

Elementaire functies gebruiken als modellen

Opgaven voor de 3^{de} graad van het secundair onderwijs

Ria Wouters
Guido Herweyers

Inhoud

Inleiding	1
Deel 1 : Functies als modellen voor functies	4
Opgave 1	
Opgave 2	
Opgave 3	
Opgave 4	
Opgave 5	
Regressie en gebruik van modellen	7
Opgave 6	
Opgave 7	
Opgave 4 bis	
Opgave 8 : de kettinglijnfunctie / de functies <i>cosh</i> en <i>sinh</i>	
Opgave 9	
Opgave 10	
Opgave 11 : groei van zonnebloemen	
Opgave 12 : tijd om een belegde som te verdubbelen	
Deel 2 : Toepassingen	
Opgave 13 : het boorplatform 'Yatzy'	13
Opgave 14 : blokrijden als remedie voor het fileprobleem ?	18
Opgave 15 : afkoelingskrommen	21
Bronnen	24
Bijlage 1 : Regressie, correlatie en modelvorming met de TI-83 (Plus) (Achtergrondinformatie voor de leerkracht)	25
Bijlage 2 : Suggesties voor het oplossen van de opgaven	34

Inleiding

Wat is de bedoeling van deze opgaven ?

Deze opgavenreeks wil een hulp zijn om

- de kennis over *reële functies* te verruimen via het gebruik van *modellen*,
- het concept '*model*' (beter) te leren kennen via het gebruik van *functies*,

en dit aan de hand van concrete opdrachten waarbij gebruik gemaakt wordt van enkele eenvoudige basisfuncties van een grafisch rekentoestel (bijvoorbeeld TI-83 plus).

Welke functies komen voor in de opgaven ?

Op welke wiskunde-voorkennis wordt gesteund ?

Hoe is het materiaal van de opgavenreeks concreet bruikbaar ?

In de opgaven die aangeboden worden, komen elementaire reële functies voor die in de tweede en de derde graad bestudeerd worden : algebraïsche, goniometrische, cyclometrische, exponentiële en logaritmische functies. Er wordt een basiskennis van deze functies verondersteld. Als dusdanig kunnen sommige van de opdrachten in het bestaande programma zonder meer ingepast worden.

Het is echter de bedoeling dat leerlingen *alle* soorten reële functies in zekere mate leren 'beheersen'. Daarom wordt getracht geleidelijk aan de 'elementaire' vorm van de functies los te laten door in sommige opgaven verschillende elementaire functies te combineren, naast elkaar te plaatsen en te vergelijken. In die optiek zijn deze opgaven uiteraard pas bruikbaar en interessant na de behandeling van de afzonderlijke 'elementaire' functies die in de opgave voorkomen.

Omdat het concept 'model' in de huidige leerplannen vrijwel niet aan bod komt, hebben we getracht om via de opbouw van de opdrachten het begrip in te voeren en diverse aspecten ervan te belichten. Hierbij wordt gesteund op de voorkennis van de hierboven vermelde elementaire functies. Bij de opgaven zal op een bepaald ogenblik gebruik gemaakt worden van de regressiemodellen die op het rekentoestel beschikbaar zijn. Een basiskennis van de elementaire functies en een beknopte uitleg over het gebruik van regressiemodellen bij het rekentoestel volstaan hier. Uitdieping van het begrip 'regressie' is facultatief en wordt behandeld in bijlage 1 als achtergrondinformatie voor de leerkracht.

Ook vanuit het oogpunt 'modellen' zijn de verschillende opgaven in te lassen in het bestaande programma indien rekening gehouden wordt met de voorkennis van de elementaire functies die in de opgave voorkomen.

Een alternatief voor het inlassen van de afzonderlijke opgaven in het bestaande programma bestaat erin om gedurende een aantal weken in de derde graad wekelijks één of twee lessen te besteden aan deze opdrachten, parallel aan het basisprogramma. Dit is ook interessant omdat deze opgaven vaak via een alternatieve werkvorm aangeboden kunnen worden (groepswork, projectwork, vakoverschrijdend project voor de toepassingen etc.)

Enkele bijbedoelingen van de opgaven

Elementaire functies leren gebruiken buiten de klassieke context.

De behandeling van deze opgaven nà en los van de systematische studie van de elementaire functies, biedt enkele voordelen : de functies moeten herhaald worden en er wordt teruggekoppeld naar vroeger opgedane kennis. Leerlingen moeten hier doelgericht de essentie van wat vroeger geleerd werd opnieuw opzoeken (verloop van de functies, eigenschappen, rekenregels etc.). Wiskunde wordt hier als werkinstrument gebruikt.

Verwerven van vakoverschrijdende vaardigheden

Het werken met het begrip ‘model’ laat een alternatieve benadering van de wiskunde toe, namelijk die van de *benadering*. Vaardigheden als *nuanceren*, *argumenteren* en *presenteren* worden eveneens aangesproken.

Het gebruik van het grafisch rekentool bij de behandeling van functies biedt enorme mogelijkheden, maar het is tegelijkertijd ook een erg ‘vluchtig’ middel. Er zal dan ook uitvoerig aandacht besteed moeten worden aan de neerslag en de presentatie van de resultaten.

Welke rekentoolfuncties worden gebruikt in de opgavenreeks ?

Grafische voorstelling van functies.

Invoeren van gegevens in tabellen.

Werken met regressiemodellen.

Oplossen van stelsels lineaire vergelijkingen m.b.v. matrices.

Hoe is deze tekst praktisch opgevat ?

De ‘*reële functie als model*’ verschijnt in de opgavenreeks onder twee gedaanten :

- Deel 1 : Elementaire functies als modellen voor (moeilijkere) functies.
- Deel 2 : Modellen als beschrijving van fenomenen (toepassingsgericht).

De tekst bestaat uit de *opgaven* zoals zij aan de leerlingen kunnen aangeboden worden en een *begeleidende tekst* waarin de specifieke doelstellingen worden toegelicht.

In een bijlage 2 worden beknopt enkele suggesties gegeven voor het oplossen van de opgaven. De concrete uitwerking van de oplossingen dient aangepast te worden aan de klasgroep.

Deel 1 : Functies als modellen voor functies

In de opgaven 1 en 2 wordt gevraagd functievoorschriften van veeltermfuncties op te stellen. De bedoeling hiervan is om kort het algemeen verloop van een n -de graadsveeltermfunctie in herinnering te brengen, o.a. door de confrontatie met het functievoorschrift (aantal termen en coëfficiënten), het opstellen van het functievoorschrift m.b.v. nulpunten etc., en dit als voorbereiding op de volgende opgaven waarbij andere reële functies zullen *benaderd* worden door veeltermfuncties.

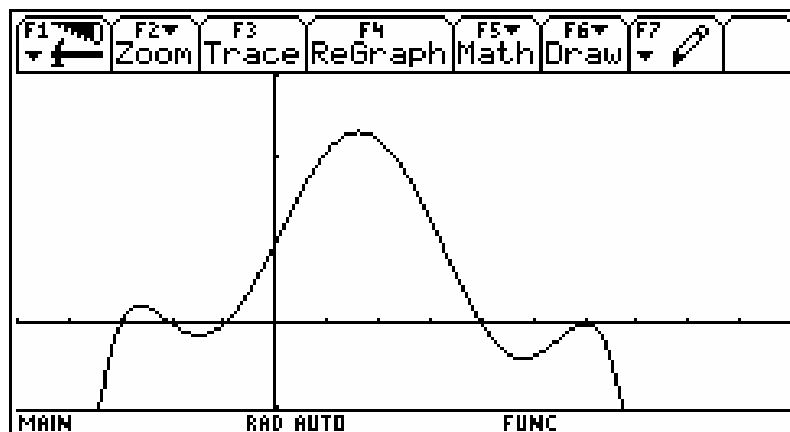
Opgave 2 geeft aanleiding tot een 5×5 - stelsel dat met het rekenoestel kan worden opgelost.

Opgave 1

Teken onderstaande grafiek op uw grafisch rekenoestel

(functievoorschrift :

)

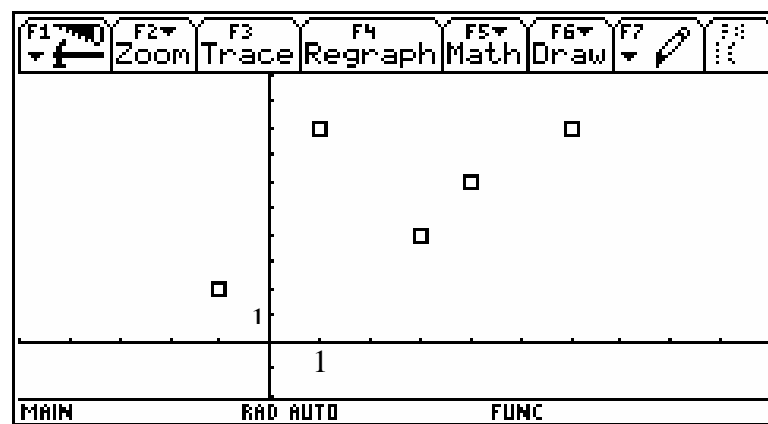


Opgave 2

Stel het voorschrift op van een veeltermfunctie door de punten

$(1,8)$, $(3,4)$, $(4,6)$, $(6,8)$ en $(-1,2)$

en teken de veeltermfunctie op uw rekenoestel.



In de opgaven 3 , 4 en 5 wordt gevraagd om enkele reële functies (met gegeven functievoorschrift) te *benaderen* d.m.v. een veeltermfunctie. De veeltermfunctie wordt hier dus gebruikt als *model* voor andere reële functies.

De eerste stap is telkens het tekenen van de grafiek van de functie. Uit de grafiek kan de graad van de veeltermfunctie afgeleid worden. Op basis van de keuze van een aantal punten kan dan het voorschrift van *een* benaderende veeltermfunctie opgesteld worden. De benadering wordt grafisch vergeleken met de oorspronkelijke functie.

In de opgave 3 wordt de cosinusfunctie benaderd in verschillende domeinintervallen. Er kan gediscussieerd worden over het aantal en de keuze van de gebruikte punten.

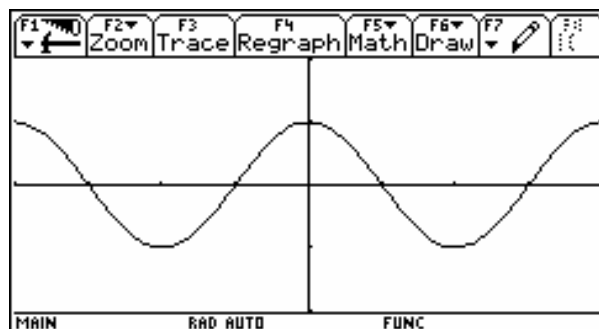
Het kan interessant zijn om hier de vergelijking te maken met de MacLaurin-veeltermbenaderingsformule indien deze vroeger behandeld werd in de Analyse. In dit licht kan men hier deze n -de graadsbenadering volgens MacLaurin her-interpreteren...

Opgave 3 : de cosinusfunctie

Benader de cosinusfunctie d.m.v. een veeltermfunctie

- a) in het domeininterval $[-\pi/2, \pi/2]$
- b) in het domeininterval $[-\pi, \pi]$

Vergelijk de beide benaderingen met de cosinusfunctie zelf.



In de opgave 4 kan de keuze van de punten (het aantal ? welke ?) een discussie op gang brengen, en achteraf kan het resultaat van de benadering gekoppeld worden aan de keuze van de punten. De discussie is een aanloop naar het gebruik van een regressiemodel voor een (3^{de} graads)veeltermbenadering.

Opgave 4

Benader $f(x) = x^2 - 2^x$ door een veeltermfunctie.

De opgave 5 is interessant als terugkoppeling naar de goniometrie ; een samengestelde goniometrisch-cyclometrische functie is een algebraïsche functie in het domeininterval, in dit geval een 5^{de} graadsveeltermfunctie. De ‘benadering’ is hier dus perfect !

Opgave 5

Benader $g(x) = \sin(5 \operatorname{Bgsin} x)$ door een veeltermfunctie.

Hoe nauwkeurig is de benadering ?

De opgaven 6 en 7 zijn opdrachten waarbij een model moet opgesteld worden op basis van gegevens in tabelvorm. Hier kunnen we beroep doen op de regressiemodellen beschikbaar op het rekentoestel. Een beknopte uitleg over deze regressiemodellen is hier ingelast. Voor een uitgebreide behandeling verwijzen we naar bijlage 1.

Hier kan reeds opgemerkt worden dat de beschikbare regressiemodellen niet beperkt blijven tot veeltermfuncties.

Regressie en het gebruik van modellen

Bij regressie wordt voor een aantal x - y -gegevens het beste (*) functievoorschrift berekend op basis van een gekozen functiemodel.

(*) de berekening gebeurt door de som van de kwadraten van de verticale afwijkingen tussen de data en het model te minimaliseren (kleinste kwadraten criterium).

De determinatiecoëfficiënt R^2 is een maat voor de kwaliteit van de benadering (in het beste geval is $R^2 = 1$, d.w.z. dat alle gegeven punten behoren tot de modelfunctie waarvan het functievoorschrift berekend werd).

