bevat de minste redundantie voor wat betreft de verwijzingen en is bovendien op zuivere theorie van het Poincaré formalisme van de simpliciale homologie gebaseerd.

3. Gebruik van het TEN model

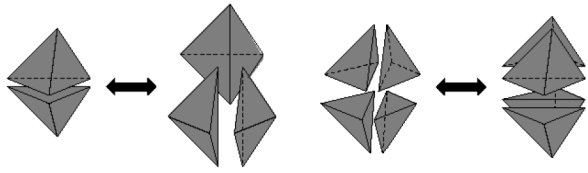
De gegevens worden in een database beheerd. Een eerste stap in het omzetten van de objecten uit de echte wereld naar een 3D TEN representatie. De feitelijke objectbegrenzingen (van b.v. de gebouwen) komen dan terug als zogenaamde verzameling van vastgeprikte driehoeken ('constraints') in het TEN, deze mogen namelijk niet zonder meer worden aangepast. Daarnaast zijn er ook driehoeken die de interne structuur beschrijven, deze kunnen worden aangepast zonder dat de buitenkant van een object in het geding komt. Bij het vormen van het tetraheder netwerk is het goed om slecht gevormde primitieven te voorkomen. Op dit gebied heeft (Shewchuk 1997) al veel onderzoek verricht. Belangrijk in een praktische setting is dat de gegevens incrementeel bijgehouden moeten kunnen worden door het toevoegen (of verwijderen) van objecten, vaak ten koste van 'lucht' of 'aarde' tetraheders. In sommige gevallen kan het ook ten koste gaan van andere objecten en zal de gebruiker goed moeten nagaan of dit inderdaad wel de bedoeling is. De meeste objecten zullen (indirect) verbonden zijn met het 'aardoppervlak' en dit moet ook tijdens het muteren goed in de gaten worden gehouden cq. worden afgedrongen door het systeem.

Het toevoegen of verwijderen van objecten wordt intern vertaald naar de volgende basis mutatieoperaties die de TEN structuur aanpassen:

1. Verplaats knoop (alleen toegestaan indien de topologische structuur correct blijft).
2. Toevoegen van een knoop en bijbehorende zijden/ driehoeken/ tetraheders (of de omgekeerde operatie 'verwijder knoop'), waarbij de volgende drie gevallen mogelijk zijn, afhankelijk van waar de knoop wordt toegevoegd:
   - Midden in een tetraheder (1 tetraheder betrokken) en toegevoegd worden +1 knoop, +4 zijden, +6 driehoeken en +3 tetraheders (respectievelijk de 0/1/2/3-simplexen).
   - Midden in een driehoek (2 tetraheders betrokken) en toegevoegd worden +1 knoop, +5 zijden, +7 driehoeken en +4 tetraheders (Figuur 4 boven).
   - Midden in een zijde (n tetraheders betrokken) en toegevoegd worden +1 knoop, $+(n+1)$ zijden, $+2n$ driehoeken en $+n$ tetraheders (Figuur 4 onder).
3. Toevoegen (of verwijderen) van constraints: in de eerste plaats driehoeken (als onderdeel van de buitenkant van een object), maar daarnaast mogelijk ook zijden en knopen (i.v.m. met lager dimensionale objecten).
4. Omklappen van tetraheders, waarbij twee situaties mogelijk zijn, afhankelijk van de configuratie (zie Figuur 5):
   - 2–3 bistellar omklappen
   - 4–4 bistellar omklappen



Figuur 4. *Toevoegen van een knoop; boven: in een driehoek (buur tetraheder niet getoond), onder: in een zijde verbonden met vier tetraheders, beide illustraties uit (van der Most 2004).*



Figuur 5. *Omklappen in 3D: 2–3 bistellar omklappen (links) en 4–4 bistellar omklappen (rechts), wederom beide illustratie uit (van der Most 2004).*

4. Conclusie

Het uitreiden van het topografische gegevensmodel is vooral relevant in het geval van grootschalige topografie. Gezien de omvang en het belang van deze gegevens zijn het garanderen van de integriteit en het bieden van een goede performance van groot belang. Daarom moeten deze gegevens dan ook in een DBMS beheerd worden in een gegevensmodel gebaseerd op het Poincaré formalisme van de simpliciale homologie. De TEN structuur is geschikt vanwege de eigenschappen als goed gedefinieerde driehoeken (per definitie vlak), topologische structuur, goede onderhoudbaarheid, visualisatie gebaseerd op driehoeken, en de mogelijkheid om meer complexe objecten te vormen (door het samenvoegen van tetraheders). Het gepresenteerde conceptuele model is de basis voor een toekomstige implementatie in Oracle en er zullen een aantal verschillende prototypen op gebaseerd worden. Bovendien zal naast de 3^{de} dimensie ook het temporele aspect in dit model worden meegenomen.

Dankwoord

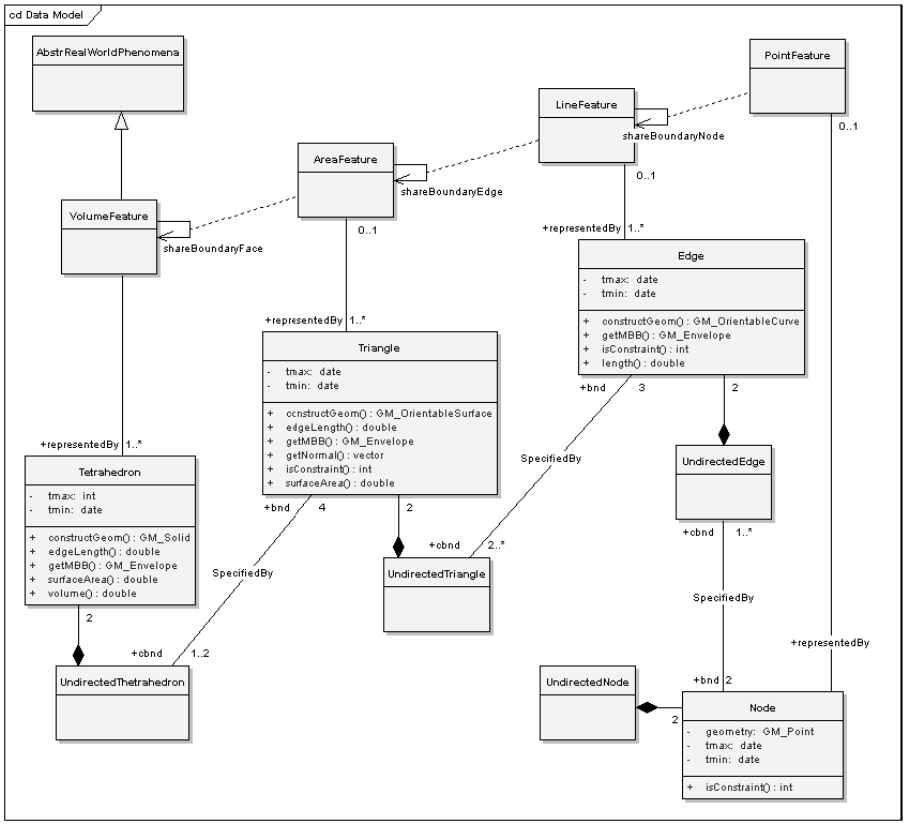
Dank aan Baris Kazar, John Herring, Siva Ravada, Ravi Kothuri en Han

Wammes voor de constructieve discussie over TEN modelleren, welke de basis voor dit artikel heeft gelegd. Het onderzoek wordt uitgevoerd in het kader van het Bsik RGI (Ruimte voor Geo-Informatie) project '3D topografie' (`http://www.gdmc.nl/3dtopo`).

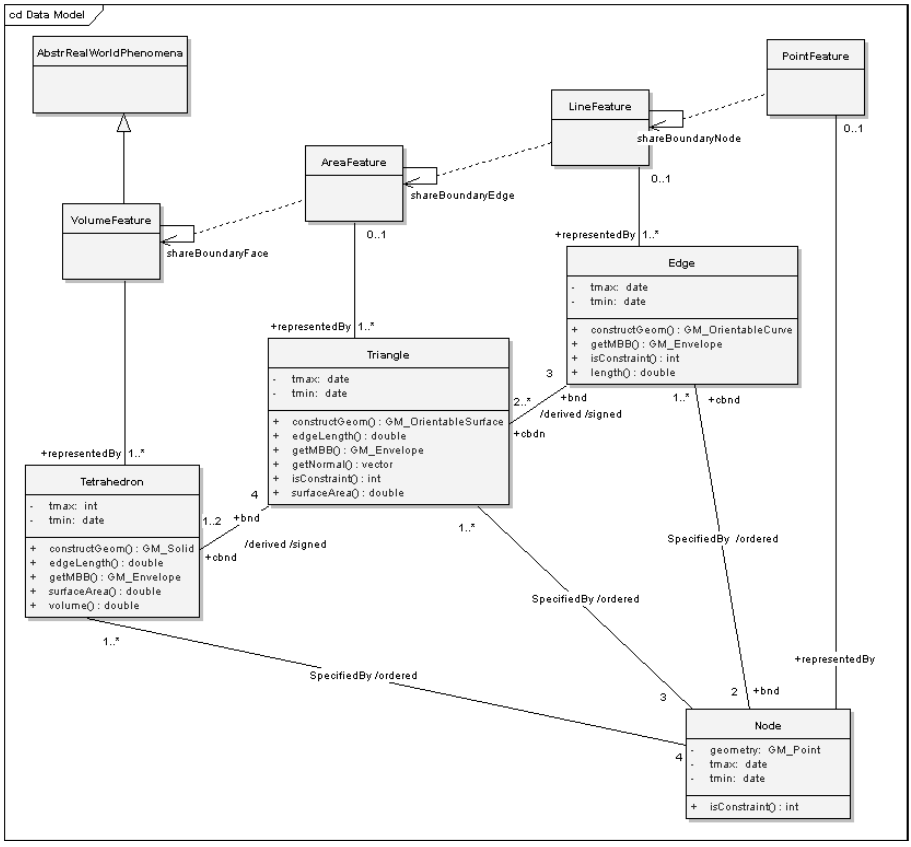
Literatuur

1. Carlson, E. (1987). Three-dimensional conceptual modeling of subsurface structures. In: *Auto-Carto* **8**, pp. 336–345.
2. Egenhofer, M. and Andrew, F. (1989). PANDA: An Extensible DBMS Supporting Object-Oriented Software Techniques Database Systems in Office, Engineering, and Science, Zurich, Switzerland, T. Harder (ed.), *Informatik Fachberichte*, Vol. **204**, Springer-Verlag, pp. 74–79.
3. Van der Most, A. (2004). An algorithm for overlaying 3D features using a tetrahedral network Master's Thesis TU Delft, 2004, 96 p.
4. Poincaré, H. (1895), Analysis Situs, Journal de l'Ecole Polytechnique ser 2, Vol. **1**, pp. 1-123.
5. Shewchuk, J.R. (1997). Delaunay refinement mesh generation, PhD thesis, Carnegie Mellon University.
6. Vosselman, G. (2005). Sensing Geo-information, Inaugural address, ITC Enschede.
7. Zlatanova, S., Rahman, A. A., Shi, W. (2002) Topology for 3D spatial objects, International Symposium and Exhibition on Geoinformation 2002, 22–24 October, Kuala Lumpur, Malaysia, 7 p.

Appendix A. Alternatieve modellen



Model 2.


cd Data Model
AbstrRealWorldPhenomena
PointFeature
LineFeature
shareBoundaryNode
0..1
AreaFeature
shareBoundaryEdge
0..1
VolumeFeature
shareBoundaryFace
0..1
+representedBy 1..*
Edge
- tmax: date
- tmin: date
+ constructGeom() : GM_OrientableCurve
+ getMBB() : GM_Envelope
+ isConstraint() : int
+ length() : double
+representedBy 1..*
Triangle
- tmax: date
- tmin: date
+ constructGeom() : GM_OrientableSurface
+ edgeLength() : double
+ getMBB() : GM_Envelope
+ getNormal() : vector
+ isConstraint() : int
+ surfaceArea() : double
3
2..*
+bnd
/derived /signed
+cbdn
1..*
+cbnd
+representedBy 1..*
Tetrahedron
- tmax: int
- tmin: date
+ constructGeom() : GM_Solid
+ edgeLength() : double
+ getMBB() : GM_Envelope
+ surfaceArea() : double
+ volume() : double
4
1..2
+bnd
/derived /signed
+cbnd
1..*
SpecifiedBy /ordered
SpecifiedBy /ordered
1..*
SpecifiedBy /ordered
+representedBy
3
2
+bnd
4
Node
- geometry: GM_Point
- tmax: date
- tmin: date
+ isConstraint() : int

Model 3.

Vakantiecursus 2006

pp. 33 – 45

Derivaten

R.M. Elkenbracht-Huizing
ABN AMRO bank N.V. Amsterdam
e-mail: marije.elkenbracht@nl.abnamro.com

1. Introductie

Derivaten is de verzamelnaam van alle financiële producten, waarvan de waarde 'afgeleid' is van de waarde van een 'onderliggend goed'. Het 'onderliggende goed' kan bijvoorbeeld een aandeel of een obligatie zijn, maar ook valutakoersen of inflatiecijfers. Derivaten kunnen gebruikt worden om te speculeren, om risico's te diversificeren, maar ook om risico's af te dekken (te 'hedgen'). Voorbeelden van het laatste zijn een boer die zijn oogst al vooraf tegen een vastgestelde prijs verkoopt, om het risico van slecht weer af te dekken. Of een gasbedrijf dat gas inkoopt tegen een vooraf vastgestelde prijs op het moment wanneer ook een prijs wordt afgesproken met de afnemers. Of een bedrijf dat met een klant een prijs voor een nog te produceren goed in Euro's afspreekt, maar nog onderdelen in Dollars moet kopen, en die het resulterende valutarisico afdekt. Opties zijn de meest bekende derivaten.

Banken verhandelen tal van soorten derivaten. Wanneer een bank derivaten wil kopen of verkopen, moeten in ieder geval de twee volgende vragen beantwoord worden:

1. Wat is de waarde van het derivaat?
2. Hoe worden de risico's gedurende de looptijd beheerd?

We zullen zien dat beide vragen sterk met elkaar verbonden zijn, omdat de waarde van een derivaat in belangrijke mate bepaald wordt door het bedrag dat nodig is om het met zo weinig mogelijk risico's gedurende zijn levensduur te beheren.

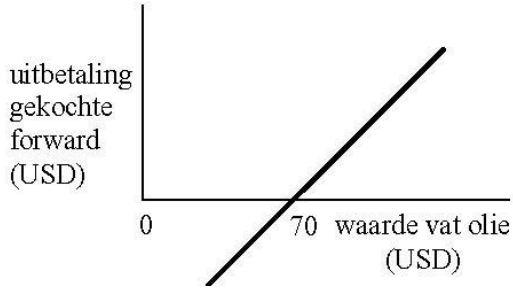
In dit artikel zullen we de belangrijkste principes bespreken die ten grondslag liggen aan het waarderen van derivaten. Tevens zullen we deze principes toepassen op een aantal concrete voorbeelden. Tot slot zullen we een indruk geven van hoe de risico's voor derivaten beheerst worden.

2. Voorbeelden van derivaten

2.1. Forwards

Bij een forward spreken twee partijen af om een bepaald goed op een later tijdstip te kopen of te verkopen tegen een vooraf vastgestelde prijs. Stel bijvoorbeeld dat je een forward koopt waarbij je afspreekt om over een jaar (de 'expiratiedatum') het goed (bijvoorbeeld een vat olie) te kopen voor USD 70.

In de volgende grafiek is aangegeven hoe op de expiratiedatum de uitbetaling van deze forward zich verhoudt tot de waarde van het vat olie:

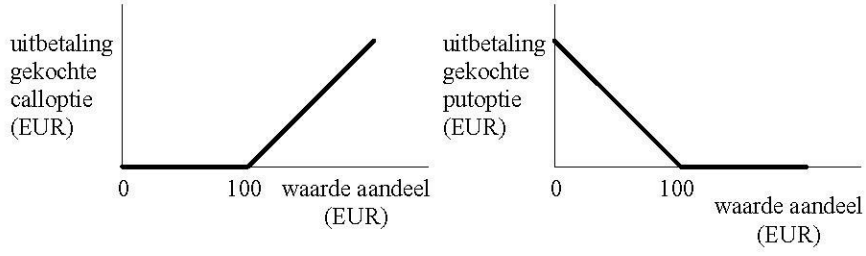


Figuur 1. *Uitbetaling gekochte forward op expiratiedatum voor aankoop van een vat olie voor USD 70.*

Behalve dat dergelijke contracten worden afgesloten tussen banken en klanten, bestaan er ook beurzen waar forwards verhandeld worden op bijvoorbeeld agrarische producten, olieachtige producten, valuta, en obligaties. Bij verhandeling op een beurs worden deze contracten 'futures' genoemd.

2.2. Opties

Wanneer je een optie koopt, koop je het *recht* (en heb je niet de plicht zoals bij een forward) om op een later tijdstip (de 'expiratiedatum') een onderliggend goed te kopen (een 'calloptie') of te verkopen (een 'putoptie') tegen een vooraf vastgestelde prijs (de 'uitoefenprijs'). Stel dat je het recht koopt om over een jaar een aandeel te kopen voor een uitoefenprijs van Euro 100. In Figuur 2a is aangegeven hoe op de expiratiedatum de uitbetaling van deze optie zich verhoudt tot de waarde van het onderliggende aandeel. In Figuur 2b geven we de uitbetaling van (de verder gelijke) putoptie weer.



Figuur 2a en 2b. *Uitbetaling gekochte call respectievelijk putoptie op de expiratiedatum voor de aankoop respectievelijk verkoop van een aandeel voor Euro 100,-.*

2.3. Gestructureerde producten

De laatste jaren zijn talloze nieuwe beleggingsproducten door beleggingsinstellingen en banken ontwikkeld. Op `http://www.abnamromarkets.nl` en op `http://www.rabobank.nl/particulieren/beleggen/beleggingsproducten/` staan diverse voorbeelden. Ook dit zijn allemaal derivaten. Vaak bieden ze participatie in één of meerdere indices, gecombineerd met een garantie op een bepaald rendement.

3. Tijdwaarde van geld

Een belangrijk principe in de financiële markten is dat iedereen geld liever vandaag heeft dan morgen. De vergoeding voor het uitlenen van geld heet 'rente'. De hoogte van de rente hangt af van de kredietwaardigheid van de debiteur, de lengte van de leenperiode en hoe vaak er rente uitgekeerd wordt. In dit artikel zullen wij alleen werken met de zogenaamde 'risicovrije' rente, dat is de term die gebruikt wordt voor de rente die financieel gezonde banken aan elkaar in rekening brengen en die een zeer laag kredietrisico veronderstelt.

Met t de tijd gemeten in jaren met $t = 0$ vandaag, V_t de waarde van 1 Euro op tijdstip t en r de rente voor één jaar die alleen aan het eind uitgekeerd wordt, kunnen we schrijven:

$$V_1 = 1 + r$$

Wanneer we voor een periode van n jaar kunnen investeren tegen rente r, die aan het eind van iedere jaar uitgekeerd wordt, geldt:

$$V_n = (1 + r)^n$$

Wanneer we meerdere keren per jaar rente bijgeschreven krijgen, bijvoorbeeld m keer per jaar, ontvangen we iedere keer r/m rente. Wanneer we weer voor n jaar investeren geldt:

$$V_n = (1 + r/m)^{nm}$$

Door uitschrijven en beredeneren kun je zien dat het eindbedrag V_n hoger is, naarmate m groter is (voor een vaste rente r en een vaste periode van n jaar). Daarom moeten we wanneer we over een rente r praten, altijd weten wat de conventie is met betrekking tot de frequentie m.

Het blijkt dat we in de praktijk van de financiële wiskunde het makkelijkst kunnen werken wanneer we met een rente werken die als het ware continu wordt bijgeschreven. We nemen dan de volgende limiet:

$$V_n = \lim_{m \to \infty} (1 + r/m)^{nm} = e^{r \cdot n}.$$

De rente r_c met deze conventie heet de continue rente. Deze rente wordt niet weergegeven bij producten die in de markt verhandeld worden, en moeten we daarom zelf vinden. Stel bijvoorbeeld dat in de markt voor een 2-jaars lening een rente r gebruikt wordt, die alleen aan het einde van ieder jaar wordt bijgeschreven. Dan moet voor de 2-jaars r_c gelden:

$$e^{r_c 2} = (1 + r)^2 \Longrightarrow r_c = \frac{1}{2} \ln(1 + r)^2 = \ln(1 + r).$$

Het proces om uit rentes die in de markten gebruikt worden de continue rentes te berekenen, wordt 'bootstrappping' genoemd. Wij zullen in het vervolg van dit artikel ervan uitgaan dat we over de continue rentes beschikken voor alle periodes.

Merk op dat omgekeerd, wanneer we de n-jaars r_c weten en over n jaar een bedrag van 1 Euro ontvangen, dit nu

$$B_0(n) = e^{-r_c n} \tag{1}$$

waard is. $B_0(n)$ wordt ook wel de *contante waarde* (in dit geval van 1 Euro op $t = n$) genoemd.

4. Geen winst zonder risico

Het tweede belangrijke principe in de financiële markten is de aanname dat er geen 'arbitrage' mogelijk is. Wij definiëren het 'geen-arbitrage principe' als:

Er bestaat geen handelsstrategie die winst oplevert zonder risico.

Hoewel er in werkelijkheid soms arbitrage mogelijkheden bestaan, blijkt dat – zeker voor veel verhandelde instrumenten – zulke mogelijkheden van korte duur zijn. Constant zijn er mensen op zoek naar dergelijke mogelijkheden, en wanneer zij handelen om van de arbitrage te profiteren, leidt dit tot een aanpassing in de prijzen dusdanig dat de arbitrage mogelijkheid verdwijnt. Daarom is deze aanname algemeen geaccepteerd. We geven een voorbeeld hoe dit principe de enige juiste prijs van een forward op een aandeel vastlegt.

Bij een forward spreekt (zoals we in paragraaf 2 gezien hebben) een partij A af om op een toekomstig moment (de expiratiedatum) van partij B één aandeel te kopen tegen een vooraf vastgestelde prijs X. Zij S_t de waarde van het aandeel op tijdstip $t, t = 0$ vandaag, $t = T$ de expiratiedatum en $r(t_1, t_2)$ de continue, risicovrije rente op t_1 tot t_2. Doel is nu X te bepalen. Met behulp van het geen-arbitrage principe laten we zien dat er maar één juiste prijs X is.

Partij B loopt op de volgende wijze in deze forward transactie geen enkel risico:

1. Hij leent bedrag S_0 voor periode T en koopt hiervan het aandeel.
2. Hij stelt X vast als het bedrag dat hij terug moet betalen voor deze lening: $X = S_0 \cdot e^{r(0,T)\cdot T}$.
3. Op de expiratiedatum ontvangt hij X van partij A, betaalt daarmee de lening af en hij geeft het aandeel aan A.

Wanneer partij B een hogere prijs $X^* > X$ van A zou vragen, kan hij dus risicoloos een winst van $X^* - X$ op de expiratiedatum maken in deze transactie. Wanneer partij B een lagere prijs $X^* < X$ zou vragen, dan zou partij A het aandeel nu alvast kunnen verkopen voor levering op tijdstip T[1], de opbrengst S_0 risicoloos investeren en op de expiratiedatum na de transactie met B een

[1] Voor particulieren is zo'n verkoop niet mogelijk, voor professionele partijen echter wel tegen de betaling van leenkosten, die we hier 0 hebben verondersteld.

risicoloze winst maken van $X - X^*$. Omdat we aangenomen hebben dat arbitrage niet mogelijk is, is daarom $X = S_0 \cdot e^{r(0,T)\cdot T}$ de enige juiste initiële prijs voor dit forward contract[2]. Merk op dat de waarde van dit forward contract bij het afsluiten voor beide partijen 0 is. Beide partijen hebben immers nog geen winst of verlies geboekt.

Wanneer het contract eenmaal afgesloten is, en X vastgesteld is, verandert in de loop van de tijd de waarde van het forward contract omdat S_0 en r veranderen en de expiratiedatum dichterbij komt. In Figuur 1 zagen we al de waarde van een forward contract op de expiratiedatum als functie van de waarde van het onderliggende goed. Op de expiratiedatum is namelijk de waarde van het forward contract gelijk aan de uitbetaling.

Nu berekenen we de waarde van een forward contract op een tussengelegen tijdstip $0 < t < T$. Zoals we gezien hebben, zal op dat moment de marktprijs van een forward contract $S_t \cdot e^{r(t,T)\cdot(T-t)}$ bedragen. Voor partij A is het prijsverschil tussen de marktprijs en de prijs X die hij eerder heeft afgesproken te betalen op tijdstip T aan partij B dus $S_t \cdot e^{r(t,T)\cdot(T-t)} - X$.

Met behulp van formule (1) kunnen we nu de waarde F_t voor partij A van het forward contract op tijdstip t berekenen als de contante waarde van het prijsverschil:

$$F_t = e^{-r(t,T)\cdot(T-t)} \left\{ S_t \cdot e^{r(t,T)\cdot(T-t)} - X \right\} = S_t - X \cdot e^{-r(t,T)\cdot(T-t)} \qquad (2)$$

Implicaties van de aanname van het geen-arbitrage principe zijn:

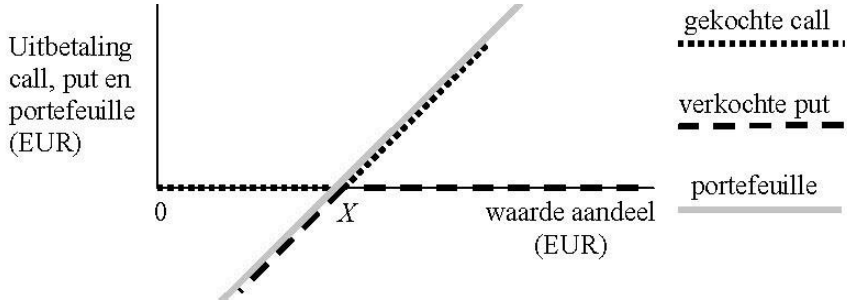
- Als er twee instrumenten zijn met dezelfde uitbetaling, moeten ze ook dezelfde waarde hebben
- Als een portfolio van instrumenten dezelfde uitbetaling heeft als een instrument, moet de waarde van het instrument gelijk zijn aan de waarde van de portfolio. Een dergelijke portefeuille heet een 'replicerende portefeuille'.
- Als we een strategie kunnen bedenken die zelffinancierend is (tijdens de strategie hoeft er geen geld bij en er is ook geen geld over), die dezelfde uitbetaling heeft als een instrument, dan is de waarde van het instrument gelijk aan de initiële kosten van de strategie. Een dergelijke strategie heet een 'dynamische replicerende strategie'.

5. De relatie tussen put- en calloptie

In deze sectie zullen we met behulp van het geen-arbitrage principe een belangrijke relatie afleiden: de zogenaamde 'put-call pariteit'. Deze relatie geeft ons de waarde van een putoptie, wanneer we de waarde van een calloptie weten en andersom.

We gaan uit van een putoptie en een calloptie op hetzelfde onderliggende goed (bijvoorbeeld een aandeel) met dezelfde expiratiedatum T en met dezelfde uitoefenprijs X. Nu beschouwen we een portefeuille bestaande uit een gekochte positie in de calloptie en een verkochte positie in de putoptie. In Figuur 3 geven we de uitbetaling van deze portefeuille op tijdstip T weer.

[2] We hebben hierbij aangenomen dat het aandeel in de periode tot T geen dividend uitkeert.



