II. DE WERKWIJZEN VAN DE ECONOMETRIE.

§ 4. De functie van de wiskundige economie; algemene opmerkingen.

Wij zagen reeds dat bij het econometrisch onderzoek wiskundige economie en statistiek hand in hand gaan. Voor een goed begrip van de werkwijzen der econometrie is het echter toch nodig, om de functie, die elk hunner uitoefent, scherp te onderkennen. Nu is het echter niet zo, dat de volgorde waarin ze worden toegepast, steeds dezelfde is. Levend onderzoek past zich voortdurend aan de omstandigheden aan. Het onderzoek is soms als een voortdurend kaatsspel tussen econoom en statisticus — ook al zijn die in één persoon verenigd —, of een discussie, waarin ieder om de beurt een bijdrage levert. In het algemeen zal een onderzoek met een economische ontleding van het vraagstuk beginnen. Daarna zal veelal een statistische toetsing van de over het beschouwde verschijnsel opgestelde theorie volgen. Na die toetsing is er vaak aanleiding om, weer met behulp van de theorie, verdere conclusies te trekken. Maar het kan ook zijn, dat de theorie op grond van de toetsing moet worden herzien. De herziene theorie moet dan opnieuw worden getoetst, alvorens de verdere conclusies kunnen worden getrokken: de kaatsbal gaat tweemaal heen en weer.

Wij beginnen met wat zich in het eerste stadium afspeelt en zich dan eventueel in de verdere tussenstadia kan herhalen.

Ook daar is echter de rol van de wiskundige economie nog een veelzijdige, die moeilijk in een vaste vorm te gieten is. Wij zullen slechts trachten een min of meer systematische bloemlezing uit de ervaringen, bij het werk opgedaan, te geven.

Zeer in 't algemeen kan men zeggen, dat het de taak van de wiskundige economie is om het verband, waarom het te doen is, af te leiden uit de premissen die het bepalen en dat verband nader te preciseren. Bij een zodanig verband gaat het er gewoonlijk om, de veranderingen in de omvang van een verschijnsel X te verklaren uit de veranderingen in de omvang van andere verschijnselen Y, Z, ... die we de oorzaken van X kunnen noemen. Wij spreken ook vaak van het *te verklaren verschijnsel* X en de *verklarende verschijnselen* Y, Z, enz. Heeft daarbij X betrekking op een groot complex van gelijksoortige verschijnselen — b.v. de totale vraag naar consumptiegoederen uitgeoefend door alle arbeiders — dan spreekt men wel van een *macro-economisch* vraagstuk tegenover een *micro-economisch*, wanneer het b.v. gaat om de vraag van één arbeider naar brood. Ook deze grens is weer niet scherp; men kan in elk geval nog graden van macro-economisch onderzoek onderscheiden, alnaarmate het gaat om enkelvoudige of meervoudige sommen van elementaire verschijnselen.

Een onderscheid van meer betekenis, waarover ook al veel gediscussieerd is, is dat tussen een *statische* en een *dynamische* behandelingswijze. Men mag deze woorden in een economisch betoog niet gebruiken zonder ze eerst te definiëren; zij zijn reeds in te veel verschillende betekenissen gebruikt. In onze beschouwingen zullen wij daarbij de definitie van Frisch volgen en van dynamisch onderzoek alleen dan spreken, als verband wordt gelegd tussen economische verschijnselen, die op verschillende tijdstippen betrekking hebben. Als dus de prijs van heden de aangeboden hoeveelheid van een maand later of een jaar later bepaalt, spreken wij van een dynamisch verband; maken wij daarvan bij ons onderzoek gebruik, dan is dat onderzoek dynamisch.

Niet-dynamisch is dus de gebruikelijke behandelingswijze van de evenwichtseconomie. Deze vooronderstelt, dat de prijs en de aangeboden hoeveelheid stationnair zijn (steeds gelijk blijven) en stelt dan vast, welk verband

er tussen die twee moet zijn, willen zij inderdaad stationnair kunnen blijven. Daartoe moet dan de prijs gelijk zijn aan de grenskosten, behorende bij die aangeboden hoeveelheid. Bij stationnaire grootheden is het niet nodig om op een eventueel tijdsverschil te letten, want het maakt geen verschil of wij de aangeboden hoeveelheid van nu of van het volgend jaar bezien.

Niet-dynamisch is echter volgens onze definitie evenzeer het verband tussen twee grootheden die niet stationnair zijn, wanneer de bewegingen van de ene zich ogenblikkelijk aanpassen aan die van de andere; wanneer dus de aangeboden hoeveelheid afhangt van de prijs op hetzelfde moment. Beide grootheden kunnen dan zeer goed veranderen in de tijd. Vele schrijvers noemen dit dan een geval van dynamica. Bij onze definitie is echter dit geval, even als het vorige, er ook een van een statisch karakter, en wel omdat de toestand op elk ogenblik niet medebepaald wordt door wat vroeger gebeurd is. Zou men aannemen, dat de aangeboden hoeveelheid ook afhangt van de aanwezige productie-installatie, die zelf weer een cumulatie is van de investeringen van vroegere perioden, dan heeft men weer een dynamisch verband. Van dynamisch verband spreken wij tenslotte ook, wanneer een verschijnsel X afhangt van de snelheid, waarmee een verschijnsel Y zich wijzigt; want een snelheid kan men slechts vaststellen door vergelijking van twee getallen op verschillende tijdstippen.

Men heeft lang gemeend, dat de statica eenvoudiger was dan de dynamica en heeft daarom de economie als statica opgebouwd, waaruit men later de dynamica wilde afleiden door afnemende abstractie. In ettelijke opzichten kan dat ook zo. De statica is echter in andere opzichten minder eenvoudig, minder doorzichtig, dan de dynamica. Zij verbergt n.l. de richting van het causale verband.

Bij een dynamische behandelingswijze kan men onderscheid maken tussen de volgende twee gevallen, die hier als voorbeeld bedoeld worden:

I. Bij een gegeven prijs y bepaalt de aanbieder zijn

aangeboden hoeveelheid x zodanig, dat de grenskosten daarvan gelijk zijn aan de prijs.

II. Bij een gegeven omvang der bestellingen, die het aanbod x van de aanbieder bepalen, stelt hij zijn prijs y zodanig vast, dat deze zijn grenskosten dekt.

In statische termen zal tussen I en II geen verschil zijn: in beide gevallen is de prijs gelijk aan de grenskosten. In dynamische termen is bij I y de oorzaak, x het gevolg, bij II is x de oorzaak en y het gevolg. Daar elke reactie in principe een — zij het misschien zeer kleine — tijd vordert, zal in I x achterlopen bij y, in II y bij x.

De statica is bij benadering juist, voor de beide gevallen; zij verarmt het beeld van de werkelijkheid echter enigszins. Wanneer de stationnaire toestand inderdaad bereikt is, doet er niet toe: I en II zijn dan niet meer te onderscheiden. Maar bij een veranderlijke toestand is er verschil en het mechanisme van de markt is verschillend. Door het dynamisch bekijken van het probleem stelt men het scherper, hetgeen in het bijzonder blijkt, wanneer men moet teruggaan tot de premissen, om het verband met een tevoren verwaarloosde factor in te zien.

In het bijzonder bij de bespreking van volledige economische stelsels zal dit nog nader blijken.

De aard van de relaties die men in een economisch onderzoek betrekt kan nog uiteenlopend zijn. Zij kunnen van een meer triviale aard zijn, wanneer zij niet anders weergeven dan de *definitie* van een in de beschouwing voorkomende grootheid, of wanneer zij relaties van *natuurlijke, technische* of *institutionele* aard zijn. Een voorbeeld van een definitie is: waarde van de omzet is gelijk aan prijs maal hoeveelheid, of winst is gelijk aan opbrengst minus kosten. Een voorbeeld van een natuurlijke of technische relatie is de samenhang tussen de hoeveelheid product en de hoeveelheden productiemiddelen, daartoe gebruikt. Een institutionele relatie is die tussen inkomen en te betalen belasting. Ofschoon in het bijzonder de productierelatie nog zeer belangrijk kan zijn

en ook zelfs een definitie soms nog voor toetsing vatbaar is (men denke aan de uiteenlopende wijzen waarop de winst kan worden bepaald), zijn deze relaties toch voor de economist minder belangrijk dan de tweede soort die wij willen onderscheiden: de *reactierelaties*. Deze relaties geven de reactie van een persoon, onderneming of ander economisch subject, of van een groep subjecten, op de gegeven omstandigheden weer. Zij weerspiegelen het psychologisch of institutioneel streven van die subjecten, dus het streven naar maximale bevrediging of naar maximale winst b.v., in afhankelijkheid van de organisatie van de volkshuishouding. In gewone economische taal zijn deze reactierelaties de *vraag*- of *aanbod*relaties; zij kunnen echter ook afwijkende vormen aannemen. Zij vormen gewoonlijk de kern van het economisch mechanisme. De meeste onzer voorbeelden zullen inderdaad gewone vraag- of aanbodvergelijkingen [1]) betreffen.

§ 5. Welke variabelen moeten in de relatie worden opgenomen?

De grootte van de verschijnselen X, Y, Z enz., die wij zullen aanduiden met x, y, z enz. noemen wij de grootheden of ook wel — naar wiskundig voorbeeld — de *variabelen* in de beschouwde relatie. De eerste vraag die moet worden beantwoord bij de opstelling van het gezochte verband is: welke variabelen komen er in voor? Gewoon uitgedrukt: wat zijn de verschijnselen die het te verklaren verschijnsel beïnvloeden? Zoals later nog in ander verband zal worden besproken, hebben wij hier het oog op de directe invloeden; wij vragen dus naar de naastbij liggende oorzaken van wijzigingen in x.

Algemene beschouwingen zijn hier moeilijk te geven en ook van weinig waarde. De concrete voorbeelden moeten de bedoelingen duidelijk maken.

---

[1]) In overeenstemming met het wiskundig gebruik is het woord „vergelijking" hier synoniem met relatie of verband genomen; men zou ook wel, meer naar economisch gebruik, van een „wet" kunnen spreken.

Een zeer belangrijk punt is of men een *vraag-* dan wel een *aanbodrelatie* wil vaststellen; dus of men de omgezette hoeveelheden wil zien vanuit het gezichtspunt van de kopers dan wel van de verkopers. In het eerste geval zal men alleen variabelen mogen opnemen die voor de kopers relevant zijn, in het tweede geval variabelen die voor de verkopers van belang zijn. De verhandelde hoeveelheid x zal b.v. vanuit de vraagzijde kunnen worden verklaard uit de prijs y, de inkomens z en de prijs van een concurrerend artikel u. Het is echter zinloos om ook nog de prijzen van de grondstoffen v er in te willen betrekken. Die interesseren de aanbieders, niet de vragers. Voorzover ze de vragers interesseren, zijn ze in de prijs y van het product zelf verwerkt.

Bij de overweging van verdere variabelen die in de *vraagrelatie* een rol kunnen spelen, zal men zich rekenschap moeten geven van de eigenschappen van het verhandelde goed, van de drijfveren van de vragers, van de inrichting van de markt.

De eigenschappen van bepaalde „goederen" zijn aan wijzigingen onderhevig. Zo b.v. die van het „goed" genaamd „aandeel". Een der belangrijkste eigenschappen voor de koper is het rendement van dat aandeel en veelal ook de te verwachten koersstijging. De waardering van de vrager en dus ook de prijs (de koers) zullen van dat rendement en de verwachte koersbeweging dan ook mede afhangen.

Het kan verder zeer goed voorkomen dat de door de vrager gezochte eigenschappen van het goed ook bij andere goederen aanwezig zijn. Er zijn dan dus *vervangende* goederen en het zal van de prijs van deze goederen afhangen, in verhouding tot die van het goed in kwestie, hoeveel er gevraagd wordt. Hetzelfde geldt wanneer het goed *complementair* is t.a.v. andere; dat is dat het steeds gecombineerd met andere wordt gebruikt. Terwijl in het eerste geval de twee prijzen een tegengestelde invloed zullen uitoefenen, zullen zij in het tweede geval een gelijk gerichte invloed uitoefenen.

Op een vraagrelatie kunnen tenslotte tal van factoren invloed hebben die voor elk goed samenhangen met speciale bijzonderheden van dat goed of die markt. Men dient in elk voorkomend geval de eigenaardigheden dier markt goed te bestuderen, alvorens de algemene theorie toe te passen. Voor een voorbeeld zie men § 28.

Wanneer het gaat om een *aanbodvergelijking*, zullen de variabelen, die in de eerste plaats een rol spelen, naast de prijs van het aangeboden product de *kosten van productie* zijn, die eventueel te splitsen zijn in loon- en grondstofkosten en die zelf weer afhangen van de prijzen van die factoren en de hoeveelheden. Deze hoeveelheden zullen evenredig zijn met de productiviteit der betrokken ondernemingen en voor de *grens*kosten — waarom het bij de aanbodfunctie gaat — zullen de grensproductiviteiten bepalend zijn. De grensproductiviteiten zullen veelal des te lager zijn, naarmate de productie hoger is en meestal van deze afhangen.

In bepaalde situaties — n.l. bij volle bezetting — zal de *productiecapaciteit* voor het aanbod bepalend zijn. Deze is zelf veelal als een vrijwel gegeven grootheid op te vatten, doch kan, indien men fijner wil onderscheiden, ook weer afhankelijk van andere factoren gesteld worden.

Voor zowel vraag- als aanbodvergelijking is nog van betekenis de *inrichting van de markt*. In het bovenstaande werd van de veronderstelling van vrije concurrentie uitgegaan, waarbij de vragers, resp. aanbieders, de prijs als gegeven en niet door hen te beïnvloeden beschouwen; bovendien is er maar één prijs. Bij beperkte concurrentie kan het anders zijn. Er kunnen soms meer dan één prijs zijn: de koper heeft soms de keuze tussen meer prijzen van verschillende aanbieders. Dan doet zich een situatie voor als bij de concurrerende goederen: de vraag, die bij elk der verkopers wordt uitgeoefend, hangt af, zowel van de prijs die hijzelf, als van de prijs die zijn concurrent vraagt. De aanbodcurve van een monopolist zal voorts

ook van de prijs op andere wijze afhangen dan de aanbodcurve van een met anderen concurrerende verkoper. (vgl. § 9).

§ 6. De wiskundige vorm der relatie.

Een vraag, die afzonderlijke overweging verdient, is: welke moet de wiskundige vorm zijn van de gezochte relatie? Moet, in het voorbeeld van de vraagfunctie, de gevraagde hoeveelheid afhangen van de eerste macht van de prijs of van de tweede macht, of van welke andere functie van de prijs? Voor het rekenprocédé maakt dit uiteraard nog enig verschil, soms zelfs veel verschil. De economische theorie vertelt ons daarover al heel weinig; zij weet het meestal zelf niet. In enkele bijzondere gevallen is het mogelijk om op grond van redenering iets naders over de vorm van de functie af te leiden. Wanneer b.v. de vraag naar een bepaald goed door een zuiver technisch verband samenhangt met de hoeveelheid van een ander goed, en daarmee b.v. evenredig is, dan staat het wiskundig verband t.o.v. deze andere hoeveelheid althans vast. Wanneer van te voren vaststaat, dat iemand aan een bepaald goed een bepaald bedrag zal besteden, dan is de hoeveelheid, die hij zal kopen, omgekeerd evenredig met de prijs. Deze gevallen zijn echter zeldzaam. Meestal weet men het niet en het is juist hier, dat de theoretische economie alleen door econometrisch onderzoek verder zal kunnen komen. Ook kan men soms op grond van nadere redenering wel bepaalde eigenschappen van het gezochte wiskundige verband op het spoor komen. De wijze, waarop de vraag naar een bepaald goed afhangt van het inkomen, vormt hiervoor een goed voorbeeld. In gedachten kan men het inkomen laten variëren van nul tot oneindig en vaststellen, dat bij een inkomen van nul nu ook eenmaal niets kan worden uitgegeven aan het bewuste artikel. Verder kan nooit meer worden uitgegeven dan het gehele inkomen bedraagt. Bij een artikel van eerste levensbehoefte zonder vervangingsmogelijkheden

van betekenis kan men aannemen, dat er een zekere verzadigingshoeveelheid is, waarboven de vraag niet uitkomt. Uit al dergelijke overwegingen kan men enkele algemene eigenschappen van de gezochte functie afleiden, waardoor de keuze beperkter wordt (vgl. § 25).