Voorbeeld :

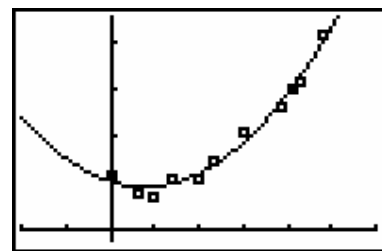
gegeven : een aantal punten (x,y)

gekozen model : 2de graadsveeltermfunctie

berekend functievoorschrift :

$$y = 0.189690 x^2 - 0.299425 x + 1.041210$$

determinatiecoëfficiënt $R^2 : 0.9788$



```

2nd [MODE] CALC TESTS
1: Edit...
2: SortA(
3: SortD(
4: ClrList
5: SetUpEditor
    
```

L1	L2	L3	Z
2.3	1.5		
3	2.1		
3.9	2.6		
4.1	3		
4.3	3.1		
4.8	4.1		

L2(12) =			

```

CATALOG
DelVar
DependAsk
DependAuto
det(
DiagnosticOff
DiagnosticOn
dim(
    
```

```

EDIT [MODE] TESTS
1: 1-Var Stats
2: 2-Var Stats
3: Med-Med
4: LinReg(ax+b)
5: QuadReg
6: CubicReg
7: QuartReg
    
```

```

DiagnosticOn
Done
QuadReg L1,L2,Y1
    
```

```

QuadReg
y=ax^2+bx+c
a=.1896901878
b=-.2994245389
c=1.041210373
R^2=.9787692481
    
```

Beschikbare regressiemodellen op het grafisch rekenoestel (TI-83 plus)

Lineair	$a.x + b$	LinReg
Kwadratisch	$a.x^2 + b.x + c$	QuadReg
Derdegraads	$a.x^3 + b.x^2 + c.x + d$	CubicReg
Vierdegraads	$a.x^4 + b.x^3 + c.x^2 + d.x + e$	QuartReg
Logaritmisch	$a + b.\ln(x)$	LnReg
Exponentieel	$a.b^x$	ExpReg
Machten	$a.x^b$	PowerReg
Sinus	$a.\sin(b.x+c)+d$	SinReg
Logistiek	$c/(1+a.e^{-b.x})$	Logistic

Bij het oplossen van de opgaven 6 en 7 worden zowel de data als de benadering in grafiek gebracht. De verschillende benaderingen kunnen vergeleken worden, grafisch en via de determinatiecoëfficiënt R^2 .

Opgave 6

De tabel hiernaast geeft het resultaat van metingen van het verbruik van een wagen bij verschillende snelheden.

Gevraagd wordt het verbruik uit te drukken als functie van de snelheid. (Maak een vergelijking tussen verschillende functiemodellen)

snelheid in km / u	verbruik in liter / 100 km
10	21
20	12
30	10
40	8
50	7
60	5.9
70	6.3
80	6.95
90	7.57
100	8.27
110	9.03
120	8.87
130	10.79
140	11.77
150	12.83

Opgave 7

Ik wil mijn auto (bouwjaar 1997) verkopen aan een vriend voor een zo correct mogelijke prijs. Ik beschik over onderstaande verkoopprijzen van hetzelfde model, afkomstig uit recente kranten en streekbladen.

Welke prijs zal ik vragen ?

jaar van productie	vraagprijs / 1000 euro
1998	6.0
1999	7.7
2000	8.8
1995	3.4
2001	9.8
1999	8.4
2001	8.9
1993	1.5
1992	1.6
1990	1.4
1994	2.0
1987	1.0

Opgave 4 bis

Benader $f(x) = x^2 - 2^x$ door een veeltermfunctie door gebruik te maken van één van de regressiemodellen die beschikbaar zijn op uw rekentoestel.

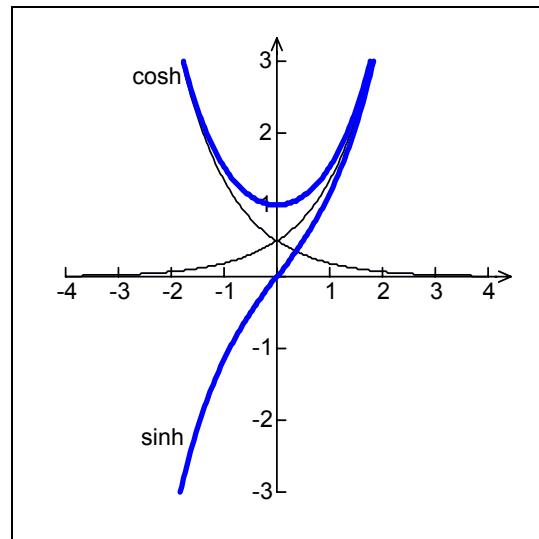
In de opgaven 8, 9 en 10 worden elementaire functies gecombineerd. Het is hier interessant om op zoek te gaan naar de 'elementaire' bestanddelen van de nieuwe functie en deze eventueel te modelleren.

Opgave 8 : de kettinglijnfunctie / de functies *cosh* en *sinh*

Een vrij opgehangen touw (ketting) heeft de vorm van de grafiek van de functie :

$$f(x) = \frac{e^{ax} + e^{-ax}}{2a}$$

waarbij a een constante is, afhankelijk van de lineaire dichtheid van het touw en van de spankracht in het touw.



De functie voor $a = 1$ wordt gedefinieerd als *cosinus hyperbolicus-functie (cosh)*

$$\cosh x = \frac{e^x + e^{-x}}{2}$$

De *sinus hyperbolicus-functie (sinh)* wordt gedefinieerd als

$$\sinh x = \frac{e^x - e^{-x}}{2}$$

Benader de kettinglijn door een parabool en vergelijk beide.

Benader de *sinh* functie door een veeltermfunctie.

Opgave 9

Is de functie $y = \log(\sin(3x) + 2)$ een *periodieke functie* ?

Is deze functie een *goniometrische functie* ?

Zo ja, geef het functievoorschrift in goniometrische vorm, zo neen, geef een benaderend goniometrisch functievoorschrift.

Opgave 10

Beschouw de functie $y = x^2 + \cos x$

Hoe verwacht je dat de grafiek van deze functie verloopt ?

Klopt de voorspelling ? Geef een verklaring voor het verloop van de functie.

Bewijs dat de grafiek geen slingerbeweging maakt.

Op welke manier(en) kan je het functievoorschrift wijzigen om toch een 'slingerende' grafiek te krijgen ?

Bepaal in dat geval het aantal relatieve extrema en het aantal buigpunten van de functie.

Totnogtoe werd in de modeleringsopgaven hoofdzakelijk beroep gedaan op veeltermbenaderingen. Om te vermijden dat leerlingen 'modeleren' steeds gaan associëren met 'veeltermfuncties' is opgave 11 ingelast. Hier wordt een voor hen 'nieuwe' functie geïntroduceerd : de *logistieke functie*. Aan de hand van de opgave zullen zij deze functie kunnen ontleden en gebruiken als groeimodel.

Opgave 11 : Groei van zonnebloemen

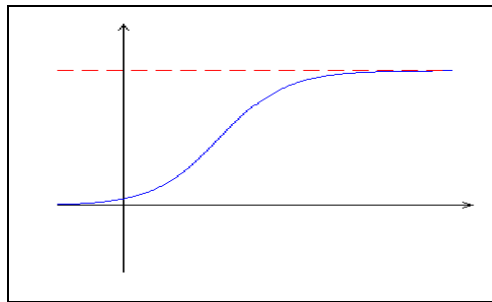
We willen in deze opgave de groei van een zonnebloem modelleren op basis van enkele meetgegevens. Bij het begin van de metingen was het plantje 10 cm hoog. Na 100 dagen mat men een hoogte van 1,27 meter en na 200 dagen constateerde men dat de zonnebloem volgroeid was ; de hoogte was op dat moment 2 meter.

- Stel een 3^{de} graadsveeltermfunctie op die deze groei zo goed mogelijk beschrijft.
- Waarom voldoet dit model niet als groeimodel voor de zonnebloem ?
- Hoe moet de grafiek van de modelfunctie verlopen om de groei te beschrijven ?

Dit soort verloop kan beschreven worden door een zogenaamde 'logistische functie' (*) waarvan het functievoorschrift als volgt luidt :

$$f(x) = \frac{a}{1 + b \cdot e^{-cx}} \quad \text{waarbij } a, b > 0 \quad \text{en } c < 0 \quad (1)$$

en waarvan de grafiek een vorm heeft zoals in onderstaande figuur



Het logistiek regressiemodel is beschikbaar op het TI-83 plus reken toestel.

- Modelleer de zonnebloemgroei met behulp van dit logistiek regressiemodel.
- Vergelijk uw gevonden model met onderstaand model uit de literatuur :

$$f(t) = \frac{2 \cdot e^{rt}}{e^{rt} + 19} \quad \text{waarbij } r > 0 \quad \text{een constante en } t \text{ de tijd in dagen} \quad (2)$$

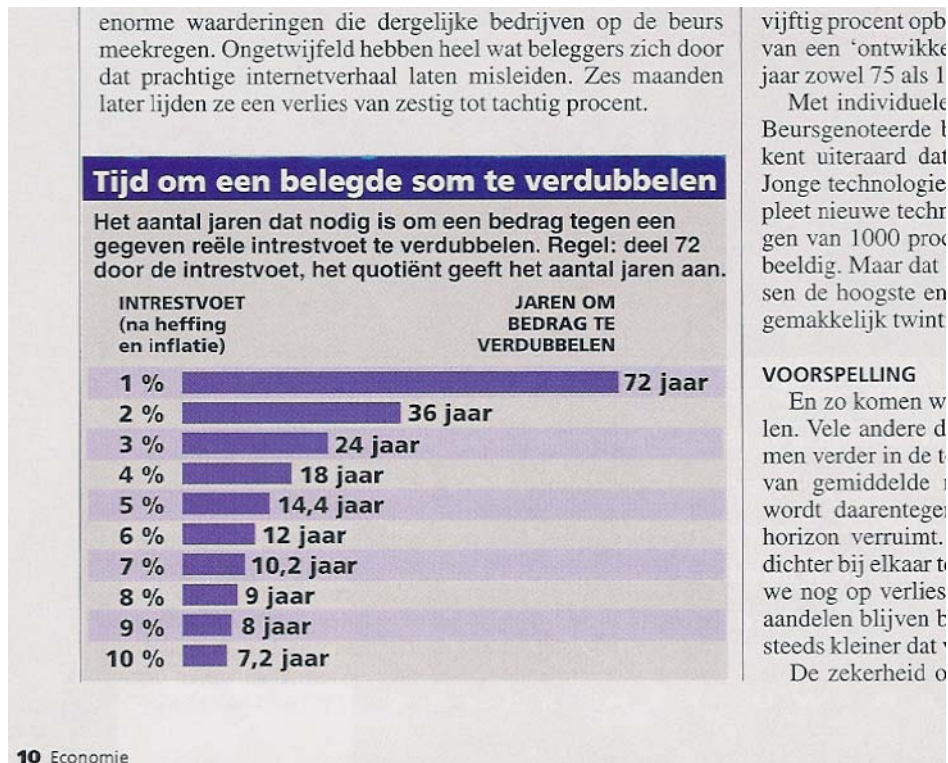
- Hoe hoog is de zonnebloem, volgens dit model, bij het begin van de metingen ?
- Hoe groot wordt de zonnebloem volgens dit model ?
- Wat is de betekenis van de parameter a in de vergelijking (1) ?
- Op welk tijdstip is de groeisnelheid het grootst ? Hoe groot is de zonnebloem dan ? Wat is de betekenis van dit tijdstip ?
- Onderzoek de invloed van r in het functievoorschrift (2) op het grafisch verloop van de groei.

(*) voor het eerst voorgesteld als model voor een populatiegroei rond 1830 door de Belgische wiskundige Pierre-François Verhulst.

In de opgave 12 worden 'functie' en 'functie als model' met elkaar geconfronteerd in een concrete toepassing.

Opgave 12 : Tijd om een belegde som te verdubbelen

In onderstaand artikel wordt een vuistregel gegeven om de tijd te berekenen om een belegde som te verdubbelen bij een bepaalde intrestvoet.



Geef het functievoorschrift van de functie die het werkelijk verband beschrijft tussen de verdubbelingstijd en de intrestvoet en vergelijk met de functie die als model dient voor de bovenstaande vuistregel.

Ga na hoe je zelf tot die eenvoudige vuistregel had kunnen komen.

Deel 2 : Toepassingen

In de opgaven 13, 14 en 15 worden concrete problemen opgelost door gebruik te maken van modellen. In de opdracht 13 (boorplatform) ligt de nadruk vooral op het omgaan met gegevens en de berekeningen; voor het opstellen van de modellen worden nog vrij nauwkeurige aanwijzingen gegeven.

De opdracht 14 (blokrijden) gaat een stap verder: de nadruk ligt hier op het *opstellen* van het model (randvoorwaarden, veronderstellingen, interpretatie, alternatieven...). Hier kan dan ook aandacht besteed worden aan argumentatie en verslaggeving.