Figuur 3. *Uitbetaling van een portefeuille met een gekochte calloptie en een verkochte putoptie, beide met uitoefenprijs X op de expiratiedatum.*

Uit een vergelijking van de uitbetaling van de portefeuille in Figuur 3 met Figuur 1, zien we dat de portefeuille in feite een forward is om het aandeel op tijdstip T te kopen tegen een vooraf vastgestelde prijs X. Uit het geen-arbitrage principe volgt nu dat de waarde van de forward gelijk moet zijn aan de waarde van de portefeuille. Zij Π_t de waarde van de portefeuille, F_t de waarde van de forward, P_t de waarde van de putoptie, en C_t de waarde van de calloptie allemaal op tijdstip t. Gebruikmakend van formule (2) geldt daarom[3]:

$$\Pi_t = F_t = C_t - P_t = S_t - Xe^{-r(t,T)\cdot(T-t)} \tag{3}$$

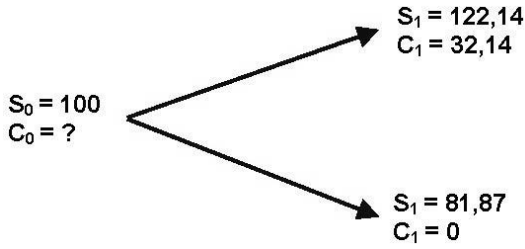
6. Het binomiale model

Het eenvoudigste model om een calloptie te waarderen is het binomiale model. Ook in dit model speelt het geen-arbitrage principe een cruciale rol.

We demonstreren dit model aan de hand van een calloptie met uitoefenprijs van 90 en een expiratiedatum over 1 jaar. We kunnen de uitbetaling in formule weergeven als $C_1 = \max(S_1 - 90, 0)$.

Om te beginnen gaan we ervan uit dat het aandeel nu op 100 genoteerd staat en op de expiratiedatum twee mogelijke waarden kan aannemen: 81,87 of 122,14. Op de keuzes van deze waardes zullen we later terugkomen. In het volgende schema geven we de beide eindwaarden van het aandeel en de bijbehorende waarden van de optie weer:

[3] C_t en P_t zijn gedefinieerd voor gekochte opties. In formule (3) komt er een minteken voor P_t omdat de portefeuille een verkochte putoptie bevat.



Figuur 4. *Eénstaps binomiale boom om de waarde van een calloptie te bepalen.*

We creëren nu een portefeuille die bestaat uit:

- Δ aandelen
- een verkochte calloptie

De huidige waarde van deze portefeuille is $\Pi_0 = \Delta \cdot S_0 - C_0$. We weten nu alleen S_0 en willen Δ, Π_0 en vooral C_0 bepalen.

Nu kiezen we de waarde van Δ zodanig dat de waarde van de portefeuille op de expiratiedatum zonder risico is, dus dezelfde waarde heeft in beide scenario's:

$$\Delta \cdot 122,14 - 32,14 = \Delta \cdot 81,87 - 0 \Longrightarrow \Delta = 0,798$$

De waarde van de portefeuille op de expiratiedatum wordt daarmee: $\Pi_1 = 65,35$ (Per constructie is de waarde van de portfolio gelijk bij zowel een stijging als een daling van het aandeel.)

Volgens het geen-arbitrage principe moet een portefeuille die geen risico loopt evenveel opbrengen als de risicovrije rente. Stel dat de continue, risicovrije rente met een looptijd van 1 jaar 4%, ofwel 0,04 is, dan volgt uit formule (1) de huidige waarde van de portefeuille:

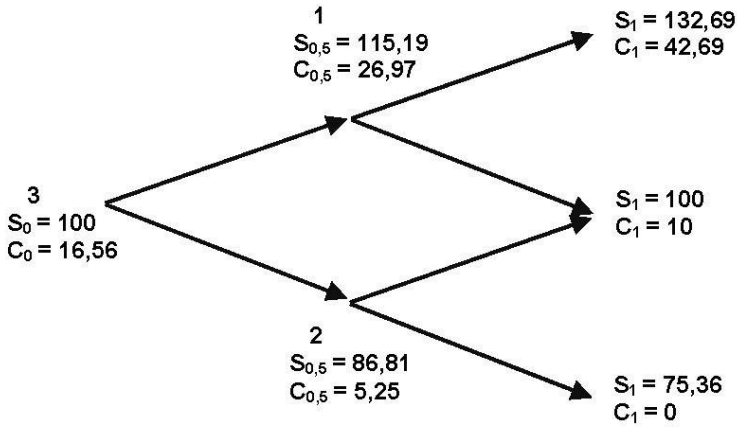
$$\Pi_0 = 65,35 \cdot e^{-0,04} = 62.79$$

Verder is bekend dat de portefeuille bestaat uit Δ aandelen en een verkochte calloptie:

$$\Pi_0 = \Delta \cdot S_0 - C_0 \Longrightarrow 62,79 = 0,798 \cdot 100 - C_0 \Longrightarrow C_0 = 17,03$$

De huidige waarde van de calloptie kan dus alleen Euro 17,03 zijn onder deze aannames.

Het is nogal een grote aanname dat de prijs van een aandeel slechts twee waarden kan aannemen aan het einde van de looptijd van de optie. Maar het binomiale model kan gegeneraliseerd worden door meerdere stappen te combineren. In onderstaande figuur is een voorbeeld gegeven van een binomiale boom met twee stappen van beide een half jaar. De mogelijke waardes die het aandeel kan aannemen zullen nog worden toegelicht.



Figuur 5. *Tweestaps binomiale boom om de waarde van een calloptie te bepalen.*

Om de optie te waarderen kan nu gebruik worden gemaakt van recursie. Op eenzelfde manier als beschreven in het éénstaps model, kan men de waarde van de optie halverwege de looptijd $C_{0,5}$ bepalen. Zoals in de figuur te zien is dienen er hier twee gevallen onderscheiden te worden. Het aandeel na één tijdstap is ofwel 115,19 of 86,81.

(1) $S_{0,5} = 115,19$
$\Rightarrow \Delta = 1; \Pi_1 = 90; \Pi_{0,5} = 90 \cdot e^{-0,5 \cdot 0,04} = 88,22$
$\Rightarrow C_{0,5} = \Delta \cdot 115,19 - \Pi_{0,5} = 26,97$

(2) $S_{0,5} = 86,81$
$\Rightarrow \Delta = 0,406; \Pi_1 = 30,59; \Pi_{0,5} = 30,59 \cdot e^{-0,5 \cdot 0,04} = 29,98$
$\Rightarrow C_{0,5} = \Delta \cdot 86,81 - \Pi_{0,5} = 5,25$

Vervolgens kan men de optie prijs op tijdstip 0 berekenen op een vergelijkbare manier:

(3) $S_0 = 100$
$\Rightarrow \Delta = 07656; \Pi_{0,5} = 61,19; \Pi_0 = 61,19 \cdot e^{-0,5 \cdot 0,04} = 59,98$
$\Rightarrow C_0 = \Delta \cdot 100 - \Pi_0 = 16,56$

Op eenzelfde manier kan men binomiale bomen maken met honderden of zelfs duizenden stappen en een computer programma schrijven om de prijs van een optie te bepalen.

Voordat hiervan een voorbeeld wordt gegeven komen we terug op de keuzes voor de mogelijke waarden van een aandeel in de boom. De reden waarom de waardes op deze manier gekozen zijn heeft te maken met de volatiliteit. Dit is de mate van beweeglijkheid van een aandeel. De waarde van een aandeel met een hoge volatiliteit verandert vaker en met grotere verschillen dan de waarde van een aandeel met een lage volatiliteit. De volatiliteit wordt aangeduid met

Aantal stappen	Vol = 10%	Vol = 20%	Vol = 30%
1	13,53	17,03	21,29
2	13,96	16,56	19,56
10	13,85	16,08	19,39
100	13,83	16,07	19,16
1000	13,83	16,06	19,15
Black Scholes	13,83	16,06	19,15

Tabel 1. *Waarden van een calloptie voor verschillende volatiliteiten (Vol) en aantal stappen in een binomiale boom vergeleken met de Black-Scholes waarde die we in paragraaf 7 zullen beschrijven ($S_0 = 100,\ X = 90,\ r = 0,04,\ T = 1$).*

de letter σ en is gedefinieerd als de standaard deviatie van

$$\frac{1}{\sqrt{\Delta t}} \cdot \ln(\frac{S_{t+\Delta t}}{S_t})$$

voor $\lim \Delta t \downarrow 0$. We zullen dit niet verder bewijzen, maar door het aandeel met een factor $\exp(\sigma \cdot \sqrt{\delta t})$ omhoog te laten gaan per stap in de boom en met een factor $\exp(-\sigma \cdot \sqrt{\delta t})$ omlaag, waarbij δt de stapgrootte in jaren is, kunnen we ervoor zorgen dat het aandeel een volatiliteit van σ krijgt[4]. In voorgaande voorbeelden hebben wij een σ van 0.2 gebruikt.

In Tabel 1 zijn optieprijzen berekend voor drie verschillende volatiliteiten en voor bomen met verschillende aantallen stappen. Verder is ook de Black-Scholes prijs vermeld, die in de volgende paragraaf zal worden behandeld.

7. Het Black-Scholes model

In 1973 is het zogenaamde Black-Scholes model afgeleid, dat gebruikt kan worden voor het waarderen van opties. Dit model kan gezien worden als het continue equivalent van het binomiale boom model zoals hierboven beschreven. De belangrijkste aanname is dat over een kleine tijdstap dt, de verandering van het aandeel dS als volgt kan worden gemodelleerd:

$$dS = \mu \cdot S \cdot dt + \sigma \cdot S \cdot dW \tag{4}$$

Hier kan μ gezien worden als het verwachte rendement van het aandeel en σ is weer de volatiliteit. dW kan gezien worden als een stochastische variabele gerelateerd aan de normale verdeling. Een andere aanname is dat het aandeel ongelimiteerd en zonder kosten verhandeld kan worden.

De afleiding van het Black-Scholes model is gebaseerd op hetzelfde principe als het binomiale boom model. Er wordt een portefeuille gecreëerd met een positie Δ in aandelen en een verkochte positie in de calloptie. Δ kan zó worden gekozen dat de portefeuille risicovrij is. Robert Merton is het geweest die beargumenteerd heeft dat een risicovrije portefeuille de risicovrije rente zou

[4] Dit is een benadering die vooral goed werkt voor δt klein.

moeten opleveren. Een aantal technische operaties leidt vervolgens tot deze waarde van een calloptie:

$$C_0 = S_0 \cdot N(d_1) - X \cdot e^{-rT} \cdot N(d_2)$$

$$d_1 = \frac{\ln(S_0/X) + (r + \frac{1}{2}\sigma^2) \cdot T}{\sigma \cdot \sqrt{T}} \qquad (5)$$

$$d_2 = d_1 - \sigma \cdot \sqrt{T}$$

Hier is $N(\cdot)$ de cumulatieve normale verdelingsfunctie, r de risicovrije, continue rente, X de uitoefenprijs van de optie en T de expiratiedatum. Wat opvalt in de optieformule is dat het verwachte rendement van het aandeel totaal geen invloed op de optieprijs heeft. Dit komt doordat men in staat is een risicovrije portefeuille te creëren. In Tabel 1 geven we een vergelijking van de waarden die we verkrijgen met de binomiale boom en met de Black-Scholes formule voor verschillende waarden van de volatiliteit.

Met behulp van de put-call pariteit (3) kunnen we nu ook de waarde van een putoptie bepalen:

$$P_0 = X \cdot e^{-rT} \cdot N(-d_2) - S_0 \cdot N(-d_1) \qquad (6)$$

waarbij d_1 en d_2 weer bepaald zijn zoals in formule (5).

In de loop der jaren zijn er vele uitbreidingen op dit model geweest. Men kan bijvoorbeeld geïnteresseerd zijn in de waarde van een optie op een aandeel dat dividend uitkeert. Of men is geïnteresseerd in de waarde van de optie wanneer de rente tijdsafhankelijk of zelfs stochastisch is.

Myron Scholes en Robert Merton hebben in 1997 de Nobelprijs voor de economie gekregen voor hun onderzoek naar optiewaardering. Fisher Black was toen al overleden.

8. Risicobeheer

Zoals we in de modellen risico's elimineren door portefeuilles samen te stellen en continu aan te passen, zo worden derivaten ook in de praktijk beheerd.

Als voorbeeld beschouwen we een bank die alle risico's volledig heeft afgedekt, dat wil zeggen de waarde van zijn portefeuille is onafhankelijk van de waarde van onderliggende aandelen en andere risicofactoren.

Vervolgens meldt een klant zich bij de bank die bang is dat de waarde van een zeker aandeel zal dalen en die zich daartegen wil verzekeren. Hiervoor koopt de klant een putoptie op dat aandeel. De putoptie biedt de verzekering die de klant wilde, aangezien de optie uitbetaalt als de waarde van het onderliggende aandeel is gedaald.

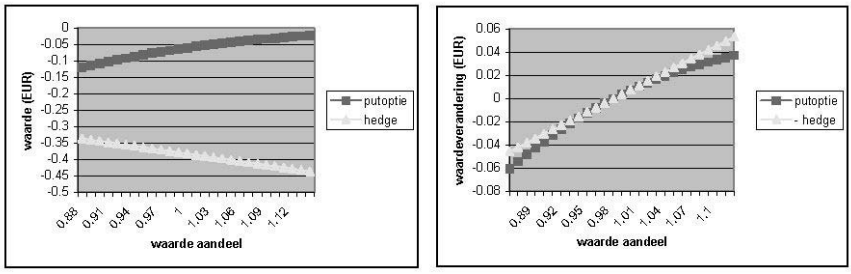
De bank heeft nu een verkochte positie in de putoptie. Dit betekent dat, wanneer het onderliggende aandeel in waarde daalt, de bank verlies lijdt. De bank heeft door de verkoop van een putoptie het risico van de klant overgenomen.

De bank die het risico heeft overgenomen zal niet afwachten of het onderliggende aandeel gaat dalen, maar zal proberen het risico te reduceren. Dit kan op analoge wijze als in het binomiale model de waarde van de optie is afgeleid. Namelijk, de bank verkoopt een hoeveelheid $-\Delta$ onderliggende aandelen. Deze hoeveelheid aandelen noemen we de hedge voor de optie. In het Black-Scholes model is de delta Δ van een putoptie de partiële afgeleide van de waarde van de putoptie (zie vergelijking (6)) naar de waarde van het onderliggende aandeel:

$$\Delta = \frac{\partial P}{\partial S} = N(d_1) - 1.$$

De portefeuille van de bank is dan

$$\Pi = \Delta \cdot S - P.$$



Figuur 6a en b. *Waarde en waardeverandering van de verkochte putoptie en de hedge met* $S_0 = 1, X = 1, \sigma = 0.2, r = 0.04, T = 1$.

In Figuur 6a staan de waarde van de verkochte putoptie en van de hedge, $\Delta \cdot S$ als functie van de waarde van het onderliggende aandeel. In Figuur 6b staat de waardeverandering van de verkochte putoptie en van de hedge ten opzichte van de huidige waarde als functie van de waarde van het onderliggende aandeel, waarbij we de waardeverandering van de hedge met -1 vermenigvuldigd hebben. Uit de figuur kunnen we zien dat voor een kleine verandering in de waarde van het onderliggende aandeel de waardeverandering van de putoptie wordt opgeheven door de waardeverandering van de hedge.

In de loop van de tijd kan de waarde van het onderliggende aandeel flink veranderen (bijvoorbeeld van 1 naar 0,8) en dan werkt de hedge minder goed. In de praktijk wordt daarom regelmatig (typisch dagelijks) het aantal aandelen in de hedge aangepast.

Door bovenstaande strategie heeft de bank het risico van waardeverandering van zijn portefeuille door waardeveranderingen van het aandeel gereduceerd. Maar de bank loopt ook nog andere risico's.

Bijvoorbeeld de volatiliteit van het aandeel σ kan anders blijken te zijn dan verwacht. Dit risico wordt aangegeven met een Griekse letter vega:

$$v = \frac{\partial P}{\partial \sigma}.$$

Dit risico kan niet gereduceerd worden door de koop en verkoop van het onderliggende aandeel, omdat de prijs van het aandeel niet van de volatiliteit afhangt. Dit risico kan wel gereduceerd worden door andere opties (maar op hetzelfde onderliggende aandeel) te kopen of verkopen zodanig dat vega van de totale portefeuille van alle opties op het betreffende aandeel

$$v_{tot} = \frac{\partial \Pi}{\partial \sigma}$$

klein is.

Een derde risico dat de bank loopt komt van veranderingen in de rente. Zoals we in het Black-Scholes model hebben gezien, hangt de waarde van een optie ook van de rente af. Als risicomaat is hiervoor de Griekse letter rho ingevoerd:

$$\rho = \frac{\partial P}{\partial r}.$$

Om het renterisico te reduceren zal een bank de rentegevoeligheid bepalen voor bijvoorbeeld een hele portefeuille van opties op verschillende aandelen en daarna renteproducten afsluiten (bijvoorbeeld leningen) zodanig dat rho voor de totale portefeuille $\rho = \frac{\partial \Pi}{\partial r}$ klein is.

Aangezien de bank continu opties koopt en verkoopt moeten de Griekse letters continu in de gaten worden gehouden. Steeds past de bank de portefeuille met onderliggende aandelen, andere opties en renteproducten aan zodat er zo weinig mogelijk risico wordt gelopen. Merk op dat in feite de bank steeds op zoek is naar een portefeuille die de verplichtingen aan de klanten zo goed mogelijk repliceert. Daarom wordt de hedge portefeuille ook wel een replicerende portefeuille genoemd.

9. Verdere ontwikkelingen

In voorgaande paragrafen hebben we een model beschreven dat er – onder bepaalde aannames – in zal slagen een dynamische strategie aan te geven zodat de uitbetaling van een standaard call of putoptie zo goed mogelijk gerepliceerd kan worden met een portefeuille van onderliggende instrumenten. Er zijn echter een aantal redenen waarom het bovenstaande model continu verder ontwikkeld wordt.

Een van de belangrijkste aannames in het Black-Scholes model is dat de volatiliteit van het onderliggende aandeel constant is. Uit historisch onderzoek is gebleken dat deze aanname niet volkomen juist is. Bijvoorbeeld in sterk dalende aandelenmarkten neemt de volatiliteit over het algemeen toe. Ook de andere aannames van het Black-Scholes model zijn maar niet volledig geldig. Dit heeft ertoe geleidt dat het Black-Scholes model in praktijk op een iets andere manier gebruikt wordt dan hierboven is beschreven. In plaats van een zo goed mogelijke schatting te doen van de verwachte volatiliteit van het aandeel, en met deze volatiliteit dan call en putopties te waarderen voor alle mogelijke expiratiedata en uitoefenprijzen, kent de markt een ‘eigen’ volatiliteit toe voor iedere afzonderlijke optie. Op deze manier is het Black-Scholes model voor

call en putopties meer een 'vertaalmechanisme' geworden tussen volatiliteit en waarde dan een model die gegeven de parameters de waarde bepaalt.

In dit artikel beperkten wij ons tot standaard, zogenaamde 'vanilla' opties. In de praktijk worden vele soorten 'exotische' opties verkocht. Een aantal voorbeelden:

- De uitbetaling is afhankelijk van een gemiddelde in plaats van een eindwaarde, of van een combinatie van beide.
- De optie kan niet alleen aan het einde van de looptijd, maar altijd, of op vooraf vastgestelde tussengelegen momenten uitgeoefend worden.
- De optie houdt op te bestaan wanneer de onderliggende een bepaalde waarde bereikt.
- De optie hangt af van meerdere onderliggende waarden, waardoor hun correlatie van belang is. Hierbij kun je denken aan een optie op meerdere aandelen, aan een optie op een aandeel met een hoofdsomgarantie, maar ook aan een optie waarbij rentes met verschillende looptijden een rol spelen.
- Bepaalde parameters worden pas op een toekomstig moment vastgesteld.

Ook voor zulke exotische opties probeert men modellen te ontwerpen die een dynamische hedge strategie aangeven. Vooral vanwege de aanname van constante volatiliteit blijkt – zelfs voor derivaten die maar afhankelijk zijn van één onderliggende waarde – model (1) soms verbeterd te kunnen worden. Bijvoorbeeld:

- De volatiliteit wordt zelf tot een stochastisch proces gemaakt, al dan niet gecorreleerd met het proces voor de onderliggende waarde.
- De onderliggende kan sprongen vertonen met een bepaalde intensiteit en een verdeling voor de spronggrootte.
- Bij meerdere onderliggenden kan niet alleen de 'gewone' correlatie, maar een hele correlatiestructuur worden meegenomen met behulp van zogenaamde 'copulas'.

Echter, ook voor zulke meer complexe modellen speelt het creëren van een risicoloze portefeuille en het geen-arbitrage principe een cruciale en leidende rol.

Vakantiecursus 2006

pp. 47 – 57

Reken mee met ABC

G.C. Geuze, B. de Smit
Mathematisch Instituut, Universiteit Leiden
e-mail: geuze@math.leidenuniv.nl desmit@math.leidenuniv.nl

Het ABC-vermoeden is één van de belangrijkste open problemen in de getaltheorie. De oorsprong van het vermoeden ligt in diepe vermoedens over elliptische krommen, die uiteindelijk ook hebben geleid tot het bewijs van de Laatste Stelling van Fermat. De formulering van het ABC-vermoeden is echter volkomen elementair en de lezer hoeft er niet meer voorkennis voor te hebben dan kennis van optellen van gehele getallen en de factorisatie van een getal in priemfactoren. De letters ABC verwijzen naar de eenvoudige vergelijking waar het vemoeden over gaat:

$$a + b = c.$$

Het ABC-vermoeden doet een uitspraak over hoe de priemfactorisaties van gehele getallen a, b, c die aan deze vergelijking voldoen met elkaar samenhangen.

In de herfst van 2006 wordt het project *Reken mee met ABC* gelanceerd door Kennislink en de *Universiteit Leiden*. Het doel is tweeledig. Een breed publiek kan via dit project kennis maken met moderne wiskunde die volop in ontwikkeling is. Bovendien kan iedereen met een computer bijdragen aan het verzamelen van experimentele gegevens over het ABC- vermoeden, die dit onderzoek weer verder kunnen helpen.

In dit hoofdstuk volgt een korte uiteenzetting over het vermoeden, de wiskundige achtergrond en geschiedenis, enige recente ontwikkelingen, met name van algorithmische soort, en een beschrijving van de activiteiten in het project *Reken mee met ABC*.

1. Het ABC-vermoeden

Om het vermoeden te kunnen formuleren moeten we eerst weten dat het *radicaal* $r(n)$ van een positief geheel getal n gedefinieerd is als het product van de priemdelers van n:

$$r(n) = \prod_{p|n \text{ priem}} p.$$

Bijvoorbeeld, als $n = 6000 = 2^4 \cdot 3 \cdot 5^3$, dan is $r(n) = 2 \cdot 3 \cdot 5 = 30$. We kunnen $r(n)$ ook definiëren zonder de priemfactorisatie van n te gebruiken: het is de grootste deler van n die zelf niet door een kwadraat groter dan 1 deelbaar is.

Stel nu dat a, b, en c positieve gehele getallen zijn waarvoor geldt $a+b=c$. We vragen ons nu af hoe klein het getal $r = r(abc)$ kan zijn in verhouding tot c. Om deze vraag preciezer te kunnen formuleren, definiëren we de *kwaliteit* $q(a,b,c)$ als het reële getal q met $c = r^q$. Met andere woorden: $q = \log(c)/\log(r)$.

De eerste observatie is enigszins flauw: als a, b en c een gemeenschappelijke factor hebben die een hoge macht is, dan kan de kwaliteit willekeurig hoog worden. Bijvoorbeeld, als $a = b = 2^{99}$, dan is $c = 2^{100}$, $r = 2$ en $q = 100$. We zullen ons daarom beperken tot het geval dat de getallen a, b en c geen gemeenschappelijke deler groter dan 1 hebben.

Met deze beperking is de hoogst bekende kwaliteit op dit moment het record uit 1987 van de Franse wiskundige Eric Reyssat:

$$2 + 3^{10} \cdot 109 = 23^5 \qquad q = \frac{5\log(23)}{\log(2 \cdot 3 \cdot 109 \cdot 23)} = 1,6929912\ldots.$$

Het ABC-vermoeden zegt dat deze kwaliteit nauwelijks boven de 1 komt, in de volgende zin:

1.1. ABC-vermoeden

Voor elk reëel getal $\varepsilon > 0$ zijn er hoogstens eindig veel drietallen positieve gehele getallen a, b, c met

$$\begin{aligned} &a + b = c;\\ &\text{ggd}(a,b,c) = 1;\\ &q(a,b,c) > 1 + \varepsilon. \end{aligned}$$

Van dit vermoeden verwachten alle deskundigen dat het waar is, en ook dat het bewijs ver buiten het bereik van bestaande wiskundige technieken ligt.

Het is daarentegen betrekkelijk eenvoudig om oneindig veel drietallen te maken waarbij de kwaliteit boven de 1 ligt. Zo'n drietal noemen we een ABC-drietal.

1.2. Definitie

Een ABC-drietal is een drietal positieve gehele getallen a, b, c waarvoor geldt

$$a + b = c;$$
$$\text{ggd}(a, b, c) = 1;$$
$$q(a, b, c) > 1.$$

De laatste van deze drie voorwaarden is equivalent met $r(abc) < c$. De ABC-drietallen die bestaan uit getallen onder de 100 zijn $1 + 8 = 9$, $5 + 27 = 32$, $1 + 48 = 49$, $32 + 49 = 81$, $1 + 63 = 64$ en $1 + 80 = 81$.

1.3. Stelling
Er zijn oneindig veel ABC-drietallen.

Bewijs. Voor elke $n \geq 1$ beschouwen we het drietal $a = 1$, $b = 9^n - 1$ en $c = 9^n$. Deze getallen voldoen duidelijk aan de eerste twee voorwaarden uit de definitie. Om te zien dat het radicaal van abc klein genoeg is, merken we eerst op dat $r(abc) = 3r(b)$. Omdat 9 een 8-voud plus 1 is, is ook 9^n een 8-voud plus 1 en is b deelbaar door 8. Hieruit volgt dat 4 een deler is van $b/r(b)$ en dat $r(b) \leq b/4$. Hieruit volgt dat $r(abc) = 3r(b) \leq \frac{3}{4}b < c$. □

De laatste ongelijkheid uit het bewijs laat zich makkelijk vertalen in een ondergrens voor de kwaliteit:

$$q(1, 9^n - 1, 9^n) > 1 + \frac{\log(4/3)}{2n \log(3)} > 1 + \frac{1}{8n}.$$

Een bovengrens voor de kwaliteit van deze drietallen is niet eenvoudig te geven: we weten niet eens dat deze begrensd is als functie van n. Maar het ABC-vermoeden impliceert dat deze kwaliteit naar 1 convergeert als $n \to \infty$.

1.4. De laatste Stelling van Fermat
Het ABC-vermoeden is een zeer sterke uitspraak, waarvan niemand verwacht dat er een eenvoudig bewijs voor zal opduiken. De kracht van de uitspraak blijkt bijvoorbeeld uit de implicaties van het ABC-vermoeden voor de notoir moeilijke vergelijking van Fermat:

$$x^n + y^n = z^n.$$

Voor elke $n \geq 3$ geeft een oplossing in positieve gehele onderling ondeelbare getallen x, y, z aanleiding tot een ABC-drietal $a = x^n$, $b = y^n$, $c = z^n$ met kwaliteit

$$\frac{\log c}{\log r(abc)} = \frac{n \log z}{\log r(xyz)} > \frac{n}{3} \geq 1.$$

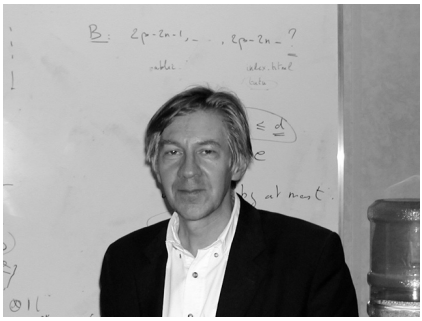
Het ABC-vermoeden impliceert dus dat voor alle n vanaf een zekere grens de Fermat-vergelijking geen oplossingen heeft.