Zeer veel nut heeft men in al deze gevallen van een zeer algemene wiskundige stelling, n.l. dat binnen een klein interval van variatie bijna elke functie mag worden benaderd door een *rechtlijnige*. Men vervangt de curve die het verband voorstelt dan voor dat interval door een stukje van de raaklijn. Wanneer men met kleine veranderingen in de betrokken variabelen te maken heeft, kan men dus steeds met een lineaire functie van die variabelen volstaan, d.w.z. dat men dus probeert om x te benaderen door een uitdrukking $x' = a + by + cz + du \ldots$, waarin a, b, c, enz. constante getallen zijn. Dit betekent overigens niet, zoals wel eens wordt gedacht, dat er evenredigheid tussen x' en y wordt verondersteld. Er wordt slechts evenredigheid verondersteld tussen een kleine toeneming van b.v. y en de daarmee corresponderende toeneming van x'. Geven we die toenemingen aan, zoals in de wiskunde gebruikelijk is, met resp. Δy en $\Delta x'$, dan geldt inderdaad krachtens onze veronderstelling:

$$\Delta x' = b \, \Delta y. \qquad (6.1)$$

Doch deze veronderstelling is veel minder vergaand dan een van evenredigheid tussen x' en y zelf.

Er zijn echter nog andere aspecten van het wiskundig verband die, ook bij kleine wijzigingen in de variabelen, niet van te voren vaststaan en daarom aandacht verdienen. Uit de economische theorie is soms niet duidelijk, of de grootheid x afhangt van y, dan wel van de *cumulatie van* y (of in meer gewone wiskundige taal: van de „integraal over de tijd" van y). Wat de cumulatie van y is, moge eerst even aan een getallenvoorbeeld worden toegelicht. Indien het verloop van de reeks y in de tijd gegeven is en b.v. is:

$$y = 5, \quad 3, \quad 6, \quad 8, \quad 4 \quad \ldots,$$

dan wordt onder de cumulatie van y de reeks verstaan, waarvan de eerste term 5 is, de tweede $5 + 3 = 8$, de derde $5 + 3 + 6 = 14$, de vierde $5 + 3 + 6 + 8 = 22$, enz. Wil men een economisch voorbeeld van de bedoelde onzekerheid, dan kan de vraag genoemd worden, hoe de rentestand op de voorraden van gegeven goederen inwerkt. Tweeërlei is denkbaar: òf de rentestand is bepalend voor de voorraad, òf hij is bepalend voor de toevoegingen, positief of negatief, aan de voorraad. De literatuur op dit punt is verre van duidelijk en men hinkt vaak op twee gedachten. Als het eerste waar is, wordt de voorraad x bepaald door de rentestand y. Als echter het tweede waar is, wordt de voorraad x, die zelf de cumulatie is van de voorraadtoevoegingen in elke voorafgaande periode, bepaald door de cumulatie van de waarden van de rentestand y.

Het maakt nu een zeer aanmerkelijk verschil of men het ene dan wel het andere veronderstelt. En alvorens men een theorie van dit type gaat toetsen, dient men dus zijn gedachten te bepalen.

§ 7. De wijze waarop de tijd in de relatie voorkomt.

Wij zijn daardoor al gekomen op een verder punt, n.l. de vraag op welke wijze de tijd in de relatie waarom het gaat een rol speelt. Het voorkomen van een gecumuleerde variabele betekent reeds een vorm van invloed van de tijd; wij hebben dan reeds, zoals gezegd, een dynamisch verband. Op eenvoudiger wijze kan dit ook optreden, wanneer een variabele z niet op x inwerkt op hetzelfde moment, maar met een zekere *vertraging* T. Wij kunnen dan zeggen, dat

$$x_t = a + c\, z_{t-T},$$

waarbij de indices bij de grootheden x en z aangeven, voor welke tijdsperiode men die variabelen moet nemen. Men noemt T ook wel de „*lag*" (een Engels woord), die

x t.o.v. z vertoont, of de „*lead*", die z t.o.v. x vertoont. Dit is misschien wel de allereenvoudigste dynamische relatie, die men zich kan denken. Een vraag is, of ook de economische theorie uitsluitsel kan geven over de wijze waarop de tijd in de relaties dient voor te komen, wanneer men de exacte samenhang der dingen beschouwt. Een dergelijke eenvoudige vertraging kàn zich dan ongetwijfeld voordoen, n.l. wanneer er een technische oorzaak is die een constante tijd verlangt alvorens een bepaald besluit uitwerkt. De lengte van een productieproces kan tot een dergelijke vertraging aanleiding zijn.

Het kan echter ook zijn, dat de dynamische relaties geheel anders van vorm zijn, wanneer zij meer psychologisch ontstaan. Denken wij ons dat, naar de omstandigheden van het ogenblik t, de vraag naar een zeker goed gelijk moet zijn aan x_1 om de koper de maximale bevrediging te schenken. Wanneer zich nu de omstandigheden plotseling wijzigen, zal de meest gewenste vraag wellicht x_2 worden. Doch er verloopt enige tijd alvorens dit tot den koper doordringt. Bij een kleine verandering in omstandigheden zal hij 't wellicht in 't geheel niet bemerken, dat een wijziging van x_1 in zijn voordeel is. Bij een grotere verandering zal dit hem langzamerhand duidelijk worden. Hij zal er nu naar streven, om van x_1 tot x_2 te komen. Allerlei „hindernissen" kunnen dat echter tegenwerken en daardoor geschiedt die aanpassing niet onmiddellijk. Er heeft een proces van aanpassing plaats. Wellicht geschiedt dat aldus, dat de snelheid waarmee hij zich van x_1 naar x_2 beweegt evenredig is met de afstand tot x_2 — de „aantrekkingskracht" zou men kunnen zeggen — en daarom langzaam afneemt. In 't allereerst kan zij ook nog toenemen, enz. Op dit terrein is nog vrijwel niets onderzocht.

Terwijl dan verder die aanpassing plaats vindt, veranderen de omstandigheden weer en wordt het aanpassingsproces afgebroken en door een nieuw vervangen. Het resultaat van een dergelijke reeks van aanpassings-

processen is een geheel andere curve dan één enkel aanpassingsproces; bij benadering kan er echter zeer goed een vertraagde lijn uitkomen.

De tijd speelt ook nog op een geheel andere wijze een rol. Bij het streven naar maximale bevrediging dat aan elk economisch handelen ten grondslag ligt maakt het verschil, over hoe lange tijd men vooruitziet. In alledaagse taal uitgedrukt, kan het onmiddellijke belang soms niet samenvallen met het belang voor later en vaak is het „onverstandig", kortzichtig te zijn. In meer wetenschappelijke discussies heeft men wel gebruik gemaakt van het begrip „*horizon*", d.i. het tijdvak waarover iemand vooruit ziet. De lengte van de horizon beïnvloedt de aanbodcurven en vraagcurven soms aanzienlijk. Het eenvoudigste voorbeeld is het geval van een overmatig grote oogst. Als de boer of de handelaar niet verder vooruit zag dan één jaar, zou hij die gehele oogst aanbieden en genoegen moeten nemen met een extra lage prijs. Als hij echter niet over één jaar, maar over twee jaar vooruitziet, zal hij rekening houden met de waarschijnlijkheid, dat er het tweede jaar een gemiddelde oogst zal zijn; en als hij nu de totale oogst van die twee jaren gelijkmatig over die twee jaren verdeelt, zal hij een grotere opbrengst verkrijgen. Hij zal daarom goed doen om niet de gehele, overvloedige, oogst van dit jaar aan te bieden, doch de helft van het exces achter te houden. De lengte van de horizon heeft hier een duidelijke en voor vereenvoudigde gevallen berekenbare invloed op het aanbod.

Hetzelfde doet zich voor bij de waardering van aandelen. Als een onderneming een jaar geen dividend uitkeert, zou een aandeelhouder die niet verder dan dat jaar ziet kunnen redeneren dat het aandeel hem niets waard was: het levert geen enkel revenu. Meestal zal hij echter over langere tijd vooruitrekenen en bij zijn waardering rekening houden met de waarschijnlijkheid dat na enige tijd wederom dividend zal worden uitgekeerd. Aan de andere kant zal men ook niet, wanneer het dividend in een jaar

het drievoudige van normaal bedraagt, de koers van dat aandeel tot het drievoudige van normaal zien stijgen. Fig. 4 B (blz. 69) doet dit zeer duidelijk zien. Daarin is het dividendpercentage langs de horizontale as afgezet, terwijl in verticale richting het product van de koers en van de rentevoet voor lang krediet (het rendement van staatsobligaties) is uitgezet. Deze laatste grootheid, aangegeven met de benaming „waardering", geeft a.h.w. aan met een obligatie van welk rentetype men het aandeel gelijkstelt. Is de koers 200 en het rendement van staatsobligaties 4, dan stelt men het aandeel blijkbaar gelijk met een (denkbeeldige) obligatie die 8% zou opleveren. Indien men zich nu door het ogenblikkelijke dividendpercentage geheel zou laten leiden, zou bij een dividend van 8% ook een „waardering" van 8% moeten optreden. In werkelijkheid ziet men een dergelijke gelijkheid niet, doch een verband dat van het ogenblikkelijke dividendpercentage veel minder sterk afhangt. Grofweg kan men zeggen, dat een stijging van het dividend met 1% slechts een verhoging van de „waardering" met ½% tot gevolg heeft, zodat bij een dividend van 4% een „waardering" van 4%, maar bij een dividend van 8% een „waardering" van 6% optreedt. Ook hier dus een voorbeeld van de invloed van de horizon op het gedrag der economische subjecten.

Er is een laatste wijze, waarop „de factor tijd" zich doet gevoelen, wanneer het verband dat wij zoeken (vraag- of aanbodverband of eventuele andere relaties) zich in de tijd geleidelijk wijzigt. Strikt genomen mag men hier eigenlijk niet spreken van een invloed van de tijd, want de directe invloed gaat hier uit van andere verschijnselen. Dit kunnen zijn een *geleidelijke* vermeerdering van de technische kennis, een *geleidelijke* wijziging van de smaak en vele andere verschijnselen. Doch wanneer men moeilijk een andere maatstaf voor deze verschijnselen kan vinden — er zijn er vaak vele tegelijk aanwezig — en wanneer men zich toch in hoofdzaak interesseert voor de snellere

fluctuaties, neemt men vaak „de tijd" als variabele op in de relatie die men zoekt. Men zegt dan ook wel, dat men een trendcomponent toevoegt om de invloed van de genoemde geleidelijk veranderde factoren te vertegenwoordigen. Vanzelfsprekend is dit slechts zolang juist, als de bedoelde factoren zich ook inderdaad geleidelijk blijven ontwikkelen en niet schoksgewijze. Een paar concrete voorbeelden zijn de invloed van de drankbestrijding — in zijn uiteenlopende directe en indirecte vormen — op het verbruik van alcoholica en de invloed van het opkomend wegverkeer op de vraag naar spoorwegvervoer.

§ 8. Verband met andere relaties.

Van bijzondere betekenis is de wiskundige formulering der economische relaties, wanneer men het verband met andere, min of meer verwante, relaties duidelijk wil maken. Ook hier kunnen voorbeelden meer doen dan algemene betogen. Kiezen wij als eerste voorbeeld daarom de verschillende relaties, die men bij het beschrijven van een enkelvoudige markt op „Walrasiaanse" wijze pleegt te onderscheiden — of soms, tot zijn schade, niet scherp onderscheidt! Wij spraken reeds over de vraag- en de aanbodrelatie. Indien p de prijs voorstelt en u de omgezette hoeveelheid, zijn deze beide relaties vergelijkingen waarin de u o.m. verklaard wordt uit de p. Doch de aanbodvergelijking verklaart u van uit de gezichtshoek der aanbieders en naast p komen er daarom grootheden in voor, die men aanbodfactoren zou kunnen noemen en die we met a_1, a_2 ... zullen aangeven. Evenzo komen in de vraagvergelijking vraagfactoren v_1, v_2 ... voor.

Volledig geschreven zijn zij dus:

$$\text{aanbodvergelijking: } u = f\,(p,\ a_1,\ a_2 \ldots) \qquad (8.1)$$

$$\text{vraagvergelijking: } u = g\,(p,\ v_1,\ v_2 \ldots) \qquad (8.2)$$

Door gelijkstelling van, zoals men het vaak uitdrukt, vraag en aanbod, verkrijgt men een nieuwe vergelijking

$$f\ (p,\ a_1,\ a_2\ ..) = g\ (p,\ v_1,\ v_2\ ...)$$

waarin u niet meer voorkomt. Daaruit kan men in de meeste voor ons belangrijke gevallen de p oplossen:

$$p = h\ (a_1,\ a_2, ...\ v_1,\ v_2\ ...) \qquad (8.3)$$

Laten wij dit de *prijsvergelijking* noemen. Zij drukt de prijs uit in de aanbod- en vraagfactoren, zonder behulp van de omgezette hoeveelheid. Men kan echter ook uit (8.1) en (8.2) p oplossen en de aldus verkregen uitdrukkingen aan elkaar gelijk stellen, m.a.w. uit (8.1) en (8.2) p elemineren; dan krijgt men een uitdrukking in u, a_1, a_2, v_1, v_2.. waaruit men u kan oplossen:

$$u = k\ (a_1,\ a_2, ...\ v_1,\ v_2\ ...) \qquad (8.4)$$

Men zou kunnen spreken van de *omzetvergelijking*. De vergelijkingen (8.3) en (8.4) zijn tezamen equivalent met (8.1) en (8.2); zij drukken de marktgrootheden — in dit beperkte vraagstuk de „onbekenden" — uit in de marktgegevens a_1, a_2... v_1, v_2... Men kan trouwens elk tweetal nemen om de markt te beschrijven. Doch het is b.v. onjuist om p te willen verklaren uit u, a_1, a_2,... v_1, v_2... allemaal; een dergelijke relatie heeft daarom geen zin, omdat u alleen reeds uit a_1, a_2, v_1, v_2... kan worden afgeleid, zoals (8.4) doet zien en het dus „dubbelop" is om in een vergelijking voor p naast a_1, a_2, v_1, v_2... ook nog u te willen zetten.

Daar, zoals gezegd, (8.3) en (8.4) de plaats van (8.1) en (8.2) kunnen innemen, moet er ook verband tussen de coëfficiënten van die vergelijkingen zijn. Het eenvoudigste ziet men dat, wanneer men van de lineaire benadering of van een exponentiële vorm der vergelijkingen gebruik maakt. Zij b.v.

$$u = v_{11}\,p + v_{12}\,a \text{ de aanbodrelatie en} \qquad (8.1')$$

$$u = v_{21}\,p + v_{22}\,v \text{ de vraagrelatie,} \qquad (8.2')$$

waarbij wij gemakshalve nu slechts één vraagfactor en één aanbodfactor hebben aangenomen. (Het geeft geen moeilijkheden, er meer in te voeren).