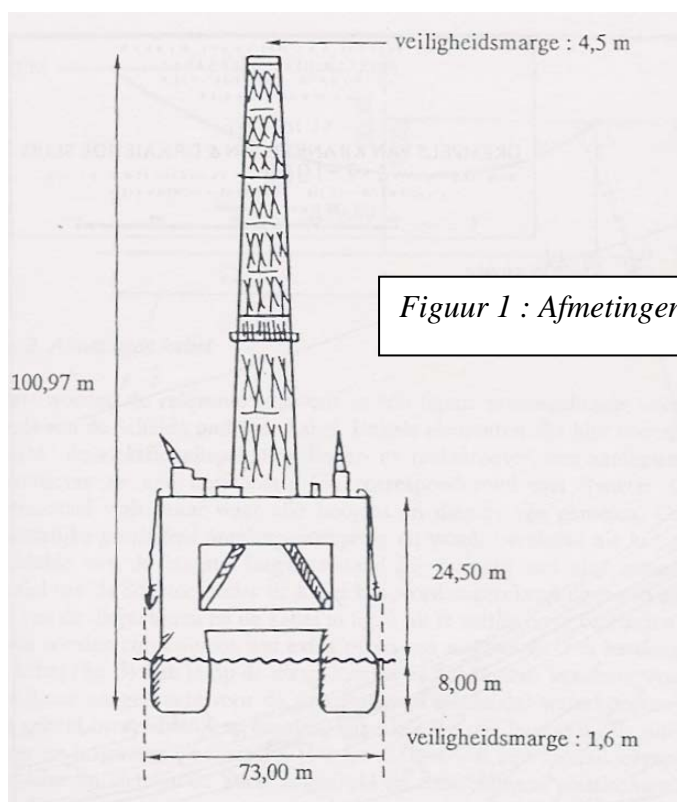
In opgave 15 (afkoelingskrommen) wordt een geschikt regressiemodel gezocht voor concrete opgemeten data bij een fysisch experiment. Hier komt tot uiting dat de *context*, gekoppeld aan de data, een belangrijke rol speelt bij de modelkeuze.

Opgave 13 : het boorplatform 'Yatzy'

1. Inleiding

In deze opgave wordt in een realistische context geschetst hoe functies als modellen kunnen gebruikt worden. Het probleem dat bekeken wordt, is de reis van 'Yatzy' op de Schelde. Yatzy is een boorplatform, door zijn ontwerpers genoemd naar een Oosters kansspel.

Het exploratie-boorplatform Yatzy werd gebouwd door de Boelwerf in Temse en is speciaal ontworpen om onderzeese aardolielagen te verkennen. Op 11 januari 1989 maakte dit 100 meter hoge gevaarte een eerste tocht op een mistige Schelde, met bestemming Rotterdam. Rond 13u30 ontmoette Yatzy zijn eerste hindernis: een hoogspanningskabel dwars over de Schelde. De Boelwerf-ingenieurs hadden aan de voorbereiding van deze doorvaart een stevige kluit; het wiskundig deel van hun analyse wordt hieronder nog eens overgedaan, weliswaar in een vereenvoudigde vorm. De vraag waar het hier om draait is of Yatzy onder de kabel kan en zo ja, met welke marges.



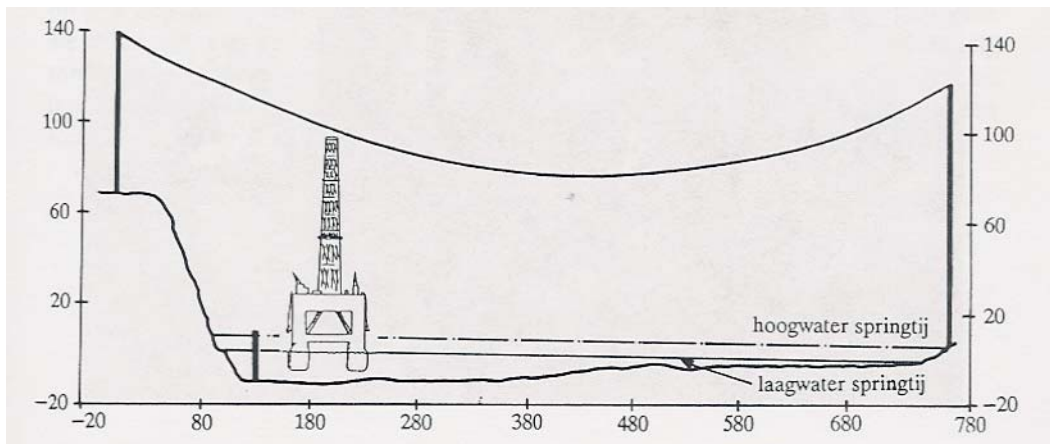
Figuur 1 : Afmetingen en diepgang van Yatzy

In figuur 1 vind je de afmetingen van Yatzy. Merk op dat voor de horizontale afmetingen een andere schaal gebruikt is dan voor de verticale : Yatzy is veel breder dan de figuur suggereert.

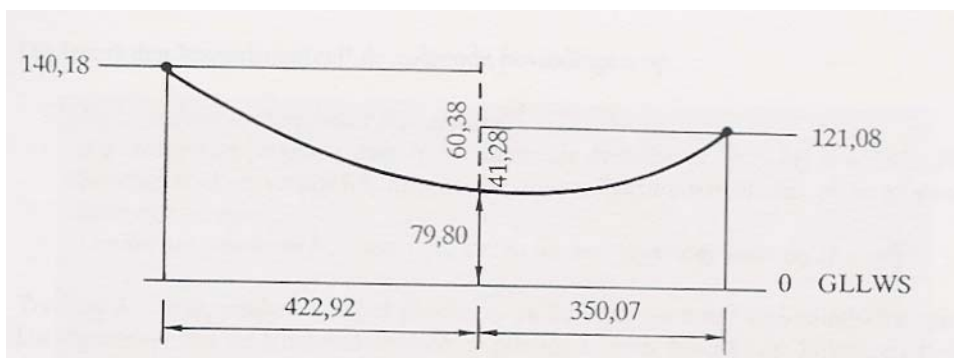
In figuur 2 zie je de dwarsdoorsnede van de Schelde onder de kabel. De figuur is op schaal getekend, al is die horizontaal een verticaal wel verschillend.

Horizontaal zijn de afstanden uitgedrukt in meter en is $x = 0$ voor de piloon op de linkeroever.

Verticaal zijn alle hoogtes en dieptes weergegeven in meter t.o.v. 0 meter GLLWS. GLLWS is het plaatselijk *gemiddeld laaglaagwaterspring* en wordt berekend als het meerjarig gemiddelde van de laagste laagwaterstand bij springtij van elke maansmaand. Het profiel van de Schelde is getekend aan de hand van een dieptekaart van de Schelde, de kabel aan de hand van metingen waarvan de resultaten in figuur 3 te vinden zijn.



Figuur 2 : Dwarsdoorsnede van de Schelde onder de kabel



Figuur 3 : Afmetingen van de kabel

Tenslotte werden horizontale lijnen aangebracht voor de maximale en minimale waterhoogte (tabel 1). Om het geheel overzichtelijk te houden, beperken we ons hierbij tot de uitersten : hoogwater en laagwater bij springtij.

Uit deze figuur kunnen we reeds het volgende besluiten :

- Het probleem zit bovenaan (bij de kabel).
- Als Yatzy kan passeren, dan is het langs linkeroever (bv. bij $x = 180$). De Schelde is daar voldoende diep en de elektriciteitsdraden bieden er de grootste doorvaarthoogte.
- Omdat het probleem bovenaan zit, kiezen we best voor laagwater bij springtij.

	Laagwater	Hoogwater	Periode
Gemiddeld springtij	0.25 m	6.00 m	12 u 10
Gemiddeld tij	0.46 m	5.60 m	12 u 15
Gemiddeld doottij	0.74 m	5.09 m	12 u 40

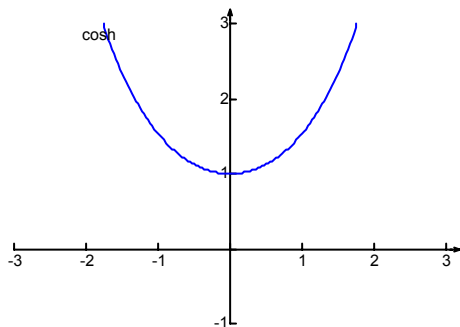
Tabel 1 : Getijden te Antwerpen t.o.v. GLLWS

2. Een model voor de kabel

Een vrij opgehangen, homogene kabel heeft, zo leert ons de natuurkunde, de vorm van een kettinglijn. Een kettinglijn is de grafiek van een functie met functievoorschrift :

$$y = a \cosh \frac{x-c}{a} + d$$

De basisfunctie $y = \cosh x = \frac{1}{2}(e^x + e^{-x})$ is getekend in figuur 4.



Figuur 4 : cosh-functie

Omdat we slechts een klein stukje rond de top van de kettinglijn nodig hebben, kunnen we de vorm van de kabel goed benaderen door een parabool.

Opdrachten :

- *Stel het functievoorschrift op voor deze parabool.*
- *Yatzy moet onder de kabel bij $x = 180$ m. Hoe hoog hangt de kabel daar ?*
- *Kan Yatzy op die plaats onder de kabel door*
 - *bij laagwater springtij ?*
 - *bij hoogwater springtij ?*

3. De horizontale speling

Is de voorwaarde om Yatzy bij $x = 180$ onder de kabel te loodsen wel realistisch ? Enkele scheepvaartdeskundigen die de operaties vanop de oever gadesloegen, beschreven het boorplatform als 'een grote onhandelbare wastobbe'. Een dergelijk gevaarte via sleepboten op een vooraf bepaalde plaats onder een kabel trekken, is alles behalve eenvoudig. Daarom is het zeker zinvol om ook *de speling in de breedte* te berekenen. De vraag is dan : *hoe ver* mag Yatzy zich *horizontaal* van de steiger verwijderen ? als men de veiligheidsmarge respecteert, moet de hoogte van de kabel dan minstens 97.72 meter bedragen.

- *Stel de ongelijkheid op waaraan x (de eerste coördinaat van het hoogste punt van Yatzy) moet voldoen.*
- *Los de ongelijkheid op en selecteer uit de wiskundige oplossingen, de realistische oplossingen voor het probleem.*

Deze speling is gezien de slechte bestuurbaarheid van Yatzy erg klein.

4. Hoeveel tijd heeft Yatzy ?

Op 11 januari 1989, de dag van de geplande doorvaart van Yatzy onder de hoogspanningskabel Zwijndrecht-Scheldelaan, was op die plaats laagwater voorzien om precies 13u 24'. Idealiter zou men de Yatzy juist op dat moment onder de kabel loodsen. Bij dichte mist op de Schelde is het volgen van een stipte timing evenwel gemakkelijker gezegd dan gedaan ! Voor de scheepsleiding is het daarom interessant te weten in welk tijdsinterval de doorvaart moet plaatsvinden. Om vat te krijgen op de veranderlijke hoogte van het water moet er een model opgesteld worden voor de getijden. Dit kan door middel van een algemene sinusfunctie :

$$h(t) = a \cdot \sin(b(t-c)) + d$$

waarbij h de waterhoogte (in meter) en t de tijd (in uur) voorstelt.

Wil men ook het effect springtij-doodtij in rekening brengen, dan kan dit door als modelfunctie een som van twee sinussen te nemen. De cyclus van spring- en doodtij is

immers ook te beschrijven met een sinusfunctie waarop men de sinus van eb en vloed kan superponeren. Voor de berekening van het tijdsinterval waarin Yatzy kan passeren, volstaat één sinus voor de waterhoogte bij springtij.

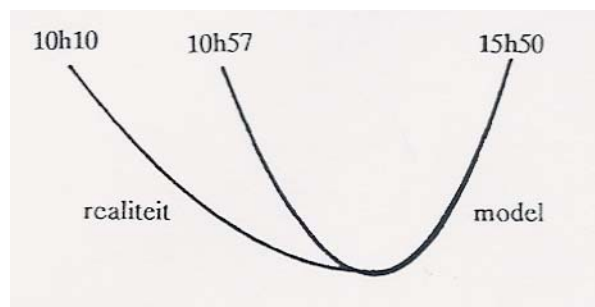
De getijdegegevens (tabel 1) plus het tijdstip waarop laagwater voorspeld werd (13u 24'), laten toe het functievoorschrift voor het sinusmodel op te stellen.

- *Stel dit functievoorschrift op.*

Voor de berekening van het tijdsinterval kiezen we een vaste plaats onder de kabel, bijvoorbeeld opnieuw $x = 180$ waar hij volgens het paraboolmodel 99.73 m hoog is. Yatzy (92.97 m boven de Scheldespiegel plus 4.5 m veiligheidsmarge) mag dus niet hoger komen dan 99.73 m. Dit geeft aanleiding tot een goniometrische ongelijkheid.

- *Los deze ongelijkheid op en bereken hiermee het tijdsinterval.*

De tijd is bijgevolg niet het probleem. In werkelijkheid had Yatzy zelfs nog drie kwartier meer tijd voor zijn doorvaart onder de kabel. Dat komt omdat bij getijden de daling langer duurt dan de stijging (zie figuur 5). Een nauwkeuriger model (maar dan noodzakelijk ook ingewikkelder) zou een betere voorspelling geven.



Figuur 5

Opgave 14 : Blokrijden als remedie voor het fileprobleem ?

Eén van de maatregelen die regelmatig door de verkeerspolitie getroffen wordt om een dreigende verzadiging van de verkeersstroom op te vangen, is het toepassen van het zogenaamde 'blokrijden'. Hierbij wordt een collectieve snelheidsvermindering opgelegd zodat de tussenruimte tussen twee opeenvolgende wagens kan afnemen en er dus meer auto's per tijdseenheid kunnen passeren. Een bijkomend voordeel is dat de veiligheid ook vergroot door minder inhaalmanoeuvres en een geringere snelheid.