2. Ontstaan van het vermoeden

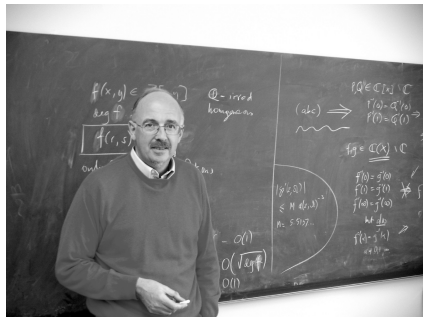
De oorsprong van het ABC-vermoeden ligt bij de Franse wiskundige Joseph Oesterlé van het *Institut de Mathématiques de Jussieu* in Parijs. In de jaren tachtig brachten zijn pogingen om het zogenaamde Taniyama-Weil vermoeden te bewijzen voor bepaalde *elliptische krommen* hem tot ongelijkheden waar die elliptisch krommen aan zouden moeten voldoen: de *discriminant* van de kromme moest begrensd zijn door een macht van de *conductor*.

We doen hier geen poging om uit te leggen wat deze begrippen betekenen; dat hoort thuis in een gevorderd doctoraalcollege. Maar voor een drietal onderling ondeelbare gehele getallen a, b, c met $a + b + c = 0$ kan men een elliptische kromme maken waarvan de discriminant $(abc)^2$ is, en de conductor $r(abc)$. Oesterlé formuleerde zijn ongelijkheid in dit geval als volgt: er zijn constanten λ en C, zodat voor al zulke drietallen a, b, c geldt:

$$(abc)^2 \leq Cr(abc)^{\lambda}.$$



Oesterlé



Masser

David Masser van de *Universität Basel* woonde in 1985 een lezing bij van Oesterlé aan het *Max-Planck-Institut für Mathematik* in Bonn. Hij begon over de ongelijkheid na te denken los van de context van elliptische krommen. Hij gaf deze equivalente formulering: er zijn constanten C' en λ' zodat voor alle onderling ondeelbare positieve gehele getallen a, b, c met $a + b = c$ geldt:

$$\max\{|a|, |b|, |c|\} \leq C'r(abc)^{\lambda'}.$$

Er is een bekende analogie tussen gehele getallen en polynomen over de complexe getallen, waarbij de graad van een polynoom de rol van absolute waarde van een getal speelt en het aantal nulpunten de rol van het radicaal. Het analogon van de laatste ongelijkheid luidt dan als volgt: er zijn getallen λ'', C'', zodat voor alle polynomen f, g, h zonder gemeenschappelijke nulpunten met $f + g + h = 0$ geldt:

$$\max\{\text{graad}(f), \text{graad}(g), \text{graad}(h)\} \leq \lambda'' Z(fgh) + C'',$$

waarbij $Z(fgh)$ het aantal nulpunten van het polynoom fgh is (geteld *zonder* multipliciteiten). Vergeleken met de vorige ongelijkheid zijn de vermenigvuldigingen vervangen door optellingen omdat de graad van een product van twee

polynomen de *som* van hun graden is, terwijl de absolute waarde van het product van twee getallen het *product* van hun absolute waarden is.

Masser herinnerde zich een vergelijkbare *bewezen* ongelijkheid van Richard Mason, die gepubliceerd is in de *proceedings* van de grote getaltheorieconferentie in Noordwijkerhout in 1983: als de polynomen niet constant zijn, dan geldt

$$\max\{\text{graad}(f), \text{graad}(g), \text{graad}(h)\} \leq Z(fgh) - 1.$$

Blijkbaar mogen we $\lambda'' = 1$ en $C'' = -1$ nemen. De rol van de constante C'' is ondergeschikt aan die van λ''. Masser vroeg zich af wat de "correcte" waarde van λ' moest zijn. Hij wist dat de letterlijke terug-vertaling $\lambda' = 1$ teveel van het goede was, en hij vroeg zich af of elke $\lambda' > 1$ wel zou voldoen. Op een symposium aan het Imperial College in London in 1985 ter ere van de wiskundige K. F. Roth formuleerde hij het als volgt in de lijst öpen problemen":

PROBLEM OF D.W. MASSER (After Oesterlé)

Disprove (or prove) that for every $\varepsilon > 0$ there exists $C(\varepsilon)$ such that

$$\max(|a|,|b|,|c|) \leq C(\varepsilon)\left(\prod_{p \mid abc} p\right)^{1+\varepsilon}$$

for all coprime integers a,b,c with a + b + c = 0.

De bewering in Masser's opgave is equivalent met onze formulering van het ABC-vermoeden. Uit de formulering blijkt dat hij eerder verwachtte dat dit weerlegd zou worden, dan bewezen. En hoewel het bewijs nog steeds mijlen ver weg lijkt te liggen, raken steeds meer experts ervan overtuigd dat het vermoeden waar is.

3. Takken van sport

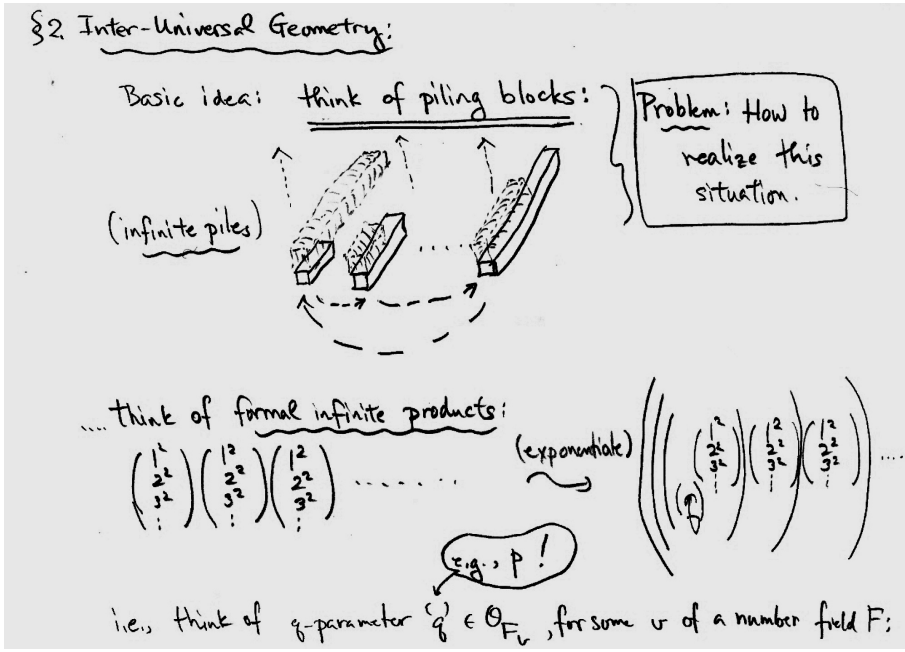
Het ABC-vermoeden heeft aanleiding gegeven tot diverse takken van sport. De kracht van het ABC-vermoeden maakt het buitengewoon geschikt om er allerlei mooie gevolgen uit af te leiden. Er zijn ook veel mogelijkheden om het vermoeden te formuleren in een algemenere context. Zie de webpagina van Abderrahmane Nitaj (Caen, Frankrijk)

`http://www.math.unicaen.fr/~nitaj/abc.html`

voor een uitgebreide opsomming. We lichten hier een paar andere onwikkelingen uit.

3.1. Op weg naar een bewijs?

De Japanse wiskundige Shinichi Mochizuki van de universiteit van Kyoto is sinds 2000 bezig met een ambitieus programma dat tot doel heeft het ABC-vermoeden te bewijzen. Hiervoor ontwikkelt hij een buitengewoon abstracte theorie die hij "inter-universele meetkunde" noemt; zie



Schets op de webpagina van Mochizuki (2006)

`http://www.kurims.kyoto-u.ac.jp/~motizuki/research-english.html`

Het idee is om het instrumentarium dat in Mason's context van polynoomringen tot een bewijs leidt, over te planten naar de gehele getallen. Het is duidelijk dat dit niet zonder slag of stoot zal lukken. Maar als Mochizuki in zijn opzet slaagt, dan zal dat de wereld van de getaltheorie danig op zijn kop zetten.

3.2. Recordjacht

Benne de Weger ontwikkelde in 1985 als bijproduct van zijn promotieonderzoek aan de *Universiteit Leiden* een methode om ABC-drietallen van hoge kwaliteit te maken. Hij ontdekte 14 ABC-drietallen met kwaliteit boven de $1,4$. De grens van $1,4$ heeft hij naar eigen zeggen uit zijn duim gezogen, maar desalniettemin staan ABC-drietallen met kwaliteit boven de $1,4$ nu internationaal bekend als "goede ABC-drietallen". Er zijn er nu zo'n 200 bekend en op internet wordt de laatste stand van zaken bijgehouden op

`http://www.minet.uni-jena.de/~aros/abc.html`

Nieuwe goede ABC-drietallen worden soms gevonden door verbeterde hardware en verbeterde implementaties van bekende zoekmethoden en soms door het ontdekken van geheel nieuwe methoden. Zo zijn er in 2003 enige tientallen gevonden door Tim Dokchitser met een methode van Jaap Top van de *Rijksuniversiteit Groningen*. De top-tien van ABC-drietallen met de hoogst bekende kwaliteit is als volgt:

a	b	c	q	ontdekker
2	$3^{10} \cdot 109$	23^5	1,630	Eric Reyssat
11^2	$3^2 \cdot 5^6 \cdot 7^3$	$2^{21} \cdot 23$	1,626	Benne de Weger
$19 \cdot 1307$	$7 \cdot 29^2 \cdot 31^8$	$2^8 \cdot 3^{22} \cdot 5^4$	1,623	Browkin Brzezinski
283	$5^{11} \cdot 13^2$	$2^8 \cdot 3^8 \cdot 17^3$	1,581	Browkin Brzezinski, Nitaj
1	$2 \cdot 3^7$	$5^4 \cdot 7$	1,568	Benne de Weger
7^3	3^{10}	$2^{11} \cdot 29$	1,547	Benne de Weger
$7^2 \cdot 41^2 \cdot 311^3$	$11^{16} \cdot 13^2 \cdot 79$	$2 \cdot 3^3 \cdot 5^{23} \cdot 953$	1,544	Nitaj
5^3	$2^9 \cdot 3^{17} \cdot 13^2$	$11^5 \cdot 17 \cdot 31^3 \cdot 137$	1,537	te Riele, Montgomery
$13 \cdot 19^6$	$2^{30} \cdot 5$	$3^{13} \cdot 11^2 \cdot 31$	1,527	Nitaj
$3^{18} \cdot 23 \cdot 2269$	$17^3 \cdot 29 \cdot 31^8$	$2^{10} \cdot 5^2 \cdot 7^{15}$	1,522	Nitaj

Als het ABC-vermoeden waar is, dan is er een ABC-drietal waarvoor geen enkel ander ABC-drietal een hogere kwaliteit heeft. Experts achten het niet onwaarschijnlijk dat Reyssat's drietal de absolute kampioen is.

3.3. Een betere maat voor kwaliteit?

Voor elk reëel getal $q > 1$ waarvoor er een ABC-drietal bestaat met kwaliteit minstens q, impliceert het ABC-vermoeden dat er een grootste ABC-drietal is met kwaliteit minstens q. Van alle ABC-drietallen a, b, c met kwaliteit dicht bij q vinden we die met de grootste c daarom de "beste".

Dit wordt tot uitdrukking gebracht in een ander kwaliteitsmaat voor ABC-drietallen. We definiëren de *verdienste* $m(a, b, c)$ van een ABC-drietal a, b, c als

$$m = m(a, b, c) = (q - 1)^2 (\log r) \log \log r,$$

waarbij $q = q(a, b, c)$ en $r = r(abc)$. Nu geldt

$$\frac{c}{r} = \exp \sqrt{m \frac{\log r}{\log \log r}}.$$

De aanleiding voor deze definitie is nog ongepubliceerd werk van Cameron L. Stewart van de *University of Waterloo* in Canada en Gérald Tenenbaum van het *Institut Élie Cartan* in Nancy in Frankrijk. Zij hebben een verfijnd ABC-vermoeden geformuleerd, dat zegt dat er voor elke $m > 48$ maar eindig veel ABC-drietallen zijn met verdienste minstens m, en dat er een oneindige rij ABC-drietallen bestaat waarvan de verdienste convergeert naar 48. De hoogst bekende verdienste is van het drietal in de tabel bij 3.2 met de grootste c:

$$m(7^2 \cdot 41^2 \cdot 311^3, 11^{16} \cdot 13^2 \cdot 79, 2 \cdot 3^3 \cdot 5^{23} \cdot 953) = 31,54.$$

Het verfijnde ABC-vermoeden is gebaseerd op een subtiele heuristiek die zegt dat $r(a + b)$ zich onafhankelijk gedraagt van $r(a)$ en $r(b)$ als a en b onderling

ondeelbaar zijn en op bewezen stellingen over het gedrag van de functie $n \mapsto r(n)$.

Deze heuristieken hebben meer interessante gevolgen. Gegeven een reëel getal $q > 1$ zegt het ABC-vermoeden dat er maar eindig veel ABC-drietallen zijn met kwaliteit minstens q. Maar hoe groot kan zo'n drietal dan maximaal zijn? Als we q van boven naar 1 laten naderen, dan voorspelt het verfijnde ABC-vermoeden hoe hard het grootste ABC-drietal met kwaliteit minstens q zal groeien.

3.4. ABC-drietallen tellen

Voor elk getal X noteren we het aantal ABC-drietallen (a, b, c) met $c < X$ als $N(X)$. We weten uit Stelling 1.3 dat $N(X)$ naar oneindig gaat als X naar oneindig gaat. Maar hoe hard gaat $N(X)$ naar oneindig? Men kan het volgende bewijzen: voor elke $\varepsilon > 0$ zijn er $C, C' > 0$ zodat

$$C \exp((\log X)^{1/2-\varepsilon}) \leq N(X) \leq C' X^{2/3+\varepsilon}.$$

Het bewijs van deze stelling hoort thuis in het deelgebied van de wiskunde dat bekend staat als "analytische getaltheorie". Het is vooralsnog niet gepubliceerd. De ondergrens is recent aangekondigd door Sander Dahmen, promovendus aan de *Universiteit Utrecht* en de bovengrens is gevonden naar aanleiding van de algoritmische inspanningen van het project *Reken mee met ABC*; zie hieronder.

Het ligt in de verwachting dat het daadwerkelijke asymptotische gedrag van $N(X)$ dichter bij de ondergrens dan bij de bovengrens zal liggen.

X	$N(X)$
10	1
10^2	6
10^3	31
10^4	120
10^5	418
10^6	1268
10^7	3499
10^8	8987
10^9	22316
10^{10}	51677
10^{11}	116978
10^{12}	252856

De 31 ABC-drietallen onder de 1000						
1+8	=	9	32+343	=	375
5+27	=	32	5+507	=	512
1+48	=	49	169+343	=	512
1+63	=	64	1+512	=	513
1+80	=	81	27+512	=	539
32+49	=	81	1+624	=	625
4+121	=	125	49+576	=	625
3+125	=	128	81+544	=	625
1+224	=	225	1+675	=	676
1+242	=	243	1+728	=	729
2+243	=	245	25+704	=	729
7+243	=	250	104+625	=	729
13+243	=	256	200+529	=	729
81+175	=	256	1+960	=	961
1+288	=	289	343+625	=	968
100+243	=	343			

3.5. Een net door de zee

Als onderdeel van het project *Reken mee met ABC* wordt deze herfst een project voor gedistribueerd rekenen gestart dat tot doel heeft om alle ABC-drietallen tot een bepaalde grens te vinden. Het gaat hierbij dus niet om speciale methoden die speciaal wenselijke ABC-drietallen moeten produceren, maar om een methode die geen enkel ABC-drietal overslaat. Hiervoor is een algoritme ontwikkeld dat op een efficiënte wijze het gebied van paren (a, b) met $a+b < X$ afzoekt. Experimenten van Jeroen Demeyer van de *Universiteit Gent* met dit algoritme hebben geleid tot de bovenstaande tabel met precieze waarden voor $N(X)$ en tot een zestal nieuwe "goede" ABC-drietallen, met kwaliteit boven $1,4$:

a	b	c	kwaliteit
$2^7 \cdot 89^2$	$5^4 \cdot 7^6 \cdot 11^2 \cdot 71^4$	$3^{13} \cdot 19^3 \cdot 4547^2$	1,4342
$2^{32} \cdot 73^3$	$3^{14} \cdot 5^3 \cdot 11 \cdot 13^5 \cdot 557$	$7^{13} \cdot 23^2 \cdot 163^2$	1,4323
2^{39}	$3^8 \cdot 13^2 \cdot 23^5$	$5^8 \cdot 151^2 \cdot 863$	1,4210
$3^{17} \cdot 89^4$	$7^3 \cdot 61 \cdot 359^5$	$2^{13} \cdot 5 \cdot 19^8 \cdot 191$	1,4068
$2^{27} \cdot 17^2$	$7^{11} \cdot 3041^2$	$3 \cdot 5^3 \cdot 13^8 \cdot 23^2 \cdot 113$	1,4062
1	$3^7 \cdot 7^5 \cdot 13^5 \cdot 17 \cdot 1831$	$2^{30} \cdot 5^2 \cdot 127 \cdot 353^2$	1,4012

Om ABC-drietallen op een grotere schaal te kunnen zoeken worden de rekentaken gedistribueerd via de Berkeley Open Infrastructure for Network Computing, waarmee iedereen thuis op zijn computer ABC-drietallen kan vinden en toevoegen aan de statistieken. Dit project zal in de herfst van 2006 van start gaan.

4. Scholen

Het project *Reken mee met ABC* probeert scholieren en docenten op verschillende manieren te betrekken bij de moderne getaltheorie. Het poogt de indruk weg te nemen dat de wiskunde al "af" is door juist de open problemen te benadrukken. Scholen kunnen meedoen met het vinden van nieuwe ABC-drietallen, en zo een steentje bijdragen.

Veel elementaire onderwerpen uit de wiskunde die in nauw verband staan met het ABC-vermoeden hebben een kleurrijke geschiedenis die doorloopt tot op de dag van vandaag. Deze zijn uitstekend geschikt om de cultuur van de wiskunde mee voor het voetlicht te brengen, zowel binnen het klaslokaal en daarbuiten. Daarom zullen deze onderwerpen op verschillende niveaus in de vorm van lesbrieven aangeboden worden op de website. We denken hierbij aan de volgende onderwerpen: De lesbrieven zijn met name bedoeld om gebruikt te worden in de les, voor een praktische opdracht of voor een profielwerkstuk vanaf groep 8 van de basisschool tot en met HBO. Op de volgende bladzijde volgt een voorbeeld ter illustratie.

	ABC-drietallen	
delers	Mersenne-priemgetallen	modulo rekenen
machten	logaritmen	Diophantische vergelijkingen
priemgetallen	Pythagoreïsche drietallen	het vermoeden van Goldbach
Fermat	de vergelijking van Pell	het vermoeden van Hall
Catalan	Vermoeden van Mordell	de stelling van Dirichlet

Voor uitgebreide achtergrondinformatie verwijzen we de lezer naar de webpagina van het project *Reken mee met ABC*:

`http://www.rekenmeemetabc.nl`.

Wat zijn delers?

We kunnen gehele getallen op elkaar delen en soms is het antwoord dan ook een geheel getal, maar soms komen er breuken voor in het antwoord.

16 : 2 = 8, 15 : 3 = 5, 124 : 4 = 32 enz, maar……,
19 : 3 = 6 met rest 1 of te wel 6 1/3.
127 : 5 = 25 met rest 2 of te wel 25 2/5.

Je kan nu bij het eerste voorbeeld zeggen: 2 is een deler van 16 omdat 2 * 8 = 16.
Het getal 16 noemen we een veelvoud van 2.
We noemen een getal een deler van een ander getal als het antwoord van de deling gelijk is aan een geheel getal.

Definitie: a is een deler van b als geldt b = a * x waarbij x een geheel getal is.

In ons voorbeeld geldt dus: a = 2, b = 16 en x = 8.

Elk getal is deelbaar door 1 en door zichzelf. Sommige getallen hebben meer delers.
De delers van 16 zijn 1, 2, 4, 8, 16.

Oefeningen:

Is er sprake van delers of niet?

a) Is 2 een deler van 5?	c) Is 24 een deler van 38?	e) Is 2^5 een deler van 2^{10}?	g) Is 3 een deler van 0?
b) Is 2 een deler van -6?	d) Is 19 een deler van 323?	f) Is 2 een deler van 5?	h) Is 0 een deler van 3?

Veelvouden
a) Als *a* en *b* veelvouden zijn van een getal *d,* geldt dan hetzelfde voor *a + b* en *a - b*? Geef een tegenvoorbeeld of probeer het te bewijzen.
b) Als *a* en *b* delers zijn van *c,* geldt dan ook dat *(a * b)* een deler is van *c*?

Eén deler
De getallen 5018569, 5561527 en 6374939 hebben precies één deler gemeenschappelijk (behalve 1). Welk getal is dat?

Alle delers
Wat zijn alle delers van:

a) 30	d) 56	g) 7	j) 360
b) 14	e) 18	h) 15	k) 1470
c) 2^{10}	f) 31	i) 53	l) 121

Vakantiecursus 2006

pp. 59 – 60

Zonnezeilen naar de polen van de zon

B.A.C. Ambrosius
Technische Universiteit Delft
e-mail: b.a.c.ambrosius@lr.tudelft.nl

Deze case is gebaseerd op het werk van Daniëlle Garot, afgestudeerd op 13 april 2006.

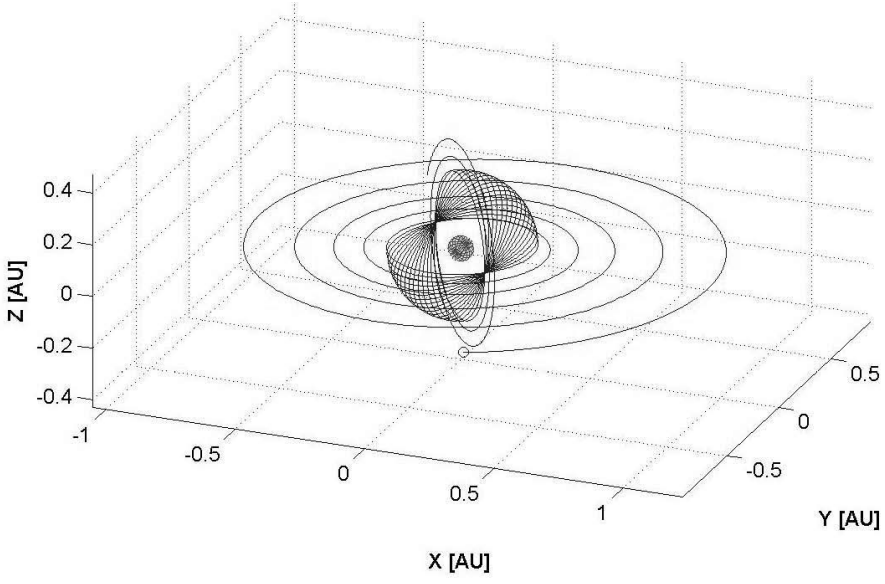
De zon is allesbepalend voor leven op aarde: zij verschaft energie, in de vorm van licht en warmte. De oorzaken en details van vele processen op en in de zon zijn ons echter nog grotendeels onbekend. Een eerste vereiste voor een beter begrip is om een volledige bemonstering van dergelijke fenomenen te maken; op dit moment zijn onze waarnemingen aan de zon vrijwel volledig beperkt tot een 2-dimensionaal plaatje en verschaft alleen de satelliet Ulysses ons enig inzicht in de derde ruimtelijke component (met een regelmaat van slechts 2 passages over de polen per 6 jaar). Deze beperkte bemonstering vindt z'n oorzaak grotendeels in de enorme hoeveelheid energie die het kost om een ruimtevoertuig in een (circulaire) baan over de polen van de zon te brengen: onze lanceerders zijn domweg niet krachtig genoeg om een dergelijke baan te bereiken. Een uitkomst voor dit probleem is de techniek van het zonnezeilen: fotonen dragen hun impuls bij botsing over aan een satelliet en veroorzaken zo een 'lichtkracht' op dat voertuig. Weliswaar is deze kracht uiterst gering, maar bij langdurige blootstelling en goede eigenschappen van het voertuig kan e.e.a. wel degelijk resulteren in snelheden die voor traditionele lanceervoertuigen onbereikbaar zijn.

De missie naar de polen van de zon heeft een aantal specifieke doelstellingen. Allereerst moet het ruimtevoertuig terechtkomen in een circulaire baan op een afstand van 0,4 Astronomische Eenheid (AE) van de zon en met een inclinatie (helling) van 90 ten opzicht van het eclipticavlak (het gemiddelde baanvlak van de beweging van de aarde om de zon). Om de zonneactiviteit onder gunstige condities te bemonsteren moet het voertuig aldaar arriveren omstreeks januari 2020, 2 jaar voor het verwachte optreden van een zogenaamd 'Solar Maximum'. Het beginpunt voor deze studie is de zgn. GTO-baan, een baan om de aarde die de overgang levert van een circulaire parkeerbaan op lage hoogte naar de (eveneens circulaire) geostationaire baan.

In de studie is de missie onderverdeeld in een aantal fases. In de eerste fase moet het voertuig op eigen gelegenheid (lees: met eigen voorstuwingstechnieken, dus zonnezeilen) ontsnappen aan het aardse zwaartekrachtsveld. In de tweede fase vliegt het voertuig tot relatief dicht bij de zon ('inspiraliseren' en recircularisatie), waarna de zgn. 'cranking' fase volgt: het 'opkrikken' van de inclinatie (helling) van het baanvlak tot de waarde van 90. Tenslotte wordt het voertuig naar een afstand van de vereiste 0,4 AE gebracht.

De studie is te beschouwen als een optimalisatieprobleem: op welke datum, met welke stuurhoeken, in welke vertrekgeometrie en met welke dichtste nadering tot de zon kan deze missie worden uitgevoerd, dusdanig dat de totale missiekosten (ontwikkeling, bouw, lancering en operaties) geminimaliseerd worden? En: wat is de invloed van de hoeveelheid nuttige lading (instrumentatie) die aan boord worden meegenomen? Deze optimalisatie is uitgevoerd met een zgn. Evolution Program, een variant van de techniek van Genetische Algoritmen. In dergelijke berekeningen worden toeval en combinatie verenigd en kunnen goede oplossingen binnen een redelijke rekentijd gevonden worden, ondanks de vele onbekenden en de sterk niet-lineaire verbanden.

In het onderhavige geval is aangetoond dat de missie in staat is om de satelliet op het gewenste tijdstip in de vereiste baan af te leveren, voor totale missiekosten van 226 miljoen dollar. Voor het instrumentarium is 5 kg beschikbaar. De totale vluchtduur (tot aan het moment van injectie in de operationele baan om de zon) bedraagt 6 jaar.



Figuur 1. *Een illustratie van de gevonden baan*

Vakantiecursus 2006

pp. 61 – 82

Hoe onstaan Tsunami's en waarom?