De prijsvergelijking wordt dan

$$(v_{11} - v)_{21} p = v_{22} v - v_{12} a$$

of

$$p = \frac{v_{22} v - v_{12} a}{v_{11} - v_{21}} \qquad (8.3')$$

en de omzetvergelijking

$$(v_{21} - v_{11}) u = v_{21} v_{12} a - v_{11} v_{22} v$$

of

$$u = \frac{v_{21} v_{12} a - v_{11} v_{22} v}{v_{21} - v_{11}} = \frac{v_{11} v_{22} v - v_{21} v_{12} a}{v_{11} - v_{21}}, \qquad (8.4')$$

waaruit duidelijk blijkt, hoe de coëfficiënten der laatste twee vergelijkingen samenhangen met die van de eerste twee. Zou men om de een of andere reden beter de vergelijkingen (8.3′) en (8.4′) statistisch kunnen toetsen dan (8.1′) en (8.2′), dan is het dus met behulp van de bovengegeven uitdrukkingen mogelijk, om daaruit de coëfficiënten van de vraag- en aanbodrelatie te bepalen.

Wij willen er tevens nog even aan herinneren, dat de coëfficiënten v_{11} en v_{21} ten nauwste samenhangen met de elasticiteitscoëfficiënten van resp. aanbod en vraag. Onder de elasticiteit verstaat men immers de verhouding tot een gegeven prijsverandering, uitgedrukt in %, van de daaruit volgende vraag- resp. aanbodverandering, ook in % uitgedrukt; strikt genomen moeten de veranderingen oneindig klein zijn, doch voor kleine veranderingen is de afwijking niet groot. Noemen we een kleine verandering van de prijs Δp, de daaruit volgende verandering van de omgezette hoeveelheid Δu, dan is de elasticiteit:

$$\varepsilon = \frac{\Delta u}{u} : \frac{\Delta p}{p} = \frac{\Delta u}{\Delta p} \frac{p}{u}$$

Voor de aanbodselasticiteit moet men Δu uit de aanbodvergelijking afleiden; dan vindt men n.l.

$$\Delta u = v_{11} \Delta p \quad \text{of} \quad \frac{\Delta u}{\Delta p} = v_{11} \quad \text{en} \quad \varepsilon_a = v_{11} \frac{p}{u}.$$

Voor de vraagelasticiteit moet men Δu uit de vraagvergelijking afleiden, waarbij men vindt:

$$\varepsilon_v = v_{21} \frac{p}{u}.$$

Daarbij moeten dan p en u steeds bij elkaar behorende waarden zijn, d.w.z. u moet uit p volgen uit de vraag- of aanbodvergelijking.

Werkt men met de *exponentiële* benadering, dan stelt men de aanbodvergelijking voor door

$$u = p^{v_{11}} a^{v_{12}} \quad \text{of} \quad \log u = v_{11} \log p + v_{12} \log a \qquad (8.1'')$$

en de vraagvergelijking door

$$u = p^{v_{21}} a^{v_{22}} \quad \text{of} \quad \log u = v_{21} \log p + v_{22} \log v \qquad (8.2'')$$

De prijs- en de omzetvergelijking krijgt men nu door in (8.3′) en (8.4′) in plaats van p, a, v en u telkens te lezen $\log p$, $\log a$, $\log v$ en $\log u$.

De elasticiteit is nu eenvoudiger; zij bedraagt voor het aanbod v_{11} en voor de vraag v_{21} en is dus onafhankelijk van de waarden van u en p, d.w.z. van de plaats op de curve waar men zich bevindt. Dit is de reden waarom men veelal vraag- en aanbodcurven van het exponentiële type aanneemt, indien men niets naders omtrent hun gedaante weet. Voor andere doeleinden, nl. voor het gebruik van dergelijke relaties in zgn. complete stelsels, is het gebruik van lineaire vergelijkingen weer eenvoudiger en daarom worden die ook veel gebruikt. Beide typen zijn, zoals reeds gezegd werd, als benaderingen over kleine intervallen bedoeld en zijn gewoonlijk bepaald fout als men zich te ver van het punt van uitgang verwijdert.

§ 9. Aanbodvergelijking en prijsfixatievergelijking.

Een tweede voorbeeld van het verband met andere relaties en van het nut, bij het onderzoek daarvan, van een wiskundige behandelingswijze, verkrijgt men wanneer

men het nodig vindt om dieper te graven en het ontstaan van de aanbod- of vraagcurve zelf te onderzoeken. [1]) Beide worden, in de Walrasiaanse gedachtengang, afgeleid uit het streven naar maximumbevrediging en de coëfficiënten van de functies (8.1) en (8.2) moeten dus samenhangen met de coëfficiënten in b.v. de nuttigheidsfunctie of in de winstfunctie. Een voorbeeld van het eerste zullen wij later bespreken (§ 25), terwijl een uiterst eenvoudig voorbeeld van het tweede nu zal worden besproken, tevens ter adstructie van een andere kwestie, nl. hoe de organisatie van de markt invloed kan hebben op de coëfficiënten van de aanbodfunctie. Denken wij ons een bedrijfstak die een product vervaardigt in een hoeveelheid u en waarvan de vraagfunctie is

$$u = V_0 - V_1 p \tag{9.1}$$

Laat de productie kosten vergen die alleen bestaan uit een constant gedeelte k_0 en een proportioneel gedeelte $k_1 u$, waarin k_0 de totale vaste kosten voor de gehele productie en k_1 de variabele kosten per eenheid product voorstellen.

Nemen wij nu eerst aan, dat alle productie zich in handen van een *monopolist* bevindt. Dan zal de winst Z van deze monopolist bij een prijs p bedragen:

$$\begin{aligned} Z &= p\,u - k_0 - k_1 u \\ &= (p - k_1)\,u - k_0 \\ &= (p - k_1)\,(V_0 - V_1 p) - k_0 \end{aligned}$$

Deze zal tot een maximum worden, indien de prijs gesteld wordt op een bedrag p_M dat voldoet aan de eis, dat Z zo groot mogelijk wordt. Men berekent op de bekende wijze dat:

$$p_M = \frac{V_0 + V_1 k_1}{2 V_1} = \frac{V_0}{2 V_1} + \frac{k_1}{2} \tag{9.2}$$

Als nu in de loop van de tijd de vraagcurve van plaats

[1]) Veelal is dit niet nodig en handelt de econometrist hierbij dus eigenlijk evenals Cassel, die de vraag- en aanbodcurven als het grondgegeven wil beschouwen.

verandert („shifts vertoont"), d.w.z. dat V_0 veranderlijk is, terwijl de V_1 onveranderd blijft, dan wordt het verloop van de prijs door (9.2) gegeven. Deze vergelijking is echter geen aanbodvergelijking, want het is geen vergelijking waarin p uitsluitend met aanbodfactoren voorkomt. De grootheid k_1 is een aanbodfactor, en V_1 is een constante; maar V_0 is een vraagfactor. Aan de andere kant komt u in vgl. (9.2) niet voor, hetgeen in een aanbodvergelijking wel het geval is. Om de aanbodfunctie in dit geval te vinden, moeten wij dus V_0 vervangen door u, als dat kan. Het kan inderdaad, met behulp van de vraagvergelijking. Wij krijgen dan:

$$p_M = \frac{u + V_1 p_M}{2 V_1} + \frac{k_1}{2}$$

hetgeen wij kunnen schrijven in de vorm:

$$u = V_1 (p_M - k) \tag{9.3}$$

Deze aanbodvergelijking mag niet opgevat worden in de strenge Walrasiaanse zin van een vergelijking, die op van de vraagzijde onafhankelijke wijze het gedrag van de aanbieders vastlegt. Maar naar de vorm is zij het wel. Het merkwaardige is, dat er een constante van de vraagzijde in voorkomt, nl. V_1. Er blijkt uit, dat de elasticiteit van het aanbod van een monopolist afhangt van V_1, en wel gelijk is aan $V_1 \frac{p}{u}$. Hierin is dus een mogelijkheid gelegen, om statistisch te toetsen of op een gegeven markt monopolistisch wordt aangeboden.

Wanneer er vrije concurrentie zou heersen, zou de aanbodfunctie er geheel anders hebben uitgezien, n.l.

$$p = k_1 \tag{9.4}$$

Indien tenslotte beperkte concurrentie zou hebben geheerst, zouden nog andere aanbodfuncties hebben gegolden. Op deze details zullen wij echter nu niet ingaan.

Dit eenvoudige voorbeeld is echter een goed uitgangspunt voor een nadere beschouwing van de aanbodfunctie

en haar betekenis. Wanneer men het marktgebeuren van nabij beschouwt en in het bijzonder de opeenvolging in de tijd der afzonderlijke „economische daden" beziet — dus dynamisch te werk gaat — zal het in vele gevallen voorkomen, dat de aanbieder de prijs niet als een gegeven beschouwt en op grond daarvan bepaalt, hoeveel hij zal produceren. Indien hij voor de markt en dus eventueel op voorraad produceert, kan dat bij benadering het geval zijn. Indien hij echter op bestelling produceert, zeker niet. Het gaat dan eerder aldus, dat hij zijn productie baseert op zijn bestellingen, d.i. op de gevraagde hoeveelheid, en de prijs in zekere zin zelfstandig vaststelt. Natuurlijk wordt hij daarbij op de lange duur gedwongen, om in zijn prijs zijn concurrenten te volgen; een te grote afwijking zou op den duur zijn klanten doen verdwijnen. Doch op korte termijn is hij in zijn prijsfixatie autonoom. Hij zal bij die prijsfixatie rekening houden met zijn kosten. Deze hangen af van de reeds vroeger genoemde aanbodfactoren, als loonvoet, prijs der grondstoffen en productiviteit van de arbeid in zijn bedrijf, en daarnaast van de omvang van zijn productie. Deze laatste bepaalt immers welke marginale machines en arbeiders nog moeten meewerken en welke arbeidsorganisatie hij kan toepassen. Zou men de reactierelaties ter beschrijving van deze situatie willen weergeven, dan zou men kunnen schrijven:

$$u = v_{21}\, p + v_{22}\, v,$$

de vroegere vraagvergelijking, die de reactie van den koper op een voor hem gegeven prijs p voorstelt en

$$p = g\ (a,\ u) \tag{9.5}$$

die de reactie van den verkoper, n.l. het *fixeren van de prijs* voorstelt, dat geschiedt op grond van de reeds genoemde aanbodfactoren, benevens op de omvang van de productie. Het enige dat nog is verwaarloosd in deze vergelijkingen is het mogelijke tijdsverschil tussen de grootheden. Dit interesseert ons op 't ogenblik minder; wat wij thans willen betogen is, dat de vergelijking (9.5)

naar de variabelen, die er in voorkomen, toch een aanbodvergelijking is: het is een relatie tussen de grootheden u, p en a. Er is dus, zodra men de tijdsverschillen verwaarloost, geen verschil en men kan dus een aanbodvergelijking ook interpreteren als een *prijsfixatievergelijking*. In veel gevallen zal dit een realistischer behandelingswijze in de hand werken.

§ 10. Macro-economische onderzoekingen.

Deze hebben betrekking, zoals wij reeds opmerkten, op *groepen* van personen of van markten, dus b.v. op de totale vraag naar een bepaald artikel, uitgeoefend door alle consumenten of op de totale vraag naar alle consumptiegoederen, uitgeoefend door alle consumenten. Als wij ons eerst tot het eerstgenoemde geval beperken, en aannemen dat er vrije concurrentie op bedoelde markt bestaat, waardoor de prijs voor elk der vragers dezelfde is, is het niet moeilijk om de totale vraag naar het beschouwde goed te vinden in afhankelijkheid van de prijs. Als de individuele vraagfuncties zijn

$$u_1 = f_1(p),\ u_2 = f_2(p) \ldots \qquad (10.1)$$

waarbij de index betrekking heeft op een bepaalde persoon, dan is de totale vraag $u = u_1 + u_2 + \ldots = f_1(p) + f_2(p) + \ldots = f(p)$, of: opnieuw een functie van de prijs.

De zaken liggen anders, wanneer het gaat om de totale vraag naar alle consumptiegoederen tezamen. Dan is er sprake van verschillende prijzen, n.l. een voor elk der goederen, zeggen wij p^1, p^2, $p^3 \ldots$ enz., waarbij de cijfers weer indices zijn, geen exponenten. De vraag naar elk goed afzonderlijk hangt af van elk dezer prijzen, of, in eerste benadering misschien, van de prijs van het goed zelf alleen. In beide gevallen hangt de totale vraag naar alle goederen af van alle genoemde prijzen. Bij macroeconomische beschouwingen pleegt men echter niet meer alle prijzen afzonderlijk te beschouwen, doch slechts een gemiddelde prijs van consumptiegoederen. De bedoeling

is n.l. juist om het beeld te vereenvoudigen. Het staat echter van te voren geenszins vast, dat een functie van de prijzen p^1, p^2, $p^3 \ldots$ ook mag worden opgevat als een functie van een of ander gemiddelde dier prijzen:

$$p = \frac{a^1 p^1 + a^2 p^2 + \ldots}{a^1 + a^2 + \ldots} \qquad (10.2)$$

Er zijn natuurlijk functies waarvoor dat wel het geval is. Zo is b.v. de functie $F(p^1, p^2 \ldots)$, die gedefiniëerd wordt als:

$$F(p^1, p^2 \ldots) = A \times (a^1 p^1 + a^2 p^2 + \ldots) \qquad (10.3)$$

natuurlijk gelijk aan $A(a^1 + a^2 + \ldots)p$. Maar een willekeurige andere functie zal als regel niet als functie van p alleen te schrijven zijn. Is daarmee dus het ontoelaatbare van dergelijke macro-economische methoden bewezen? Neen; er zijn enige voorwaarden op te noemen, waaronder de macro-economische procedures als benadering zeer goed bruikbaar zijn.

Wanneer het gaat om een enkele relatie, is de zaak het eenvoudigst. Men mag n.l. bij kleine veranderingen in de variabelen steeds met lineaire functies werken. De totale vraag naar alle goederen mag men dus ook aldus voorstellen:

$$u = b^1 p^1 + b^2 p^2 + \ldots + c \qquad (10.4)$$

Het is alleen niet zeker, dat de coëfficiënten b^1, b^2, enz., die dan voor de termen met p^1, $p^2 \ldots$ komen te staan, juist dezelfde verhouding hebben als de coëfficiënten a^1, $a^2 \ldots$ in de definitie van p (10.2). Als dat echter niet zo zou zijn, dan is er geen enkel bezwaar tegen om p op andere wijze te definiëren, n.l. met de coëfficiënten b^1, $b^2 \ldots$ in plaats van a^1, $a^2 \ldots$. Integendeel, dit is er een voorbeeld van, dat men indexcijfers juist steeds zodanig moet construeren, dat zij voor het vraagstuk, dat men er mee wil behandelen, geschikt zijn.