In de onderstaande opgave wordt een model opgesteld voor dit blokrijden, en zullen we trachten na te gaan wat de 'beste' snelheid is voor het blokrijden, m.a.w. *bij welke snelheid het debiet van de verkeersstroom het grootst is*.

De remweg die de tussenruimte tussen twee wagens bepaalt, is afhankelijk van de snelheid, en dus ook het debiet op een weg met verkeerssaturatie is afhankelijk van de snelheid.

1. Eerste model voor de remweg als functie van de snelheid

In de tabel hieronder staan gegevens van gemiddelde remafstanden bij enkele snelheden (auto in perfecte staat, droog wegdek).

snelheid (km/u)	remafstand (m)
50	13
70	25
90	41
120	72
140	98

Tabel 2 : Remafstandgegevens

Stel op basis van deze gegevens een model op voor de remweg (remafstand) als functie van de snelheid (gebruik hierbij een regressiemodel dat beschikbaar is op uw rekentoestel).

Een mogelijke (eenvoudige) modelaannname zou kunnen zijn : de remweg r is evenredig met het kwadraat van de snelheid v

$$r(v) = k \cdot v^2 \quad \text{waarbij } k \text{ een evenredigheidsconstante is.} \quad (1)$$

Argumentatie voor dit model :

de remweg is te vinden als een integraal van de snelheid :

$$r = \int_0^{t_r} v(t) dt$$

waarbij : t_r de remtijd is,
 $v(t)$ de snelheid bij het remmen,
 t de tijd vanaf het begin van het remmen,

Als we een constante versnelling a veronderstellen tijdens het remmen, dan wordt

$$v(t) = v_0 + a.t \quad (a < 0)$$

waarbij v_0 de snelheid is op het ogenblik dat men begint te remmen

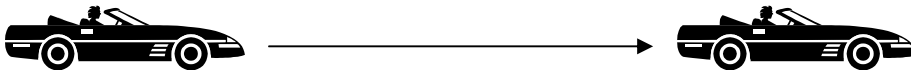
en dus $v = 0$ als $t = -\frac{v_0}{a}$

zodat $r = \int_0^{-v_0/a} (v_0 + at) dt = \dots = -\frac{v_0^2}{2a}$

Bereken op basis van de gegevens uit de tabel 2 de waarde van de evenredigheidsconstante k in het kwadratisch model (1) voor de remafstand.

Vergelijk dit kwadratisch model met het model dat je berekend hebt met regressie. Is de vereenvoudiging verantwoord ?

2. Model voor het debiet



Figuur 6

Stel een model op voor het debiet (bv. aantal passerende auto's per minuut) als functie van de snelheid (functievoorschrift, grafische voorstelling). Veronderstel dat de file enkel uit personenauto's bestaat (naar de kust..., gemiddeld 4 meter lengte)

Bij welke snelheid is het debiet maximaal ?

Wat is de onderlinge afstand tussen de wagens bij deze optimale snelheid ?

3. Evaluatie en aanpassingen van het model

3.1. Had je zulke (lage) snelheid verwacht ?

Waarom zouden de gangbare bloksnelheden hoger liggen ? (m.a.w. wat is het effect van een hogere snelheid op het debiet ?

3.2. Onderzoek nu in welke mate een grotere of kleinere remweg (een grotere of kleinere evenredigheidsconstante k) dit resultaat beïnvloedt.

3.3. Pas het model aan door rekening te houden met de reactietijd van de chauffeur.

De reactietijd van een oplettende chauffeur in goede conditie bedraagt ongeveer 1 seconde.

Wat is het effect van het in rekening brengen van de reactietijd op de optimale snelheid van de file, en op het debiet ?

3.4. Wat is het effect van de lengte van de voertuigen op de resultaten ?

Toegepast op vrachtwagens ?

3.5. Alternatief model voor de remafstand.

In rijsscholen leert men de toekomstige chauffeurs een vuistregel voor een veilige afstand tot de voorligger : veilige afstand in meter = snelheid per uur delen door twee.

Pas je model voor het debiet aan aan deze vuistregel.

Vergelijk de beide modellen (het verloop, de resultaten, de modelaannamen).

Waarom zou de Federale Politie 110 km /u als bloksnelheid gebruiken ?

Opgave 15 : Afkoelingskrommen

Guido Herweyers (met medewerking van Hans Bekaert)

Probleemstelling

Als je je koffie te lang laat staan, koelt die af. In het begin gaat dat zeer snel, nadien wat trager. Wanneer is de koffie te koud om te drinken? Hoe lang duurt het vooraleer die temperatuur bereikt is? Met welk wiskundig model zou je de afkoeling kunnen beschrijven?

Het onderzoek

Om een antwoord te zoeken op de wijze waarop materie afkoelt, doen we een experiment met een temperatuursensor. We stoppen de temperatuursensor eerst in een beker met heet water. Daarna brengen we de sensor in een beker met koud water en we starten de meting om na te gaan hoe snel de afkoeling gebeurt.

Wat heb je nodig?

- Een temperatuursensor
- Een CBL 2 en een grafische rekenmachine TI-83 Plus
- 2 bekers
- Koud en heet water, waterkoker

Uitvoering van het experiment

Voorbereiding

1. Plug de sensor in kanaal 1 van de CBL 2.
2. Verbind de CBL 2 met je rekentoestel.
3. Maak wat water heet in de waterkoker.

Voorbereiding op de rekenmachine

1. Schakel je rekentoestel aan en start het DATAMATE¹-programma. Druk op om het programma te initialiseren.
2. Kies in het hoofdmenu voor SETUP.
3. Lees je achter CH 1 : TEMP ga dan verder met stap 7.
4. Gebruik en en om CH 1 te selecteren.
5. Kies TEMPERATURE uit het SELECT SENSOR menu.
6. Kies voor STAINLESS TEMP (C°)
7. Gebruik en om MODE te selecteren en druk .
8. Kies voor TIME GRAPH uit het SELECT MODE menu.
9. Kies voor CHANGE TIME SETTINGS
10. Achter ENTER TIME BETWEEN SAMPLES...; geef je 1 in en druk
11. Achter NUMBER OF SAMPLES, geef je 30 in en druk op
12. Kies voor OK en nogmaals OK om naar het hoofdmenu terug te keren.

Meting

13. Vul een eerste beker met heet water. Dompel de temperatuursensor hierin.
14. Vul een tweede beker met koud water.
15. Volg de temperatuurstijging op je rekentoestel. Laat de temperatuur zo hoog mogelijk oplopen.
16. Dompel de sensor nu in de beker met koud water en START de meting. De meting stopt automatisch na 30 seconden.
17. De grafiek verschijnt op je display.
18. Druk om naar het hoofdmenu terug te keren.

¹ Het Datamate programma moet eerst op je rekentoestel geïnstalleerd staan. Zet je rekentoestel in de recieve-mode (**y** LINK – [RECEIVE] – 1:Recieve – **f**) en druk op de transfer-knop van de CBL 2. Het programma wordt nu naar je rekentoestel gedownload.

Analyse van de meetgegevens

Om de meetgegevens te analyseren verlaten we het Datamate-programma en bestuderen we de resultaten via de statistische functies van de TI-83 Plus. Verlaat het hoofdmenu door te kiezen voor QUIT. De meetresultaten van de laatste meting zijn opgeslagen in **L1** en **L2**. De tijd x (eenheid s) vind je terug in **L1**, de temperatuur y (eenheid °C) in **L2**.

De omgevingstemperatuur van het experiment is 19.5 °C (de temperatuur van de beker met koud water).

Hier volgen de meetresultaten:

tijd	temperatuur	tijd	temperatuur	tijd	temperatuur
0	56.1	11	26.9767	21	21.4048
1	51.0606	12	26.0465	22	21.119
2	46.7778	13	25.1364	23	20.8372
3	43.1053	14	24.4884	24	20.6512
4	39.9	15	23.8333	25	20.4651
5	37.122	16	23.1667	26	20.3721
6	34.6905	17	22.7955	27	20.186
7	32.5952	18	22.3409	28	20.093
8	30.907	19	21.9762	29	20
9	29.4048	20	21.5952	30	19.9048
10	28.093				

Opgave:

- Maak een grafiek van de data.
- Stel enkele functies voor als kandidaat regressiemodel voor de data, onderzoek je vermoeden door de concrete regressiekromme(n) te tekenen.
- Ga na dat een vierdegraadsregressie een prima model levert.
- Bij nader inzien is de vierdegraadsregressie een goed model voor tijden tussen 0 en 30 seconden (interpolatie), is dit ook het geval voor $x=50$ s (extrapolatie)? Hou rekening met de context of concrete betekenis van de data.
- Waarom is geen enkele van de regressiekrommen, beschikbaar op de TI-83 (zie ook pagina 6), geschikt om temperatuursvoorspellingen te doen op lange termijn?
- Wat verwacht je van het temperatuursverloop op lange termijn? Heeft de kromme asymptoten? Suggereer een waarde voor c in het model $y = a \cdot x^b + c$. Welk teken verwacht je voor a , welk teken verwacht je voor b ?
- Zelfde vraag als (f) voor het model $y = a \cdot b^x + c$.

Volgens de *afkoelingswet van Newton* voor een afkoelend lichaam is de snelheid waarmee de temperatuur van een lichaam afneemt evenredig met het verschil van de temperatuur van het lichaam en de omgevingstemperatuur. Hieruit volgt (door oplossing van een differentiaalvergelijking) het wiskundig model volgens Newton:

$$y = a \cdot e^{-k \cdot x} + 19.5 \quad (\text{met } k > 0)$$

Dit kan ook worden geschreven in de vorm

$$y = a \cdot b^x + 19.5 \quad \text{met} \quad b = e^{-k}$$

De constanten a en b kunnen worden gevonden door te eisen dat de kromme door 2 meetpunten gaat, bijvoorbeeld $(0, 56.1)$ en $(10, 28.093)$. Dit levert het model

$$y = 36.6 \cdot 0.865^x + 19.5$$

Dit model houdt echter enkel rekening met twee meetpunten.

Door exponentiële regressie verkrijgen we een (beter) model dat rekening houdt met alle data. Het model kan ook worden geschreven als

$$y - 19.5 = a \cdot b^x$$

of $Y = a \cdot b^x$ met $Y = y - 19.5$

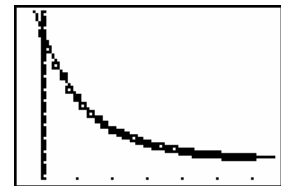
Voor de exponentiële regressie moeten we dus werken met de data x_i en $Y_i = y_i - 19.5$.

De data x_i bevinden zich in lijst **L1** en de data y_i in lijst **L2**. Van lijst **L2** trekken we 19.5 af, dit levert de lijst der getallen Y_i die we opslaan in **L3**.

```
L1
{0 1 2 3 4 5 6 ...
L2
{56.1 51.0606 4...
L3={L2-19.5+L3}
{36.6 31.5606 2...
```

```
ExpReg
y=a*b^x
a=37.31843131
b=.8642045201
r^2=.9987188831
r=-.9993592363
```

```
Y1=37.318431312
571*.86420452005
941^X
Y2=Y1+19.5
Y3=
Y4=
Y5=
```



De exponentiële regressie levert $a = 37.32$ en $b = 0.864$.

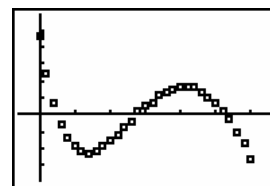
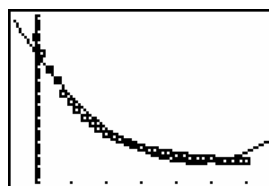
Opgave (voor de leerkracht):

1) Uit $Y = a \cdot b^x$ volgt $\ln Y = \ln a + x \cdot \ln b$ of $Z = A + B \cdot x$ met $Z = \ln Y$, $A = \ln a$, $B = \ln b$.

Ga na dat de resultaten van de exponentiële regressie verkregen werden via een lineaire regressie op de data $(x_i, Z_i) = (x_i, \ln Y_i)$; lineaire regressie op deze data levert de getallen A en B waaruit $a = e^A$ en $b = e^B$ volgen. Merk op dat de determinatiecoëfficiënt r^2 afkomstig is van deze lineaire regressie.

2) Voer een kwadratische regressie uit op de oorspronkelijke data (x_i, y_i) en teken een residuplot. De residu's vormen een duidelijk patroon. Dit wijst op het bestaan van een beter model.

```
QuadReg
y=ax^2+bx+c
a=.0608326152
b=-2.787733011
c=51.42802174
R^2=.9713759546
```



BRONNEN

Deel 1 : Functies als modellen voor functies

Opgave 1

JANSSENS, D., Seminariemap Wiskunde, Acad. Lerarenopleiding Wiskunde , K.U.Leuven, 1998

Regressie en gebruik van modellen (opgaven 6 en 7, bijlage 1)

HERWEYERS, G., STULENS, K., Statistiek met een grafisch rekentoestel, Acco Leuven, 2000

KÖHLER, A., Mathematikabitur im Technologiezeitalter, Praxis der Mathematik in der Schule, 6/43, Jg.2001, pp 278-279.

Opgave 10

KÖHLER, A., Mathematikabitur im Technologiezeitalter, Praxis der Mathematik in der Schule, 6/43, Jg.2001, pp 278-279.