C.M. Dohmen-Janssen
Universiteit Twente
e-mail: c.m.dohmen-janssen@ctw.utwente.nl

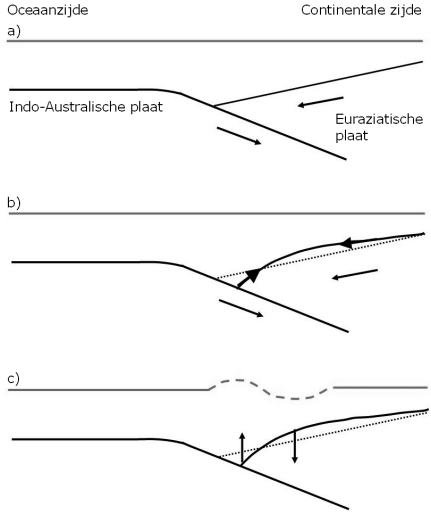
1. Inleiding

Op zondag 26 december 2004 werd Zuidoost-Azië getroffen door een van de zwaarste aardbevingen uit de geschiedenis. De magnitude was tenminste 9,0 [1]. Deze aardbeving, waarvan het epicentrum ongeveer 150 km ten westen van het Indonesische eiland Sumatra lag, veroorzaakte een vloedgolf – een *tsunami* – die tot enorme overstromingen leidde in de kustgebieden langs de Indische Oceaan: van Sumatra en Thailand tot aan Sri Lanka, India en zelfs Somalië. Ongeveer 300.000 mensen kwamen om het leven en miljoenen raakten gewond en/of dakloos. In dit artikel leggen we kort uit hoe een tsunami ontstaat (Paragraaf 2). Vervolgens gaan we uitgebreider in op de voortplanting van tsunami's over de oceaan en de veranderingen die optreden wanneer een tsunami de kust nadert. Daartoe laten we eerst zien hoe golven wiskundig te beschrijven zijn (Paragraaf 3). Een tsunami is een voorbeeld van een *lopende lange* golf. We hebben ervoor gekozen geen volledig overzicht van de theorie van (lange) golven te geven (zie hiervoor bijvoorbeeld [2]). In plaats daarvan presenteren we enkele karakteristieken van lopende golven. Lopende golven zijn er in vele verschijningsvormen. De meest bekende zijn de 'gewone' golven op zee. Dit zijn korte golven. Het getij en een tsunami zijn voorbeelden van lange golven. We zullen laten zien dat de karakteristieken van 'gewone' (korte) golven en van lange golven (zoals tsunami's) twee uitersten vormen. Aan de hand hiervan worden de overeenkomsten en de verschillen duidelijk tussen korte en lange golven (en dus tussen 'gewone' golven en tsunami's) en wordt duidelijk hoe tsunami's zich voortplanten en hoe ze veranderen nabij de kust. In Paragraaf 4 gaan we nog iets specifieker in op de voortplanting van de tsunami van 26 december 2004 over de oceaan om vervolgens in Paragraaf 5 te laten zien waarom tsunami's zulke desastreuze gevolgen kunnen hebben (in tegenstelling tot gewone golven terwijl die vaak veel hoger zijn). Tot slot staan we kort stil bij de (on)mogelijkheden om de gevolgen van een dergelijk natuurverschijnsel te beperken (Paragraaf 6).

2. Oorsprong tsunami's

Om een tsunami te genereren is een mechanisme nodig dat in korte tijd een grote hoeveelheid water verplaatst. Een voorbeeld van zo'n mechanisme is een aardbeving die plaatsvindt in de aardkorst onder water. Niet elke aardbeving onder water veroorzaakt een tsunami. Twee platen die horizontaal langs elkaar

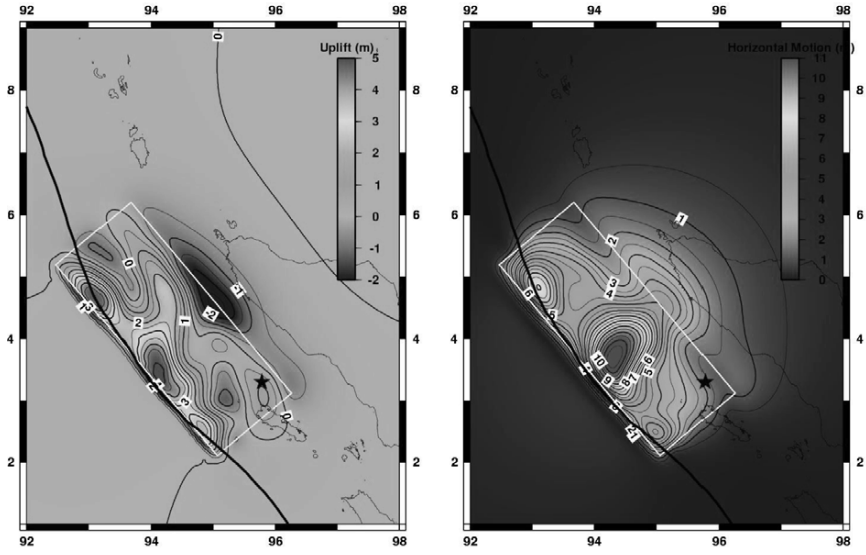
schuiven kunnen wel een aardbeving veroorzaken, maar hierbij zal nauwelijks water verplaatst worden. Er moet een verticale beweging plaatsvinden van de zeebodem. Bovendien moet de aardbeving behoorlijk sterk zijn (magnitude > 7,0) en niet te diep onder de zeebodem plaatsvinden. Áls een tsunami wordt opgewekt, hangt het ook nog af van de diepte van het water waaronder de aardbeving plaatsvindt of de tsunami zo hoog wordt als in december 2004. Als de aardbeving onder een ondiepe zee of oceaan plaatsvindt, is het verschil in waterdiepte tussen de plaats van ontstaan en het kustgebied klein. Daardoor is ook het verschil in voortplantingssnelheid van de tsunami klein en zal de hoogte van de golf weinig toenemen (zie Paragraaf 3.9). De aardbeving van 26 december 2004 vond plaats onder een ruim 1 kilometer diep zeegebied.



Figuur 1. *Schematische weergave van de tektoniek van het gebied rond Sumatra en de beweging van de aarde vóór en tijdens de aardbeving. De Indo-Australische plaat duikt onder de Euraziatische plaat (a). Door de schuifweerstand tussen deze twee platen trekt de onderduikende plaat de overliggende plaat een stukje mee naar beneden en buigt de laatste op (b). Als de spanning te groot wordt 'schiet' de bovenliggende plaat los en strekt weer. Als gevolg hiervan wordt het water aan de oceaanzijde omhoog geduwd en daalt het waterniveau aan de continentale zijde (c).*

De aardkorst bestaat uit verschillende platen die ten opzichte van elkaar bewegen. Voor de westkust van Sumatra schuift de Indo-Australische plaat geleidelijk naar het noordoosten onder de Euraziatische plaat (Figuur 1a). Door de schuifweerstand tussen de twee platen wordt de Euraziatische plaat meegetrokken naar beneden en buigt de plaat op (Figuur 1b). Bij de aardbeving is de Euraziatische plaat 'losgeschoten' en ongeveer twintig meter langs het breukvlak verplaatst. Het breukvlak maakt daar een hoek van 13° met het horizontale vlak, wat resulteert in een verticale verplaatsing van ongeveer 4,5

meter. Door de ontspanning is het opgebogen deel weer gestrekt. Dit losschieten en strekken heeft een stijging van de zeebodem en dus van de waterspiegel aan de oceaanzijde en een daling aan de continentale zijde veroorzaakt (Figuren 1c en 2). Deze verstoring plant zich vervolgens voort als een lopende golf. Door de specifieke configuratie zoals hierboven beschreven was het front van de tsunami-golven aan de westkant een top en aan de oostkant een dal, het laatste merkbaar als een zich aanvankelijk terugtrekken van de zee aan de kusten van Sumatra en Thailand. Het brongebied van de aardbeving en dus van de waterspiegelverstoring strekte zich uit over een gebied van circa 400×200 km (zie Figuur 2). Dit betekent dat een tsunami gegenereerd wordt met golflengtes van deze grootteorde.

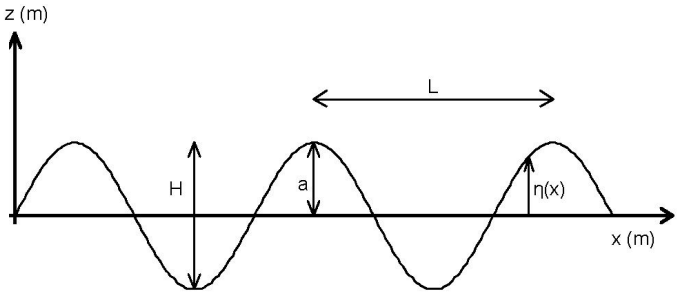


Figuur 2. *(Voor kleurenillustratie zie pagina 86–90.) Berekende verplaatsing van de zeebodem als gevolg van de aardbeving, Links: verticale verplaatsing, rechts: horizontale verplaatsing [7]. De dikke zwarte lijn geeft de rand van de Indo-Australische plaat weer. De dunne zwarte lijn geeft de kustlijn van noordwest-Sumatra en de omliggende eilanden aan. De getallen langs de horizontale en verticale assen geven respectievelijk de graden noorderbreedte en oosterlengte.*

3. Karakteristieken van lopende golven

3.1. Definities en voorbeelden korte en lange golven

Om te beginnen presenteert Figuur 3 een definitieschets van een lopende golf.



Figuur 3. *Definitieschets van een lopende golf*

Tabel 1 geeft een overzicht van de gebruikte terminologie.

Grootheid	**symbool**	**eenheid**	**omschrijving**	**Engels**
Golfhoogte	H	m	afstand tussen golftop en golfdal	wave height
golflengte	L	m	afstand tussen twee golftoppen	wave length
amplitude	a	m	maximale uitwijking wateroppervlak t.o.v. gemiddeld zeeniveau	amplitude
golfperiode	T	s	tijdsduur tussen passeren twee golftoppen	wave period
voorplantingssnelheid	$c = L/T$	m/s	snelheid waarmee individuele golf zich voortplant	wave celerity
hoekfrequentie	$\omega = 2\pi/T$	rad/s		angular frequency
golfgetal	$k = 2\pi/L$	rad/m		wave number
oppervlakteuitwijking	$\eta(x,t)$	m	niveau van de waterspiegel t.o.v. de gemiddelde waterstand	
golfsteilheid	H/L	–	verhouding tussen golfhoogte en golflengte	wave steepness

Tabel 1: *Gebruikte terminologie bij lopende golven*

Lopende golven zijn het gevolg van een overdracht van energie. De 'gewone' golven worden veroorzaakt door de wind en worden daarom windgolven genoemd. Het getij wordt veroorzaakt door de aantrekkingskracht tussen aarde en maan. Tsunami's worden veroorzaakt door bijvoorbeeld een aardbeving of door een landverschuiving of een vulkaanuitbarsting onder water.

- Golven worden 'kort' genoemd, of 'diepwatergolven' wanneer de *waterdiepte* waarover de golven zich voortplanten *groter* is dan de halve *golflengte*: $h \geq \frac{1}{2}L$ of $L \leq 2h$
- Golven worden 'lang' genoemd, of 'ondiepwatergolven' wanneer de *waterdiepte* waarover de golven zich voortplanten *kleiner* is dan $\frac{1}{20}$ van de *golflengte*: $h \leq \frac{1}{20}L$ of $L \geq 20h$

Voor een typische windgolf op de Noordzee met een golfperiode van 7 s is de golflengte ongeveer 75 m (zie later). Dit betekent dat dit een korte of diepwatergolf is in waterdieptes van ongeveer 35 m of meer. De diepte van de Noordzee is 30–40 meter en dus is dit inderdaad een korte of diepwatergolf; tenminste wanneer de golf zich voldoende ver uit de kust bevindt. Deze golf wordt een 'lange' of 'ondiepwatergolf' in waterdieptes van ongeveer 4 meter. Dichtbij de kust gedragen windgolven zich dus als ondiepwatergolven. Tsunami's hebben golflengtes in de orde van honderden kilometers. Zelfs op oceanen van enkele kilometers diep gedragen deze zich dus altijd als lange of ondiepwatergolven (voor $L = 200$ km: $h \leq$ 10 km).

3.2. Oppervlakte-uitwijking

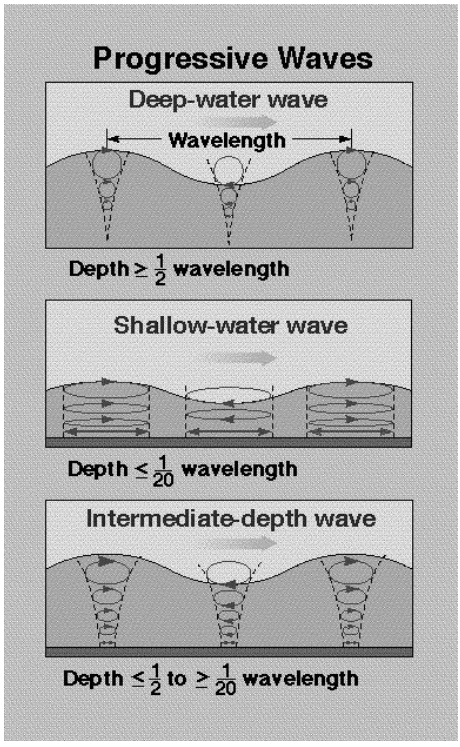
Voor een eerste benadering van de voortplanting van golven wordt gebruik gemaakt van de lineaire golftheorie. Deze gaat ervan uitgegaan dat de verstoringen t.g.v. de golven klein zijn; oftewel dat de golfhoogte klein is t.o.v. de waterdiepte en de golflengte ($H/h \ll 1, H/L \ll 1$). In dat geval is de uitwijking van het wateroppervlak te schrijven als:

$$\eta(x,t) = a \ \sin(\omega t - kx) \tag{1}$$

3.3. Orbitaalbeweging

Golven planten zich voort met snelheid $c = L/T$. Bovendien zetten golven het water in beweging. De snelheid van de waterdeeltjes is echter niet gelijk aan de voortplantingssnelheid van golven. Deze *orbitaalbeweging* hangt af van het feit of we te maken hebben met diep- of ondiepwatergolven:

- Diepwatergolven veroorzaken een cirkelvormige orbitaalbeweging. Hoe verder onder het wateroppervlak, hoe minder de golven worden gevoeld en hoe kleiner de orbitaalbeweging is. Uiteindelijk is op een diepte van een halve golflengte onder het wateroppervlak de invloed van golven geheel verdwenen (zie voor een animatie van de orbitaalbeweging op diep water [8]).
- Het kenmerk van ondiepwatergolven is dat de waterdiepte zo klein is dat de drukfluctuaties en de snelheden van de waterdeeltjes ten gevolge van de golf vrijwel onverminderd doordringen tot de bodem. Dit leidt tot een ellipsvormige orbitaalbeweging, die vlak boven de bodem afgevlakt wordt tot een heen-en-weer gaande orbitaalbeweging evenwijdig aan de bodem. Onder lange golven beweegt de hele waterkolom dus (enigszins) heen en weer tijdens het passeren van de golftoppen en dalen.
- Voor tussenliggende waterdieptes is de orbitaalbeweging een combinatie van deze uitersten.



Figuur 4. *(Voor kleurenillustratie zie pagina 86–90.) Schematische weergave van de waterbeweging (orbitaalbeweging) onder lopende golven. (uit:* `http://www.geneseo.edu/~gsci205/waves/waves.htm`*)*

De orbitaalbeweging is schematisch weergegeven in Figuur 4 voor diepwatergolven (bovenste figuur), ondiepwatergolven (middelste figuur) en golven op tussenliggende waterdieptes (onderste figuur). Merk op dat de figuren niet op schaal zijn. De diameter van de cirkel die de waterdeeltjes beschrijven op diep water is gelijk aan de golfhoogte H. Deze is zeer veel kleiner dan de golflengte (typische waarden voor windgolven zijn $H = 2$ m, $L = 75$ m).

3.4. Orbitaalsnelheid

De horizontale en verticale snelheid van een waterdeeltje (horizontale en verticale orbitaalsnelheid) u en w worden gegeven door de volgende vergelijkingen:

$$u(x,t) = \omega a \frac{\cosh[k(h+z)]}{\sinh[kh]} \sin(\omega t - kx) = \hat{u} \sin(\omega t - kx) \tag{2}$$

$$w(x,t) = \omega a \frac{\sinh[k(h+z)]}{\sinh[kh]} \cos(\omega t - kx) = \hat{w} \cos(\omega t - kx) \tag{3}$$

De horizontale en verticale snelheid verschillen op ieder moment 90° in fase, overeenkomend met een cirkelvormige (of ellipsvormige) beweging. Bovendien

is de horizontale snelheid u in fase met de oppervlakteuitwijking η: onder de top van de golf is de horizontale snelheid van een waterdeeltje maximaal (en in de richting van de golfvoortplanting) en onder het dal van de golf is de horizontale snelheid van een waterdeeltje maximaal tegen de richting van de golfvoortplanting in. De verticale snelheid w is *uit* fase met de oppervlakteuitwijking η: onder de top en het dal van de golf is de verticale snelheid van een waterdeeltje nul. De verticale snelheid is maximaal (positief/negatief) bij de overgang van golftop naar golfdal en andersom.

Wat betreft de grootte van de horizontale en verticale watersnelheid kijken we naar de amplitudes $\hat{u}$ en $\hat{w}$ aan het wateroppervlak ($z = 0$) en aan de bodem ($z = -h$). Dit levert het volgende resultaat:

$$\text{Aan het wateroppervlak} : z = 0 : \quad \hat{u} = \omega a \frac{\cosh[kh]}{\sinh[kh]} = \frac{\omega a}{\tanh[kh]} \tag{4}$$

$$\hat{w} = \omega a \frac{\sinh[kh]}{\sinh[kh]} = \omega a \tag{5}$$

$$\text{Aan de zeebodem} : z = -h : \quad \hat{u} = \omega a \frac{\cosh[0]}{\sinh[kh]} = \frac{\omega a}{\sinh[kh]} \tag{6}$$

$$\hat{w} = \omega a \frac{\sinh[0]}{\sinh[kh]} = 0 \tag{7}$$

De amplitude van de verticale watersnelheid is dus altijd gelijk aan nul aan de zeebodem (direct boven de bodem kan het water alleen evenwijdig aan de bodem stromen) en altijd gelijk aan $\omega a (= \pi H/T)$ aan het oppervlak.

3.5. Verschil in orbitaalsnelheid tussen korte en lange golven

De amplitude van de horizontale snelheid daarentegen hangt ook nog af van $kh (= 2\pi h/L)$ en dus van de waterdiepte/golflengte verhouding. Met andere woorden deze horizontale watersnelheid is anders voor korte of diepwatergolven ($kh \gg 1$) dan voor lange of ondiepwatergolven ($kh \ll 1$). We maken gebruik van de volgende gegevens voor het maken van een schatting van $\hat{u}$ en $\hat{w}$ in de verschillende situaties:

$$\lim_{x \to \infty} (\sinh x) = \frac{1}{2} e^x \tag{8}$$

$$\lim_{x \to \infty} (\cosh x) = \frac{1}{2} e^x \tag{9}$$

$$\lim_{x \to 0} (\sinh x) = x \tag{10}$$

$$\lim_{x \to 0} (\cosh x) = 1 \tag{11}$$

- Voor korte of diepwatergolven ($kh \gg 1$) op voldoende afstand boven de bodem ($k(h + z) \ll 1$) geldt:

$$\hat{u} \approx \omega a \frac{\frac{1}{2} e^{k(h+z)}}{\frac{1}{2} e^{kh}} = \omega a \frac{e^{kh} e^{kz}}{e^{kh}} = \omega a e^{kz} \tag{12}$$

$$\hat{w} \approx \omega a \frac{\frac{1}{2}e^{k(h+z)}}{\frac{1}{2}e^{kh}} = \omega a \frac{e^{kh}e^{kz}}{e^{kh}} = \omega a e^{kz} \tag{13}$$

Dus $\hat{u}$ en $\hat{w}$ zijn altijd aan elkaar gelijk (cirkelvormige beweging), zijn maximaal aan het wateroppervlak ($\hat{u} = \hat{w} = \omega a$) en nemen exponentieel af met de diepte.

- Voor lange of ondiepwatergolven ($kh \ll 1$) geldt:

$$\hat{u} \approx \omega a \frac{1}{kh} = \frac{\omega a}{kh} \tag{14}$$

$$\hat{w} \approx \omega a \frac{k(h+z)}{kh} = \omega a \frac{h+z}{h} = \omega a \left(1 + \frac{z}{h}\right) \tag{15}$$

Hieruit blijkt dat voor lange golven de horizontale watersnelheid constant is over de diepte (de hele waterkolom is in beweging), terwijl de verticale watersnelheid lineair toeneemt van nul aan de zeebodem tot ωa aan het wateroppervlak.

3.6. Golfvoortplantingssnelheid en relatie tussen golfperiode en golflengte

Voor vrije zwaartekrachts-oppervlaktegolven bestaat een relatie tussen de frequentie en het golfgetal (ofwel tussen de golfperiode en de golflengte). Deze wordt gegeven door de volgende *dispersierelatie:*

$$\omega^2 = gk \tanh(kh) \tag{16}$$

Waarbij g de zwaartekrachtsversnelling is. Deze vergelijking is gelijk aan:

$$L = \frac{gT^2}{2\pi} \tanh(kh) \tag{17}$$

Omdat per definitie de voortplantingssnelheid gelijk is aan ω/k geldt:

$$c = \sqrt{\frac{g}{k} \tanh(kh)} \tag{18}$$

Omdat $\omega/k = L/T$ geldt ook:

$$c = \frac{gT}{2\pi} \tanh(kh) \tag{19}$$

3.7. Verschil in voortplantingssnelheid en golflengte tussen korte en lange golven

We maken gebruik van:

$$\lim_{x \to \infty} (\tanh x) = 1 \tag{20}$$

$$\lim_{x \to 0} (\tanh x) = x \tag{21}$$

- Voor korte of diepwatergolven ($kh \gg 1$) geldt dus:

$$c = \frac{gT}{2\pi} \tag{22}$$

$$L = \frac{gT^2}{2\pi} \tag{23}$$

  Zowel de golflengte als de voortplantingssnelheid zijn slechts afhankelijk van de golfperiode (en de zwaartekrachtsversnelling) en blijven onveranderd terwijl de golf zich voortplant (aangezien de golfperiode altijd constant is).
- Voor lange of ondiepwatergolven ($kh \ll 1$) geldt:

$$c = \sqrt{\frac{g}{k}kh} = \sqrt{gh} \tag{24}$$

$$L = \sqrt{gh}T \tag{25}$$

  Vergelijking (24) laat zien dat voor lange golven de voortplantingssnelheid slechts afhankelijk is van de waterdiepte (en de zwaartekrachtsversnelling) en dat deze afneemt met afnemende waterdiepte. De golflengte is daarnaast ook nog afhankelijk van de golfperiode, maar omdat deze niet verandert, neemt ook de golflengte af bij afnemende waterdiepte. Dit alles heeft gevolgen voor de golfhoogte zoals we in Paragraaf 3.9 zullen zien.

3.8. Dispersie

Uit vergelijking (22) blijkt dat de voortplantingssnelheid van korte golven slechts afhankelijk is van de golfperiode. Dit betekent dat naarmate de golfperiode langer is, de golven zich sneller voortplanten. Tijdens het voortplanten op de oceaan halen de langere golven de kortere dus in en na verloop van tijd bevinden de langere golven zich vooraan in het golfveld en de kortere golven achteraan. Dit wordt golfdispersie genoemd. Op het moment dat de langere golven de kortere golven inhalen verandert de oppervlakteuitwijking. Deze is namelijk het resultaat van beide golven (zie voor een animatie [9]). Door golfdispersie kan de oppervlakteuitwijking dus groter zijn dan ten gevolge van de individuele golven. Vergelijking (24) laat zien dat de voortplantingssnelheid van lange golven slechts afhankelijk is van de waterdiepte. Dit betekent dat – ongeacht de golfperiode – deze golven op een bepaalde lokatie allemaal dezelfde voortplantingssnelheid hebben. Dit betekent dat golfdispersie zich bij lange golven niet voordoet en dat de oppervlakteuitwijking slechts bepaald wordt door de individuele golven. Omdat tsunami's lange golven zijn, is in eerste instantie de verwachting dat dispersie verwaarloosbaar is en dat de oppervlakteuitwijking slechts bepaald wordt door de individuele tsunamigolven. Er is echter gebleken dat voor de lange afstand waarover tsunami's zicht verplaatsen dispersie toch een rol speelt en daardoor de oppervlakteuitwijking groter kan zijn dan die ten gevolge van de individuele tsunamigolven (zie [3], [4]).

3.9. Energie en energieoverdracht

Per eenheid van oppervlak bezitten golven een hoeveelheid energie gelijk aan:

$$E = \frac{1}{8}\rho g H^2 \tag{26}$$

Waarbij ρ de dichtheid van water is. De hoeveelheid energie hangt dus slechts af van de golfhoogte (en wel kwadratisch) en de dichtheid van water en de zwaartekrachtversnelling. Belangrijker dan de hoeveelheid energie die de golven bezitten is de hoeveelheid energie die ze over kunnen dragen (het vermogen). De energie-overdracht (energieflux) in de voortplantingsrichting per tijd en per breedte (d.w.z. per lengte golfkam) is gelijk aan:

$$F = Enc = \frac{1}{8}\rho g H^2 nc \tag{27}$$

$$n = \frac{1}{2} + \frac{kh}{\sinh(2kh)} \tag{28}$$

$$\text{Op diep water } (kh \gg 1) \text{ geldt} : n = 1/2 \text{ en dus} : F = \frac{1}{16}\rho g H^2 c \tag{29}$$

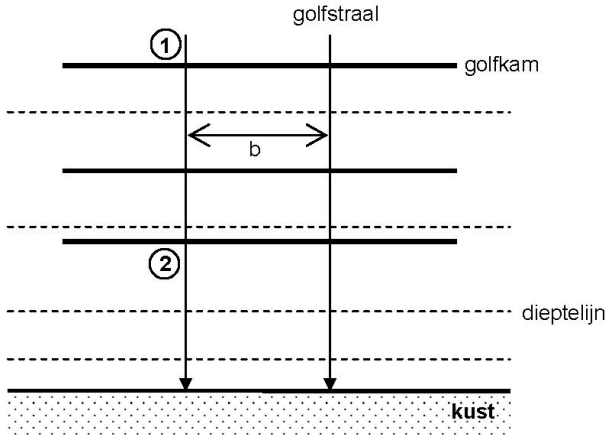
$$\text{Op ondiep water } (kh \ll 1) \text{ geldt} : n = 1 \text{ en dus} : F = \frac{1}{8}\rho g H^2 c \tag{30}$$

3.10. Shoaling en refractie

Zolang golven niet breken, vindt energieverlies voornamelijk plaats door wrijving langs de bodem. In een eerste benadering is dit energieverlies verwaarloosbaar. De energieoverdracht vindt plaats in de richting van de golfvoortplanting. Er is dus geen energieoverdracht in de richting loodrecht hierop (langs een golfkam). Dit alles houdt in dat de energieoverdracht constant is tussen twee golfstralen (loodlijnen op de golfkammen). Dit leidt tot twee belangrijke fenomenen wanneer golven de kust naderen: shoaling en refractie. Shoaling is het vervormen van de golven door het ondieper worden van het water, refractie is het bijdraaien van golven naar de dieptelijnen toe met eveneens een vervorming van de golf tot gevolg.

Wanneer golven de kust naderen, gedragen uiteindelijk alle golven zich als ondiepwatergolven. Het enige verschil is dat de positie van waaraf golven beschouwd kunnen worden als ondiepwatergolven dichterbij de kust ligt, naarmate de invallende golf korter is: Zoals we in Paragraaf 3.1 hebben gezien gedraagt een tsunami zich zelfs op de oceaan als een lange golf. Een typische windgolf zal zich pas als lange golf gedragen in waterdiepte van enkele, dus enkele honderden meters uit de kust. Omdat we voor dit artikel met name geïnteresseerd zijn in de effecten van shoaling en refractie voor tsunami's, gaan we er in het vervolg van uit dat de golf zich steeds gedraagt als lange golf (zoals voor en tsunami het geval is).