Wij willen dit nog even met een concreet voorbeeld toelichten. Nemen wij aan, dat de vraag naar ieder afzonderlijk goed alleen afhangt van zijn eigen prijs en

nemen wij ook bij deze afhankelijkheden de lineaire vorm aan; b.v.

$$u^1 = b^1 p^1 + c^1,\ u^2 = b^2 p^2 + c^2,\ u^3 = b^3 p^3 + c^3 \qquad (10.5)$$

De grootheden b^1, b^2 ... hangen van twee omstandigheden af. Een bepaalde b^i is groter, naarmate de hoeveelheid, die van dat goed gebruikt wordt, groter is en hij is ook groter, naarmate de vraag naar dat goed meer elastisch is. Zou men de elasticiteit aangeven met ε^1, dan is, zoals wij al eerder hebben gezien (§ 8)

$$\varepsilon^1 = b^1 \frac{p^1}{u^1} \qquad (10.6)$$

waaruit volgt, dat

$$b^1 = \frac{\varepsilon^1\ u^1}{p^1} \qquad (10.7)$$

In plaats van (10.5) kan men daarom ook schrijven, dat

$$u^1 = \frac{\varepsilon^1\ u^1}{p^1} p^1 + c^1,\ u^2 = \frac{\varepsilon^2\ u^2}{p^2} p^2 + c^2 \ldots.$$

Daaruit ziet men, dat de totale vraag

$$u = \frac{\varepsilon^1\ u^1}{p^1} p^1 + \frac{\varepsilon^2\ u^2}{p^2} p^2 + \ldots + c^1 + c^2 + \ldots \qquad (10.8)$$

en dat dus de definitie van het algemene prijsniveau voor dit vraagstuk zo moet zijn, dat daarin de gewichten gelijk zijn aan $\frac{\varepsilon^1\ u^1}{p^1}$, $\frac{\varepsilon^2\ u^2}{p^2}$, enz. Die gewichten zijn daarmee anders, dan men het gewoonlijk in een prijsindexcijfer doet. Daar neemt men gewoonlijk de u^1, u^2 enz. als gewichten. Hier krijgen dus artikelen met een sterk elastische vraag een extra zwaar gewicht. [1])

Zoals gezegd, hiertegen is geen enkel bezwaar, zolang men met slechts één relatie heeft te maken. Wanneer

[1]) Voor een practische toepassing zie men mijn „Business Cycles in the United States of America, 1919—1932" (uitgave v. d. Volkenbond), § 3.4, welke bewerkt is door Dr. J. J. Polak.

nu echter in een andere relatie ook het prijsniveau p een rol moet spelen, kan het gebeuren dat de weging, die daartoe de meest geschikte zou zijn, niet overeenkomt met de voor de eerste relatie nodige weging. Feitelijk zou men dan dus voor de tweede relatie een ander indexcijfer p' nodig hebben, dat met p wellicht niet overeenstemt, en dan is de vereenvoudiging die men beoogde, ten dele illusoir. Wanneer nu echter de bewegingen van de afzonderlijke prijzen niet sterk uiteenlopen, zal het verschil tussen p en p' binnen zeer enge grenzen blijven, waardoor men bij benadering toch $p = p'$ mag stellen. In vele gevallen is dit parallelisme der afzonderlijke prijzen inderdaad aanwezig. In het bijzonder is dat het geval, wanneer de systematische factoren, die de oorzaak van de prijsbeweging zijn, in alle prijzen aanwezig zijn, b.v. het loon als kostenelement of algemene overproductie als oorzaak van prijsdaling. Het historisch verloop van de onderdelen der prijsindices is ook veelal vrij sterk evenwijdig. Dit neemt niet weg, dat het in principe steeds gewenst is om te onderzoeken, of aan de gestelde voorwaarden is voldaan.

Een fraai voorbeeld van de complicaties, die een macroeconomisch begrip kan meebrengen, doet zich ook voor als men de afhankelijkheid, die de vraag vertoont van het *inkomen*, nader beschouwt. [1])

Laten voor een afzonderlijk gezin het inkomen i en de vraag naar consumptiegoederen c zijn; zij verder het totale inkomen van alle gezinnen I en de totale vraag naar consumptiegoederen C. Alleen als c een lineaire functie van i is, zal ook C op eenvoudige wijze van I afhangen; is n.l.

$$c = a + b\,i, \quad a \text{ en } b \text{ constant,}$$

en denkt men zich dit opgeteld voor alle gezinnen, dan is

$$C = a\,N + b\,I,$$

[1]) Het volgende is ontleend aan een artikel van P. de Wolff: Income Elasticity of Demand, a micro-economic and a macro-economic interpretation, Economic Journal 1940.

waarbij N het aantal gezinnen is. Doch zodra c op ingewikkelder wijze van i afhangt, is de som van alle c's niet langer alleen van I afhankelijk. Daardoor is het dan b.v. al een moeilijk te beantwoorden vraag, hoe een vermeerdering van het inkomen I op de vraag naar consumptiegoederen zal inwerken. Een vermeerdering van I kan op zovele verschillende wijzen plaats vinden en het is niet langer onverschillig, hoe dat geschiedt. Nadere berekeningen in het geciteerde artikel tonen aan, dat slechts dan een ondubbelzinnig verband tussen I en C bestaat, wanneer de inkomensverdeling slechts twee „parameters" bevat en dat het dan ook geen zin heeft, om te veronderstellen dat C, behalve van I, ook afhangt van de verdeling der inkomens. Onder „parameters" verstaat men de grootheden, die een frequentieverdeling „beschrijven"; het bekende voorbeeld is dat van de inkomensverdeling volgens de wet van Pareto, waarbij als een der parameters de bekende α optreedt. Ofschoon de finesses van het bewijs moeilijk zonder hogere wiskunde zijn duidelijk te maken, zij de aandacht er op gevestigd als een fraai voorbeeld van wat wiskundige analyse hier vermag te helpen.

Ook in het reeds genoemde geval, dat c van i lineair afhangt, heeft men er een voorbeeld van, dat dan de afhankelijkheid tussen I en C ondubbelzinnig is en dat de verdeling der inkomens op de totale vraag geen invloed uitoefent. Een gelukkige omstandigheid voor het macro-economisch onderzoek is, dat inderdaad binnen wijde grenzen c lineair afhangt van i (vgl. § 25).

§ 11. De beschrijving van volledige stelsels.

Wij gaan ons thans wat meer verdiepen in de bijdrage van de wiskundige economie in wat boven is genoemd het laatste stadium van het onderzoek, d.w.z. nadat de toetsing der hiervoor besproken afzonderlijke relaties heeft plaats gevonden. Om de zin van hetgeen volgt in te zien, moet men zich voor de geest brengen wat het uiteindelijk

doel van alle economische analyse is. Wij menen dit tweeledig te mogen formuleren: a) het historisch verloop van de economische verschijnselen te verklaren uit zijn buiten-economische oorzaken of „data" en b) vast te stellen hoe een wijziging in deze „data" — b.v. een bepaalde vorm van economische politiek — op dat verloop zal inwerken. In het voorgaande hebben wij ons steeds bezig gehouden met één enkele relatie tussen economische verschijnselen. Daarbij werden dan veranderingen in één bepaald verschijnsel teruggebracht tot veranderingen in andere economische verschijnselen en tot bepaalde „data" van natuurlijke, technische of institutionele aard. Wil men het gehele economisch gebeuren echter verklaren, dan moet men niet één dergelijke relatie, doch een aantal tezamen beschouwen, n.l. zovele, dat het aantal te verklaren verschijnselen gelijk is aan het aantal relaties. Eerst dan kan men hopen, de economische verschijnselen werkelijk uit de buiten-economische te verklaren. In meer concrete termen: het is niet voldoende om de vraag naar consumptiegoederen af te leiden uit de inkomens en de prijzen, doch men moet ook deze laatste weer verklaren.

Daartoe moet men het netwerk van causale relaties, die tezamen de economische structuur van een gebied vormen, systematisch onderzoeken. Als voorwerk is daarbij de hulp van enige begripsafbakening niet te vermijden. Wij menen deze het best uit te voeren door bijgaand (zie fig. 1) „pijlenschema" te gebruiken. Elke stip vertegenwoordigt daarbij een gepaald economisch verschijnsel gedurende een bepaalde tijdseenheid, zeggen wij gedurende een bepaald

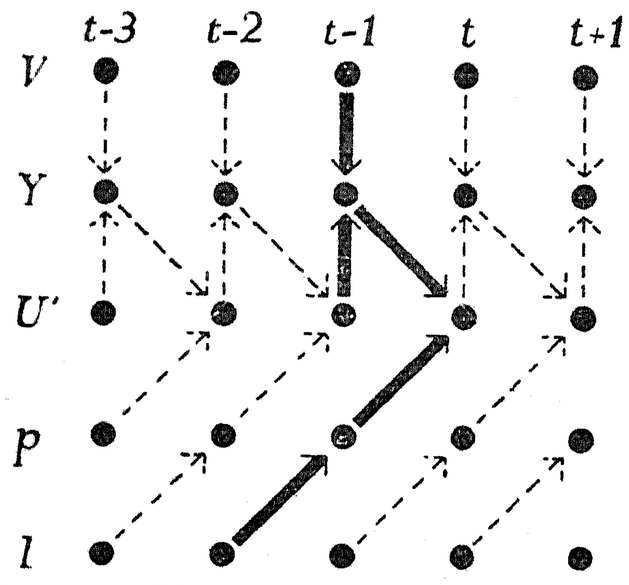

Fig. 1. „Pijlenschema".

kwartaal. Horizontaal is hetzelfde verschijnsel in de opeenvolgende kwartalen gesymboliseerd. Verticaal vindt men voor éénzelfde kwartaal de verschillende verschijnselen, die beschouwd worden. Nu zal de waarde van een bepaald verschijnsel in een bepaald kwartaal, zeggen wij van de vraag naar consumptiegoederen U_t' in het kwartaal t, bepaald worden door de waarden van enkele andere verschijnselen in een ander kwartaal. Om de gedachten te bepalen zeggen wij, dat zij wordt bepaald door de grootte van de inkomens in het voorafgaande kwartaal Y_{t-1} en de prijzen in datzelfde kwartaal p_{t-1}. Wij geven dit aan door de pijlen die van Y_{t-1} en p_{t-1} naar U_t' lopen. Deze pijlen geven een direct causaal verband aan. Zij herhalen zich, wanneer de structuur van de volkshuishouding constant wordt verondersteld, elk kwartaal. De waarden van U', Y' en p behoeven zich geenszins te herhalen — dat zou alleen zo zijn in een stationnaire toestand —, doch de causale relatie herhaalt zich.

Op hun beurt zijn de inkomens Y weer afhankelijk van andere grootheden. Zeggen wij van de waarde van het verbruik U' en van de waarde der investeringen V' in hetzelfde kwartaal. Dat zal dan worden aangegeven door pijlen van U' en V' naar Y, zonder vertraging. De prijzen p en de investeringen V' zullen elk ook weer hun bepalende factoren hebben. Zo ontstaat een geheel schema van zich telkens herhalende pijlen. Door steeds opnieuw te vragen naar de directe oorzaken van de nog niet verklaarde verschijnselen, krijgen wij een groeiend aantal relaties. Aanvankelijk zullen wij ook een groeiend aantal verschijnselen moeten betrekken in onze beschouwingen. Doch op een gegeven moment is dat niet meer nodig: dan treden reeds beschouwde verschijnselen weer als oorzaken op; op dat moment hebben wij de nog onverklaarde verschijnselen alle teruggebracht tot reeds verklaarde. Als het aantal verschijnselen dan gelijk is aan het aantal relaties, hebben wij een volledig systeem. Daarin kunnen dan als „data" wel nog onverklaarde buiten-economische verschijnselen voorkomen, doch het

is niet de taak der economie om deze te verklaren.

Het geheel van alle pijlen is nu de *logische structuur* van ons stelsel. Die pijlen kunnen op tweeërlei wijze worden gerangschikt. Men kan telkens alle pijlen tezamen beschouwen die in één punt uitkomen; dan beziet men tegelijkertijd de verschillende *oorzaken* van één bepaald verschijnsel, door dat punt gekarakteriseerd. Men kan ook alle pijlen tezamen nemen, die van één punt uitgaan; dan bekijkt men de *werking* van één gegeven verschijnsel op de andere. Deze laatste rangschikking is a.h.w. voorwaarts gericht; in de meeste verbale betogen wordt zij gebruikt. De eerste rangschikking is degene, die met onze vergelijkingen correspondeert: de verschillende termen daarvan geven de verschillende pijlen aan, die in de te verklaren grootheid hun eindpunt hebben. Voor beide wijzen van rangschikking geldt natuurlijk, dat achtereenvolgens elk punt in één verticale rij moet worden beschouwd, d.w.z. alle verschijnselen moeten worden beschouwd. De gehele aldus verkregen figuur herhaalt zich dan bovendien in de tijd, zoals reeds werd opgemerkt.

Bij de beide genoemde rangschikkingen wordt uitsluitend met directe causale relaties gewerkt. Dat is ook voldoende, want elke pijl krijgt zijn beurt. Men denke echter niet, dat de *indirecte* — de „diepere" — oorzaken daarmee dan verwaarloosd zijn. Doch elke directe oorzaak is tegelijkertijd ook indirecte oorzaak. Als p directe invloed heeft op U' en het loonpeil l heeft directe invloed op p, dan heeft l indirecte invloed op U'. Teruglopend langs de pijl die in U' uitkomt van uit p, komt men eerst in p en kan vervolgens verder teruglopen naar l, enz. Men komt dan de tweede, derde, enz. oorzaken te weten. Het stelsel is dus volkomen gedefinieerd door de door ons ingevoerde vergelijkingen, die de „elementaire wetten" weergeven, die de variabelen aan elkaar binden. Wij zullen ook spreken van de *elementaire vergelijkingen.*

Op de gedaante van deze vergelijkingen moeten wij ook nog iets nader ingaan: wat voor soort van grootheden komen daarin voor? Nemen we als voorbeeld nogmaals

de vergelijking van de vraag naar consumptiegoederen:

$$U'_t = v_1 Y_{t-1} + v_2 p_{t-1} + v_3 (t) + v_4 + R_t^{U'}$$

Wij hebben nog enige termen toegevoegd, die zo dadelijk zullen worden besproken. De volgende grootheden komen nu voor:

1. de *te verklaren variabele* U' (vraag naar consumptiegoederen);
2. de *verklarende variabelen* Y' en p (inkomen en prijsniveau);
3. een aantal *constante* grootheden: coëfficiënten v_1, v_2, v_4 en lags (van één tijdseenheid in dit geval), die tezamen de „*structuur*" van het stelsel bepalen (over deze term zullen wij nog nader spreken);
4. een van de tijd afhankelijke term v_3 (t), die de verandering van *systematisch* veranderende *data* weerspiegelt, zoals veranderingen in techniek, seizoensinvloeden of invloeden van het buitenland uitgaande;
5. een nog niet eerder besproken term $R_t^{U'}$, die de onsystematische veranderingen in *data* representeert, die zich kunnen voordoen. Hierbij heeft men te denken aan een aantal elk op zich zelf kleine invloeden die bij de vereenvoudiging zijn weggelaten en aan de invloed van de minder belangrijke „data", die wij gewoonlijk als van toevallige aard beschouwen. De vraag naar consumptiegoederen kan lager zijn, omdat er toevallig meer mensen ziek waren of omdat er meer regen is gevallen dan gewoonlijk, enz. Dergelijke grootheden hebben een onregelmatige invloed, maar bovendien een kleine en ze interesseren ons als regel niet. Er kunnen een enkele keer echter ook iets belangrijker invloeden in opgenomen worden, zoals b.v. de oogstopbrengsten en eventuele technische verbeteringen van grote allure of de invloed van een staking, een politieke gebeurtenis van middelgrote betekenis, e.d.

Wij hebben het woord *structuur* in het bijzonder onder 3) gebezigd. Inderdaad zal men met de structuur van een volkshuishouding in het bijzonder bedoelen de „blijvende"

kenmerken ervan. Toch is er ook veel voor te zeggen om ook de onder 4) genoemde, systematisch veranderende, gegeven grootheden daaronder te begrijpen, in het bijzonder de natuurlijke, technische en institutionele grootheden, die zich daaronder bevinden. Bij een „open" volkshuishouding heeft men onder 4) echter ook gegeven grootheden, die men voor de economische wetenschap als geheel allicht niet als data beschouwt en daarom misschien ook niet als structuurgrootheden beschouwt.

In ons voorbeeld zijn de v_1 en v_2 als structuurconstanten beschouwd, d.w.z. de coëfficiënten waarmee resp. het inkomen en het prijsniveau de consumptie beïnvloeden. Men kan hier, zoals reeds werd opgemerkt, ook dieper graven en als structuurgrootheden beschouwen de „daarachter liggende" grootheden, als b.v. de coëfficiënten in de nuttigheidsfuncties en de technische coëfficiënten. Het is bij wijze van stylering, dat de coëfficiënten van de vraagcurve zelf als data zijn beschouwd; een stylering, te vergelijken met Cassel's werkwijze.

§ 12. De bewegingen van volledige stelsels.

De elementaire vergelijkingen tezamen stellen ons nu in staat om de bewegingen van alle economische verschijnselen te leren kennen, als wij maar kennen:

(a) de structuur van het stelsel, belichaamd in de gegeven coëfficiënten, lags en trends;
(b) de waarden van de overige systematisch veranderende data voor elke tijdsperiode;
(c) de waarden van de onsystematisch veranderende data R, en
(d) bepaalde uitgangswaarden van de economische variabelen; welke, dat kan slechts in elk concreet geval worden vastgesteld.