Opgave 11 : groei van zonnebloemen

T3-Regionaltagung in Aachen am 08.11.2003, Workshop 3 : Ausgewählte Beispiele aus der Analysis für Klausur, Abitur und Unterricht mit dem TI 83 und Voyage 200

Deel 2 : Toepassingen

Opgave 13 : het boorplatform ‘ Yatzy’

ROELS, J., e.a., Wiskunde vanuit Toepassingen, Aggregatie HSO Wiskunde, Leuven, 1990

Opgave 14 : Blokkrijden

DE BOCK, D., e.a., Afgeleiden en integralen, Acco Leuven, 1994

Bijlage 1: Regressie, correlatie en modelvorming met de TI-83 (Plus) (Achtergrondinformatie voor de leerkracht)

Guido Herweyers

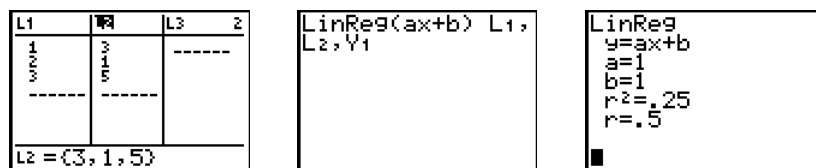
1. Inleidend voorbeeld

De punten (1,3), (2,1) en (3,5) liggen niet op één rechte. Hoe vind je met een TI -83(Plus) de rechte die het “best” aansluit bij die punten?

Plaats de coördinaten in de lijsten **L1** en **L2** m.b.v. **STAT 1:Edit** en druk vervolgens **STAT<CALC> 4:LinReg(ax+b)**. Op het basisscherm verschijnt **LinReg(ax+b)**.

Vul dit aan met **2nd[L1]**, **2nd[L2]**, **VARS<Y-VARS> 1:Function 1:Y1**. Door het toevoegen van **Y1** wordt de vergelijking van de beste rechte weggeschreven in **Y1**.

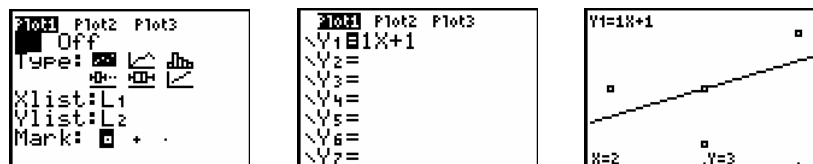
Het drukken op **ENTER** levert de beste rechte $y = x + 1$.



De extra gegevens r en r^2 verkrijg je indien je eerst via **2nd[CATALOG] DiagnosticOn** hebt uitgevoerd vanuit het basisscherm.

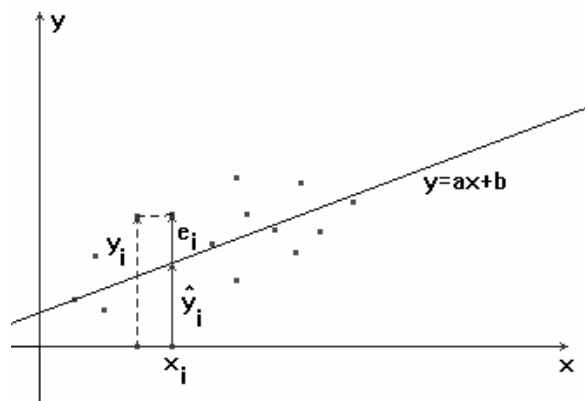
Met **2nd[STAT PLOT] 1: Plot1** definieer je **Plot1** zoals hieronder. Controleer de functies (druk **Y=**) die zullen worden getekend. Met **ZOOM 9:ZoomStat** verkrijg je de grafieken van de puntenwolk en de beste rechte.

Via **TRACE** kunnen we de rechte doorlopen om eventueel waarden van y horende bij x te voorspellen.



2. Hoe wordt de beste rechte berekend ?

Gegeven een *puntenwolk* of *spreidingsdiagram* van n punten $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$, die min of meer een lineaire trend vertonen (zie volgende figuur).



Beschouw een rechte $y = ax + b$ “door” deze punten waarmee we de grootheid y wensen te voorspellen bij gegeven x .

De voorspelling van y horende bij x_i noteren we met \hat{y}_i , er geldt dus dat $\hat{y}_i = ax_i + b$.

We definiëren voor elk punt het *residu* e_i (zie figuur) als volgt :

$$\begin{aligned}
 e_i &= \text{observatie} - \text{voorspelling} \\
 &= y_i - \hat{y}_i \\
 &= y_i - (ax_i + b)
 \end{aligned}$$

Merk op dat een residu positief is wanneer het punt boven de rechte gelegen is en negatief wanneer het onder de rechte gelegen is.

Om de “beste” rechte door de puntenwolk te zoeken gebruiken we het *kleinste kwadraten criterium* : bepaal a en b zodanig dat $\sum_{i=1}^n e_i^2$ minimaal is.

Deze beste rechte wordt de *kleinste kwadraten rechte* of de *lineaire regressie* van y op x genoemd, met y de afhankelijke of *verklaarde* veranderlijke en x de onafhankelijke of *verklarende* veranderlijke.

Om de beste a en b te bepalen gebruiken we de afwijkingen

$$u_i = x_i - \bar{x} \text{ en } v_i = y_i - \bar{y}$$

t.o.v de rekenkundige gemiddelden $\bar{x} = \frac{\sum_{i=1}^n x_i}{n}$ en $\bar{y} = \frac{\sum_{i=1}^n y_i}{n}$.

Voor de beste rechte $y = ax + b$ door de punten $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$ geldt :

$$a = \frac{\sum_{i=1}^n u_i v_i}{\sum_{i=1}^n u_i^2} \quad \text{en} \quad b = \bar{y} - a\bar{x}$$

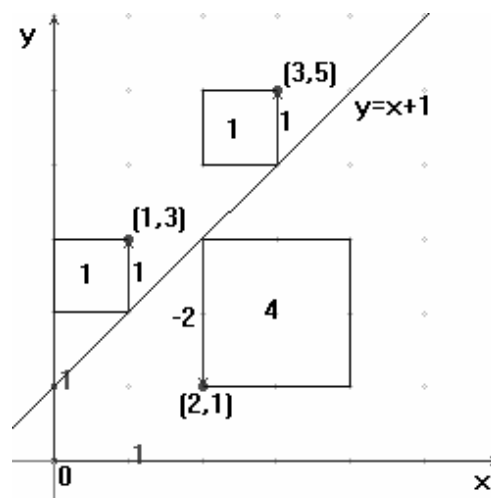
We berekenen hiermee de beste rechte door de punten (1,3), (2,1) en (3,5) van het inleidende voorbeeld :

x_i	y_i	$u_i = x_i - \bar{x}$	$v_i = y_i - \bar{y}$	$u_i v_i$	u_i^2
1	3	-1	0	0	1
2	1	0	-2	0	0
3	5	1	2	2	1
\sum_i	6	9	0	0	2
	$\bar{x} = 2$	$\bar{y} = 3$		$a = 2/2 = 1$ $b = \bar{y} - a\bar{x} = 1$	

Klaarblijkelijk is $y = x + 1$ de beste rechte.

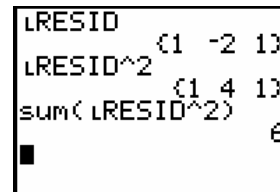
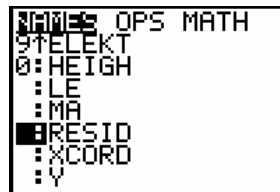
De som van de kwadraten der residu's, $\sum_{i=1}^3 e_i^2 = 6$, is minimaal voor deze rechte.

Dit is te interpreteren als een som van oppervlakten van vierkanten die minimaal is:



De **TI-83** maakt na een regressieberekening automatisch de lijst van de residu's aan. Deze lijst vind je met **2nd[LIST] RESID**.

Voor het berekenen van de som van de kwadraten der residu's gebruik je **2nd[LIST]<MATH> 5:sum(.**

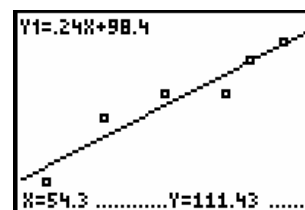
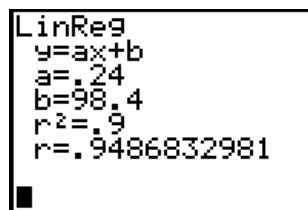


3. Correlatie

De lengte l van een metalen staaf wordt gemeten bij verschillende temperaturen t :

t (°C)	20	30	40	50	60
l (mm)	102	107	109	109	113

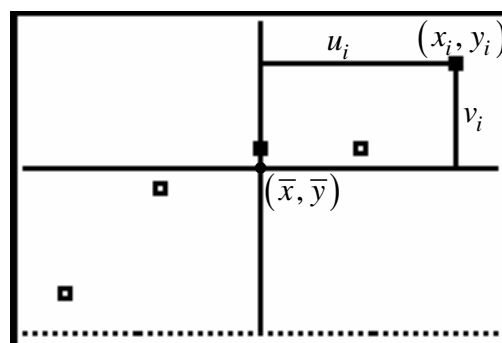
In de fysica gebruikt men als model het lineair verband $l = a.t + b$. We zoeken de beste rechte uitgaande van de data (reken ook manueel na met een tabel zoals op de vorige pagina). Vervang hiertoe t door x en l door y .



Wat is de betekenis van de *correlatiecoëfficiënt* r ?

Hiertoe tekenen we eerst de rechten $y = \bar{y}$ (met $\bar{y} = 108$) en $x = \bar{x}$ (met $\bar{x} = 40$). Teken de verticale rechte met de opdracht **2nd[DRAW] 4:Vertical 40** in het basisscherm.

Beschouw nu we de verticale afwijking $v_i = y_i - \bar{y}$ en de horizontale afwijking $u_i = x_i - \bar{x}$ van een punt (x_i, y_i) t.o.v. het zwaartepunt (\bar{x}, \bar{y}) van de puntenwolk (zie figuur):



Voor bovenstaande puntenwolk spreken we over een *positieve correlatie* tussen de grootheden x en y aangezien de meeste punten (u_i, v_i) in het eerste of derde kwadrant gelegen zijn t.o.v. het assenstelsel met oorsprong (\bar{x}, \bar{y}) . De beste rechte heeft een positieve richtingscoëfficiënt.

Merk op dat de beste rechte $y = ax + b$ steeds door het zwaartepunt (\bar{x}, \bar{y}) van de puntenwolk gaat aangezien $\bar{y} = a \cdot \bar{x} + b$. Controleer dit door met je rekenmachine het snijpunt te zoeken van de beste rechte met de rechte $y = \bar{y}$.

Indien de meeste punten gelegen zijn in het tweede of vierde kwadrant dan is er sprake van een negatieve correlatie. De beste rechte heeft een negatieve richtingscoëfficiënt.

Indien de punten eerder verdeeld zijn over de vier kwadranten, bestaat er geen *lineair* verband tussen de grootheden x en y . Een beste rechte berekenen en tekenen is nog steeds mogelijk, maar weinig zinvol.

In onze figuur is het duidelijk dat $\sum_{i=1}^n u_i v_i$ positief is, aangezien de bijdragen van de punten in het eerste en derde kwadrant tot die som steeds positief zijn.

De som $\sum_{i=1}^n u_i v_i$ is echter afhankelijk van de gekozen eenheden voor x en y , zodat we tot de volgende definitie komen van de *correlatiecoëfficiënt* r , onafhankelijk van de gekozen eenheden voor x en y :

$$r = \frac{\sum_{i=1}^n u_i v_i}{\sqrt{\left(\sum_{i=1}^n u_i^2\right) \cdot \left(\sum_{i=1}^n v_i^2\right)}}.$$

Bovenstaande formule is symmetrisch in x en y en maakt dus geen onderscheid tussen de onafhankelijke en afhankelijke veranderlijke.

We berekenen r voor de gegeven data m.b.v. onderstaande tabel (vergelijk met het resultaat van het rekentoestel).

x_i	y_i	$u_i = x_i - 40$	$v_i = y_i - 108$	$u_i \cdot v_i$	u_i^2	v_i^2
20	102	-20	-6	120	400	36
30	107	-10	-1	10	100	1
40	109	0	1	0	0	1
50	109	10	1	10	100	1
60	113	20	5	100	400	25

\sum_i	240	1000	64
----------	-----	------	----

$$r = \frac{240}{\sqrt{1000 \cdot 64}} \approx 0,9487$$

Er kan worden aangetoond dat $r^2 \leq 1$ en dus $-1 \leq r \leq 1$.

De correlatiecoëfficiënt r is een maat voor het *lineaire* verband tussen de grootheden x en y . Hoe dichter de absolute waarde van r bij 1 gelegen is, hoe beter het lineaire verband.

Merk op dat r en de richtingscoëfficiënt a van de beste rechte het teken hebben van $\sum_{i=1}^n u_i v_i$.

Voor positieve r is er een positieve correlatie en voor negatieve r een negatieve correlatie.

Als $r = 1$ of $r = -1$, dan liggen alle punten op één (beste) rechte $y = ax + b$.

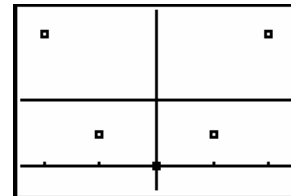
Voor waarden van r dicht bij 0 is er nagenoeg geen *lineair* verband. Er kan echter wel een ander functioneel verband zijn tussen x en y !