Beschouw ten eerste een situatie van een rechte kust met evenwijdige dieptelijnen en een golf die loodrecht op de kust invalt, d.w.z. met de golfkammen evenwijdig aan de kust en de dieptelijnen (zie Figuur 5).



Figuur 5. *Bovenaanzicht van golven die loodrecht invallen op een kust met parallelle dieptelijnen.*

F is de energieoverdracht per eenheid van breedte. De energieoverdracht tussen twee golfstralen is dus Fb, waarbij b de afstand tussen twee golfstralen is. Overigens is – bij loodrecht invallende golven – deze afstand constant (zie Figuur 5). Het feit dat de energieoverdracht tussen twee golfstralen constant blijft, levert de volgende vergelijking:

$$F_1 b = F_2 b$$

$$\frac{1}{8}\rho g H_1^2 c_1 b = \frac{1}{8}\rho g H_2^2 c_2 b \tag{31}$$

$$H_1^2 c_1 = H_2^2 c_2$$

Waarbij F_1 de energieoverdracht op locatie 1 is (ver uit de kust) en F_2 de energieoverdracht op locatie 2 (dichtbij de kust). Evenzo zijn H_1, H_2 en c_1 en c_2 de golfhoogtes en voortplantingssnelheden op de verschillende locaties. Omdat we een lange golf beschouwen is de voortplantingssnelheid gelijk aan $\sqrt{gh}$ en dus:

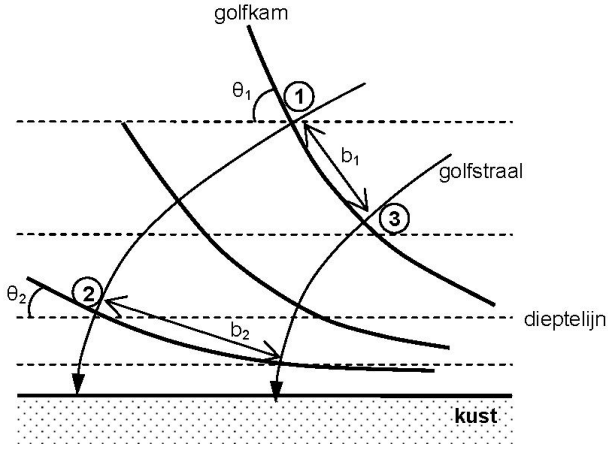
$$H_1^2 c_1 = H_2^2 c_2$$

$$H_1^2 \sqrt{g h_1} = H_2^2 \sqrt{g h_2} \tag{32}$$

$$H_2 = \left(\frac{h_1}{h_2}\right)^{\frac{1}{4}} H_1$$

De golfhoogte dichtbij de kust is dus een factor $(h_1/h_2)^{1/4}$ groter dan de golfhoogte ver uit de kust. Een tsunami die opgewekt wordt op een diepte van 1

km en daar een hoogte heeft van een halve meter, heeft dus op een waterdiepte van 3 m een hoogte van ruim 4 m. Deze toename in golfhoogte is het gevolg van het ondieper worden van het water en wordt shoaling genoemd.



Figuur 6. *Bovenaanzicht van golven die schuin invallen op een kust met parallelle dieptelijnen.*

Beschouw nu een situatie waarbij golven schuin invallen op een rechte kust met evenwijdige dieptelijnen (Figuur 6). Het deel van de golf dat zich bevindt bij locatie 1, bevindt zich in dieper water dan het deel dat zich bevindt bij locatie 3. Omdat de voortplantingssnelheid afhangt van de waterdiepte, heeft het deel van de golf dat zich bevindt bij locatie 1 een hogere voortplantingssnelheid dan het deel dat zich bevindt bij locatie 3. Hierdoor verplaatst de golfkam zich gedurende een bepaalde tijdsduur op locatie 1 over een grotere afstand dan bij locatie 3. Het gevolg is dat de golven bijdraaien naar de ondiepten, dus naar de kust, toe. Dit proces heet refractie. Dit bijdraaien van de golven heeft eveneens gevolgen voor de golfhoogte.

Constante energieoverdracht tussen twee golfstralen leidt nu tot de volgende vergelijking:

$$F_1 b_1 = F_2 b_2$$

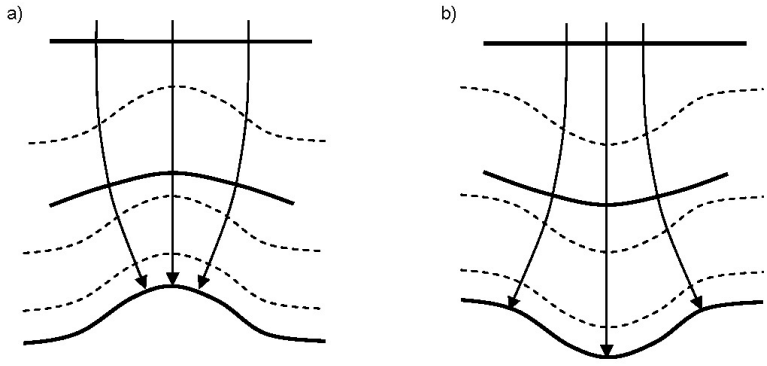
$$\begin{aligned} \frac{1}{8}\rho g H_1^2 c_1 \cos\theta_1 &= \frac{1}{8}\rho g H_2^2 c_2 \cos\theta \\ H_1^2 \sqrt{g h_1}\cos\theta_1 &= H_2^2 \sqrt{g h_2}\cos\theta_2 \end{aligned} \tag{33}$$

En dus:

$$H_2 = \left(\frac{h_1}{h_2}\right)^{\frac{1}{4}} \sqrt{\frac{\cos\theta_1}{\cos\theta_2}} \tag{34}$$

Waneer θ_2 kleiner is dan θ_1 (zoals in de situatie van schuin-invallende golven op een rechte kust), leidt refractie tot een afname in de golfhoogte. Visueel is dit voor te stellen doordat de aanwezige energie tussen de golfstralen op locatie

1, bij locatie 2 over een groter lengte van de golfkam wordt verspreid. Dit gaat dus samen met een afname van de golfhoogte. Bij gekromde dieptelijnen, hangt het van de situatie af, of refractie leidt tot een toe- of een afname van de golfhoogte, zoals geïllustreerd in Figuur 7.



Figuur 7. *Golfrefractie bij gekromde dieptelijnen. Situatie a) Kapen - convergentie van golfstralen: refractie leidt tot toename van golfhoogte. Situatie b) Baaien - divergentie van golfstralen: refractie leidt tot afname van golfhoogte.*

3.11. Golfbreking en golfoploop

De verhouding tussen de golfhoogte en de golflengte geeft de golfsteilheid. Wanneer de golfsteilheid te groot wordt, worden golven instabiel en zullen ze breken. De maximale steilheid die golven kunnen bereiken voordat ze breken wordt gegeven door het volgende criterium:

$$\left(\frac{H}{L}\right)_{\max} \cong 0.14 \tanh(kh) \tag{35}$$

Op diep water ($kh \gg 1$) kunnen windgolven breken tijdens een storm, wanneer door sterke wind de golven groeien. Dit gebeurt bij een steilheid van 0.14. Golven breken echter voornamelijk op stranden. In de vorige paragraaf is gebleken dat door shoaling de voortplantingssnelheid en de lengte van de golven afnemen, terwijl de golfhoogte toeneemt. Hierdoor worden golven dus steeds steiler. Bovendien treedt er nog een extra vervorming op doordat de hogere delen sneller gaan lopen dan de lagere, waardoor met name de voorzijde van de golven (gezien vanuit de kust) extra steil wordt. Op ondiep water leidt het steilheidscriterium tot de volgende vergelijking:

$$\left(\frac{H}{L}\right)_{\max} \cong 0.14kh = 0.14\frac{2\pi h}{L} = 0.88\frac{h}{L} \tag{36}$$

$$\left(\frac{H}{h}\right)_{\max} = 0.88 \tag{37}$$

In de praktijk blijkt de maximale golfhoogte/waterdiepte verhouding meer in de buurt van 0.6 te liggen. Dit betekent dat golven als gevolg van shoaling steeds hoger worden, totdat de golfhoogte ongeveer gelijk is aan 60% van de waterdiepte en de golven breken. Bij het breken verliezen ze een groot deel van de energie en dus daalt de golfhoogte.

Of golven werkelijk breken – en hoe precies – hangt af van de helling van het bodemprofiel (hellingshoek β) t.o.v. de golfsteilheid; de brekerparameter ξ:

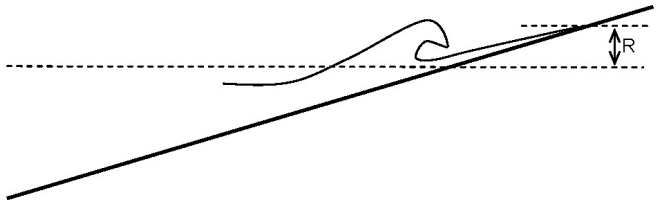
$$\xi = \frac{\tan\beta}{\sqrt{H/L_0}} \tag{38}$$

Voor relatief steile bodemhellingen ($\xi > 3$) breken golven niet, maar worden ze teruggekaatst. Voor relatief flauwe bodemhellingen ($\xi < 3$) breken de golven op het strand.

In het laatste stadium van de nadering van een hellende kust treedt oploop op, waarbij de impuls van het aankomende water het water tegen de helling opstuwt. De golfoploophoogte R boven het ongestoorde zeeniveau (zie Figuur 8) die daarbij bereikt kan worden, kan wel tot tien keer de golfhoogte op dieper water bedragen. De oploophoogte wordt sterk door de brekerparameter bepaald. Voor brekende golven met $\xi <$ ca. 2.5 geldt:

$$\frac{R}{H} \cong \xi = \frac{\tan\beta}{\sqrt{H/L_0}} \tag{39}$$

$$R \cong \sqrt{\frac{H}{L_0}}\tan\beta \cong 0.4T\sqrt{gH}\tan\beta \tag{40}$$



Figuur 8. *Definitieschets golfoploophoogte R*

4. Voortplanting van de tsunami van 26 december 2004 over de oceaan

Het breukvlak van de aardbeving strekte zich uit over circa vierhonderd kilometer vanuit het hypocentrum in een ongeveer noord-noordwestelijke richting (zie Figuur 2). Vanwege de relatief grote lengte was de uitstraling van energie in de tsunami niet alzijdig maar meer in de vorm van twee gerichte bundels, in ongeveer oostelijke respectievelijk westelijke richting. De gebieden in het pad van deze bundels (Noord-Sumatra/Thailand respectievelijk Sri Lanka/Somalië)

zijn daarom veel zwaarder getroffen dan de meer zuidelijk of noordelijk gelegen gebieden.

In oost-west-richting strekte het brongebied van de aardbeving zich uit over circa tweehonderd kilometer (zie Figuur 2). Het grootste deel van de golfenergie zit daardoor in golflengten van deze grootteorde. De waterdiepte in de Golf van Bengalen is ongeveer 3 kilometer. De golflengte is dus ruimschoots langer dan $20h$ (= 60 km), waardoor de tsunami-golven inderdaad in redelijke benadering wiskundig beschreven kunnen worden als lange golven (Paragraaf 3.1). De voortplantingssnelheid c van de tsunami kan bepaald worden met behulp van Vgl. (24). Met waterdiepten van ongeveer 3 kilometer is c dus ruim 600 km/uur. Ter vergelijking: voor windgolven op de oceaan wordt de voortplantingssnelheid slechts bepaald door de golfperiode (ordegrootte 10 s; zie Vgl. (22)). Windgolven hebben dus een voortplantingssnelheid in de orde van 15 m/s of 50 km/uur.

Zoals uitgelegd in Paragraaf 3.9, veroorzaken variaties van de diepten in het zeegebied en de kustzones waarin de tsunami beweegt refractie, waardoor de golven naar de ondiepte toe draaien. Bij het passeren van kapen en kleine eilanden of openingen treedt uiteraard ook buiging op. Bij flauw hellende bodems wordt dit effect versterkt of zelfs gedomineerd door refractie. Vandaar dat ook kustgebieden aan de lijzijde toch zwaar getroffen kunnen worden, zoals de zuidwestkust van Sri Lanka.

5. Waarom zo desastreus?

Een tsunami is op de oceaan niet extreem hoog (twee uur na de aardbeving had de tsunami van 26 december 2004 op de oceaan een maximale hoogte van ongeveer 60 cm [10]). Gewone windgolven kunnen vele malen hoger zijn. Toch leiden dat soort windgolven niet tot dergelijke grote overstromingen. Het verschil zit voornamelijk in de periode en de lengte van de golf. Windgolven hebben een periode in de orde van 5–15 s en een lengte van maximaal enkele honderden meters. Tsunami's hebben een veel grotere periode (in de orde van een kwartier à een half uur) en een veel grotere lengte (de tsunami van 26 december had golflengten in de grootte-orde van ongeveer 150 kilometer).

5.1. Shoaling

Zoals uitgelegd in Paragraaf 3.9 worden golven door shoaling hoger als ze de kust naderen en in ondieper water komen. Shoaling treedt dus pas op wanneer de golven zich in ondiep water bevinden ($h < 1/20L$). Door de relatief korte lengte van windgolven begint een windgolf pas te groeien wanneer de golf relatief dichtbij de kust is. Al voorbeeld: een windgolf met een golfperiode van 10 s (en een golflengte op diep water van 156 m; zie Vgl.(23)) begint te groeien als gevolg van shoaling in waterdieptes van ongeveer 7.5 m. Door de grote golflengte van een tsunami is zelfs de oceaan 'ondiep water'. Dit betekent dat een tsunami begint te groeien zo gauw de waterdiepte afneemt. De toename in golfhoogte als gevolg van shoaling is gerelateerd aan de verhouding van de waterdiepte op diep water (of de diepte waar voor het eerst shoaling optreedt)

en de locale waterdiepte (zie Vgl.(32)). De tsunami van 26 december 2004 werd gegenereerd in een zeegebied met een diepte van ongeveer 1 km. Op een waterdiepte van bijvoorbeeld 3 m zou de tsunami dus ruim 4× zo hoog kunnen worden $(= 1000/3)^{0.25}$, terwijl een windgolf van 10 s dan slechts ongeveer 25% hoger is dan op diep water $(= 7.5/3)^{0.25}$.

5.2. Breking

De golfhoogte van een tsunami neemt dus veel sterker toe dan die van windgolven wanneer deze in ondieper water komt. Windgolven zijn echter op diep water vaak veel hoger dan tsunami's. Waarom tsunami's toch zo veel destructiever kunnen zijn, heeft weer te maken met de grote golflengte van een tsunami. Zoals in Paragraaf 3.10 is uitgelegd worden golven instabiel wanneer de golfsteilheid te groot wordt. Als gevolg van shoaling neemt de golfhoogte toe en de golflengte af. Door de relatief korte lengte van windgolven, bereiken zij daarbij veel eerder dan tsunami's de maximale steilheid. Windgolven breken dus voordat ze de kust bereiken. Daarmee verliezen ze een groot deel van hun energie. Door de veel grotere lengte wordt een tsunami lang niet zo steil, waardoor een tsunami niet of pas veel later breekt, maar het land oploopt; in sommige gevallen 'als een muur van water'.

5.3. Invloed golfperiode

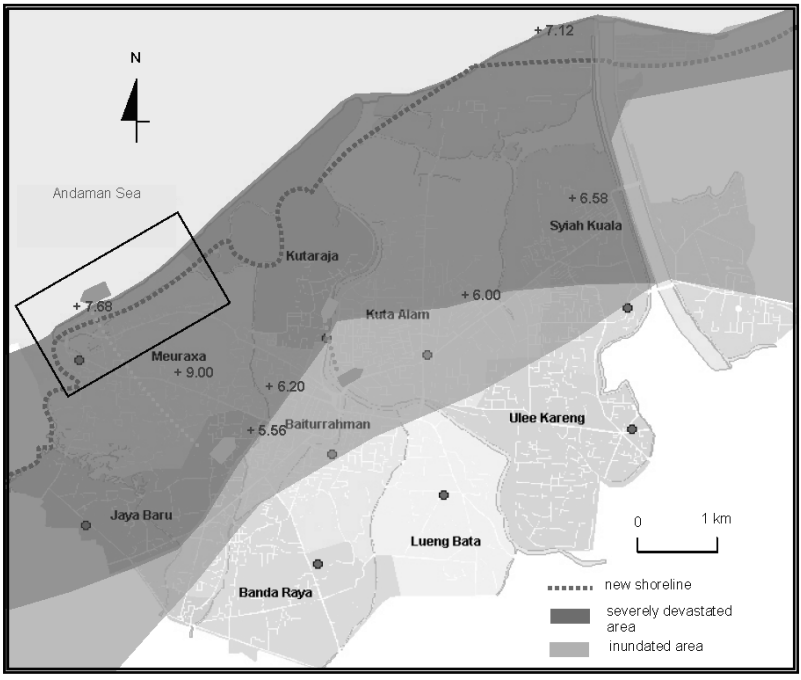
Tot slot is een tsunami veel destructiever dan een windgolf doordat de lange golfperiode van een tsunami resulteert in een vrij lange duur van elke aanval. Om enig idee te geven van de hoeveelheid water die door een tsunami aangevoerd kan worden, is een vergelijking gemaakt met de afvoer van de Rijn in Nederland. Tijdens het hoogwater van 1995 voerde de Rijn (die bij hoog water een breedte heeft in de orde van 1 km) ongeveer 12.000 m^3/s af. Een grove schatting laat zien dat over een strook kust met een lengte van een kilometer een tsunami met een hoogte op de oceaan van 60 cm in 7 minuten (de halve golfperiode ofwel de duur van de waterstandsverhoging door de tsunami) bijna 30 miljoen m^3 water meevoert. Dit komt overeen met ruim 70.000 m^3/s (ongeveer zes keer de hoogwaterafvoer van de Rijn) over deze strook kust van een kilometer.

5.4. Gevolgen van de tsunami in Atjeh

Dit alles heeft grote gevolgen. In geval van lager gelegen land treedt overstroming op. De hoge stroomsnelheden die zich daarbij kunnen voordoen (tot wel 5 m/s), gecombineerd met de soms aanzienlijke overstromingsdiepten, kunnen leiden tot grote schade en veel slachtoffers. Uit sporen in het landschap heeft een Japans onderzoekteam geprobeerd in kaart te brengen hoe ver landinwaarts het water in en rond Banda Atjeh is gekomen [11]. Hieruit blijkt dat in sommige laaggelegen gebieden het water 4 km de stad in is gestroomd en dat op sommige plekken, die 2.5 – 3 km uit de kust lagen, het water 12 m boven gemiddeld zeeniveau heeft gestaan (Figuur 9). Aan de westkust ten zuiden van Banda Atjeh, waar de vlakke kuststrook slechts enkele honderden meters breed

is en een heuvelrug aan de kust grenst, zijn oploophoogten opgetreden van circa 35 m (Figuur 10).

Als gevolg van de tsunami zijn grote stukken kust weggeslagen en is bijna alle vegetatie en bebouwing verwoest. Satellietfoto's [12] laten zien dat de kust van Banda Atjeh nu enkele honderden meters verder landinwaarts ligt dan voor de tsunami (Figuur 9 en Figuur 11). Volgens het Japanse onderzoekteam is dit deels veroorzaakt door bodemdaling als gevolg van de aardbeving en deels door erosie van het strand als gevolg van de tsunami.

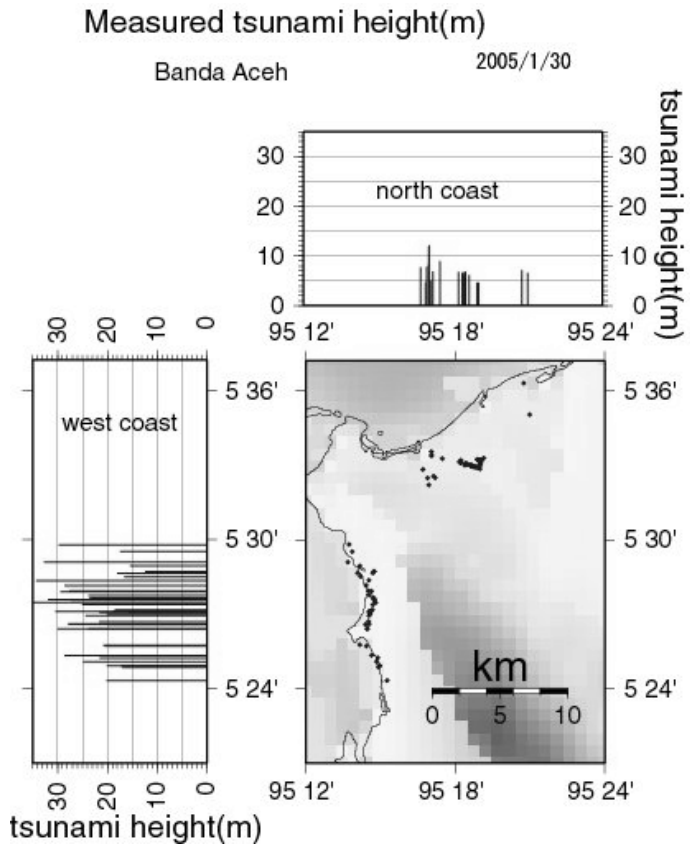


Figuur 9. *(Voor kleurenillustratie zie pagina 86–90.) Kaart van Banda Atjeh met de oorspronkelijke en de nieuwe kustlijn, het gebied dat grotendeels verwoest is en het gebied dat overstroomd is. De oranje symbolen geven de verplaatsing weer van twee schepen door de tsunami. (Het linker schip, van meer dan 1000 ton, werd door de tsunami ruim 3 km landinwaarts meegesleept.) Het kader geeft de locatie aan van de satellietfoto van figuur 6. De getallen geven de waterdiepte in meters ten gevolge van de tsunami. De werkelijke overstromingsdiepte was nog ongeveer een meter hoger als gevolg van het getij op het moment van de tsunami (uit [5] op basis van data uit [11]).*

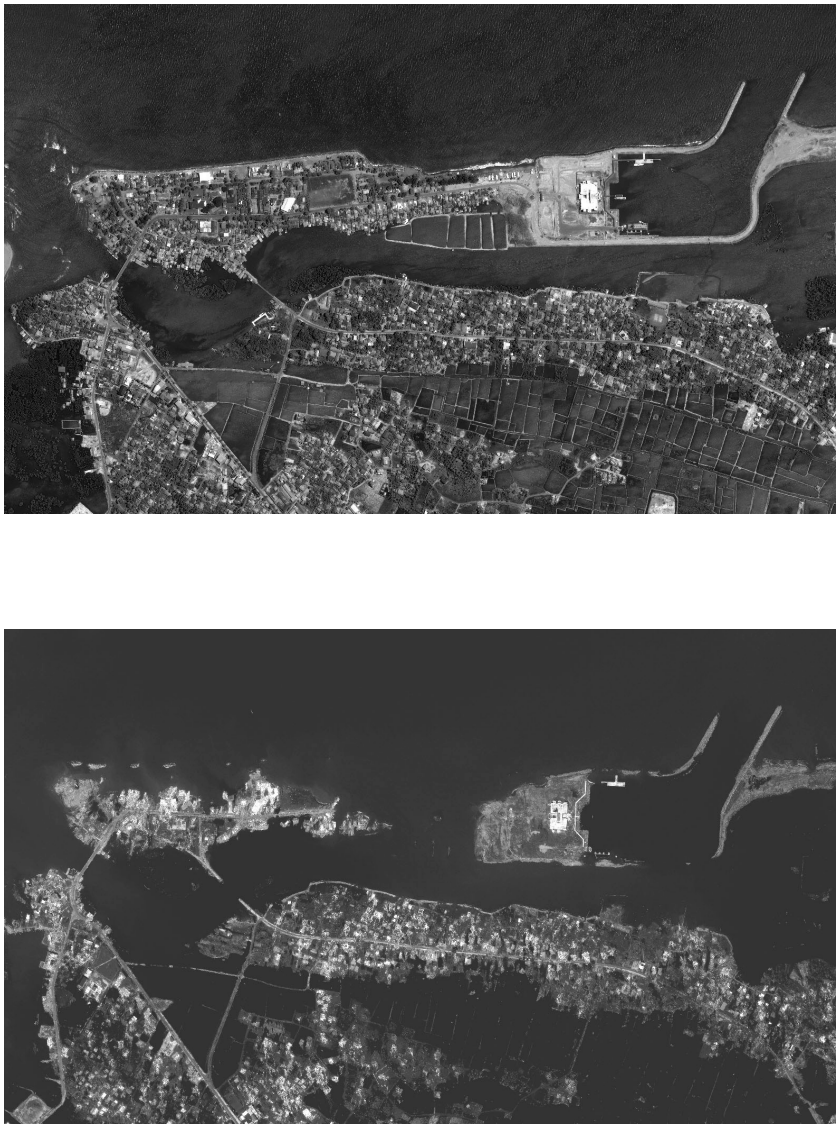
6. Voorkómen of schade beperken?

Het voorkómen van een tsunami is onmogelijk: de aardbeving die de tsunami veroorzaakt is niet te voorkomen. Wel kan er worden nagedacht over hoe de schade valt te beperken. Daarbij valt te denken aan de volgende zaken:

1) Voorlichting van de lokale bevolking over het vóórkomen van tsunami's en de te nemen acties. Bijvoorbeeld mensen bewust laten worden van het feit dat het terugtrekken van de zee een voorteken van een tsunami kan zijn en mensen laten weten waar ze op dat moment naartoe moeten (in ieder geval niet naar zee!).
2) Een tsunami-waarschuwingssysteem. Belangrijke aspecten hierbij zijn de vraag hoe de gehele bevolking bereikt kan worden – ook in afgelegen gebieden – en hoe de evacuatie geregeld moet worden. Overigens had een tsunami-waarschuwingssysteem voor Banda Atjeh nauwelijks zin gehad: er zat minder dan een half uur [13] tussen het optreden van de aardbeving en het moment waarop de tsunami deze stad van meer dan 300.000 inwoners bereikte.
3) Regels voor het landgebruik in de kustzone. Bijvoorbeeld het niet langer toestaan dat mensen of bedrijven zich vestigen dicht bij de kust. Dit zal heel moeilijk blijken omdat in veel gevallen de kust de bron van inkomsten vormt voor de bevolking.
4) Maatregelen die ervoor zorgen dat de tsunami zijn energie verliest voordat hij bebouwd gebied bereikt. Er is wel gesuggereerd om mangrove bossen aan te planten. Het is echter de verwachting dat de strook mangrove onrealistisch breed moet zijn (enkele tot honderden kilometers) om de tsunami af te remmen.
5) Aangepaste bebouwing. Bijvoorbeeld gebouwen van meerdere verdiepingen met een open constructie op de begane grond: alleen kolommen of dragende muren evenwijdig aan de verwachte stroomrichting.



Figuur 10. *(Voor kleurenillustratie zie pagina 86–90.) Gemeten golfoploophoogte (in de figuur "tsunami height" genoemd) ten gevolge van de tsunami in meters boven lokaal getijniveau [11].*



.

Figuur 11. *(Voor kleurenillustratie zie pagina 86–90.) Noordkust van Banda Atjeh, vóór en na de tsunami van 26 december 2004 [12].*

7. Conclusies

Tsunami's worden veroorzaakt door zware aardbevingen waarbij een verticale beweging plaatsvindt van de zeebodem. De verticale waterverplaatsing die hierbij optreedt, leidt tot een lange golf die zich voortplant over de oceaan. Wanneer de golf in ondieper water komt en de kustgebieden bereikt, neemt de golfhoogte zeer sterk toe (zie voor een illustratieve animatie [14]). Dit kan leiden

tot grote overstromingen met alle gevolgen van dien. Door het bijdraaien van de golf naar de ondiepten toe, kunnen ook gebieden aan de lijzijde zwaar getroffen worden. Tsunami's zijn niet te voorkomen. Wel zouden het aantal slachtoffers en de schade beperkt kunnen worden, bijvoorbeeld door goede voorlichting van de bevolking, het gebruik van een tsunami-waarschuwingssysteem, door regels op te stellen voor het landgebruik in de kustzone of door gebouwen zo te ontwerpen dat ze beter bestand zijn tegen overstromingen.