Inderdaad kan men, wanneer men over deze gegevens beschikt, het verloop steeds voortzetten door berekening in getallen. Wij willen dit aan een zeer eenvoudig voorbeeld demonstreren, dat op een enkele markt betrekking

heeft, die door zijn isolement een stelsel op zichzelf vormt. Zij de vraag u_t alleen afhankelijk van de prijs p_t en wel zonder vertraging:

$$u_t = -p_t + 200 + t \qquad (12.1)$$

Zij voorts het aanbod ook alleen van de prijs afhankelijk, maar met een vertraging van 3 tijdseenheden:

$$u_t = p_{t-3} \qquad (12.2)$$

Wij hebben in de vraagcurve nog een langzaam stijgende term aangenomen, die door de stijging van de bevolking veroorzaakt wordt. De positieve constante 200 is nodig, omdat anders voor elke positieve waarde van de prijs p op het tijdstip 0 een negatieve vraag zou optreden, hetgeen onaannemelijk is.

Zoals gezegd, dienen wij nog te beschikken over enige initiale waarden van een der grootheden om het verdere verloop te kunnen berekenen. In dit geval is het gemakkelijk in te zien, dat wij daartoe moeten kennen het verloop van een der grootheden voor de eerste drie perioden, zeggen wij dus voor $t = 1$, 2 en 3. Wij nemen aan, dat wij u_1, u_2 en u_3 kennen, en dat ze resp. bedragen:

$$u_1 = 100 \quad u_2 = 120 \quad u_3 = 130$$

Inderdaad kunnen wij nu met behulp van (12.1) de p berekenen voor $t = 1$, 2 en 3 en met behulp daarvan uit (12.2) de waarden van u voor $t = 4$, 5 en 6; vervolgens weer met (12.1) de p voor $t = 4$, 5 en 6, enz. Het resultaat van deze *numerieke extrapolatie* is:

$t =$	1	2	3	4	5	6	7	8	9	10	11	12
$u_t =$	100	120	130	101	82	73	103	123	133	104	85	76
$p_t =$	101	82	73	103	123	133	104	85	76	106	126	136

Op de eigenaardigheden van deze beweging komen wij nog terug. De berekeningen boden hier geen bijzondere moeilijkheden, doch de methode biedt in het algemeen twee nadelen. Ten eerste moet men, om de beweging van een der grootheden te leren kennen, ook steeds die van de andere er bij berekenen. Bij stelsels met een groot

aantal variabelen is dat een aanzienlijk werk. Ten tweede leert men op deze wijze het type van de beweging en zijn afhankelijkheid van de gemaakte veronderstellingen niet dan zeer moeizaam, n.l. eigenlijk slechts experimenteel, kennen. De wiskunde doet ons middelen aan de hand om beide bezwaren te overwinnen.

Om alle niet-gewenste variabelen buiten beschouwing te kunnen laten, moeten en kunnen wij ze *elimineren*, een uit de lagere algebra bekende bewerking. In het eenvoudige geval van ons voorbeeld hebben wij daartoe niets anders te doen dan (12.2) in te vullen in (12.1):

$$p_{t-3} = -p_t + 200 + t \qquad (12.3)$$

Dit betekent, dat wij van de beschouwing van „eerste oorzaken" overgaan op de studie van „tweede oorzaken"; of, anders gezegd, dat wij van de elementaire economische wetten overgaan op „*afgeleide economische wetten*". Er is een hele hiërarchie van dergelijke afgeleide wetten, n.l. afgeleide wetten, die men krijgt door het combineren van 2, van 3, enz. elementaire vergelijkingen. Het is niet gemakkelijk om over deze wetten in woorden te spreken — in plaats van in formules — omdat de meeste hunner geen afzonderlijke namen hebben gekregen. Daarom is deze methode wel eens gedisqualificeerd als „nachttrein-analyse". Het enige wat daaraan te doen is, is, dat men het stelsel zo eenvoudig mogelijk kiest, maar van zelfsprekend lijdt daaronder dan weer het werkelijkheidsgehalte.

Het einde van dit eliminatieproces is bereikt, wanneer er één vergelijking over is met maar één variabele. Niet altijd is het mogelijk om zover te komen; doch met lineaire vergelijkingen wel en met een aantal van ander type ook. In deze „*eindvergelijking*" komt de enig overgebleven variabele in 't algemeen voor verschillende perioden voor; zij is dus *dynamisch*. Haar coëfficiënten zijn combinaties van de coëfficiënten der elementaire vergelijkingen. Bovendien zullen in deze eindvergelijking ook allerhande toevallige storingen uit de elementaire

vergelijkingen voorkomen. In ons voorbeeld hebben wij die verwaarloosd, doch men kan het gemakkelijk in die richting completeren.

Van de eindvergelijking geldt evenals van het stelsel als geheel dat men, indien men enige beginwaarden van de variabele kent, het verder verloop kan berekenen door numeriek extrapoleren. Wij willen thans echter de daarbij te verwachten uitkomsten meer systematisch gaan beschouwen. Daartoe schrijven wij de vergelijking (12.3) eerst nog wat anders:

$$p_t = -p_{t-3} + 200 + t \qquad (12.4)$$

Afgezien van de hier verwaarloosde toevallige storingstermen zijn er aan de rechterzijde nu drie soorten termen.

De eerste soort termen, waarvoor wij de aandacht vragen, zijn de constante; in ons voorbeeld de term 200. Indien deze er alleen stond, zou ook de waarde van de variabele p_t constant zijn. D.w.z. er zou een stationnaire toestand heersen, en wij zouden met de *statica* in de engere zin van het woord te maken hebben. Deze constante waarde van p_t zou afhangen van de constante additieve termen in de elementaire vergelijkingen en van de constante coëfficiënten (b.v. elasticiteitscoëfficiënten e.d.). Het zou wat wij gewoonlijk noemen de *evenwichtsstand* van p zijn. Indien er ook nog toevallige afwijkingen R in de eindvergelijking (12.4) voorkomen, zal p toevallige afwijkingen vertonen, die zich echter om deze evenwichtswaarde heen scharen.

De tweede soort termen zijn die, *waarin de tijd op „expliciete" wijze voorkomt*; die het gevolg zijn van de systematisch veranderende „data", zoals toenemende bevolking en technische kennis of seizoenschommelingen of ook wijzigingen in het buitenland. Indien deze — al dan niet in vereniging met de eerste soort — voorkomen, zal de variabele p niet constant meer kunnen zijn — een statisch evenwicht is dan onmogelijk —, doch zich in de tijd veranderen en daardoor systematische bewegingen vertonen. Deze bewegingen geschieden *gelijktijdig met*

de bewegingen in de data, die er de oorzaak van zijn. Het stelsel past zich n.l. wegens het ontbreken van vertragingstermen als p_{t-3} onmiddellijk aan aan de veranderingen in de data. Als deze veranderingen in de data er niet waren, zouden geen bewegingen mogelijk zijn, d.w.z. *endogene bewegingen zijn dan niet mogelijk*. Dit stelsel is niet statisch in de beperktere zin; het is echter wel statisch volgens onze definitie en het wordt ook wel *comparatief-statisch* genoemd. Men kan nl. dezelfde formules ook gebruiken om twee stationnaire toestanden met elkaar te vergelijken, waarin de data verschillende waarden hebben.

Beschouwen wij tenslotte de invloed van de derde soort termen, nl. de termen waarin de variabele voorkomt voor andere tijdstippen. Denken wij ons eerst dat deze termen — in ons voorbeeld vertegenwoordigd door de term p_{t-3} — alleen voorkomen, met of zonder de constante termen. Wij kunnen dan ten eerste vaststellen, dat er in dat geval bewegingen mogelijk zijn, zonder dat de data zich wijzigen, d.w.z. *endogene bewegingen*. Daartoe is echter de aanwezigheid dezer derde soort van termen noodzakelijk, d.w.z. dat het stelsel dynamisch moet zijn.

Wiskundig is nu aan te tonen, dat de bewegingen die voortspruiten uit de aanwezigheid der termen met de variabele voor verschillende tijdstippen een speciaal type hebben. Zij bestaan uit een of meer van de volgende bewegingen, die in verschillende relatieve sterkte kunnen optreden [1]):

1) gedempte, ongedempte of anti-gedempte periodieke bewegingen,
2) eenzijdige bewegingen die steeds dichter tot nul naderen (convergente eenzijdige bewegingen),
3) eenzijdige bewegingen die steeds verder van de begintoestand wegvoeren.

Deze eenzijdige bewegingen zijn daarbij van bepaalde

---

1) Voor exacte bewijzen kan verwezen worden naar „Einige Grundfragen der mathematischen Konjunkturtheorie", Archiv für mathematische Wirtschafts- und Sozialforschung III (1937), blz. 1 en 83.

wiskundige typen, doch daarop willen wij hier niet nader ingaan. De periodieke bewegingen kunnen nog verschillende *perioden* hebben. Deze perioden zijn echter bepaald door de structuurconstanten; het is er mee als met de z.g.n. eigentrillingen van allerlei mechanische stelsels, waarvan de perioden ook bepaald zijn door de massa's, hun afmetingen, elasticiteiten, buigingsweerstanden, e.d.

In ons voorbeeld is gemakkelijk in te zien, dat de bedoelde beweging een zuiver periodieke beweging is (ongedempt dus), met een periode van 6 tijdseenheden. De vergelijking wordt toch immers onder weglating der laatste twee termen:

$$p_t = -p_{t-3}$$

Daaruit volgt:

$$p_{t+3} = -p_t = p_{t-3},$$

dus dat de waarde van p_{t-3} gelijk is aan die van p_{t+3}. Dat geldt voor elke willekeurige t, dus voor twee tijdsperioden, die 6 eenheden uiteenliggen, is de p gelijk. Zou men in plaats van $p_t = -p_{t-3}$ gehad hebben

$$p_t = -\alpha\, p_{t-3}$$

en dus

$$p_{t+3} = \alpha^2\, p_{t-3},$$

dan zou bij $\alpha < 1$ een gedempte en bij $\alpha > 1$ een antigedempte beweging zijn ontstaan, zoals men door proberen gemakkelijk zal zien. In dit eenvoudige geval hangt de periode slechts af van één der structuurgegevens, n.l. de „lag" van 3 tijdseenheden, die b.v. de duur van een productieproces kan zijn. In de meeste gevallen zal de periode van andere structuurgegevens tevens afhangen en zal het verband ook minder eenvoudig zijn.

De *amplitude* van de schommelingen — d.i. de grootte der afwijkingen die maximaal optreden — hangt echter af van de afwijking van het evenwicht, die in het begin bestaat; iets dat dus een erfenis van de geschiedenis kan genoemd worden, ofschoon dat niet altijd een geschiedenis van lang geleden behoeft te zijn. In 't algemeen zullen hier

van invloed zijn de intensiteiten waarmee de toevallige afwijkingen — dus b.v. de gevolgen van oogstwisselingen, eventueel oorlogen, e.d. — zich doen gevoelen.

Wat gebeurt er nu, wanneer alle termen tegelijk aanwezig zijn? In het geval, dat wij met een lineaire vergelijking te maken hebben, en men kan dat in eerste benadering vaak aannemen, laat zich bewijzen dat de beweging dan de *som* is van:

1) een evenwichtsniveau,
2) een endogene beweging als bovenbeschreven en
3) door de verandering der data veroorzaakte „exogene" bewegingen, die dezelfde periode hebben als de bewegingen der data, voorzover ze periodiek zijn, en die een rechtlijnige groei in de tijd vertonen, wanneer de data een rechtlijnige groei vertonen.

Daarbij komen dan eventueel nog de toevallige afwijkingen, die op een bepaalde wijze gecumuleerd worden. Dit laatste moge nog even nader worden toegelicht. Denken wij ons daartoe eenvoudigheidshalve alleen de eerste term van de rechterzijde van (12.4) en een toevallige afwijking aanwezig, dan hebben wij:

$$p_t = -p_{t-3} + R''_t$$

Denken wij ons p_{t-3} gegeven, dan volgt p_t uit deze vergelijking. Voor p_{t+3} hebben wij:

$$p_{t+3} = -p_t + R''_{t+3} = p_{t-3} - R''_t + R''_{t+3}$$

Hierin is dus reeds de gezamenlijke invloed van twee toevallige afwijkingen te bemerken. Voor p_{t+6} zouden wij hebben:

$$p_{t+6} = -p_{t-3} + R''_t - R''_{t+3} + R''_{t+6}$$

enzovoort.

Overzien wij het resultaat, dan blijkt dus, dat de bewegingen, die de economische verschijnselen vertonen, onder de invloed van de veranderingen in de data en van het dynamisch karakter der relaties de volgende componenten vertonen:

a) eenzijdige bewegingen als gevolg van eenzijdige v
anderingen in de data, b.v. bevolkingsgroei;
b) eenzijdige bewegingen als gevolg van de inter
dynamica, b.v. de kapitaalvorming;
c) periodieke bewegingen als gevolg van periodie
bewegingen in de data, b.v. buitenlandsche cc
juncturele invloeden en seizoenschommelingen;
d) periodieke bewegingen als gevolg van de inter
dynamica, b.v. de conjunctuurbeweging;
e) toevallige bewegingen als gevolg van toevalli
(onregelmatige) veranderingen in de data, b.v. oog
wijzigingen en vele kleine structuurwijzigingen.

§ 13. Stabiele en labiele evenwichten. Het doel v de economische politiek.

Wij hebben reeds gezien, dat de economische stels
waarmede wij ons in het voorgaande hebben bezig
houden geen statisch evenwicht kennen. Niettemin hoc
men als doel van de economische politiek, inzonderhe
van de conjunctuurpolitiek vaak formuleren: een me
evenwichtige ontwikkeling van het economisch lev
Daaronder hebben wij in het licht van het voorgaan
te verstaan een ontwikkeling waarbij de bewegingen onc
c), d) en *e)* zo veel mogelijk worden beperkt. Doel v
het econometrisch onderzoek in deze is dan ook, va
te stellen of door de maatregelen die van verschillen
zijden worden voorgesteld een dergelijk doel kan word
bereikt. Het onderzoek staat hier eerst in zijn begi
stadium; enige voorlopige resultaten zullen in § 40 word
medegedeeld. In dit methodische hoofdstuk is het op z
plaats de begrippen *evenwicht* en *stabiliteit* en *labilit*
van evenwichten kort aan te duiden. Verstond men onc
het evenwicht in de beperkte statische zin een toesta
die gehandhaafd kon blijven, voor onze stelsels heeft e
zodanige definitie weinig waarde, aangezien zij een onk
staanbare toestand eist. Meer zin heeft de definitie v
evenwichtsontwikkeling, die deze als een zo geleidel

mogelijke beweging aangeeft, of als men het nauwkeuriger wil: de beweging, onder alle mogelijke bewegingen van een stelsel, die zo dicht mogelijk aansluit bij een rechte lijn. Uit het voorgaande laat zich afleiden, dat dit voor een gegeven stelsel heel bepaalde bewegingen zijn en wel de componenten *a)* en *b)* der aan het einde van § 12 genoemde reeks. De in de werkelijkheid optredende bewegingen, die eveneens de componenten *c), d)* en *e)* bevatten, zijn schommelingen daaromheen.