Beschouw hiertoe volgende data, gelegen op de parabool $y = x^2$. De beste rechte is $y = 2$, maar deze benadering heeft uiteraard geen zin.

```
LinReg
y=ax+b
a=0
b=2
r^2=0
r=0
```

L1	L2	L3	Z
-2	4	-----	
-1	1		
0	0		
1	1		
2	4		

L2(6) =			



4. Modelvorming

4.1 De determinatiecoëfficiënt

De *determinatiecoëfficiënt* R^2 is een maat voor de kwaliteit van een regressiemodel dat *niet noodzakelijk lineair* is.

Als introductie nemen we de data (1,3), (2,1) en (3,5) die we reeds vroeger hebben gebruikt, met als beste rechte door deze drie punten $y = x + 1$.

Beschouw de volgende tabel met $\hat{y}_i = x_i + 1$, dit is de voorspelde waarde van y m.b.v. het regressiemodel.

x_i	y_i	$(y_i - \bar{y})^2$	$(y_i - \hat{y}_i)^2$
1	3	0	1
2	1	4	4
3	5	4	1
	\sum_i	8	6

Als er geen verband zou bestaan tussen de variabelen x en y , dan zouden de opeenvolgende data y_i lukraak schommelen rond hun gemiddelde \bar{y} .

Voor de totale variatie van de data y_i t.o.v. \bar{y} nemen we $\sum_{i=1}^n (y_i - \bar{y})^2$ als maat, deze waarde geeft aan in welke mate de puntenwolk verticaal afwijkt van de rechte $y = \bar{y}$.

Met het regressiemodel kunnen we y beter voorspellen daar de som van de kwadraten der residu's (of de variatie van de data y_i t.o.v. het regressiemodel) $\sum_{i=1}^n (y_i - \hat{y}_i)^2$ minimaal is,

bijgevolg is $\sum_{i=1}^n (y_i - \hat{y}_i)^2 \leq \sum_{i=1}^n (y_i - \bar{y})^2$. Er geldt :

$$\sum_{i=1}^n (y_i - \bar{y})^2 = \sum_{i=1}^n (y_i - \hat{y}_i)^2 + \left(\sum_{i=1}^n (y_i - \bar{y})^2 - \sum_{i=1}^n (y_i - \hat{y}_i)^2 \right)$$

totale variatie = onverklaarde variatie + verklaarde variatie

Het regressiemodel is goed als de onverklaarde variatie veel kleiner is dan de totale variatie, d.w.z. dat de verklaarde variatie niet veel kleiner is dan de totale variatie.

De determinatiecoëfficiënt R^2 is de fractie van de variatie van de data y_i t.o.v. \bar{y} die verklaard wordt door het regressiemodel :

$$R^2 = \frac{\text{verklaarde variatie}}{\text{totale variatie}}$$

$$R^2 = \frac{\sum_{i=1}^n (y_i - \bar{y})^2 - \sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2}$$

In bovenstaand voorbeeld is $R^2 = 1 - \frac{6}{8} = \frac{1}{4}$, er wordt dus slechts 25% van de variatie van de data y_i t.o.v. \bar{y} verklaard door het regressiemodel.

Voor de correlatiecoëfficiënt $r = 0.5$ geldt dat $r^2 = R^2$.

Dit is geen toeval, voor een *lineair* regressiemodel geldt steeds dat $r^2 = R^2$. Derhalve noemt men r^2 ook de lineaire determinatiecoëfficiënt.

Vermeld steeds r^2 bij een lineaire regressie als maat voor hoe succesvol het lineaire model is om y te voorspellen.

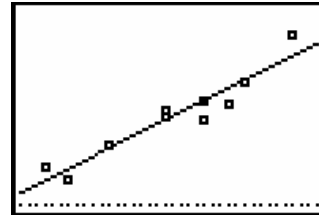
4.2 De residuplot

Als voorbeeld beschouwen we de lengte x in cm en de massa y in kg van 10 lukraak gekozen studenten :

x_i	163	185	180	175	168	175	191	180	160	183
y_i	60	90	78	81	71	79	104	84	64	83

We brengen deze data in de lijsten **L1** en **L2** en tekenen de puntenwolk. Zo te zien is een lineair model een goede benadering.

```
LinReg
y=ax+b
a=1.215261959
b=-134.4861048
r^2=.9002252918
r=.9488020298
```



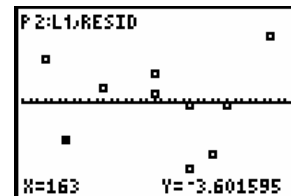
Bij lineaire regressie is de som van de residu's steeds nul : $\sum_{i=1}^n (y_i - \hat{y}_i) = \sum_{i=1}^n e_i = 0$, er geldt

immers dat $\sum_{i=1}^n (y_i - \hat{y}_i) = \sum_{i=1}^n (y_i - ax_i - b) = n(\bar{y} - a\bar{x} - b) = 0$.

Voor een goed model moet de *residuplot*, d.i. een grafiek van de residu's uitgezet tegen de data x_i , een lukrake verdeling tonen t.o.v de x -as die hier de regressielijn voorstelt.

```
Plot1 Plot3
Off
Type:
Xlist:L1
Ylist:RESID
Mark: +
```

```
Plot1 Plot3
Y1=1.2152619589
977X+-134.486104
7836
Y2=
Y3=
Y4=
Y5=
```



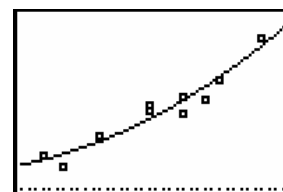
De determinatiecoëfficiënt r^2 is 0.9. Er wordt 90% van de variatie van de data y_i t.o.v. \bar{y} verklaard door het lineaire regressiemodel.

Met **sum(LRESID²)** vind je de som van de kwadraten der residu's: $\sum_{i=1}^{10} e_i^2 = 143.7$ (ga na).

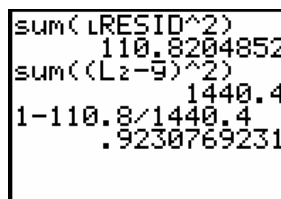
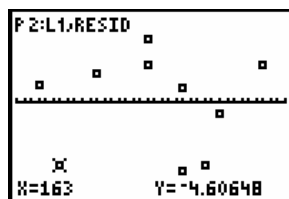
We onderzoeken of het kwadratisch model beter is.

```
QuadReg
y=ax^2+bx+c
a=.0208689599
b=-6.077560068
c=500.7813767
R^2=.9230627012
```

```
Plot1 Plot2 Plot3
Y1=1.2152619589
977X+-134.486104
7836
Y2=.02086895985
862X^2+-6.077560
0679524X+500.781
37670333
```



We maken nog een residuplot en berekenen de som van de kwadraten der residu's. Tevens rekenen we na dat de determinatiecoëfficiënt $R^2 = 0.923$. Het kwadratisch model is dus iets beter dan het lineair model.



Als modelkeuze kan men hier gerust kiezen voor het eenvoudigere lineair model.

Uit het gevonden verband tussen lengte en massa niet mogen besluiten dat dit verband causaal (oorzaak – gevolg) is. We mogen dus niet besluiten dat een grotere lengte een grotere massa veroorzaakt, er kunnen immers nog andere factoren zijn die een invloed hebben op de massa.

Onderzoek van andere modellen die de TI-83 berekent , geeft :

Model	R^2	r^2	r
LinReg(ax+b)		0.900	0.949
QuadReg ($ax^2 + bx + c$)	0.923		
CubicReg ($ax^3 + bx^2 + cx + d$)	0.938		
QuartReg ($ax^4 + bx^3 + cx^2 + dx + e$)	0.963		
LnReg ($a + b \cdot \ln(x)$)		0.892	0.945
ExpReg ($a \cdot b^x$)		0.915	0.957
PwrReg ($a \cdot x^b$)		0.911	0.955

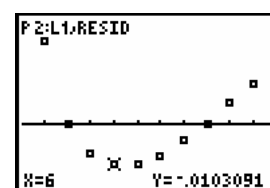
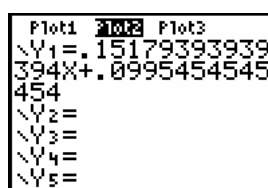
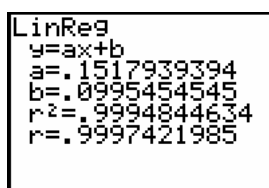
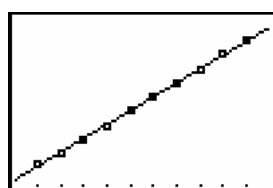
De *logaritmische regressie* LnReg wordt berekend met lineaire regressie van de data $(\ln(x_i), y_i)$, de *exponentiële regressie* ExpReg met lineaire regressie van de data $(x_i, \ln(y_i))$ en de *machtsregressie* PwrReg met lineaire regressie van de data $(\ln(x_i), \ln(y_i))$.

Dat er lineaire regressie werd gebruikt zie je aan het verschijnen van een correlatiecoëfficiënt en de notatie van de determinatiecoëfficiënt. Verklaar de transformaties van de data.

Als laatste voorbeeld beschouwen we een aantal gelijkmatige veelhoeken met zijden gelijk aan de lengte-eenheid en de corresponderende straal van de omgeschreven cirkel :

Aantal zijden	3	4	5	6	7	8	9	10	11	12
straal	0.577	0.707	0.851	1.000	1.152	1.306	1.462	1.618	1.775	1.932

We brengen deze data in de lijsten L1 en L2 en tekenen de puntenwolk. Een lineair model is uitstekend (zie r^2). De residuplot werpt echter een andere blik op de zaak.



De residu's vormen een duidelijk patroon. Dit wijst op het bestaan van een beter model. We laten het aan de lezer over om dit model te bepalen.

Bijlage 2 : Suggesties voor het oplossen van de opgaven

Opgave 1

$$f(x) = -\frac{1}{864}(x+1)(x+2)(x+3)(x-4)(x-6)^2$$

Opgave2

$$f(x) = ax^4 + bx^3 + cx^2 + dx + e$$

waarbij a, b, c, d en e kunnen gevonden worden door onderstaand stelsel lineaire vergelijkingen op te lossen via matrixrekenen (`rref`) :

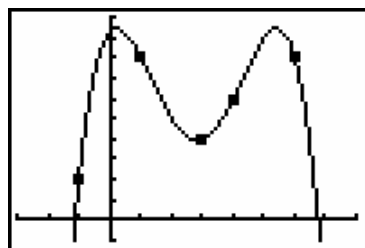
$$\begin{aligned} a \cdot 1^4 + b \cdot 1^3 + c \cdot 1^2 + d \cdot 1 + e &= 8 \\ a \cdot 3^4 + b \cdot 3^3 + c \cdot 3^2 + d \cdot 3 + e &= 4 \\ a \cdot 4^4 + b \cdot 4^3 + c \cdot 4^2 + d \cdot 4 + e &= 6 \\ a \cdot 6^4 + b \cdot 6^3 + c \cdot 6^2 + d \cdot 6 + e &= 8 \\ a \cdot (-1)^4 + b \cdot (-1)^3 + c \cdot (-1)^2 + d \cdot (-1) + e &= 2 \end{aligned}$$

[B]	
... 0 0 0 -.12142...	$\approx a$
... 0 0 0 1.36666...	$\approx b$
... 1 0 0 -4.1357...	$\approx c$
... 0 1 0 1.63333...	$\approx d$
... 0 0 1 9.25714...	$\approx e$

laatste kolom van de gereduceerde rijvorm van de matrix.