Dankwoord

Het huidige artikel is deels gebaseerd op een artikel dat verschenen is in het Nederlands Tijdschrift voor Natuurkunde [6], waarvan prof.dr. Jurjen Battjes van de TU Delft en drs. Femke Goutbeek van het KNMI mede-auteurs waren. Hierdoor hebben zij indirect ook aan dit artikel bijgedragen, waarvoor ik hen hartelijk wil bedanken. Daarnaast wil ik dr. Hein Haak en van het KNMI bedanken voor zijn nuttige commentaar op dit artikel.

Referenties en websites

1. Stein, S. en E.A. Okal, 2005. Speed and size of the Sumatra earthquake. *Nature*, 434, pp. 581–582.
2. Mei, C.C., 1989. The applied dynamics of ocean surface waves. *Advanced series on ocean engineering*; Vol. 1, Singapore [etc.]: *World Scientific*, 740 pp.
3. Van Groesen, E. and G. Klopman, 2005. Dispersive effects in tsunami generation. *Proceedings of the Workshop on The Indonesian Ocean Tsunami 2005*, July, Bali Indonesia.
4. Mirchina, N. and E. Pelinovsky, 2001. Dispersive intensification of tsunami waves. *Proc. ITS 2001*, pp. 789–794.
5. Meilianda, E., C.M. Dohmen-Janssen, S.J.M.H. Hulscher en J.P.M. Mulder, 2005. Towards coastal zone management of Banda Aceh beach: the coastal system before and after the tsunami disaster of December 26th 2004. *Proc. of the Int. Conf. on Coastal Conservation and Management* (17–20 April, Tavira – Algarve, Portugal).
6. Dohmen-Janssen, C.M., J.A. Battjes and F.H. Goutbeek, 2005. Hoe ontstaan tsunami's en waarom? *Nederlands Tijdschrift voor Natuurkunde*, Vol. 71, No. 12, pp. 380–385, invited contribution (in Dutch).
7. `http://www.gps.caltech.edu/%7Ejichen/Earthquake/2004/aceh/ aceh.html`.
8. `http://serc.carleton.edu/NAGTWorkshops/visualization/ collections/coastal_wave_mechanics.html`
9. `http://physics.usask.ca/~hirose/ep225/animation/dispersion/anim-dispersion.html`
10. `http://www.noaanews.noaa.gov/stories2005/images/tsunami-2hrs2.jpg`.
11. `http://www.eri.u-tokyo.ac.jp/namegaya/sumatera/surveylog/ eindex.htm`.

12. http://www.digitalglobe.com/.
13. http://www.eeri.org/lfe/pdf/indonesia_sumatra_tsunami_surveys.pdf.
14. http://www.solcomhouse.com/tsunamis.htm

Woordenlijst seismologie

Hypocentrum = locatie van de aardbevingshaard

Epicentrum = de plaats aan het aardoppervlak recht boven het hypocentrum

Magnitude = de maat waarmee de sterkte van een aardbeving wordt weergegeven. Omdat aardbevingen grote variaties in sterkte hebben, zijn de magnitudeschalen logaritmisch. Er zijn verschillende magnitudeschalen. De bekendste is die van Richter, die gebaseerd is op de grootste amplitude van het seismogram, gemeten met een standaard seismometer en gecorrigeerd voor de afstand. Andere schalen zijn bijvoorbeeld gebaseerd op oppervlaktegolven *(surfacewave magnitude)* of ruimtegolven *(bodywave magnitude)*. De meest recente schaal is de momentmagnitudeschaal, die afgeleid is van het seismisch moment. De in dit artikel genoemde magnitude is de momentmagnitude M. Tot $M = 6$ komt de schaal van Richter ongeveer overeen met deze schaal. Daarboven wijken deze schalen van elkaar af en geeft de schaal van Richter een te lage waarde (dit heet verzadiging).

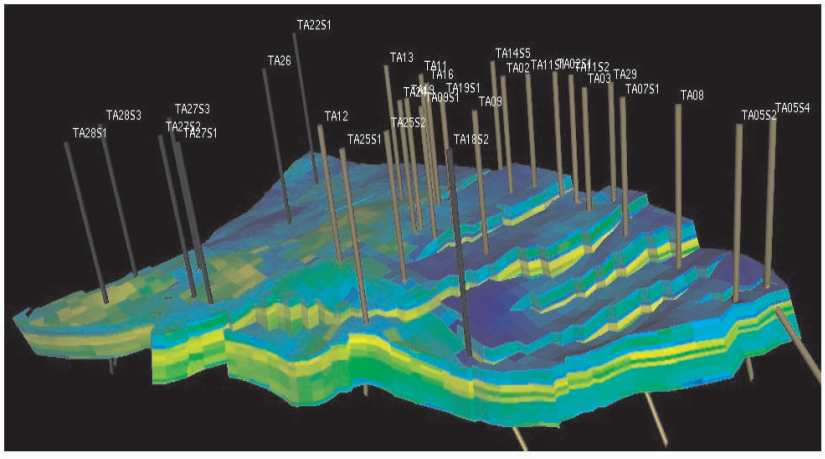
Seismisch moment = maat voor de kracht van een aardbeving waarin het totale breukoppervlak, de gemiddelde verplaatsing van het breukvlak en de elasticiteitseigenschappen van het gesteente tot uitdrukking komen. Het is een maat voor de energie die vrijkomt uit de bron van de aardbeving.

Vakantiecursus 2006

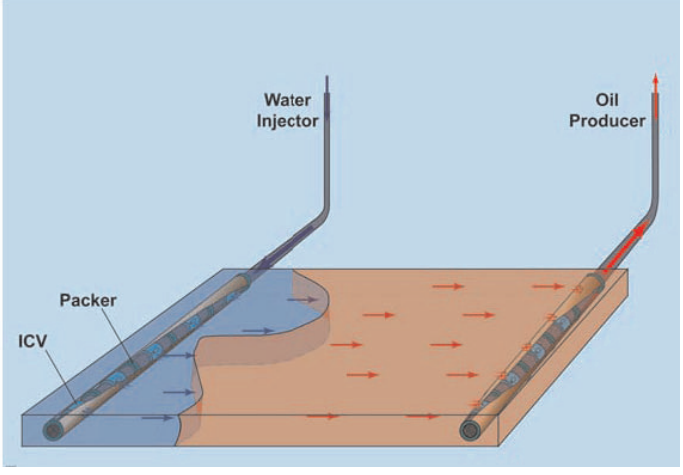
pp. 83 – 90

Kleurenillustraties

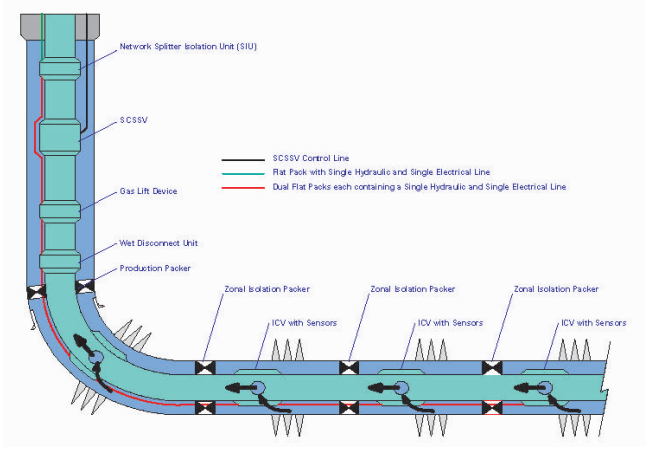
1. J.D. Jansen (Slimme olievelden, pagina 91–108)



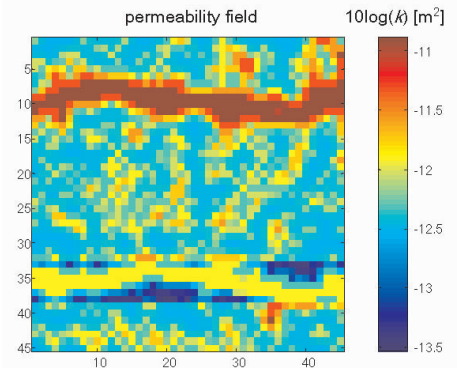
Figuur 2



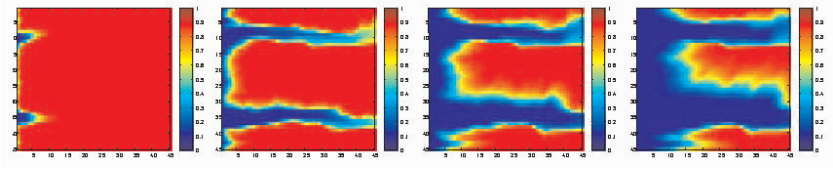
Figuur 3



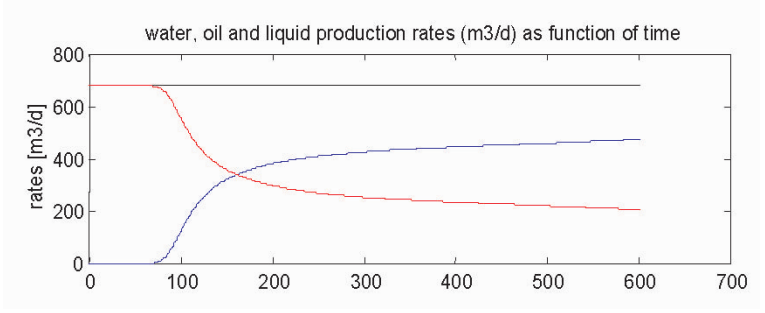
Figuur 4



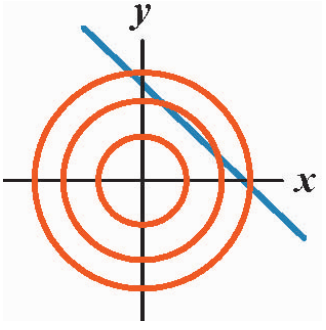
Figuur 5



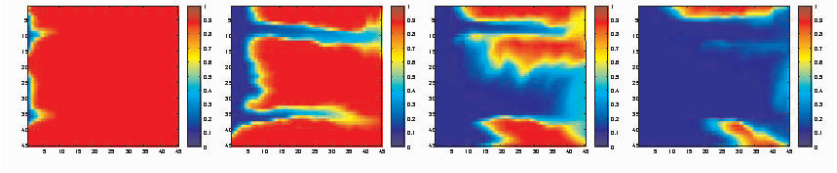
Figuur 6



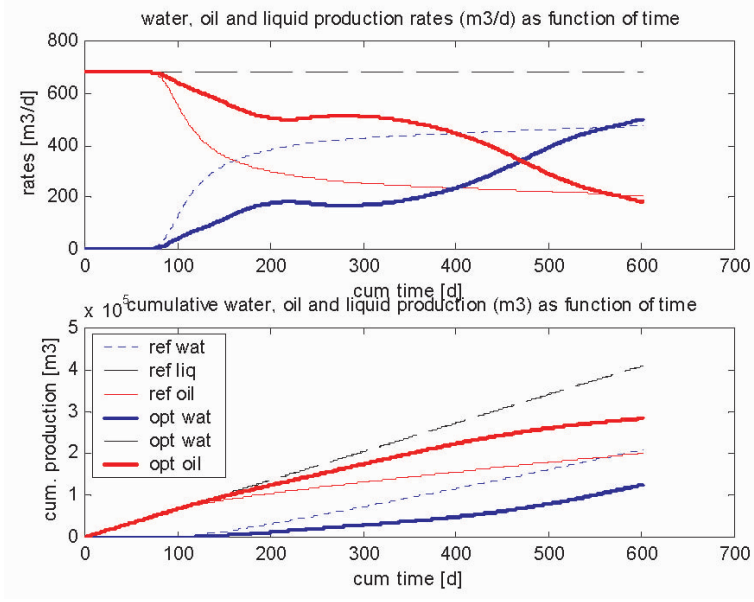
Figuur 7



Figuur 8

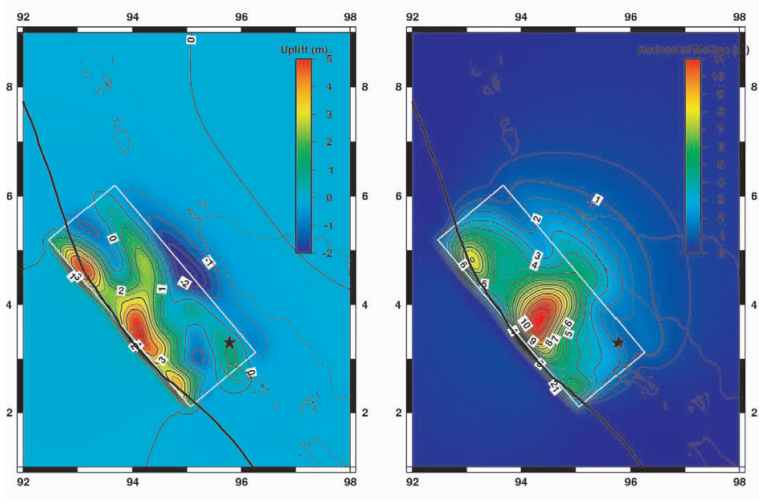


Figuur 9

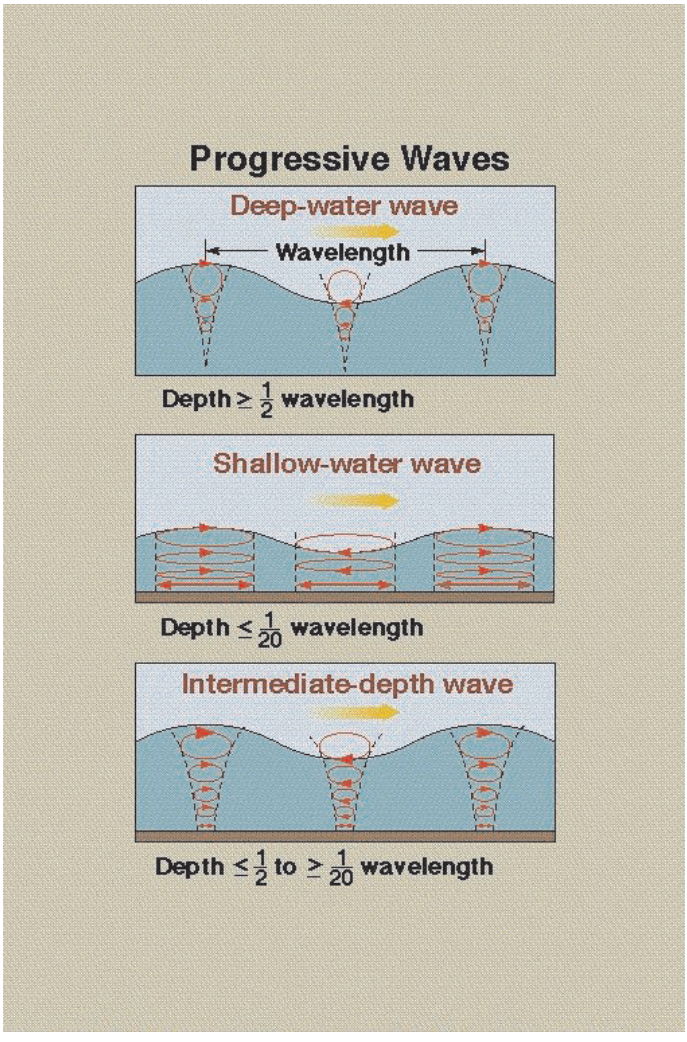


Figuur 10

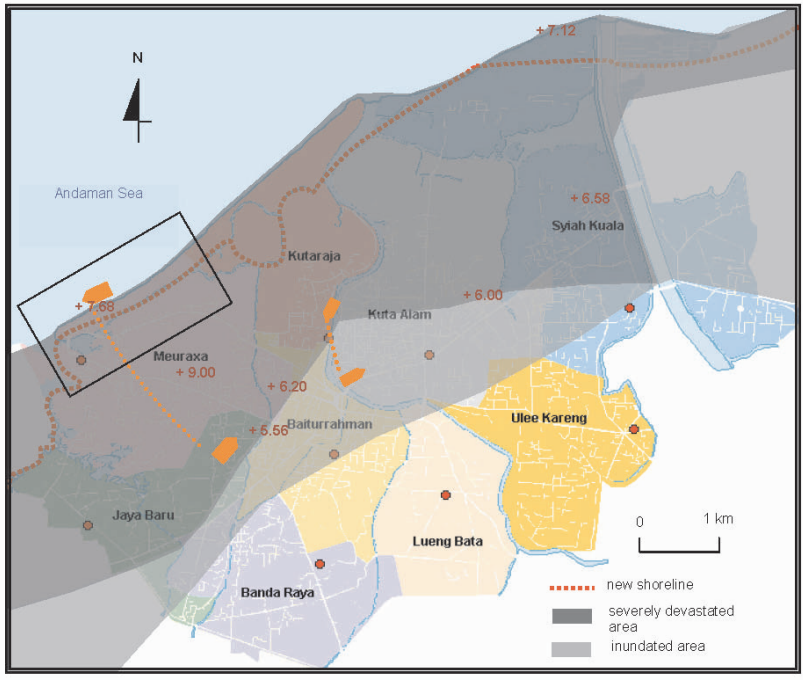
2. C.M. Dohmen-Janssen (Hoe onstaan Tsunami's en waarom?,pagina 61–82)



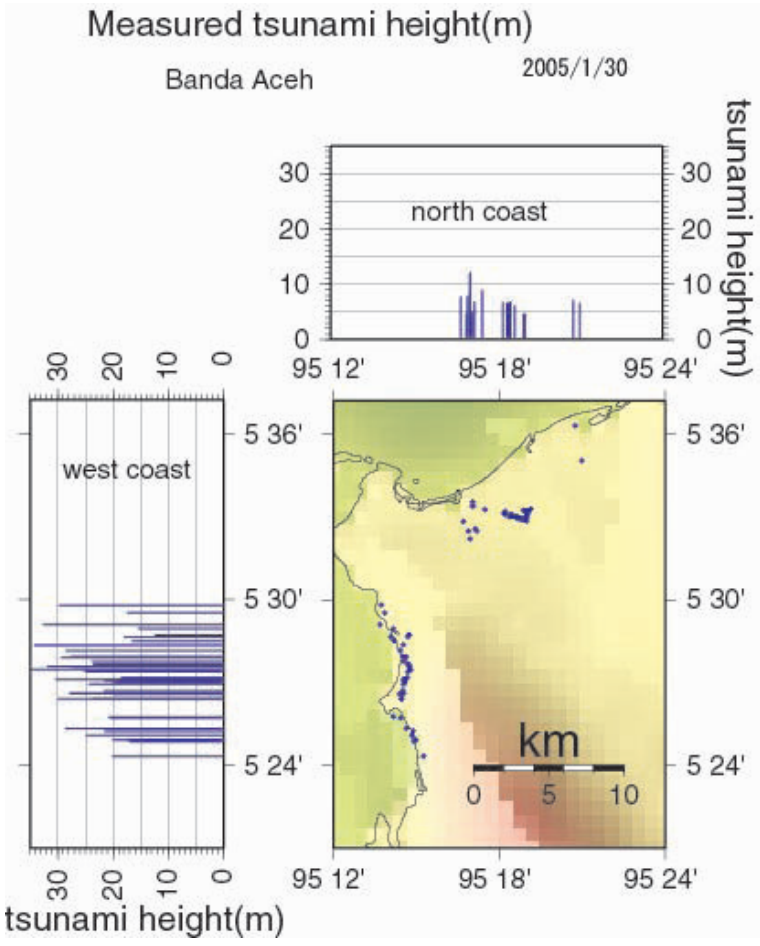
Figuur 2


Progressive Waves
Deep-water wave
Wavelength
Depth ≥ 1/2 wavelength
Shallow-water wave
Depth ≤ 1/20 wavelength
Intermediate-depth wave
Depth ≤ 1/2 to ≥ 1/20 wavelength

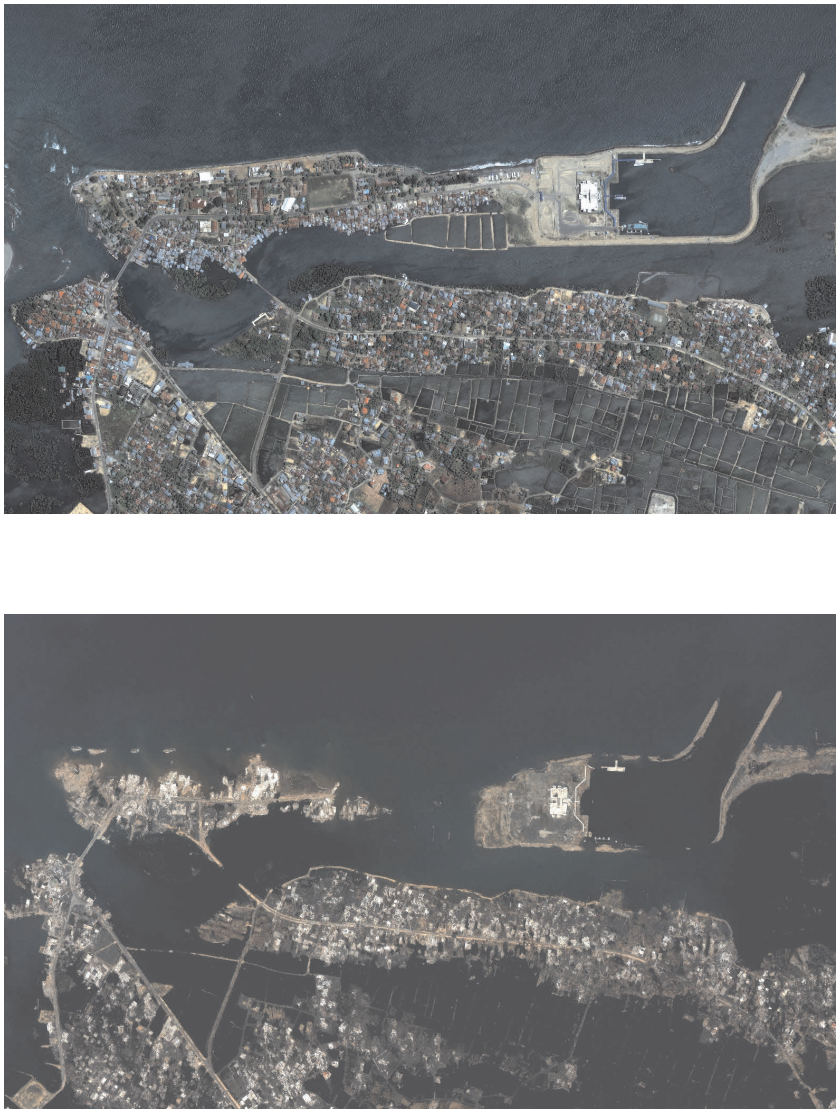
Figuur 4


N
Andaman Sea
+ 7.12
+ 6.58
Syiah Kuala
Kutaraja
+ 6.00
Kuta Alam
+ 7.68
Meuraxa
+ 9.00
+ 6.20
+ 5.56
Baiturrahman
Ulee Kareng
Jaya Baru
Lueng Bata
Banda Raya
0
1 km
new shoreline
severely devastated area
inundated area

Figuur 9


Measured tsunami height(m)
Banda Aceh
2005/1/30
north coast
tsunami height(m)
30
20
10
0
95 12'
95 18'
95 24'
west coast
5 36'
5 30'
5 24'
km
0
5
10
tsunami height(m)

Figuur 10



Figuur 11

Vakantiecursus 2006

pp. 91 – 108

Slimme olievelden

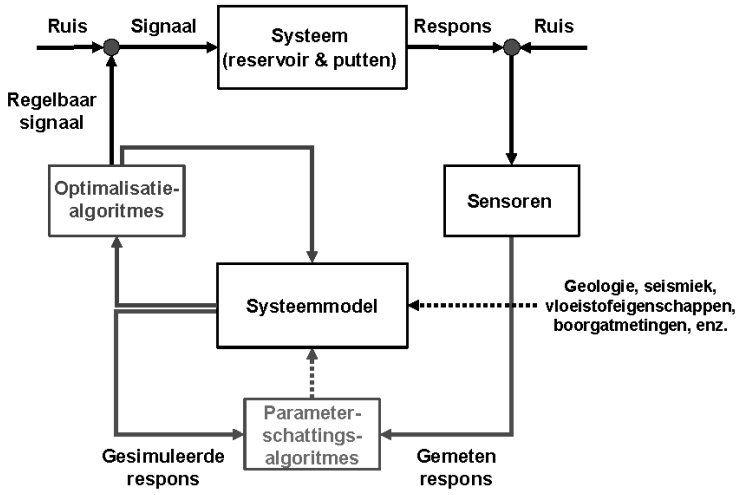
J.D. Jansen
Technische Universiteit Delft, Afdeling Geotechnologie
e-mail: j.d.jansen@citg.tudelft.nl

1. Inleiding

Meer welvaart voor meer mensen vergt meer energie. En hoewel het aandeel van duurzame bronnen, zoals wind en zon, in onze energievoorziening langzaam stijgt, zal het gebruik van olie en gas minstens tot aan het eind van deze eeuw een belangrijke rol blijven spelen. Een toenemend probleem daarbij is dat de grote 'makkelijk' vindbare olievelden inmiddels voor het merendeel gevonden zijn (tenminste dat denken we). De meeste van die olievelden bestaan uit een soort platte stenen sponzen op honderden tot duizenden meters diepte onder het aardoppervlak. Na het boren van putten stroomt de olie aanvankelijk vaak op eigen kracht omhoog, maar doorgaans is het al snel nodig om water of gas in het reservoir te injecteren om de druk op peil te houden en om de olie zo goed en zo kwaad als het gaat naar de productieputten te duwen. De meeste olie blijft echter plakken in de poriën van het reservoir, en vaak wordt daarom slechts zo'n 10 tot 50 % van de olie geproduceerd. Het verhogen van de 'winningsfactor' is daarom een uitstekend alternatief voor het vinden van nieuwe velden. Gelukkig hebben zich zowel binnen als buiten de olie-industrie ontwikkelingen voorgedaan die het mogelijk maken om in de nabije toekomst meer olie uit bestaande velden te halen. In het bijzonder wordt er op het moment op verschillende plaatsen onderzoek gedaan naar het gecombineerde gebruik van modellen en ondergrondse sensoren en kleppen. Binnen onze sectie Geotechnologie in Delft waren wij een van de eersten om deze zogenaamde 'smart fields' of 'closed-loop reservoir management' technieken systematisch te onderzoeken. Inspiratiebronnen voor ons werk vormen daarbij de 'data assimilatie methodes' die door meteorologen wordt gebruikt voor het aanpassen van weermodellen, en de 'meet- en regeltechniek' die door ingenieurs wordt gebruikt voor het beheersen van industriële processen.