Als deze schommelingen gedempt zijn, kunnen wij de evenwichtsontwikkeling *stabiel* noemen. Zijn zij ongedempt of anti-gedempt, dan noemen wij die evenwichtsontwikkeling *labiel.* De taak van de conjunctuurpolitiek is om de structuur van de volkshuishouding — c.q. wereldhuishouding — aldus te wijzigen, dat de stabiliteit zo groot mogelijk worde, dus de *dempingsgraad zo groot mogelijk zij.* In iets meer alledaagse, maar wellicht aanschouwelijker termen uitgedrukt: men moet leren om de schommelingen in de kiem te smoren en liefst niet incidenteel, maar systematisch, d.w.z. door verandering van de structuur. Een dergelijke structuurverandering zal men wellicht moeten zoeken in de financiële en monetaire overheidspolitiek, doch mogelijk ook elders. Hier heeft de econometrist geen stem: hij heeft de hem voorgezette problemen te onderzoeken.

Men kan dit probleem echter niet losmaken van de andere doelstellingen der algemene economische politiek, nl. de eis een zo hoog mogelijk gemiddeld welvaartsniveau te bereiken. Indien zou blijken, dat een groter stabiliteit alleen ten koste van een lager gemiddeld welvaartspeil zou zijn te bereiken, zou een keus moeten worden gedaan.

§ 14. Statistische toetsing; meting van een verschijnsel.

Wij merkten reeds op, dat bij het econometrisch onderzoek begonnen moet worden met het opstellen — met behulp van wiskundig-economische methoden — van een

of meer relaties tussen de te onderzoeken grootheden. Indien dit is geschied — wij spraken van het eerste stadium van het onderzoek — is het woord aan den statisticus. Hij zal trachten na te gaan of het veronderstelde verband met de feiten overeenstemt en zo ja, of dan metingen van bepaalde grootheden het mogelijk maken om concreter inlichtingen omtrent dat verband te verkrijgen.

Een eerste vraag waarvoor ons de taak van den statisticus stelt is, of alle verschijnselen waarover de econoom handelt, dan *meetbaar* zijn. Spelen niet, zo hoort men vaak vragen, zovele imponderabilia een rol in het economisch gebeuren, dat het principieel onjuist is te verwachten, dat hier door meting de werkelijkheid kan worden bloot gelegd?

Deze vraag is niet nieuw. Eerder dan voor de economie is zij voor de natuurwetenschappen gesteld. Men heeft de natuurkundigen verweten, dat zij over warmte als over iets meetbaars spraken. Ofschoon dit pleit naar vrijwel aller overtuiging in zoverre door de „meters" is gewonnen, dat zij door hun meetlust tot een ontzaglijke ontwikkeling van de warmteleer hebben bijgedragen en ofschoon ook de hedendaagse psychologen niet terugschrikken voor het meten van vele „imponderabilia", willen wij niet met analogieredeneringen werken, doch liever rechtstreeks op de vraag ingaan. De voornaamste problemen van de economie hebben betrekking op meetbare verschijnselen. Men spreekt van productie, van werkloosheid, reële lonen e.d. Voorzover het nu gaat om de invloed van „imponderabilia" op deze meetbare verschijnselen moet men onderscheid maken tussen deze imponderabilia zelf en hun invloed op de meetbare verschijnselen. Al zijn verwachtingen zelf misschien onmeetbaar, hun invloed op de investeringen kan niet anders zijn dan een zekere verhoging of verlaging van die investeringen, dus een meetbaar verschijnsel. Alleen dit laatste interesseert ons tenslotte. Daarbij komt nog, dat de verwachtingen zelf gewoonlijk gebaseerd zijn op zeer concrete, evenzeer meetbare, andere verschijnselen. Het ligt daarom voor de hand

om, over het eventueel onmeetbare van de „verwachtingen" heen, te trachten de invloed dier verschijnselen op de investeringen te onderzoeken.

Een ander voorbeeld heeft betrekking op de „bevredigingsgraad". Deze wordt in 't algemeen voor onmeetbaar gehouden. Doch het blijkt bij nader onderzoek, dat die bevredigingsgraad zelf ook niet gekend behoeft te worden, om het gedrag der economische subjecten te verklaren. Slechts de goederencombinaties die eenzelfde bevredigingsgraad veroorzaken — alle gelegen op de „indifferentieoppervlakken" — behoeven te worden gekend en zij zijn in principe meetbaar.

Wij geloven dus dat men zich omtrent deze kwestie geen grote zorgen behoeft te maken.

Een soortgelijke vraag is: zijn psychologische reacties wel zo *wetmatig*, dat men ze in wiskundige vergelijkingen kan uitdrukken? Vaak wordt daaraan twijfel uitgesproken; men wijst op de vrije wilsbeschikking, die in dezelfde omstandigheden de ene maal kan voeren tot een andere reactie dan de andere maal. Daarop wordt van andere zijde geantwoord, dat wanneer dan die andere maal een andere reactie volgt, daarvoor toch allicht een reden is geweest en men dus bij volledige omschrijving der omstandigheden van verschillende situaties moet spreken. Bovendien — en dat is voor macro-economisch onderzoek in het bijzonder van belang — is er een grote kans, dat het gezamenlijk gedrag van grote groepen personen een regelmatiger beeld vertoont dan het gedrag van afzonderlijke personen, omdat de individuele afwijkingen elkaar grotendeels opheffen. Dit wordt in veel gevallen door de waarneming ook bevestigd; zoals uit enkele onzer voorbeelden zal blijken, is er vaak een verrassend grote regelmatigheid in de reacties.

De eenvoudigste vorm van meting waarmee men in de statistiek te maken heeft is de meting van een afzonderlijk verschijnsel: het vaststellen van een prijs, van

de verbruikte hoeveelheid; van het nationale inkomen, van het gedeelte daarvan dat aan de arbeiders toevalt. Van econometrisch standpunt is hierbij niet veel toe te voegen aan wat de ambtelijke statistiek, de statistische schattingsmethoden en de theoretische statistiek op dit terrein sinds lang hebben gepresteerd. Van betekenis is vooral, dat men zich een indruk verschaft van de onzekerheidsmarge in de meetresultaten. Dit kan men veelal bereiken door verschillende alternatieve waarnemingen of berekeningen met elkaar te vergelijken; gaat het om een gemiddelde prijs, dan bestudere men de spreiding in de samenstellende prijsgegevens en passe de formule toe voor de standaarddeviatie van een gemiddelde. Soortgelijke, zij het grovere, methoden kunnen toegepast worden in andere gevallen, waarvan minder waarnemingen beschikbaar zijn.

De *standaarddeviatie* σ van een aantal cijfers berekent men door voor elk hunner de afwijking van het gemiddelde vast te stellen, deze afwijkingen in het kwadraat te brengen, de kwadraten op te tellen, deze som te delen door het aantal en uit het aldus verkregen gemiddelde van de kwadraten der afwijkingen de wortel te trekken. In formule:

$$\sigma = \sqrt{x_1^2 + x_2^2 + \ldots + x_n^2},$$

als $x_1 \ldots x_n$ de afwijkingen van het gemiddelde zijn.

De eigenlijke econometrische problemen beginnen, wanneer het gaat om gehele reeksen van waarnemingen, in het bijzonder reeksen die het verloop van een bepaalde grootheid in de tijd weergeven of zgn. *historische reeksen*. Naar Duits en Engels voorbeeld spreekt men ook wel van *tijdreeksen*.

§ 15. Vaststelling van de componenten van historische reeksen; algemene opmerkingen.

Slechts zelden worden alle bijzonderheden van het verloop ener historische reeks tegelijkertijd tot onderwerp

van econometrische studie gemaakt. Meestal onderscheidt men verschillende componenten in dit verloop; de traditionele indeling van de conjunctuuronderzoekers is een indeling in vieren, n.l.:

a) de trend of algemene bewegingsrichting van de reeks, resp. de haar voorstellende curve;
b) de conjunctuurschommelingen;
c) de seizoenschommelingen, dat zijn de schommelingen met een periode van meestal één jaar, die het gevolg zijn van de natuurlijke seizoenswisseling of van conventies, gebonden aan de kalender en
d) de toevallige en incidentele veranderingen.

Het eigenlijke indelingsbeginsel dat hieraan ten grondslag ligt is gelegen in de lengte der golven. Men zou kunnen zeggen, dat de golven van langer dan 15 jaar en de eenzijdige bewegingen (die men een oneindig grote golflengte zou kunnen toekennen) tot a), die van 3—15 jaar tot b), die van 1 jaar tot c) en de nog kortere tot d) behoren. Daarop is dan echter aanstonds het een en ander af te dingen.

Door de onregelmatige vorm van de meeste golven is de lengte daarvan vaak niet met zekerheid vast te stellen of wisselend. Sommige golven zijn oneigenlijke golven, althans naar de mening van sommige schrijvers. De z.g.n. „lange golven van de conjunctuur", die een periode van omstreeks 40 jaar hebben, zijn volgens sommigen slechts min of meer toevallige afwisselingen van op- en neergaande bewegingen. Ook de groep d) vertoont wel op- en neergangen, maar van een werkelijke periodiciteit of een neiging daartoe kan men niet goed spreken. De *eenmaal* optredende veranderingen (*structuurbreuken*, zoals men ook wel zegt) zijn moeilijk onder te brengen. Voorts hebben golven van verschillende lengte niet altijd verschillende oorzaken, waardoor het dan niet logisch is, ze te scheiden.

Voorts is op de concrete indeling — d.w.z. op de keuze der grenzen — het een en ander af te dingen. Er zijn schommelingen, die in hun ontstaan volkomen analoog

zijn aan de seizoenschommelingen, doch een veel korter periode hebben, n.l. een maand, een week of een dag. Voorbeelden zijn de schommelingen in de omzetten van de kleinhandel, in de electriciteitsproductie en het verkeer. Zij behoren dus tot groep c). Aan de andere kant zijn er toevallige schommelingen met een veel langere „periode" dan 1 jaar. De afwisselingen in de oogstopbrengsten zijn een zeer belangrijk voorbeeld van een toevallige invloed. Zou men deze een periode willen toekennen, dan zou men wellicht aan een tijdvak van drie jaar kunnen denken. [1]) Toch moeten zij bepaald tot de categorie d) worden gerekend. Wagemann komt aan deze bezwaren slechts ten dele tegemoet door zijn overigens goed gevonden onderscheiding tussen „vrije" en „gebonden" schommelingen.

Er is veel verwarring in het gebruik van de woorden *lange* en *korte* conjunctuurgolven. Onder lange conjunctuurgolven verstaat men soms de zoëven reeds genoemde golven van 40 jaar lengte, soms ook de ongeveer 9-jarige golven; onder de korte soms deze laatste en soms de in de Verenigde Staten optredende ongeveer 3-jarige golven. Schumpeter heeft voorgesteld om, ter voorkoming van deze moeilijkheid, de drie genoemde soorten te noemen naar hun ontdekkers en ze te noemen *kondratiëfs*, *juglars* en *kitchins*. Ten aanzien van de conjunctuurschommelingen is er nog de verdere moeilijkheid, dat men het over de oorzaken daarvan nog niet geheel eens is en daarom een causaal-gefundeerde definitie moeilijk kan geven. Gewoonlijk denkt men daarbij aan door endogene oorzaken opgewekte bewegingen. Voor bepaalde afzonderlijke markten bestaan er zodoende ook nog „conjunctuurgolven", die in lengte afwijken van de gewone, b.v. een golf van de huizenconjunctuur in de Ver. Staten (15 jaar), een golf op de koffiemarkt (ook ongeveer 15 jaar), een op de varkensmarkt (3 jaar), enz.

Bij de *seizoenschommelingen* is er onderscheid te maken

---

1) Een toevallige reeks vertoont n.l. gemiddeld eenmaal per drie keer een top-

tussen de normale of *gemiddelde* seizoenbeweging, die als door nauwkeurige wetten bepaald wordt beschouwd — b.v. zuiver periodiek gedacht wordt — en de aan de werkelijk in een bepaald jaar aan seizoenswisselingen toe te schrijven schommelingen. In een zeer koude winter zal b.v. het scheepvaartverkeer een dieper dal vertonen dan normaal. Dit dal behoort dan tot de individuele seizoensinvloed van het betrokken jaar, maar niet tot de normale seizoenbeweging.

Deze opmerkingen geven een aantal feiten weer, waarvan de kennis voor de econometrist nuttig kan zijn. Zij kunnen hem tevens behoeden voor mechanische toepassing van sommige der rekenschema's, die voor het vaststellen der componenten zijn opgesteld. Elk geval worde afzonderlijk bekeken. Met al deze reserves rekening houdend, menen wij toch, dat de indeling in componenten fundamenteel gezond is; in het bijzonder het afzonderen van de seizoenschommelingen, van de korte toevalsschommelingen en van een theoretisch verantwoorde trend, omdat afzonderlijke oorzakencomplexen voor hun aanwezigheid kunnen worden verantwoordelijk gesteld. Wij menen ook, dat wij in § 12 een, zij het dan voorwaardelijke, fundering van de methode hebben gegeven, omdat wij daar hebben aangetoond, dat de totale beweging van een economische variabele in een aantal additieve componenten uiteenvalt.

In het volgende zullen wij op de berekening der afzonderlijke componenten nog iets verder ingaan; het is echter niet de bedoeling de eigenlijke rekenrecepten uitvoerig te bespreken en aan voorbeelden te illustreren. Daarvoor verwijzen wij naar de literatuur.

Voorzover de onderscheiding naar de lengte der perioden zin heeft — en in veel gevallen is dat zo — is er een algemene methode mogelijk ter bepaling van alle componenten, n.l. die van de *voortschrijdende gemiddelden* (*moving averages*). Men schakelt b.v. de seizoenschommelingen en alle kortere schommelingen uit, door in plaats van een reeks van maandcijfers te nemen de reeks

der gemiddelden van de termen N°. 1—12, N°. 2—13, N°. 3—14, enz. Deze reeks heet een 12-maands-voortschrijdend gemiddelde. Elk harer termen is opgebouwd uit cijfers van alle seizoenen en daarom is in deze reeks geen seizoenwisseling meer aanwezig: de seizoenschommelingen zijn uitgeschakeld. Tegelijkertijd zijn alle kortere schommelingen ook (nagenoeg) uitgeschakeld; volledig weliswaar alleen de schommelingen waarvan de periode een geheel aantal keren op 12 maanden is begrepen, doch bij benadering alle andere in de praktijk optredende ook. Wil men de kortste der uitgeschakelde schommelingen — nemen wij aan dat deze een periode van 3 maanden hebben — zelf leren kennen, dan bepale men de 3-maandsvoortschrijdende gemiddelden en trekke deze af van de oorspronkelijke reeks: men houdt de kortste schommelingen over. Indien de eerstvolgende periode die van 12 maanden is, dan kan men de daarmee corresponderende component vinden, door van de oorspronkelijke reeks af te trekken zowel de zojuist bepaalde kortste schommelingen als het voortschrijdend twaalfmaandsgemiddelde. Op deze en soortgelijke wijzen is elk der componenten te berekenen. Deze methode is overigens zeer grof. En voor de seizoenbeweging doet zij ons niet kennen de gemiddelde of normale seizoenbeweging. Daartoe kan men dan weer de gemiddelden van alle Januari-cijfers, van alle Februari-cijfers, enz. van de boven bepaalde seizoencomponent nemen. Er bestaan fijnere methoden. Daarover willen wij thans iets naders mededelen.

§ 16. Fijnere methoden; de trend.

Beschouwen we eerst de bepaling van de *trend* nader. Indien de golflengte van de langste schommelingen, die men niet meer tot de trend wil rekenen, b.v. 9 jaar bedraagt, zal men deze en de korte schommelingen elimineren — en dus de trendreeks overhouden — indien men voortschrijdende 9-jaarsgemiddelden bepaalt. Als men dit doet, maakt men dus gebruik van de eigenschap

der kortere schommelingen, dat de som van hun uitwijkingen, over een of meer perioden genomen, nul is. Deze eigenschap geldt strikt genomen alleen voor zuiver periodieke bewegingen, die niet gestoord zijn. Zowel voor gedempte als antigedempte bewegingen als voor gestoorde bewegingen geldt zij niet nauwkeurig. Men kan, om deze onnauwkeurigheid te ontlopen, nu ook trachten om de berekening te baseren op een eigenschap van de trend zelf. Doet men dat, dan spreekt men van een *wiskundige trend,* in tegenstelling tot de zojuist besproken methode.