De exacte coëfficiënten →

... 0 0 0 1 9.2571...
Ans ▶ Frac
... 0 0 -17/140 1
... 0 0 41/30 1
... 0 0 -579/140 1
... 1 0 49/30 1
... 0 1 324/35 1 1

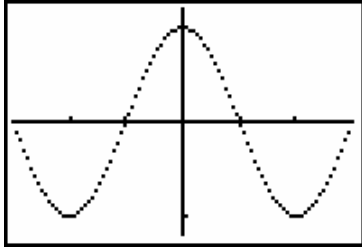
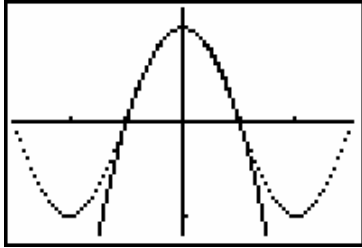
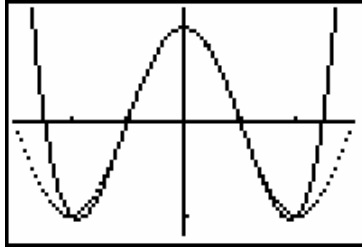


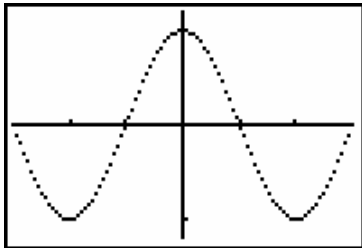
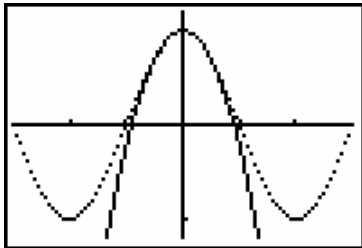
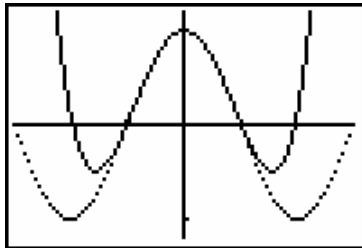
```

201 Plot2 Plot3
\Y2=[B](1,6))*X
^4+[B](2,6))*X^
3+[B](3,6))*X^2
+[B](4,6))*X+[
B](5,6)
\Y3=
\Y4=

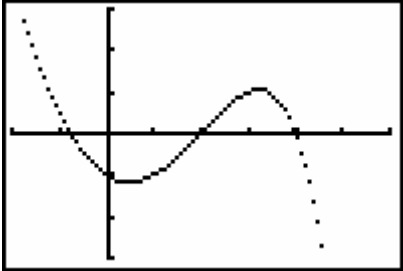
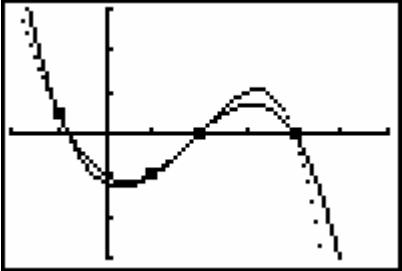
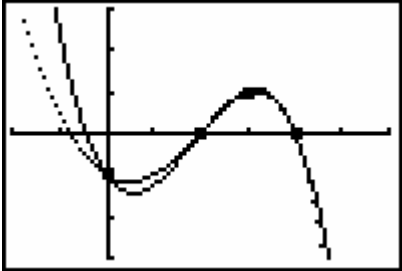
```

Opgave 3

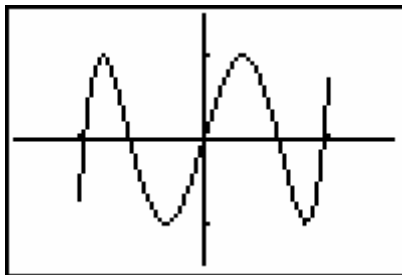
 <p style="text-align: center;">$\cos x$</p>	<p>2^{de} graadsveeltermbenadering door de punten $(-\pi/2, 0)$, $(\pi/2, 0)$ en $(0, 1)$</p>  <p style="text-align: center;">$\cos x$</p> $f_1(x) = 1 - \frac{4x^2}{\pi^2}$	<p>4^{de} graadsveeltermbenadering door de punten $(-\pi/2, 0)$, $(\pi/2, 0)$, $(-\pi, -1)$, $(\pi, -1)$ en $(0, 1)$</p>  <p style="text-align: center;">$\cos x$</p> $f_2(x) = 1 - \frac{x^2}{2.115} + \frac{x^4}{36.528}$
--	---	--

 <p style="text-align: center;">$\cos x$</p>	<p>2^{de} graadsveeltermbenadering volgens MacLaurin</p>  <p style="text-align: center;">$\cos x$</p> $f_1(x) = 1 - \frac{x^2}{2}$	<p>4^{de} graadsveeltermbenadering volgens MacLaurin</p>  <p style="text-align: center;">$\cos x$</p> $f_2(x) = 1 - \frac{x^2}{2} + \frac{x^4}{24}$
--	--	---

Opgave 4

		
$y_1 = x^2 - 2^x$	$y_1 = x^2 - 2^x$ $y_2 = (x-2) \cdot (x-4) \cdot \left(-\frac{11}{60}x - \frac{3}{20}\right)$ door de punten (2,0) , (4,0) , (-1, 1/2) en (1,-1)	$y_1 = x^2 - 2^x$ $y_3 = (x-2) \cdot (x-4) \cdot \left(-\frac{7}{24}x - \frac{1}{8}\right)$ door de punten (2,0) , (4,0) , (0,-1) en (3,1)

Opgave 5

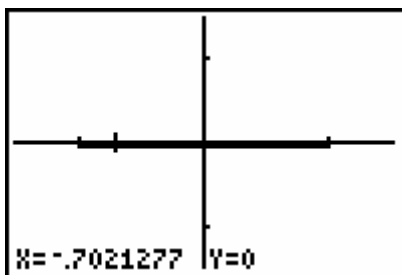


$$g(x) = \sin(5Bgsin x)$$

De functie g kan in het domeininterval benaderd worden door een 5^{de} graadsveeltermfunctie f door gebruik te maken van de 5 nulpunten en een bijkomend punt, bijvoorbeeld $(\sin(\pi/10), 1)$

$$f(x) = 16x \cdot (x^2 - \sin^2(\frac{2\pi}{5})) \cdot (x^2 - \sin^2(\frac{\pi}{5}))$$

of
$$f(x) = 16 \cdot x^5 - 20 \cdot x^3 + 5 \cdot x$$

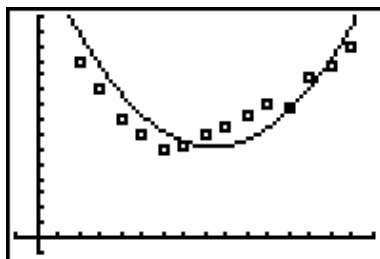


$$g(x) - f(x)$$

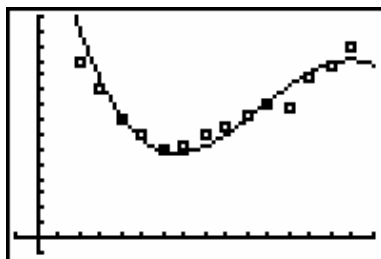
De goniometrische berekening van $\sin(5Bgsin x)$ bevestigt dat

$$g(x) = f(x)$$

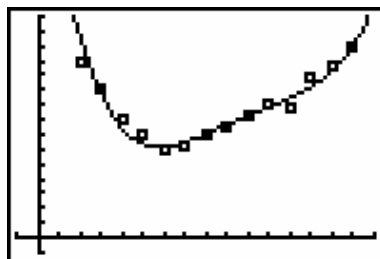
Opgave 6



QuadReg
 $y = ax^2 + bx + c$
 $a = .0019085003$
 $b = -.3193029089$
 $c = 19.5192967$
 $R^2 = .7794464181$



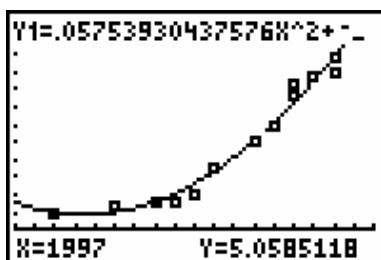
CubicReg
 $y = ax^3 + bx^2 + cx + d$
 $a = -2.301869E-5$
 $b = .007432985$
 $c = -.6843792705$
 $d = 25.15427106$
 $R^2 = .93138801$



QuarticReg
 $y = ax^4 + bx^3 + \dots + e$
 $\uparrow b = -1.263071E-4$
 $c = .0183013196$
 $d = -1.101221634$
 $e = 29.44368964$
 $R^2 = .9710326297$

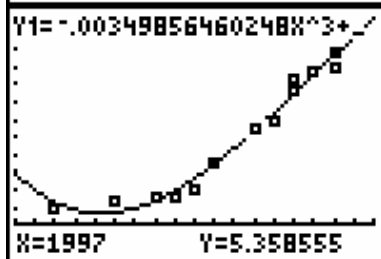
Opgave 7

QuadReg
 $y = ax^2 + bx + c$
 $a = .0575393044$
 $b = -228.8280093$
 $c = 227507.3294$
 $R^2 = .974483742$



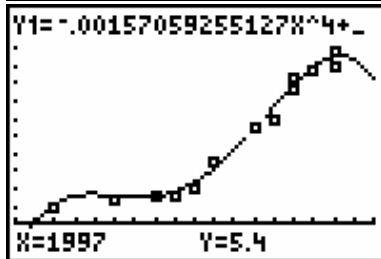
Vraagprijs volgens het
kwadratisch model :
5059 euro

CubicReg
 $y = ax^3 + bx^2 + cx + d$
 $a = -.0034985646$
 $b = 20.98626365$
 $c = -41961.19214$
 $d = 27965855.06$
 $R^2 = .9788841148$



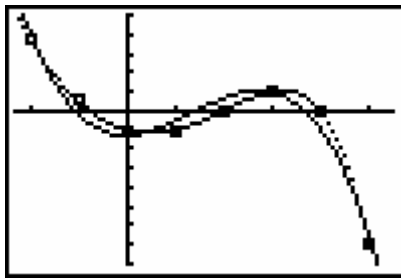
Vraagprijs volgens het
derdegraadsmodel :
5359 euro

QuarticReg
 $y = ax^4 + bx^3 + \dots + e$
 $\uparrow b = 12.52581787$
 $c = -37460.90951$
 $d = 49792786.03$
 $e = -2.4819E10$
 $R^2 = .990083998$



Vraagprijs volgens het
vierdegraadsmodel :
5400 euro

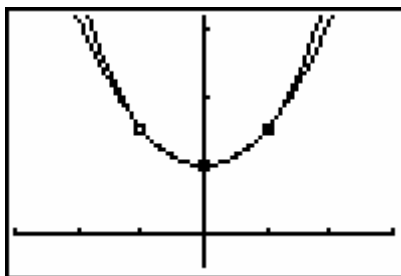
Opgave 4 bis



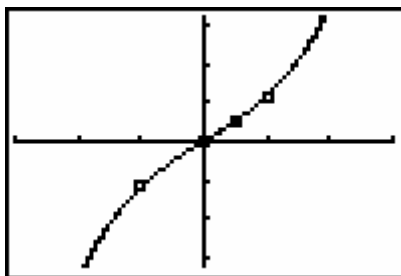
```
CubicReg
y=ax3+bx2+cx+d
a=-.2266414141
b=.9172077922
c=.0323773449
d=-1.277056277
R2=.9616314945
```

3^{de} graadsfunctie als benadering voor de functie $f(x) = x^2 - 2^x$ door gebruik te maken van 8 punten van f

Opgave 8

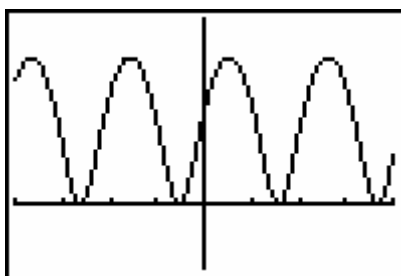


```
QuadReg
y=ax2+bx+c
a=.5430806348
b=0
c=1
R2=1
```

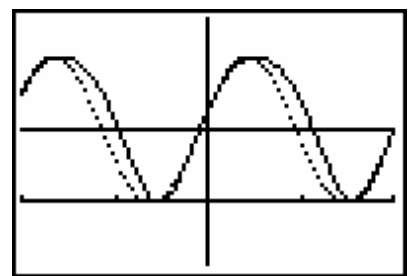


```
CubicReg
y=ax3+bx2+cx+d
a=.1773470457
b=5.9682184E-7
c=.9978535512
d=1.154E-14
R2=1
```

Opgave 9



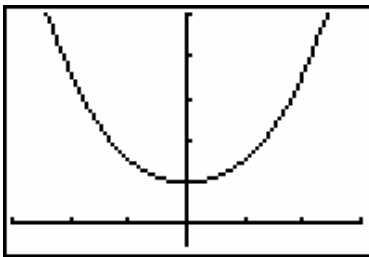
```
Intersection
X=-.0904213 Y=.23856063
```



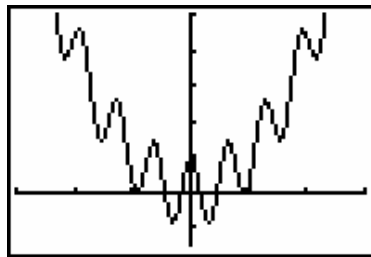
De functie $y = \log(\sin 3x + 2)$ is een periodieke functie die kan benaderd worden door de

sinusfunctie $y = \frac{\log 3}{2} \sin(3 \cdot (x + 0.090421)) + \frac{\log 3}{2}$