2. Reservoirbeheer

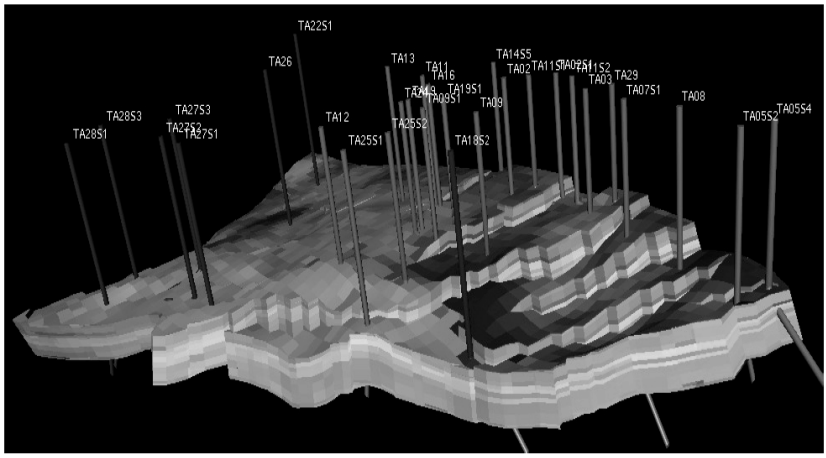
Figuur 1 geeft een schematische weergave van het olieproductieproces. De bovenste rechthoek in de figuur geeft het werkelijke reservoir weer met de bijbehorende putten (boorgaten). Tijdens de ontwerpfase van de veldontwikkeling worden één of meerdere computermodellen gebouwd voor simulatie van stroming in het reservoir, die zijn weergegeven door de centrale rechthoek. De geometrie van het reservoir wordt afgeleid uit seismische metingen, waarbij door



Figuur 1. *Reservoirbeheer weergegeven als een modelgebaseerd gesloten-lus proces.*

reflecties van akoestische signalen een beeld van de ondergrond wordt gevormd. De gesteente-eigenschappen die van belang zijn voor de stroming van olie, gas en water worden afgeleid uit geologische waarnemingen en een heel scala van fysische boorgatmetingen. Een probleem daarbij is de grote heterogeniteit van de ondergrond die er voor zorgt dat metingen in een put een beperkte ruimtelijke geldigheid hebben. Onze ondergrondse stromingsmodellen kennen daarom een grote onzekerheid in de parameterwaarden. In het bijzonder betreft dat de porositeit, en de permeabiliteit, dat wil zeggen de doorlatendheid van het gesteente. In praktijk worden daarom vaak enige tot enige tientallen verschillende modellen gebouwd, zie Figuur 2.

Op grond van deze reservoirmodellen is het mogelijk om het productieproces te optimaliseren. Dit betreft bijvoorbeeld het bepalen van het aantal en de plaats van de putten, of het bepalen van de optimale waterinjectie- en olieproductiedebieten in het verloop van de tijd. Gedurende de afgelopen jaren zijn de mogelijkheden om het productieproces te sturen aanzienlijk toegenomen. Daarbij gaat het zowel om meer complexe putconfiguraties (gekromd, horizontaal of met zijtakken), als wel om de installatie van automatische regelkleppen aan het oppervlak of onder in de put. Het optimalisatieproces is in Figuur 1 weergeven in blauw. Gedurende de olieproductie worden regelmatig metingen uitgevoerd, zowel in de putten als in de bovengrondse installaties voor het scheiden van olie, gas en water, die een indruk geven van de drukken en debieten in de putten. Traditioneel werden dergelijke metingen slechts periodiek uitgevoerd, bijvoorbeeld maandelijks of eens per kwartaal, en doorgaans met een beperkte nauwkeurigheid. Gedurende de afgelopen jaren echter, worden in toenemende maten sensoren geïnstalleerd die op bijna continue basis



Figuur 2. *(Voor kleurenillustratie zie pagina 83–86.) Reservoirmodel met injectieputten (zwart) en productieputten (bruin). De kleuren van het reservoirmodel geven de doorlatendheid van de verschillende gesteentes aan.*

informatie over de drukken en debieten in de putten kunnen verstrekken. Bovendien is er een ontwikkeling van een aantal andere meetmethoden op gang gekomen die ook een indruk van de vloeistofverzadiging tussen de putten geven. De meest veelbelovende methode is de zogenaamde 'vierdimensionale' seismiek waarbij seismische metingen worden herhaald in de tijd. Door het combineren van de gemeten respons van de sensoren en de gesimuleerde respons van de reservoir modellen krijgen we een indruk van de mate waarin de modellen de werkelijkheid juist weergeven. Met behulp van systematische algoritmes voor 'data-assimilatie' kunnen we vervolgens de parameters in onze modellen aanpassen om een betere weergave van de productiegeschiedenis te krijgen, en, belangrijker, hopelijk ook een betere voorspelling van de productie in de toekomst. Het parameterschattingsproces is in Figuur 1 weergegeven in rood. Olievelden met uitgebreide mogelijkheden om het productieproces te meten en te regelen worden tegenwoordig vaak aangeduid als 'slimme velden' ('smart fields'). Uiteraard zijn het niet zozeer de sensoren en kleppen die het veld slim maken, maar gaat het om het slimme gebruik daarvan. Veel van het huidige onderzoek op het gebied van slimme olievelden richt zich op het verhogen van de momentane olieproductie. Ons onderzoek naar 'closed-loop reservoir management' daarentegen richt zich in de eerste plaats op het verhogen van de 'winningsfactor', dat wil zeggen de fractie van de totale hoeveelheid olie in het reservoir die uiteindelijk wordt geproduceerd; zie ook Jansen et al. (2005). In het vervolg van dit artikel zal worden ingegaan op de wiskunde die vereist is voor het optimaliseren van de winningsfactor uitgaande van een bekend reservoir model, dat wil zeggen op het in Figuur 1 met blauw aangegeven proces.

3. Reservoir model

De stroming van gas en vloeistoffen door een poreus medium kan worden beschreven door een stelsel niet-lineaire partiële differentiaalvergelijkingen, zoals nader is toegelicht in de Appendix. Voor het eenvoudige geval van een reservoir dat alleen olie en water bevat kan de toestand worden beschreven met twee toestandsgrootheden, bijvoorbeeld de oliedruk en de waterverzadiging. De belangrijkste parameters in de vergelijkingen zijn de permeabiliteit (de doorlatendheid) en de porositeit van het gesteente, die beiden doorgaans sterk variëren over het reservoir. Voor het oplossen van de vergelijkingen wordt daarom altijd gebruik gemaakt van een numerieke benadering. Daartoe worden de differentiaalvergelijkingen gediscretiseerd in ruimte en tijd, wat leidt tot een stelsel niet-lineaire algebraïsche vergelijkingen dat kan worden weergegeven als

$$\mathbf{x}(k+1) = \mathbf{f}\left(\mathbf{x}(k), \mathbf{u}(k)\right) \tag{1}$$

Hierin is de vector $\mathbf{x}$ de discrete benadering van de toestandsvector van het systeem die voor het simpele voorbeeld van olie- en waterstroming bestaat uit de oliedrukken en de waterverzadigingen in de punten van een ruimtelijk rooster. De variabele k representeert een discrete tijd $t(k), k = 0, 1, 2, \cdots, K$, waarbij we er voor het gemak van uit gaan dat alle tijdstappen Δt even groot zijn. De vector $\mathbf{u}$ representeert de stuursignalen die worden gebruikt om het systeem te controleren. Daarbij valt te denken aan klepstanden, drukken of debieten in injectie- en productieputten. De vectorfunctie $\mathbf{f}$ is een abstracte beschrijving van de algebraïsche vergelijkingen die voortkomen uit de discretisatie; zie Appendix A. De structuur van vergelijking (1) geeft aan dat de vergelijkingen recursief kunnen worden opgelost. Uitgaande van een begintoestand

$$\mathbf{x}(0) = \mathbf{x}_0, \tag{2}$$

kan de toestand op elk volgend tijdstip worden berekend. Daarbij moet worden bedacht dat deze berekening niet eenvoudig hoeft te zijn, en doorgaans voor elke tijstap het meerdere malen oplossen van een groot stelsel lineaire vergelijkingen vereist om de niet-lineaire systeemvergelijking f voldoende nauwkeurig te benaderen. In het vervolg zullen we vergelijking (1) weergeven in impliciete vorm als

$$\mathbf{g}\left(\mathbf{x}(k), \mathbf{x}(k+1), \mathbf{u}(k)\right) = \mathbf{0}. \tag{3}$$

Gezien het recursieve karakter kunnen we vergelijkingen (1) en (3) ook interpreteren als vector-differentievergelijkingen waarin de discrete tijd k de onafhankelijke variabele is, $\mathbf{x}$ de afhankelijke variabele, en $\mathbf{u}$ de niet-homogene term. Vergelijking (2) is dan de bijbehorende beginvoorwaarde. We definiëren ook een vectorfunctie $\mathbf{h}$ die het verband weergeeft tussen de gemeten respons $\mathbf{y}$ en de toestandsvector $\mathbf{x}$,

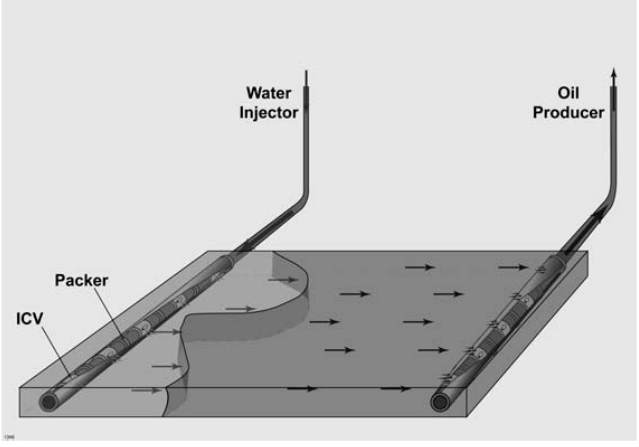
$$\mathbf{y}(k) = \mathbf{h}\left(\mathbf{x}(k)\right). \tag{4}$$

waarbij de gemeten respons bijvoorbeeld kan bestaan uit de drukken of debieten in een put. In het reservoirmodel kunnen verder ook nog andere nevenvoorwaarden (gelijkheden of ongelijkheden) voorkomen die bijvoorbeeld beperking

opleggen aan de drukken of de debieten in de putten, en die schematisch kunnen worden weergegeven als

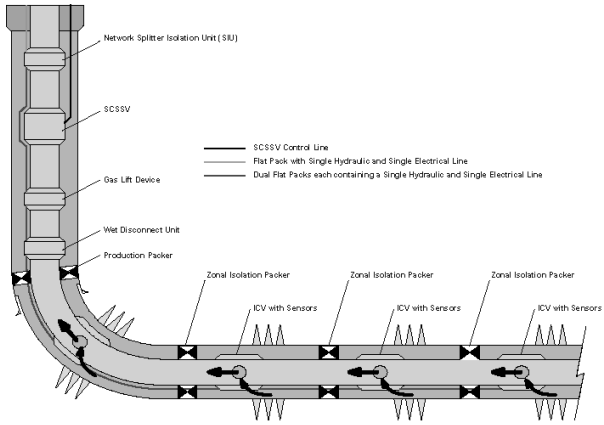
$$\mathbf{c}\left(\mathbf{x}(k), \mathbf{u}\right) \leq \mathbf{0}. \tag{5}$$

Vergelijkingen (3) tot en met (5) vormen een abstracte representatie van een reservoirsimulator. Voor realistische reservoirmodellen wordt gebruik gemaakt van zo'n 104 tot 106 roosterpunten, waarin de toestand wordt gesimuleerd voor enige honderden tot duizenden tijdstappen. Een typische reservoirsimulatie neemt dan ook uren tot dagen in beslag.



Figuur 3. *(Voor kleurenillustratie zie pagina 83–86.) Eenvoudig tweedimensionaal reservoirmodel met een horizontale injectieput en een horizontale productieput.*

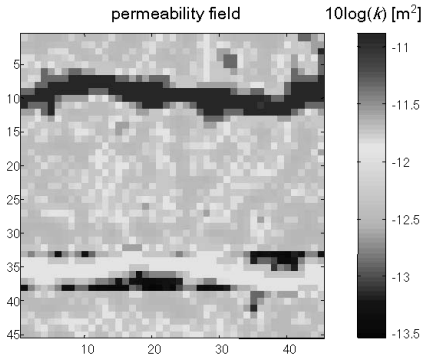
Figuur 2 geeft een beeld van een realistisch reservoirmodel met zo'n 40000 roosterpunten. In het verdere verloop van dit artikel zullen we gebruik maken van een eenvoudiger, tweedimensionaal voorbeeld met 45×45 roosterpunten dat is weergeven in Figuur 3. Aan de rechterkant van het reservoir wordt olie geproduceerd met behulp van een 'slimme' horizontale productieput die is voorzien van een binnenbuis met een groot aantal van elkaar gescheiden segmenten waarvan de instroming uit het reservoir afzonderlijk kan worden gecontroleerd; zie ook Figuur 4. In werkelijkheid bestaan er van dit soort slimme putten met een maximum van 5 segmenten. Aan de linkerkant van het reservoir wordt water geïnjecteerd met een eveneens slimme horizontale put. In beide putten is het bovendien mogelijk om een redelijk nauwkeurig meting te doen van de druk en de debieten in elk segment. Figuur 5 geeft een indruk van de doorlatendheid (permeabiliteit) k van de roosterpunten van het model. De eenheden van k zijn



Figuur 4. *(Voor kleurenillustratie zie pagina 83–86.) Schematische weergave van een 'slimme' horizontale put voorzien van een binnenbuis (groen) en een buitenbuis, en kleppen om de instroom vanuit het reservoir in afzonderlijke segmenten van de put te kunnen controleren. De grijze driehoekjes stellen openingen voor die het reservoir verbinden met de (blauwe) ruimte tussen de binnen- en buitenbuis. De cirkeltjes stellen de op afstand bedienbare kleppen voor.*

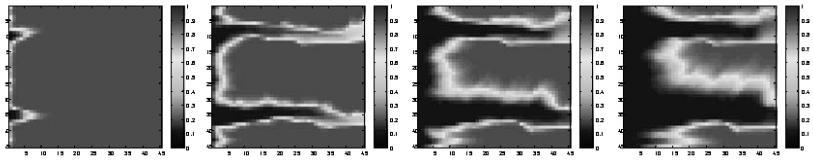
m^2 en zijn weergegeven op een logaritmische schaal. Figuur 5 maakt duidelijk dat het reservoir twee hoogdoorlatende zones bevat. Hierbij valt te denk aan twee voormalige met zand gevulde rivierbeddingen in een verder met klei gevulde omgeving. Het reservoir wordt geproduceerd onder nevenvoorwaarden die voorschrijven dat de totale debieten in de injectie- en productieputten aan elkaar gelijk zijn, en constant zijn in de tijd. In de individuele injectie- en productieputten kan wel een herverdeling van de debieten over de verschillende segmenten plaats vinden. Het valt niet moeilijk voor te stellen hoe de stroming zou verlopen als het reservoir met twee 'conventionele' horizontale putten zou worden geproduceerd, zodat in elk segment van de injectie- en productieputten de debieten zich zo zouden instellen dat er een bijna identieke druk zou heersen[5]. Het water zou dan de weg van de minste weestand kiezen zoals weergegeven in de vier plaatjes in Figuur 6. In theorie zou, voor een homogeen reservoir en een stabiel verdringingsproces, na het injecteren van één porievolume water (bijna) alle olie uit het reservoir verdwenen moeten zijn, corresponderend met een winningsfactor van (bijna) één. Door de heterogene doorlatendheid blijft er echter een aanzienlijke deel van de olie achter. Figuur 7 toont het bijbehorende productieprofiel, waarin olie- en waterproductie zijn weergeven als functie van de tijd. Het is duidelijk te zien dat na een korte

[5] Uiteraard moet er in de injectieput een hogere druk heersen dan in de productieput, anders zou er geen stroming door het reservoir zijn.



Figuur 5. *(Voor kleurenillustratie zie pagina 83–86.) Doorlatendheid (permeabiliteit) k in de roosterpunten van het reservoirmodel van Figuur 3.*

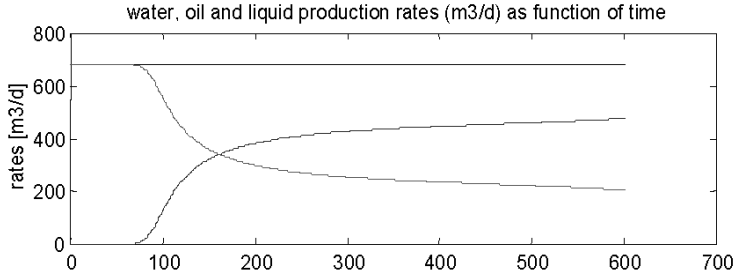
periode van 'droge' olieproductie, een doorbraak van water in de productieput plaats vind, hetgeen, aangezien de totale productie gelijk moet blijven, leidt tot een daling van de olieproductie.



Figuur 6. *(Voor kleurenillustratie zie pagina 83–86.) Waterverzadiging van het reservoir op vier verschillende tijstippen voor een conventionele productiestrategie. De vier plaatjes corresponderen met vier geïnjecteerde porievolumes. Van links naar rechts: 0.1, 0.4, 0.7 en 1.0 porievolumes geïnjecteerd. Rood: olie; blauw:water.*

4. Optimalisatie

De vraag is nu of het mogelijk is om voor de gegeven rand- en nevenvoorwaarden een betere verdeling van de debieten over de segmenten van de injectie- en productieputten te bepalen. Het doel is daarbij om een zo hoog mogelijke winningsfactor te bereiken over een van te voren gedefinieerd tijdsinterval. Als we ook de kosten van waterproductie meerekenen is een economische doelfunctie



Figuur 7. *(Voor kleurenillustratie zie pagina 83–86.) Olie- en waterproductie als functie van de tijd voor een conventionele productiestrategie.*

meer geschikt. In ons geval kiezen we een doelfunctie J gedefinieerd als

$$J = \sum_{k=0}^{K-1} J_k \left(\mathbf{u}(k), \mathbf{x}(k) \right), \tag{6}$$

waarbij

$$J_k = \sum_{j=1}^{N_{prod}} \left(r_o \mid q_{o,j} \mid \Delta t + r_w \mid q_{w,j} \mid \Delta t \right), \tag{7}$$

waarin r_o de (positieve) waarde van de olie per volume-eenheid is, en r_w de (negatieve) waarde van het water. In vergelijking (7) wordt gesommeerd over het aantal segmenten N_{prod} van de productieput. De bijdrage aan J_k van elk segment bestaat uit een positieve olieproductieterm en een negatieve waterproductieterm[6]. Het optimaliseringprobleem kan nu worden gedefinieerd als

$$\max_{\mathbf{u}} J, \tag{8}$$

waarbij voldaan dient te worden aan vergelijking (3), beginvoorwaarde (4), en nevenvoorwaarden (5). We hebben hierbij dus te maken met een optimalisatieprobleem onder nevenvoorwaarden. Ook als de expliciete voorwaarden (5) niet van toepassing zouden zijn, zouden we toch met nevenvoorwaarden te maken hebben vanwege het recursieve karakter van de dynamische vergelijkingen, zoals als volgt valt in te zien.

Voor een optimum is een noodzakelijke voorwaarde dat alle afgeleiden van J_k naar de stuursignalen $\mathbf{u}(k)$ gelijk aan nul zijn voor alle $k = 0, 1, \cdots, K-1$[7]. Doorgaans is het niet mogelijk om het optimum direct te berekenen, maar moeten we daarvoor een iteratieve methode gebruiken. Efficiënte methodes maken

[6] De absoluutstrepen rond de olie- en waterdebieten zijn nodig omdat het teken van productiedebieten in onze formulering als negatief is gekozen.

[7] De afgeleide van een scalaire grootheid naar een kolomvector is gedefinieerd als een rijvector met componenten die de afgeleiden naar de elementen van de van de kolomvector bevatten. Bijvoorbeeld: Voor $\mathbf{u} = (u_1 u_2 \cdots u_n)^T$ geldt $\frac{\partial J}{\partial \mathbf{u}} = \left(\frac{\partial J}{\partial u_1} \frac{\partial J}{\partial u_2} \cdots \frac{\partial J}{\partial u_n} \right)$.

gebruik van het feit dat de afgeleiden $\partial J_k/\partial \mathbf{u}(k)$, die buiten het optimum natuurlijk ongelijk aan nul zijn, de richting aangeven waarin de componenten van $\mathbf{u}(k)$ veranderd moeten worden om dichter bij het optimum te komen. Een probleem bij het berekenen van $\partial J_k/\partial \mathbf{u}(k)$ op een willekeurig tijdstip $k = \kappa$ is echter dat J_k niet alleen direct afhankelijk is van $\mathbf{u}(\kappa)$, maar dat J_k ook een functie is van $\mathbf{x}(k)$ voor $k = \kappa, \kappa+1, \kappa, \cdots, K-1$, waarbij al die toestanden $\mathbf{x}(k)$ op hun beurt een functie zijn van de signaalvector $\mathbf{u}(\kappa)$. Wat we dus zouden willen uitrekenen is de variatie.[8]

$$\delta J_k = \left\{ \frac{\partial J(\kappa)}{\partial \mathbf{u}(\kappa)} + \sum_{k=\kappa}^{K-1} \left[\frac{\partial J(k)}{\partial \mathbf{x}(k)} \frac{\partial \mathbf{x}(k)}{\partial \mathbf{u}(\kappa)} \right] \right\} \delta \mathbf{u}(\kappa). \tag{9}$$

De tweede term in de sommatie, $\partial \mathbf{x}(k)/\partial \mathbf{u}(\kappa)$, kan echter niet eenvoudig berekend worden want om het effect van $\mathbf{u}(\kappa)$ op $\mathbf{x}(k)$ te bepalen is het nodig om een reservoirsimulatie uit te voeren van tijdstip κ tot tijdstip k.

5. Optimalisatie met nevenvoorwaarden

Het bepalen van het minimum (of maximum) van een functie met nevenvoorwaarden is een klassiek wiskundig probleem, dat doorgaans wordt opgelost met behulp van de multiplicatoren van Lagrange, een techniek die ook bekend staat onder de naam variatierekening. Beschouw de vergelijking van een kwadratische doelfunctie van twee variabelen:

$$J = x^2 + y^2. \tag{10}$$

De vorm van het oppervlak $J(x, y)$ is een paraboloïde, en het bepalen van het minimum is een triviaal probleem. Formeel vinden we de oplossing door de afgeleiden naar $\mathbf{x}$ en y nul te stellen, wat leidt tot $2x = 0$ en $2y = 0$. Uit deze twee vergelijkingen volgt dan de oplossing voor het minimum: $(x_0, y_0) = (0, 0)$. Dit proces kan ook worden weergegeven als het bepalen van de totale differentiaal van J:

$$\delta J = \frac{\partial J}{\partial x}\delta x + \frac{\partial J}{\partial y}\delta y = 2x\delta x + 2y\delta y. \tag{11}$$

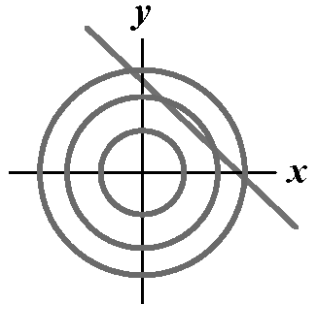
In het minimum moet gelden $\delta J = 0$ voor elke willekeurig combinatie van δx en δy, zodat we wederom vinden dat moet gelden $2x = 0$ en $2y = 0$. Beschouw nu de nevenvoorwaarde

$$x + y = k. \tag{12}$$

Dit is de vergelijking voor een rechte lijn die x-as en de y-as snijdt in respectievelijk de punten $(k, 0)$ en $(0, k)$.

Het bepalen van het minimum van J onder de eis dat aan nevenvoorwaarde (12) wordt voldaan kan worden geïnterpreteerd als het vinden van het laagste

[8] Een variatie (aangegeven met δ) is een differentiaal van een functionaal. Een functionaal is een functie van een functie. Een variatie is dus een differentiaal van een functie van een functie.



Figuur 8. *(Voor kleurenillustratie zie pagina 83–86.) Hoogtelijnen van een paraboloïde (in rood), doorsneden door een vertikaal vlak (in blauw).*

punt van de doorsnijding van de paraboloïde met het verticale vlak door de lijn $x + y = k$; zie Figuur 8. Hiertoe definiëren we een gemodificeerde doelfunctie:

$$\bar{J} = x^2 + y^2 + \lambda(x + y - k), \tag{13}$$

waarbij de nader te bepalen vermenigvuldigingsfactor λ bekend staat als een multiplicator van Lagrange. Merk op dat de term binnen de haakjes gelijk aan nul is als aan de nevenvoorwaarde is voldaan, zodat in dat geval de waarde van $\bar{J}$ gelijk is aan de waarde van J. De gemodificeerde doelfunctie is nu een functie van drie variabelen, x, y ,en λ, en de totale differentiaal kan worden bepaald als:

$$\delta\bar{J} = \frac{\partial\bar{J}}{\partial x}\delta x + \frac{\partial\bar{J}}{\partial y}\delta y + \frac{\partial\bar{J}}{\partial\lambda}\delta\lambda = (2x + \lambda)\delta x + (2y + \lambda)\delta y(x + y - k)\delta\lambda. \tag{14}$$

Hieruit volgt

$$\begin{aligned} 2x + \lambda = 0 \rightarrow x = -\lambda/2, \\ 2y - \lambda = 0 \rightarrow y = -\lambda/2, \\ x + y - k = 0 \rightarrow \lambda = -k, \end{aligned} \tag{15}$$

zodat we voor het minimum vinden:

$$(x_0, y_0) = (k/2, k/2). \tag{16}$$

6. Reservoir optimalisatie

Dezelfde techniek kan ook worden toegepast voor het vinden van het maximum van J zoals gedefinieerd in vergelijking (6) onder de nevenvoorwaarden (3). NB. We beschouwen nu alleen de nevenvoorwaarden die worden gegeven door de recursieve systeemfunctie $\mathbf{g}$. Eventuele extra nevenvoorwaarden zoals weergegeven in vergelijking (5) kunnen op eenzelfde wijze worden meegenomen.