Bij de berekening van de *wiskundige trend* neemt men a priori aan, dat de trend een wiskundige figuur, een curve, van gegeven vorm is. De eenvoudigste vorm is de *rechte lijn*; eenvoudig is ook de *exponentiële lijn*, waarvan de ordinaten op onderling gelijke afstanden een meetkundige reeks vormen. Gaat men van de gegeven reeks over op de reeks der logarithmen dier getallen, dan zal de trend, indien aanvankelijk een exponentiële lijn, een rechte lijn worden. Deze eigenschap wordt vaak benut om de oorspronkelijke reeks op een logarithmische schaal weer te geven: de exponentiële trend is dan een rechte lijn. Terwijl de rechte lijn gekenmerkt is door een constante absolute toeneming per jaar, is de exponentiële lijn er een van constante procentuele toeneming. Men heeft nog ingewikkelder lijnen als trend gebruikt, n.l. zogenaamde *parabolen* (van willekeurige orde) en dan de *logistische curve.* Deze laatste verdient wegens haar principiële en praktische betekenis nadere toelichting. Het eenvoudigste voorbeeld in getallen verkrijgt men als men de trend x_t voor elk tijdstip t berekent volgens de formule:

$$x_t = \frac{120}{1 + 2^{-t}}, \tag{16.1}$$

waarbij men t moet laten lopen van $-\infty$ tot $+\infty$. Men verkrijgt o.m. de volgende waarden:

$t =$	$-\infty$	....	-3	-2	-1	0	$+1$	$+2$	$+3$	.... $+\infty$
	0		$13\frac{1}{3}$	24	40	60	80	96	$106\frac{2}{3}$	120

Grafisch voorgesteld (zie fig. 2, curve B), verkrijgt men een curve die aanvankelijk langzaam stijgt, vervolgens sneller, om daarna te naderen tot een zekere verzadigingswaarde — 120. De stijging heeft in het eerste gedeelte wel langzaam plaats, naar absolute maatstaf gemeten, doch procentueel het snelst.

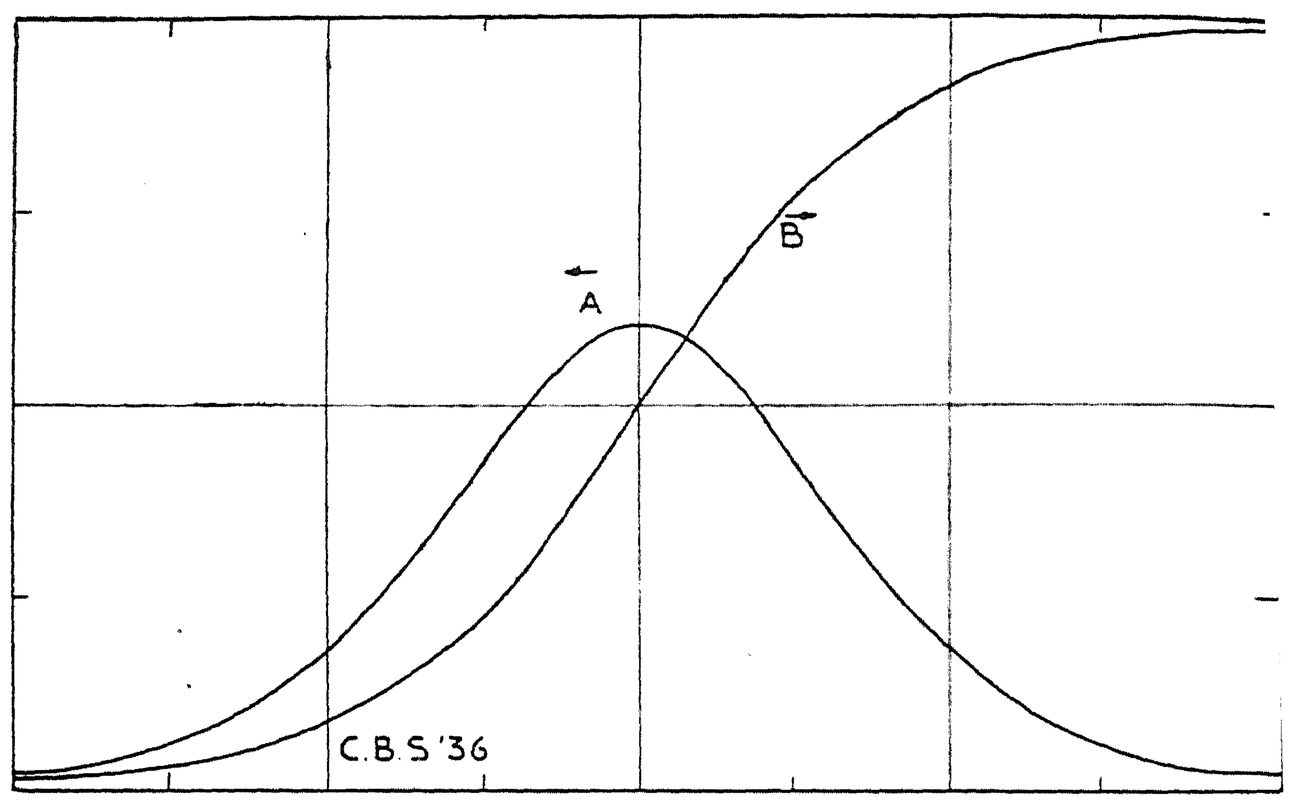

Fig. 2.
Verloop van de curve van Gauss (*A*) en de logistische curve (*B*).

Welke van deze verschillende lijnen moet nu in een concreet geval worden gebruikt — wat is daartoe het criterium? In principe moet dit worden beslist door de wet die het ontstaan van de trend bepaalt. Dat wil dus zeggen, dat men een verklaring, een theorie, van die beweging heeft. In bepaalde eenvoudige gevallen is dat ook zo. Denken wij aan de groei van de bevolking. Onder bepaalde vereenvoudigende omstandigheden zullen daar de sterfte zowel als de geboorte evenredig met de grootte van de bevolking zijn, dus ook de aanwas per jaar. Als dat zo is, zal de lijn van de bevolking een exponentiële

lijn zijn. Iets dergelijks geldt, wanneer de groei van de kapitaalgoederenvoorraad onder vereenvoudigde omstandigheden wordt bestudeerd: de toeneming van die voorraad wordt door de besparingen bepaald en deze in hoofdzaak door het niet-arbeidsinkomen, hetwelk veelal weer evenredig is met de bestaande kapitaalgoederenvoorraad. Ook in dat geval verkrijgt men een exponentiële groei. De logistische curve ontstaat, indien een iets ingewikkelder groeiproces plaats heeft; n.l. indien de toeneming van een bevolking met omvang x evenredig is zowel met x zelf als met de afstand van x tot een hoger vast getal a. Deze groeiwet heeft men aangetroffen bij koloniën van lagere dieren, die een vaste voedselvoorraad of voedselbron tot hun beschikking hadden. Het getal a komt dan overeen met de bevolking die men met die voedselbron kan onderhouden. De groei verlangzaamt zich dan wanneer x tot a nadert en a wordt niet overschreden, slechts benaderd. Het getal 120 in ons voorbeeld komt overeen met het getal a. Men heeft de wet ook voor menselijke bevolkingen vastgesteld, ofschoon daarbij afwijkingen voorkomen. Meestal is de voedselbron ook niet constant bij menselijke bevolkingen. In 't algemeen zal men een groei volgens ongeveer deze logistische curve kunnen verwachten indien er na enige tijd een zekere vertraging, een zekere verzadiging, optreedt, en bij vele economische verschijnselen is deze ook waar te nemen. In het bijzonder de ontwikkeling van de hoeveelheid van een nieuw artikel dat in gebruik komt vertoont vaak ongeveer deze curve, zoals b.v. het aantal auto's of radio's, de lengte van het spoorwegnet of het telefoonnet en de omvang van de productie van kunstzijde. Men kan deze ontwikkeling ook nog iets anders beschouwen. Men kan zeggen dat sommige mensen iets onmiddellijk willen hebben, als het nieuw op de markt komt. Anderen willen eerst de kat uit de boom kijken; weer anderen moeten een zeer grote traagheid overwinnen alvorens er toe te komen. De snelheid van reageren loopt dus voor verschillende mensen zeer uiteen en men kan aannemen dat zij net als andere menselijke

eigenschappen ongeveer verdeeld is volgens de „normale wet" of de wet van Gauss (zie ook § 17). Door de getallen-waarden van de nieuwe aankopen, die aldus in de tijd Gaussiaans verdeeld zijn, te cumuleren, verkrijgt men het totaal aantal voorwerpen in gebruik en het verloop van de gecumuleerde Gauss-curve gelijkt zeer op dat van de logistische curve.

Men kan nu de exponentiële curve als een benadering over een gedeelte van het eerste stadium van de logistische curve beschouwen; daar waar de verzadigingstendentie nog praktisch onmerkbaar is, omdat x zo klein is in vergelijking tot a, dat $a - x$ als constant mag worden beschouwd. Dan is de toeneming alleen evenredig met x. En daarnaast kan men de rechte lijn als een benadering over nog kleinere intervallen van elk dezer lijnen beschouwen, ook van de hoger gelegen delen van de logistische curve: elke curve kan nu eenmaal door een rechte lijn worden benaderd. De parabolen zijn meestal niet theoretisch te funderen en verdienen daarom minder aanbeveling; gewoonlijk worden zij gekozen op grond van een meer uiterlijke gelijkenis met de algemene loop van de te behandelen reeks.

De aanwezigheid van een exponentiële trend in een reeks x heeft tot gevolg, dat in de reeks x^1 van de procentuele toenemingen van x geen trend aanwezig is; slechts bewegen zich de getallen x^1 alle om een getal afwijkend van nul heen. In plaats van de trend uit te schakelen (te werken met trendafwijkingen) werkt men daarom ook wel eens met de procentuele toenemingen, in de Amerikaanse literatuur bekend als „link relatives".

§ 17. Bepaling en uitschakeling van de toevallige component.

Onder een toevallige variabele in ruimere zin verstaat men een variabele die meer dan één waarde kan aannemen, terwijl van te voren niet vaststaat welke dezer waarden

in gegeven omstandigheden — b.v. op een gegeven tijdstip — wordt aangenomen. Een aanschouwelijk voorbeeld dat meer zegt dan abstracte definities, is b.v. het aantal keren munt dat men werpt met een of meer munten. Werpt men één munt, dan kan dat aantal keren zijn 0 of 1; werpt men met 3 munten, dan kan het zijn 0, 1, 2 of 3. Het aantal keren dat elk dezer waarden voorkomt is de frequentie er van, de tabel dezer frequenties is de frequentieverdeling. Drukt men deze in relatieve cijfers uit, d.w.z. in cijfers die tezamen 1 zijn en evenredig met de frequenties, dan spreekt men van de *waarschijnlijkheidsverdeling*. De waarschijnlijkheid of relatieve frequentie van het optreden van één bepaalde uitkomst geeft dus aan, welk gedeelte der waarneming juist die uitkomst oplevert.

De waarschijnlijkheidsverdeling van de uitkomsten van een zeer vaak herhaald munt- of kruisspel met één munt is gegeven door de tabel:

Aantal keren munt	Waarschijnlijkheid
0	½
1	½

Er zijn tal van andere waarschijnlijkheidsverdelingen van toevallige variabelen mogelijk. Een zeer bekende verdeling is die van Gauss. Deze heeft betrekking op een continue toevallige variabele, d.w.z. een toevallige variabele die alle waarden kan aannemen in een zeker interval en niet b.v. slechts gehele waarden. De *wet van Gauss* wordt grafisch weergegeven door de bekende klokvormige curve (fig. 2, curve A); de waarschijnlijkheid dat een waarneming ligt in een klein gedacht eenheidsinterval in de omgeving van een waarde x is:

$$y = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{x^2}{2\sigma^2}}$$

Als een toevallige variabele deze verdelingswet vertoont, noemt men ze wel toevallig in de engere zin. Een dergelijke verdeling kan door allerlei mechanismen tot stand worden

gebracht, o.m. bij benadering wanneer de waarde van de variabele op elk tijdstip het resultaat is van een zeer groot aantal van elkaar onafhankelijk werkende krachten, die ieder op zich zelf een geringe uitwijking veroorzaken die zelf toevallig in de ruimere zin is.

Een reeks cijfers die aan de een of andere frequentieverdeling voldoet, b.v. aan de wet van Gauss, kan nog verschillende eigenschappen vertonen voor wat de *volgorde* der cijfers betreft. Men zou zich kunnen denken, dat de cijfers in volgorde van grootte gerangschikt zijn. Men kan zich ook denken dat er in deze volgorde geen enkele regelmatigheid te bespeuren valt. In het bijzonder kan men nagaan, of er afhankelijkheid bestaat tussen twee opeenvolgende termen, door na te gaan of de frequentieverdeling die het tweede getal van een aantal willekeurige paren vertoont anders is wanneer men het eerste getal van die paren gelijk neemt aan een vast getal a dan wanneer men dat eerste getal gelijk neemt aan een ander vast getal b. Als de frequentieverdeling der tweede getallen steeds dezelfde is, welke waarde ook het eerste getal heeft, dan zegt men dat de opeenvolgende termen van elkaar *onafhankelijk* zijn. Wanneer de reeks cijfers waarover wij het hebben, ontstaan is op de hierboven genoemde wijze, n.l. als het resultaat van een zeer groot aantal onafhankelijk werkende krachten, die ieder op zichzelf een geringe uitwijking veroorzaken die zelf toevallig in ruimere zin is, zullen de opeenvolgende termen deze onafhankelijkheid ook vertonen. In het algemeen zullen de opeenvolgende termen van een reeks cijfers onafhankelijk van elkaar zijn, wanneer de wetten die het tot stand komen van elk cijfer beheersen niet afhangen van het voorafgaande cijfer.

Het vaststellen van de toevallige component in een reeks waargenomen cijfers geschiedt nu gewoonlijk aldus, dat men eerst een hypothese opstelt omtrent het gedrag van alle andere componenten. Op grond daarvan bepaalt men de toevallige component en controleert nu achteraf of

deze laatste aan de eisen voldoet, die men aan een toevallige reeks moet stellen. De hypothese die men ten aanzien van de som der overige componenten maakt is gewoonlijk dat deze een „geleidelijk verloop" heeft. Wiskundig kan men dat preciseren door aan te nemen, dat de curve dezer som van de eerste graad (recht), van de tweede, 3e, 4e ... graad is. Deze curven zijn in afnemende mate „geleidelijk", worden n.l. steeds ingewikkelder en zijn in staat steeds meer heen en weer te slingeren. Men kan daarbij nog een integrale en een differentiële methode onderscheiden. Bij de eerste neemt men aan, dat over de gehele beschouwde periode de som der andere componenten een rechte lijn of een der curven van hogere graad is. Men stelt deze som dan dus eigenlijk gelijk met een trend en wel een wiskundige trend. Bij de differentiële methode neemt men slechts aan, dat voor kleine gedeelten van de periode aan deze eisen is voldaan, dus dat de systematische componenten tezamen in de buurt van elk punt een rechte lijn, of een der hogere curven vormen. Berekening leert, dat men dan niets anders doet dan aan te nemen dat de som dezer componenten kan worden benaderd door een voortschrijdend gemiddelde. Neemt men een rechte lijn als benadering aan, dan is het een gewoon voortschrijdend gemiddelde; neemt men daarentegen curven van hogere graad aan, dan blijkt het een gewogen voortschrijdend gemiddelde te zijn, waarbij aan de middelste termen hogere gewichten worden toegekend dan aan de uiterste termen. Het aantal termen dezer voortschrijdende — ongewogen of gewogen — gemiddelden hangt weer af van de lengte der kleinere perioden, waarvoor men de veronderstelling maakt. Als men nu b.v. aanneemt, dat over telkens 5 punten de som der systematische componenten een curve van de 2e graad (gewone parabool) vormt, dan krijgt men een gewogen voortschrijdend gemiddelde van de volgende gedaante:

$$\frac{1}{35}(-3x_{-2}+12x_{-1}+17x_0+12x_{+1}-3x_{+2}).$$

Men bepaalt dus de som der overige componenten hetzij

als een wiskundige, hetzij als een niet-wiskundige trend, vindt door aftrekking dezer trendcijfers van de gegeven reeks het overblijvende gedeelte, dat dus de toevallige component pretendeert te zijn en gaat nu controleren of deze voldoet aan de te stellen eisen.