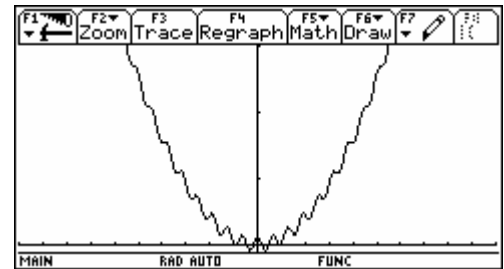
Opgave 10



$$y = x^2 + \cos x$$



$$y = x^2 + \cos 10x$$

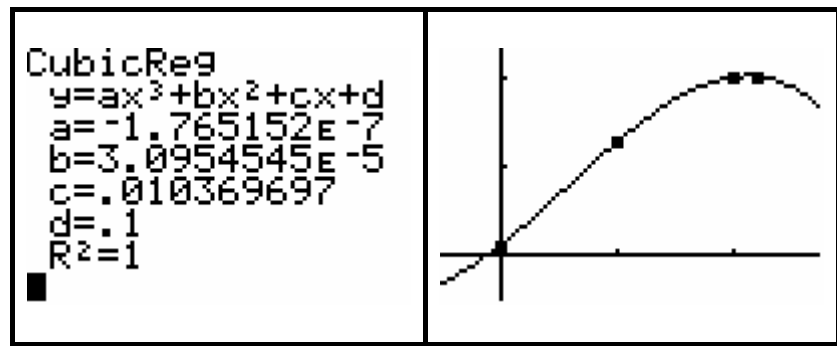


$$y = x^2 + \cos 10x$$

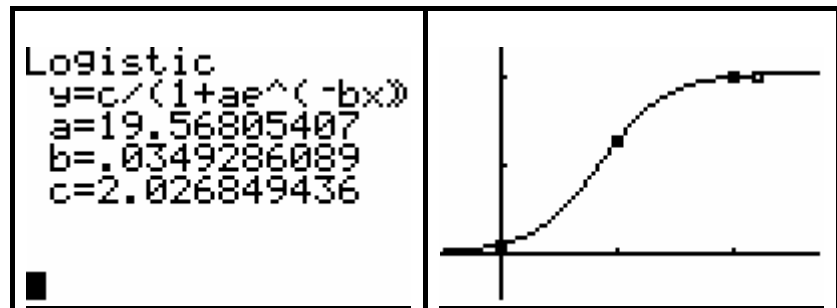
Opgave 11

3^{de} graadsveeltermfunctie
door de punten

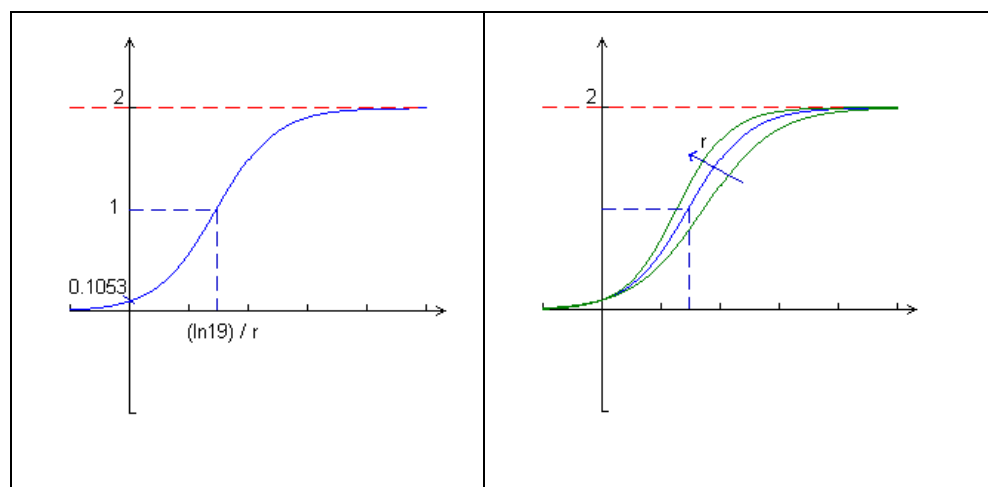
- (0 ; 0.1)
- (100 ; 1.27)
- (200 ; 2)
- (220 ; 2)



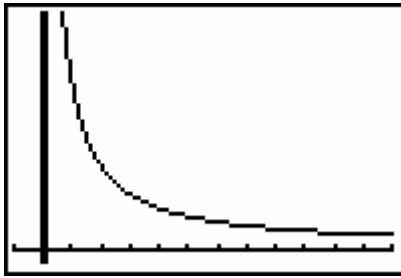
logistiek regressiemodel
op basis van dezelfde punten



$$f(t) = \frac{2 \cdot e^r}{e^r + 19}$$

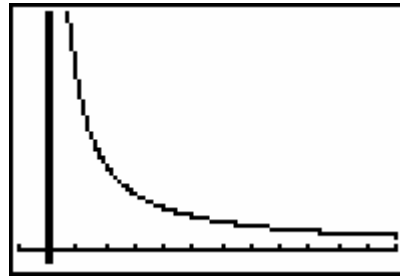


Opgave 12



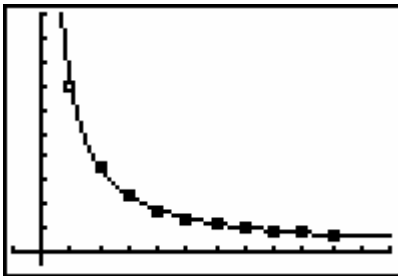
Verdubbelingstijd als functie van de intrest :

$$y = \frac{\log 2}{\log(1 + \frac{x}{100})}$$

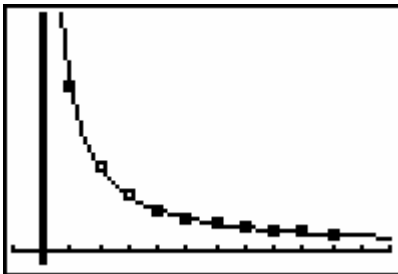


Verdubbelingstijd als functie van de intrest volgens de vuistregel (model):

$$y = \frac{72}{x}$$



```
PwrReg
y=a*x^b
a=69.19351619
b=-.9811420927
r^2=.9999639218
r=-.9999819608
```



```
LinReg
y=ax+b
a=69.43790099
b=.2794752201
r^2=.9999742439
r=.9999871219
```

lineaire regressie toegepast op het verband

$$y = \frac{a}{x} + b$$

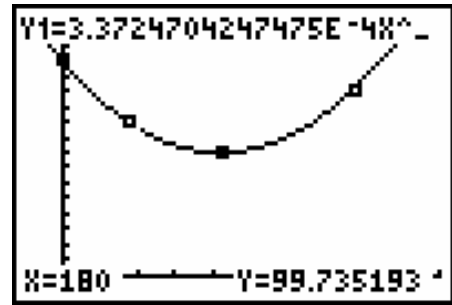
Opgave 13 : Boorplatform Yatzy

Funcievoorschrift van de parabool in het coördinatenstelsel van figuur 2 :

$$y = 3.3725 \times 10^{-4} \cdot x^2 - 0.2854 \cdot x + 140.18$$

Hoogte van de kabel op 180 meter van de piloon :

99.73 meter t.o.v. GLLWS



Kan Yatzy op die plaats onder de kabel door ?

Hoogte van Yatzy boven de Scheldespiegel : 93 m , met veiligheidsmarge : 97.5 m.

Hoogte van Yatzy in het coördinatenstelsel van figuur 2

bij laagwaterspringtij : 97.75 m passage bij $x=180$ is dus mogelijk
bij hoogwaterspringtij : 103.5 m passage bij $x=180$ is niet mogelijk

Horizontale speling

op te lossen ongelijkheid : $3.3725 \times 10^{-4} \cdot x^2 - 0.2854 \cdot x + 140.18 > 97.72$
oplossing : $x < 192.62$ (of $x > 653.64$)

de totale horizontale speling, rekening houdend met de ligging van de steiger op $x=130$ en de breedte van het boorplatform, is ongeveer 25 meter !

Hoeveel tijd heeft Yatzy ?

op te lossen ongelijkheid : $\sin(0.516 \cdot (t - 4.275)) \leq -0.301$

oplossing : $10 \text{ u } 58' \leq t \leq 15 \text{ u } 52'$

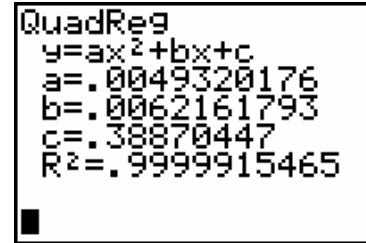
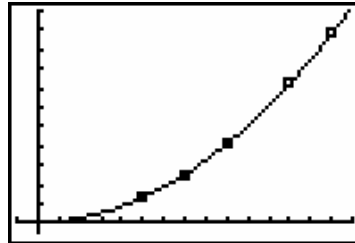
Yatzy heeft dus bijna 5 uur de tijd om de kabelhindernis te nemen.

Opgave 14 : Blokrijden

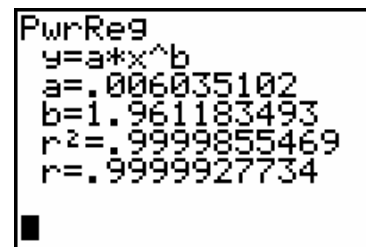
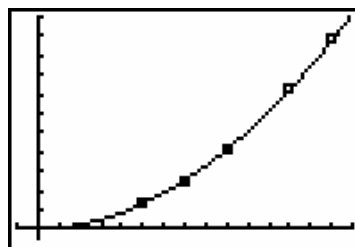
1. Model voor de remweg als functie van de snelheid

met behulp van regressie :

model 1



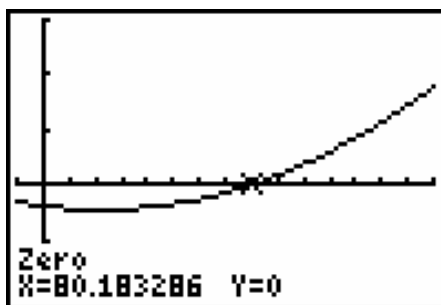
model 2



volgens het vereenvoudigd kwadratisch model :

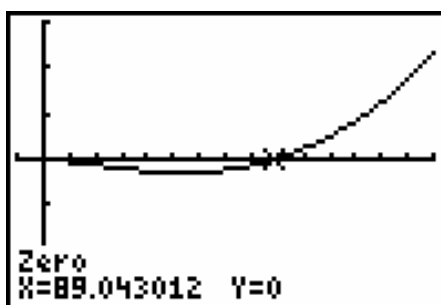
$$r(v) = k \cdot v^2 \quad \text{met } k = 0.00507 \pm 0.00007$$

vergelijking van de modellen :



$$r(v) - r_{\text{model1}}(v)$$

r in meter , v in km/u



$$r(v) - r_{\text{model2}}(v)$$

r in meter , v in km/u

2. Model voor het debiet

$$D(v) = \frac{1000 \cdot v}{60 \cdot (k \cdot v^2 + 4)}$$

met D het aantal auto's per minuut

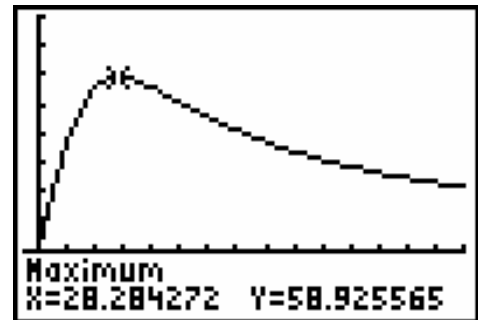
v de snelheid in km/u

k de evenredigheidsconstante $0.005 \text{ m.u}^2/\text{km}^2$

de optimale snelheid voor een maximaal debiet :

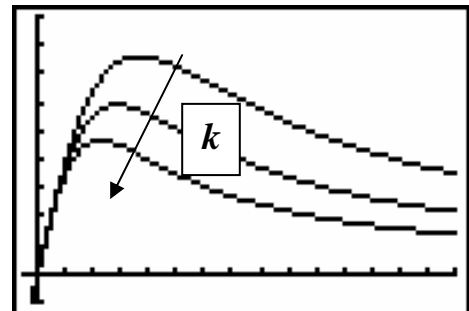
28.3 km/u (debiet : $59 \text{ auto's per minuut}$)

de onderlinge afstand tussen de wagens bij deze optimale snelheid is 4 m



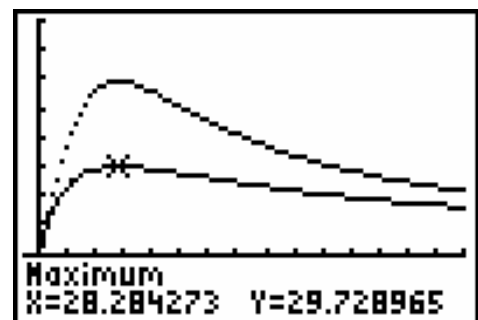
3. Aanpassingen van het model

invloed van de evenredigheidsconstante k



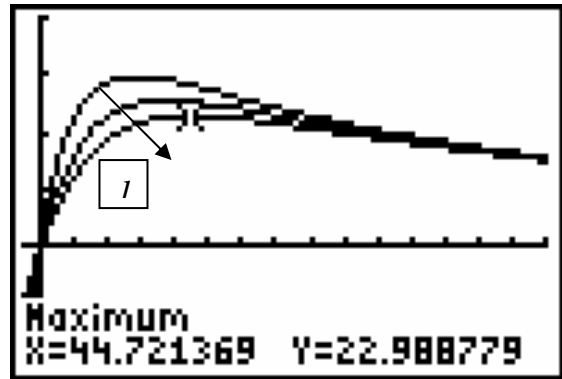
rekening houdend met een reactietijd van 1 sec :

$$D(v) = \frac{1000 \cdot v}{60 \cdot (k \cdot v^2 + 4 + \frac{10 \cdot v}{36})}$$



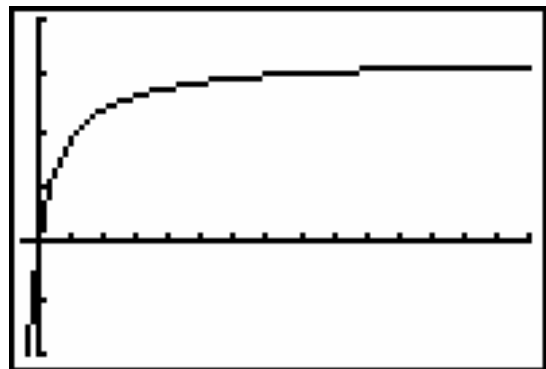
invloed van de lengte l van de voertuigen

$$D(v) = \frac{1000 \cdot v}{60 \cdot (k \cdot v^2 + l + \frac{10 \cdot v}{36})}$$



alternatief model voor de remafstand :

$$D(v) = \frac{1000 \cdot v}{60 \cdot (\frac{v}{2} + 4)}$$



vergelijking tussen de beide modellen