Ook nu definiëren we een gemodificeerde doelfunctie waarbij de nevenvoorwaarden $\mathbf{g}$ worden meegenomen met behulp van een vector $\boldsymbol{\lambda}$ van multiplicatoren:

$$\bar{J} = \sum_{k=0}^{K-1} \left\{ J_k \left(\mathbf{u}(k), \mathbf{h}(k) \right) + \boldsymbol{\lambda}(k+1)^T \mathbf{g} \left(\mathbf{x}(k), \mathbf{x}(k+1), \mathbf{u}(k) \right) \right\}. \tag{17}$$

Merk op dat zowel $\mathbf{g}$ als $\boldsymbol{\lambda}$ een functie van de discrete tijd k zijn. Vergelijking (17) kan compact worden weergegeven met behulp van de definitie

$$L(k) = J_k \left(\mathbf{u}(k), \mathbf{h}(k) \right) + \boldsymbol{\lambda}(k+1)^T \mathbf{g} \left(\mathbf{x}(k), \mathbf{x}(k+1), \mathbf{u}(k) \right), \tag{18}$$

waarbij L bekend staat als de Lagrangiaan, een naam die afkomstig is uit de klassieke mechanica. Met behulp van definitie (18) reduceert vergelijking (17) tot

$$\bar{J} = \sum_{k=0}^{K-1} L(k) \tag{19}$$

waarbij L een functie is van de toestandsvectoren $\mathbf{x}(k)$ en $\mathbf{x}(k+1)$, de signaalvector $\mathbf{u}(k)$ en de multiplicator vector $\boldsymbol{\lambda}(k+1)$. De totale differentiaal (variatie) van $\bar{J}$ is dus

$$\delta \bar{J} = \sum_{k=0}^{K-1} \left[\frac{\partial L(k)}{\partial \mathbf{x}(k)} \right] \delta \mathbf{x}(k) + \sum_{k=0}^{K-1} \left[\frac{\partial L(k)}{\partial \mathbf{x}(k+1)} \right] \delta \mathbf{x}(k+1) +$$

$$\sum_{k=0}^{K-1} \left[\frac{\partial L(k)}{\partial \mathbf{u}(k)} \right] \delta \mathbf{u}(k) + \sum_{k=0}^{K-1} \left[\frac{\partial L(k)}{\partial \boldsymbol{\lambda}(k+1)} \right] \delta \boldsymbol{\lambda}(k+1). \tag{20}$$

Vergelijking (20) geeft de veranderingen in de gemodificeerde doelfunctie weer ten gevolge van storingen in $\mathbf{x}(k)$, $\mathbf{x}(k+1)$, $\mathbf{u}(k)$ en $\boldsymbol{\lambda}(k+1)$. Op het eerste gezicht lijkt dit ons probleem niet echt te vereenvoudigen want het enige wat we willen weten is hoe de signaalvector $\mathbf{u}(k)$ de doelfunctie uiteindelijk beïnvloedt. Het blijkt echter mogelijk te zijn om vergelijking (20) zodanig te manipuleren dat de meeste termen gelijk aan nul worden. Door middel van het splitsen van de sommaties en het verschuiven van de indices kan de vergelijking worden herschreven als

$$\delta \bar{J} = \left[\frac{\partial L(0)}{\partial \mathbf{x}(0)} \right] \delta \mathbf{x}(0) + \sum_{k=0}^{K-1} \left[\frac{\partial L(k-1)}{\partial \mathbf{x}(k)} + \frac{\partial L(k)}{\partial \mathbf{x}(k)} \right] \delta \mathbf{x}(k) +$$

$$\left[\frac{\partial LK-1}{\partial \mathbf{x}K} \right] \delta \mathbf{x}(K) + \sum_{k=0}^{K-1} \left[\frac{\partial L(k)}{\partial \mathbf{u}(k)} \right] \delta \mathbf{u}(k) + \sum_{k=0}^{K-1} \left[\frac{\partial L(k)}{\partial \boldsymbol{\lambda}(k+1)} \right] \delta \boldsymbol{\lambda}(k+1). \tag{21}$$

De eerste term in vergelijking (21) is gelijk aan nul omdat de beginvoorwaarde $\mathbf{x}(0)$ vast ligt volgens vergelijking (2) zodat geldt $\delta \mathbf{x}(0) = \mathbf{0}$. De laatste term in vergelijking (21) is ook gelijk aan nul omdat uit definitie (18) volgt dat

$$\frac{\partial L(k)}{\partial \boldsymbol{\lambda}(k+1)} = \mathbf{g} \left(\mathbf{x}(k), \mathbf{x}(k+1), \mathbf{u}(k) \right), \tag{22}$$

hetgeen per definitie gelijk is aan nul, zie vergelijking (3). Om vergelijking (21) verder te vereenvoudigen eisen we nu dat de tweede en de derde term ook gelijk aan nul worden, dat wil zeggen dat

$$\frac{\partial L(k-1)}{\partial \mathbf{x}(k)} + \frac{\partial L(k)}{\partial \mathbf{x}(k)} = \mathbf{0}, \tag{23}$$

$$\frac{\partial L(K-1)}{\partial \mathbf{x}(K)} = \mathbf{0}. \tag{24}$$

In dat geval blijft alleen de vierde term over zodat we kunnen schrijven

$$\delta \bar{J} = \sum_{k=0}^{K-1} \left[\frac{\partial L(k)}{\partial \mathbf{u}(k)} \right] \delta \mathbf{u}(k), \tag{25}$$

hetgeen precies is waar we naar op zoek waren, namelijk de verandering van de gemodificeerde doelfunctie ten gevolge van een storing in het stuursignaal; zie ook vergelijking (9). Daartoe moeten we dan echter wel eerst nagaan hoe vergelijkingen (23) en (24) verwezenlijkt kunnen worden. Met behulp van definitie (18) vinden we dat vergelijking (24) impliceert dat

$$\boldsymbol{\lambda}(K)^T = \mathbf{0}^T, \tag{26}$$

en dat vergelijking (23) impliceert dat

$$\boldsymbol{\lambda}(k)^T \frac{\partial \mathbf{g}(k-1)}{\partial \mathbf{x}(k)} + \frac{\partial J(k)}{\partial \mathbf{x}(k)} + \boldsymbol{\lambda}(k+1)^T \frac{\partial \mathbf{g}(k)}{\partial \mathbf{x}(k)} = \mathbf{0}, \tag{27}$$

ofwel dat

$$\boldsymbol{\lambda}(k)^T = -\left[\boldsymbol{\lambda}(k+1)^T \frac{\partial \mathbf{g}(k)}{\partial \mathbf{x}(k)} + \frac{\partial J(k)}{\partial \mathbf{x}(k)} \right] \left[\frac{\partial \mathbf{g}(k-1)}{\partial \mathbf{x}(k)} \right]^{-1}. \tag{28}$$

Vergelijking (28) is een recursieve uitdrukking voor $\boldsymbol{\lambda}(\mathbf{k})$ die terugloopt in de tijd, en die het mogelijk maakt om alle waarden van de multiplicatorvector te berekenen uitgaande van de beginvoorwaarde, of beter gezegd de eindvoorwaarde (26). Het blijkt dus dat we om de gewenste variaties (25) te verkrijgen twee vector-differentievergelijkingen moeten oplossen: de systeemvergelijking (3) met beginvoorwaarde (2) en de zogenaamde toegevoegde, of geadjungeerde vergelijking (28) met eindvoorwaarde (26). Optimalisatie van de debieten in de segmenten van de injectie- en productieputten verloopt daarom als volgt:

1. Simuleer de olieproductie voor een eerste schatting van het signaal $\mathbf{u}(k)$ door het numeriek oplossen van vergelijking (3), uitgaande van beginvoorwaarde(2).
2. Bereken de doelfunctie (6).
3. Bereken de multiplicatoren door het numeriek oplossen van de geadjungeerde vergelijking (28), uitgaande van eindvoorwaarde(26).

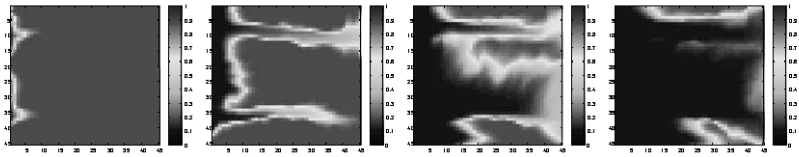
4. Bereken de variatie (25).
5. Gebruik een numerieke optimalisatietechniek voor het berekenen van een verbeterde schatting van het signaal $\mathbf{u}(k)$, waarbij gebruik wordt gemaakt van de gradiënten $\partial L(k)/\partial(\mathbf{u})$.

Stappen 1 tot en met 5 worden herhaald totdat de veranderingen in de gemodificeerde doelfunctie beneden een vooraf vastgestelde drempelwaarde zakken. De laatstgevonden waarde van $\mathbf{u}(k)$ wordt dan beschouwd als een numerieke benadering van het optimale stuursignaal. Doorgaans zijn zo'n 10 tot 20 iteratieslagen nodig om een voldoende nauwkeurig resultaat te bereiken. We gaan hier verder niet in op de details van de optimalisatietechniek, noch op de noodzakelijke voorwaarden om aan te tonen dat $\mathbf{u}(k)$ ook inderdaad een optimum is. In het algemeen kunnen we slechts aantonen dat we een lokaal optimum hebben gevonden. Om het globale optimum te vinden zou de procedure herhaald moeten worden voor een groot aantal verschillende startwaarden van $\mathbf{u}(k)$. In onze ervaring blijkt echter dat het vinden van een lokaal optimum, uitgaande van een conventionele beginstrategie voor $\mathbf{u}(k)$, meestal bevredigende resultaten geeft. Voor verdere informatie over optimalisatie met behulp van geadjungeerde functies (in het Engels bekend als 'optimal control') verwijzen we naar, bijvoorbeeld, het boek van Stengel (1994). Voor specifieke toepassing op het optimaliseren van reservoirstroming met slimme putten, zie Brouwer (2004) of Brouwer en Jansen (2004).

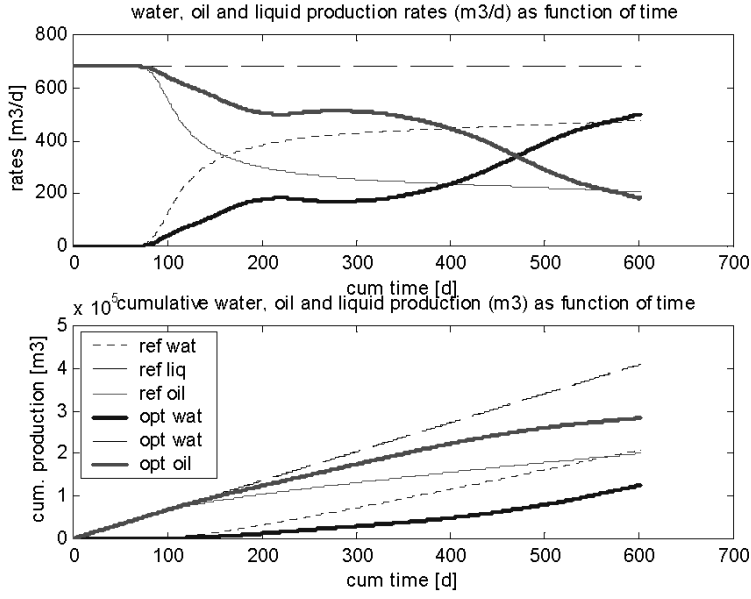
7. Voorbeeld en discussie

Als voorbeeld beschouwen we de geoptimaliseerde injectie- en productiestrategie voor het eerder beschouwde tweedimensionale reservoir. Figuur 9 toont de waterverzadiging in het reservoir voor het geoptimaliseerde geval. Uit vergelijking met Figuur 6 blijkt dat na injectie van één porievolume water een aanzienlijk grotere hoeveelheid olie uit het reservoir is verdrongen. Dit blijkt ook uit het bijbehorende productieprofiel in Figuur 10 (boven), waarin duidelijk zichtbaar is dat de waterproductie fors is gereduceerd en de olieproductie toegenomen. Figuur 10 (beneden) toont de cumulatieve geproduceerde volumes olie en water, en maakt eveneens duidelijk dat de winningsfactor is toegenomen. Wat niet zichtbaar is in deze figuren, maar wat een essentieel aspect van onze aanpak vormt, is de variatie van de injectie- en productiedebieten in de putsegmenten over de tijd. Dit is anders dan in de conventionele strategie waarbij de debieten constant blijven.

Uiterst aanvechtbaar in onze optimalisatie is de veronderstelling dat het reservoirmodel de werkelijkheid voldoende nauwkeurig weergeeft. We weten namelijk vrij zeker dat het model niet klopt. Een meer betrouwbare aanpak maakt daarom gebruik van een model dat regelmatig wordt aangepast, dat wil zeggen van 'closed-loop reservoir management' zoals geschematiseerd in Figuur 1. Wij zullen hier verder niet ingaan op deze uitgebreide aanpak van het optimalisatieprobleem in een gesloten-lus configuratie, maar vermelden slechts dat



Figuur 9. *(Voor kleurenillustratie zie pagina 83–86.) Waterverzadiging van het reservoir op vier verschillende tijstippen voor een geoptimaliseerde productiestrategie. De vier plaatjes corresponderen met vier geïnjecteerde porievolumes. Van links naar rechts: 0.1, 0.4, 0.7 en 1.0 porievolumes geïnjecteerd. Rood: olie; blauw:water.*



Figuur 10. *(Voor kleurenillustratie zie pagina 83–86.) Boven: Olie- en waterproductie als functie van de tijd voor een conventionele productiestrategie (dunne lijnen) en een geoptimaliseerde productiestrategie (dikke lijnen). Beneden: Bijbehorende cumulatieve geproduceerde volumes.*

de resultaten, hoewel minder spectaculair, nog steeds goed zijn. Voor verder informatie, zie bijvoorbeeld Brouwer et. al (2004) of Jansen et al. (2005). Ook andere aannames in ons voorbeeld maken dat de bereikte resultaten veel beter zijn dan die voor een realistisch reservoir model verwacht kunnen worden. Dat betreft bijvoorbeeld het tweedimensionale karakter van het reservoir, zodat zwaartekracht geen rol speelt, de korte tussenruimte tussen de putten, en de mogelijkheid om segmenten in te sluiten en het volledige debiet over de andere kleppen te verdelen. Hoewel in realistische situaties de mogelijkheden

voor optimalisatie daarom veel beperkter zullen zijn, hopen we toch te hebben duidelijk gemaakt dat het optimaliseren van de injectie- en productiestrategie in 'slimme olievelden' een veelbelovende aanpak is voor het verhogen van de winningsfactor.

Appendix - Reservoirmodellen

De wiskundige formulering van modellen voor stroming door poreuze media is gebaseerd op een aantal fysische principes. Allereerst maken we gebruik van de wet van behoud van massa voor de verschillende chemische componenten van de reservoirvloeistoffen. Om de zaak eenvoudig te houden beschouwen we hier de stroming van olie en water door een ééndimensionaal horizontaal reservoir met constante doorsnede. Voor een infinitesimaal klein controle volume kunnen we de massabalansen voor water en olie dan weergeven als

$$\frac{\partial}{\partial x}(\rho_w v_w) + \frac{\partial(\rho_w \phi S_w)}{\partial t} - \rho_w q_w = 0, \tag{A.1}$$

$$\frac{\partial}{\partial x}(\rho_o v_o) + \frac{\partial(\rho_o \phi S_o)}{\partial t} - \rho_o q_o = 0, \tag{A.2}$$

waarin x de ruimtelijke coördinaat langs het reservoir is, t de tijd, ϕ de porositeit van het gesteente, ρ_o en ρ_w de dichtheden van olie en water, v_o en v_w de stroomsnelheden, S_o en S_w de verzadigingen, en q_o en q_w de brontermen die alleen van belang zijn als er een put in het controle volume uitmondt. De verzadigingen zijn getallen met een waarde tussen nul en één, die bovendien tezamen altijd één moeten zijn omdat we aannemen dat de porieruimte altijd geheel gevuld is:

$$S_w + S_o = 1, \tag{A.3}$$

Het tweede fysische principe dat een rol speelt staat bekend als de 'wet van Darcy'. Het gaat daarbij om een empirisch verband tussen de drukval $\partial p/\partial x$ en de stroomsnelheid v in een poreus medium. Voor olie- en waterstroming kunnen we schrijven:

$$v_w = -\frac{k k_{rw}}{\mu_w}\frac{\partial p_w}{\partial x}, \tag{A.4}$$

$$v_o = -\frac{k k_{ro}}{\mu_o}\frac{\partial p_o}{\partial x}, \tag{A.5}$$

waarbij p_o en p_w de olie- en waterdrukken zijn, k de permeabiliteit, dat wil zeggen de doorlatendheid van het gesteente, ρ_o en ρ_w de viscositeiten, en k_o en k_w de relatieve permeabiliteiten, reductiefactoren met een waarde tussen nul en één die de invloed van de aanwezigheid van de ene vloeistof op het stromingsgedrag van de andere weergegeven. De drukken in olie en water zijn niet aan elkaar gelijk vanwege capillaire effecten. Doorgaans wordt verondersteld dat de capillaire druk (dat wil zeggen het drukverschil tussen water en olie) en de relatieve permeabiliteiten een functie zijn van de waterverzadiging:

$$p_o - p_w = p_e(S_w) \tag{A.6}$$

$$k_o = k_w(S_w) \tag{A.7}$$

$$k_w = k_w(S_w) \tag{A.8}$$

Het laatste fysische principe betreft het verband tussen enerzijds de olie- en waterdrukken en anderzijds de respectievelijke dichtheden, en wordt uitgedrukt met behulp van 'toestandsvergelijkingen':

$$c_o = \left.\frac{1}{\rho_o}\frac{\partial \rho_o}{\partial p_o}\right|_{T_0}, \tag{A.9}$$

$$c_w = \left.\frac{1}{\rho_w}\frac{\partial \rho_w}{\partial p_w}\right|_{T_0} \approx \left.\frac{1}{\rho_w}\frac{\partial \rho_w}{\partial p_o}\right|_{T_0}, \tag{A.10}$$

waarbij T_0 de (constante) reservoir temperatuur voorstelt, en c_o en c_w de samendrukbaarheid van olie en water. Een vergelijkbare uitdrukking valt op te stellen voor de samendrukbaarheid van het gesteente:

$$c_r = \frac{1}{\phi}\frac{\partial \phi}{\partial p_o}, \tag{A.11}$$

Combinatie van vergelijkingen (A.1) tot en met (A.11) resulteert in twee partiële differentiaalvergelijkingen:

$$-\frac{\partial}{\partial x}\left[\frac{\rho_w k k_{rw}}{\mu_w}\left(\nabla p_o - \frac{\partial p_c}{\partial S_w}\frac{\partial S_w}{\partial x}\right)\right] + \rho_w \phi\left[S_w(c_w + c_r)\frac{\partial p_o}{\partial t} + \frac{\partial S_w}{\partial t}\right] \tag{A.12}$$
$$-\rho_w q_w = 0,$$

$$-\frac{\partial}{\partial x}\left[\frac{\rho_w k k_{ro}}{\mu_o}\frac{\partial}{\partial x}p_o\right] + \rho_o \phi\left[(1 - S_w)(c_o + c_r)\frac{\partial p_o}{\partial t} - \frac{\partial S_w}{\partial t}\right] - \rho_o q_o = 0. \tag{A.13}$$

De onafhankelijke variabelen in vergelijkingen (A.12) en (A.13) zijn de ruimtelijke coördinaat x en de tijd t, de afhankelijke variabelen de oliedruk p_o en de waterverzadiging S_w. De vergelijkingen zijn sterk niet-lineair vanwege de saturatie-afhankelijkheid van de relatieve permeabiliteiten k_o en k_w en de capillaire druk p_c. Samen met begin- en randvoorwaarden geven de vergelijkingen een beschrijving voor de gelijktijdige horizontale stroming van olie en water door een poreus gesteente met een bij benadering ééndimensionale geometrie. Voor twee- en driedimensionale stroming kunnen met behulp van vectordifferentiaalrekening vergelijkingen met eenzelfde karakter worden afgeleid. Nadere beschouwing van de vergelijkingen (A.12) en (A.13) leert dat ze kunnen worden herschreven als twee afzonderlijke vergelijkingen voor de oliedruk en de waterverzadiging, gekoppeld door de coëfficiënten. De drukvergelijking blijkt dan een lineaire of licht niet-lineaire parabolische (diffusie) vergelijking te zijn met coëfficiënten die een functie zijn van de waterverzadiging. De verzadigingsvergelijking is een sterk niet-lineaire parabolisch-hyperbolische (niet-lineaire

convectie) vergelijking met een convectieterm die afhankelijk is van de drukgradiënten. Voor verdere details, zie Aziz en Settari (1979) en Ewing (1983).

Discretisatie in de ruimtelijke coördinaten van de partiële differentiaalvergelijkingen geschiedt doorgaans met behulp van eindige differenties of eindige elementen, resulterend in een stelsel van gewone differentiaalvergelijkingen dat met behulp van matrix-vector notatie kan worden uitgedrukt als

$$\begin{bmatrix} \mathbf{V}_{wp} & \mathbf{V}_{ws} \\ \mathbf{V}_{op} & \mathbf{V}_{os} \end{bmatrix} \begin{bmatrix} \dot{\mathbf{p}} \\ \dot{\mathbf{s}} \end{bmatrix} + \begin{bmatrix} \mathbf{T}_{wp} & \mathbf{T}_{ws} \\ \mathbf{T}_{op} & \mathbf{T}_{os} \end{bmatrix} \begin{bmatrix} \mathbf{p} \\ \mathbf{s} \end{bmatrix} = \begin{bmatrix} \mathbf{q}_w \\ \mathbf{q}_o \end{bmatrix} \tag{A.14}$$

Hierin representeren de vectoren $\mathbf{p}$ en $\mathbf{s}$ de oliedrukken en waterverzadigingen in de roosterpunten. De vectoren $\mathbf{q}_w$ en $\mathbf{q}_o$ zijn brontermen met elementen die alleen nul zijn voor de roosterpunten die corresponderen met een put. De matrices $\mathbf{V}_{wp}$ etc. zijn accumulatietermen die een lichte afhankelijkheid van $\mathbf{p}$ en $\mathbf{s}$ vertonen. De matrices $\mathbf{T}_w$ etc. staan bekend als de doorlatendheidsmatrices en zijn doorgaans sterk afhankelijk van $\mathbf{s}$. Ondanks de ogenschijnlijk lineaire structuur van vector-differentiaalvergelijking (A.14) is er dus wel degelijk sprake van een niet-lineaire vergelijking. Een meer compacte schrijfwijze voor vergelijking (A.14) kan worden verkregen door gebruik te maken van een 'toestandsruimtebeschrijving' zoals die wordt gebruikt in de systeemtheorie:

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}(t), \mathbf{u}(t)). \tag{A.15}$$

Hierbij bevat de toestandsvector $\mathbf{x}$ zowel de drukken $\mathbf{p}$ als de verzadigingen $\mathbf{s}$, en de signaalvector $\mathbf{u}$ de ongelijk-nul elementen van brontermen $\mathbf{q}_o$ en $\mathbf{q}_w$, terwijl de vectorfunctie $\mathbf{f}$ de niet-lineaire accumulatie en doorlatendheidseffecten representeert[9]. De dimensie van de toestandsruimte is gelijk aan de lengte van de toestandsvector $\mathbf{x}$, en is dus gelijk aan twee maal het aantal roosterpunten. Indien naast olie en water ook de stroming van gas wordt meegenomen neemt de dimensie van de toestandsruimte toe tot drie maal het aantal roosterpunten. Discretisatie in de tijd van vergelijking (A.15) gebeurt doorgaans met een impliciete Euler discretisatie wat leidt tot een stelsel niet-lineaire algebraïsche vergelijkingen dat kan worden weergegeven als

$$\mathbf{x}(k+1) = \mathbf{f}(\mathbf{x}(k), \mathbf{u}(k+1)). \tag{A.16}$$

waarbij k de discrete tijd $t(k)$ representeert.

Litteratuurverwijzing

1. Aziz, K. and Settari, A., 1979: *Petroleum Reservoir Simulation*, Applied Science Publishers.
2. Brouwer, D.R., 2004: Dynamic water flood optimization with smart wells using optimal control theory. *Proefschrift*, Technische Universiteit Delft.
3. Brouwer, D.R. and Jansen, J.D., 2004: Dynamic optimisation of water flooding with smart wells using optimal control theory. *SPE Journal*, December, 391-402.

[9] De vector $\mathbf{x}$ heeft niets te maken met de ruimtelijke coördinaat x die eerder werd gebruikt.

4. Brouwer, D.R., Naevdal, G., Jansen, J.D., Vefring, E. and van Kruijsdijk, C.P.J.W., 2004: Improved reservoir management through optimal control and continuous model updating, paper SPE 90149, SPE Annual Technical Conference and Exhibition, Houston, USA, September.
5. Ewing, R.E., 1983: Problems arising in the modeling of processes for hydrocarbon recovery, in: *The mathematics of reservoir simulation*, Ewing, R.E. (ed.), SIAM, Philadelphia.
6. Jansen, J.D., Brouwer, D.R., Nævdal, G. and van Kruijsdijk, C.P.J.W., 2005: Closed-loop reservoir management. *First Break*, January, 23, 43-48.
7. Stengel, R.F., 1994: *Optimal control and estimation*, Dover.

CWI SYLLABI

1 Vakantiecursus 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 Vakantiecursus 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 Vakantiecursus 1987: *De personal computer en de wiskunde op school.* 1987.
15 Vakantiecursus 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 Vakantiecursus 1988. *Differentierekening.* 1988.
19 R. de Bruin, C.G. van der Laan, J. 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 Vakantiecursus 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 Vakantiecursus 1990: *Getallentheorie en haar toepassingen.* 1990.
28 Vakantiecursus 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 Vakantiecursus 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 Vakantiecursus 1993: *Het reële getal.* 1993.
36 Vakantiecursus 1994: *Computeralgebra.* 1994.
37 G. Alberts. *Wiskunde en praktijk in historisch perspectief. Syllabus.* 1994.
38 G. Alberts, J. Schut (eds.). *Wiskunde en praktijk in historisch perspectief. Reader.* 1994.
39 E.A. de Kerf, H.G.J. Pijls (eds.). *Proceedings Seminar 1989–1990. Mathematical structures in field theory.* 1996.
40 Vakantiecursus 1995: *Kegelsneden en kwadratische vormen.* 1995.
41 Vakantiecursus 1996: *Chaos.* 1996.
42 H.C. Doets. *Wijzer in Wiskunde.* 1996.
43 Vakantiecursus 1997: *Rekenen op het Toeval.* 1997.
44 Vakantiecursus 1998: *Meetkunde, Oud en Nieuw.* 1998.
45 Vakantiecursus 1999: *Onbewezen Vermoedens.* 1999.
46 P.W. Hemker, B.W. van de Fliert (eds.). *Proceedings of the 33^{rd} European Study Group with Industry.* 1999.
47 K.O. Dzhaparidze. *Introduction to Option Pricing in a Securities Market.* 2000.
48 Vakantiecursus 2000: *Is wiskunde nog wel mensenwerk?* 2000.
49 Vakantiecursus 2001: *Experimentele wiskunde.* 2001.
50 Vakantiecursus 2002: *Wiskunde en gezondheid.* 2002.
51 G.M. Hek (ed.). *Proceedings of the 42^{nd} European Study Group with Industry.* 2002.
52 Vakantiecursus 2003: *Wiskunde in het dagelijks leven.* 2003.
53 Vakantiecursus 2004: *Structuur in schoonheid.* 2004.
54 Vakantiecursus 2005: *De schijf van vijf – meetkunde, algebra, analyse, discrete wiskunde, stochastiek.* 2005.
55 J. Hulshof (ed.). *Proceedings of the 52^{nd} European Study Group with Industry.* 2006.
56 Vakantiecursus 2006: *Actuele wiskunde.* 2006.

MC SYLLABI

1.1 F. Göbel, J. van de Lune. Leergang besliskunde, deel l: 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. 1967.
1.6b B. Dorhout, J. Kriens, J.Th. van Lieshout. Leergang besliskunde deel 6b: wiskundige programmering. 1967
1.7a G. de Leve. Leergang besliskunde, deel 7a: dynamische programmering 1. 1969
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 7e: 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, RP. 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 rekenmachines, deel 1. 1969.
7.2 Colloquium moderne rekenmachines, 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, EM. de Jager, N.M. Temme, R.J. de Vogelaere. Colloquium elliptische differentiaalvergelijkingen, deel 2. 1970.
10 l. 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, H.C. 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 Iongh, 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 programma-correctheid. 1975.
22 R. Helmers, J. Oosterhoff, F.H. Ruymgaart, M.C.A. van Zuylen. Asymptotische methoden in de toe-tsingstheorie; toepassingen van naburigheid. 1976.
23.1 J.W. de Roever (red.). Colloquium onderwerpen uit de biomathematica, deel l. 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: eenstapsmethaden. 1974.
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, SJ. Polak, M, van Veldhuizen. Colloquium discretiseringsmethoden. 1976.
28 O. Diekmann, N.M. Temme (eds.). Nonlinear diffusion problems. 1976.
29.1 J.C.P. Bus (red.). Colloquium numerieke programmatuur, deel 1A, deel 1 B. 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 bedrijfsystemen. 1976.
33 P..J. van der Houwen. Berekening van waerstanden 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 l. 1981.
46.2 P.G.M. 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 procedures in ALGOL 60, vol. l: elementary procedures; vol. 2: algebraic evaluatons. 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, procedures in ALGOL 60, vol. 4: analytical evaluations; vol. 5A: analytical problems. part I. 1981
47.6 P.W. Hemker (ed.). NUMAL, procedures in ALGOL 60, vol 5B: analytical problems, part II. 1981
47.7 P.W. Hemker (ed.). NUMAL, 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 II. 1982.
49 T.H. Koomwinder (ed.) The structure of real semisimple Lie groups. 1982
50 H. Nijmeijer. Inleiding systeemtheorie. 1982.
51 P..J. Hoogendoorn (red.). Cursus cryptografie. 1983.