Deze eisen kunnen tweeërlei zijn, n.l. die van onafhankelijkheid der opeenvolgende termen en die van de frequentieverdeling van Gauss. De onafhankelijkheid controleert men door de *autocorrelatie* te bepalen, d.i. de correlatie (zie § 19) die er bestaat tussen de opeenvolgende termen der reeks. De controle van de frequentieverdeling kan òf betrekkelijk grof geschieden door opstelling dezer verdeling en grafische vergelijking, òf meer nauwkeurig door gebruik te maken van een der daartoe bestaande wiskundig-statistische methoden (de χ^2-test).

§ 18. Bepaling van de seizoenschommelingen en de conjunctuurcomponent.

Daar bij de seizoenschommelingen als regel een vaststaande periode bestaat, zijn hier soms fijnere rekenmethoden mogelijk dan bij de overige componenten. Getrouw aan de opzet van dit werkje willen wij hier slechts de principiële betekenis van enige der belangrijkste rekenmethoden nader belichten.

Bepalen wij ons eerst tot de vaststelling van de normale seizoenbeweging (zie § 15), dan is een eerste vraag die men zich nogal eens stelt: is de seizoeninvloed *additief* of *multiplicatief?* Daarmee bedoelt men het volgende: als het algemene peil waarop zich een variabele bevindt, zich wijzigt, blijft dan de toevoeging in het drukke seizoen even groot of is deze evenredig met de hoogte van dat algemene peil? Nader onderzoek doet zien, dat het van het mechanisme der seizoenwerking afhangt of het ene dan wel het andere geschiedt. Bovendien zijn er nog vele andere mogelijkheden. Meestal echter laat het materiaal niet toe, om daaromtrent een exacte beslissing te nemen. Het ideaal van wetenschappelijk onderzoek zou hier een

nauwkeurig onderzoek vereisen, dat echter in het algemeen nog tot de vrome wensen behoort.

Men kan zich voorts een *star* en een *systematisch veranderend* seizoenpatroon denken. Bij een star patroon wordt de seizoeninvloed voor elk jaar hetzelfde verondersteld — hetzij dan additief of multiplicatief. Men kan zich echter denken dat de toppen langzamerhand minder hoog, de dalen minder diep worden; een practisch voorbeeld is dat van de eierprijzen: door de nachtverlichting in de kippenhokken heeft men het leggen in de winter kunnen stimuleren en daarmee de seizoenverschillen verminderd. Deze verlaging van de toppen kan men zich regelmatig denken op uiteenlopende manieren: gewoonlijk zal men aannemen dat er elk jaar een even groot bedrag afgaat en men kan dan statistisch vaststellen, hoe groot men dat bedrag moet nemen, om een zo goed mogelijke aansluiting bij het materiaal te verkrijgen.

Men kan zich de systematische verandering nog wel ingewikkelder denken, zoals bij de methode van Wald het geval is, waar men de hoogte der toppen en de diepte der dalen als evenredig met de standaarddeviatie der seizoenschommelingen denkt; deze laatste dan bepaald uit de afwijkingen van voortschrijdende twaalfmaands-gemiddelden.

Een bijzondere moeilijkheid bieden de verschuivende seizoenen. Soms kunnen deze, wanneer ze systematisch geschieden, als een bijzonder geval van het eerstgenoemde voorbeeld van veranderend seizoenpatroon worden beschouwd: tegenover een daling van een der maanden kan een stijging van een andere staan. Wanneer echter de verschuiving niet steeds in dezelfde richting doorgaat, is deze methode niet toe te passen. Een voorbeeld is de verschuiving van seizoenen als gevolg van het verspringen van de data van Pasen en Pinksteren. Daarvoor kan men natuurlijk incidenteel trachten te corrigeeren. Een ander geval is een veranderend seizoen als gevolg van b.v. weersomstandigheden. Dit brengt ons langzamerhand op het gebied van de individuele seizoenbeweging: de seizoen-

beweging die van jaar tot jaar verschilt op grond van een aanwijsbare oorzaak.

Het ideaal van wetenschappelijke behandeling is hier de causale verklaring, dus b.v. het verklaren van seizoenbewegingen uit de temperatuur, de neerslag, enz. Het worden dan vraagstukken van correlatie, die wij thans zullen gaan bespreken.

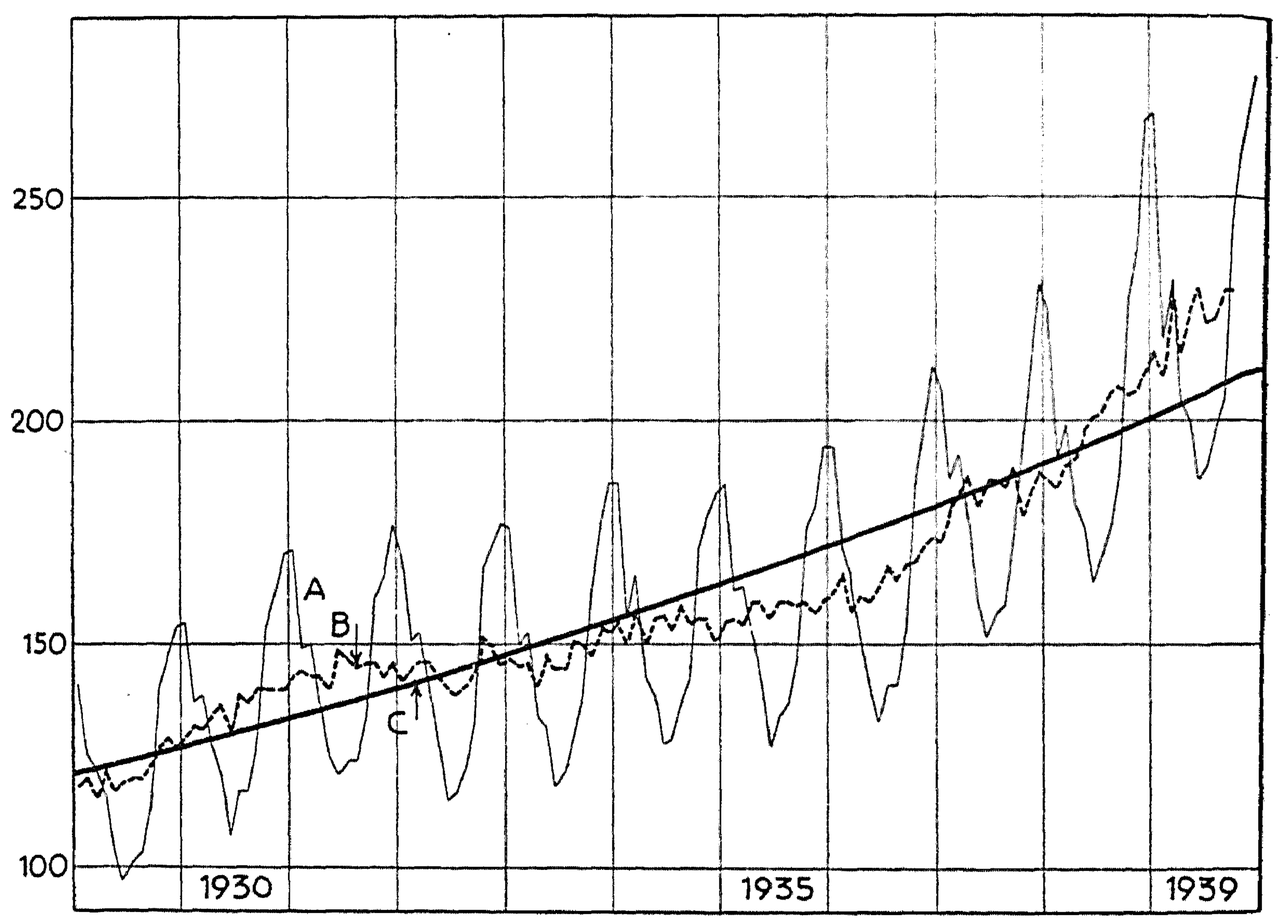

Fig. 3.
Voorbeeld van de berekening van componenten van een reeks (*A*) (electriciteitsverbruik in Nederland in millioenen kWh, dunne lijn). De zware getrokken lijn (*C*) geeft de trend aan, berekend volgens de exponentiële methode. De stippellijn (*B*) is de reeks na uitschakeling van de seizoenschommelingen volgens de multiplicatieve methode, die hier de aangewezen methode is, omdat de seizoenschommelingen evenredig zijn met de jaargemiddelden. De stippellijn geeft dus de som van trend, conjunctuurcomponent en toevallige component.

In fig. 3 geven wij tenslotte een voorbeeld van een reeks *A*, de voor seizoeninvloeden gecorrigeerde reeks (*B*) en de exponentiële trend (*C*).

§ 19. Enkelvoudige correlatie.

Wij zijn hiermee gekomen tot een verdere stap van econometrisch onderzoek, n.l. de vergelijking van twee reeksen waarnemingen, betrekking hebbend op twee verschijnselen, waarvan men op grond van redenering aanneemt, dat zij causaal verband vertonen. Dit geval doet zich klaarblijkelijk voor, wanneer men, op grond van de economische analyse, meent dat de schommelingen in een verschijnsel x in hoofdzaak worden bepaald door de schommelingen in één ander verschijnsel y.

Zoals wij nog nader zullen uiteenzetten, heeft de statistiek hier een dubbele taak, n.l. die van *controle* en die van *meting*.

Om te beginnen is het steeds aanbevelenswaardig van *grafische methoden* gebruik te maken en wel op twee wijzen: door het construeren van historische grafieken en van een zgn. spreidingsdiagram. Een historische grafiek geeft het verloop der reeksen x en y in de tijd aan, een spreidingsdiagram geeft een serie punten, waarvan elk als zijn twee coördinaten reeds y en x heeft; gewoonlijk zal men daarbij de onafhankelijk veranderlijke (de verklarende reeks, bij ons de y) langs de horizontale as afzetten en de afhankelijk veranderlijke (de te verklaren reeks, bij ons x) langs de verticale. Als er een duidelijk verband is tussen de twee reeksen, is het spreidingsdiagram een fijner hulpmiddel dan de historische grafieken. Doch deze laatste doen gemakkelijker zien, of er soms een verschil in *trend* is of een *lag*. In die gevallen lijken de spreidingsdiagrammen soms reeds zeer slecht, zodat men niet ziet, op welke wijze het verband verbeterd kan worden. Deze laatste opmerking heeft inmiddels reeds betrekking op pogingen om, als de eerste statistische controle eigenlijk reeds geschied is — en negatief was — suggesties tot wijziging der te toetsen economische hypothese aan den econometrist te geven.

Als de twee verschijnselen inderdaad nauw verband vertonen, zal zich dat manifesteren doordat de punten

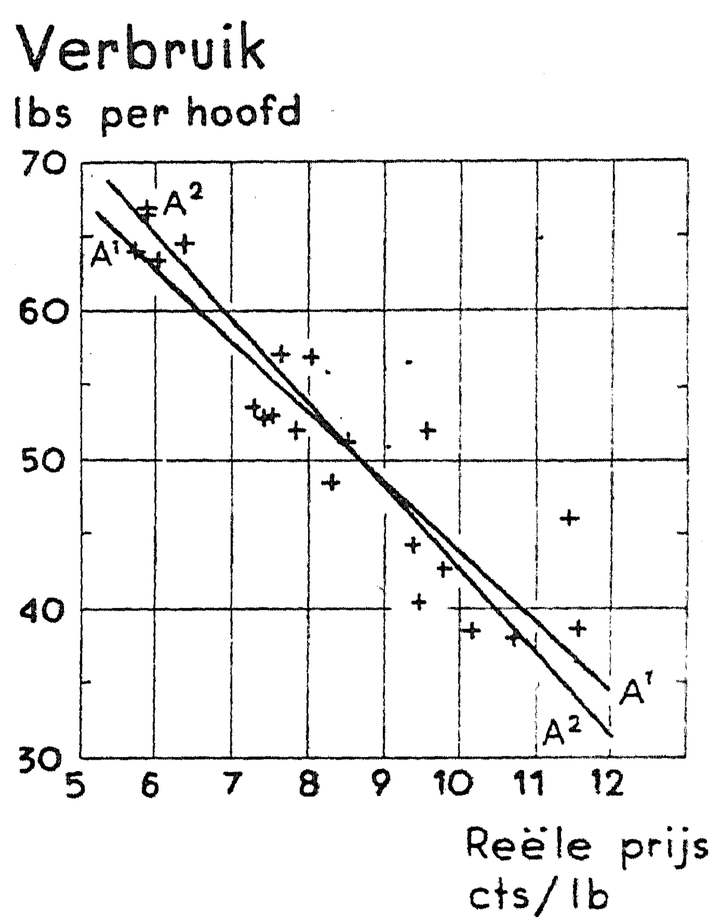

Fig. 4A.
Voorbeeld van een enkelvoudige correlatie (de vraaglijn voor suiker, ontleend Schultz, The Theory and Measurement of Demand). De twee rechte lijnen geven de eerste (A1) en tweede (A2) regressielijn aan.

die men verkrijgt (vgl. b.v. fig. 4 A) vrijwel op één lijn liggen. Als dat het geval is, zegt men dat de *correlatie* hoog is, eventueel volledig. Liggen de punten meer verspreid, dan is er onvolledige of lage correlatie. Al naar de vorm van de lijn waaromheen zich de punten scharen heeft men *rechtlijnige* of *kromlijnige* correlatie. In het geval van rechtlijnige correlatie spreekt men ook nog van *normale correlatie*, wanneer bij elke waarde van y de afwijkingen die de x-waarden vertonen de frequentieverdeling van Gauss volgen en eveneens bij elke waarde van x de afwijkingen van y.

Wanneer de grafische methode heeft doen zien, dat er een zeker verband is — en dus kans van slagen —, kan een verdere verfijning der analyse met algebraïsche methoden plaats vinden.

De *controlerende* functie van de statistiek moet antwoord geven op de vraag: is er verband tussen y en x? Hierop past geen antwoord ja of neen, doch een quantitatief antwoord, omdat er toch immers alle mogelijke tussenvormen zijn. Het verband is dus nauw of minder nauw, en dit kan worden uitgedrukt door de *correlatiecoëfficiënt*, door de *correlatie-index* of de *correlatie-verhouding*. Voor alle drie geldt dat:

$$r^2 = 1 - \frac{\sigma^2_{x,y}}{\sigma^2_x}$$

Hierbij is r een dezer drie maatstaven van de correlatie, σ_x is de standaarddeviatie van de te verklaren reeks x (zie § 14), terwijl $\sigma_{x,y}$ de standaarddeviatie is van een maatstaf voor de niet door y verklaarde afwijkingen of de zgn. residuen. Zijn deze residuen alle nul, dan is elke waarde van x dus precies verklaard door het aangenomen verband tusschen y en x; er is volledige correlatie. Zijn de residuen in het gemiddeld even groot als de afwijkingen die x van zijn eigen gemiddelde vertoont, dan heeft de verklaring door y een minimum van succes bereikt en is de correlatie afwezig. De verschillen tussen de drie genoemde maatstaven van de correlatie zijn gelegen in de wijze waarop men de residuen berekent. Ten behoeve van de correlatiecoëfficiënt berekent men de afwijkingen die de punten in het spreidingsdiagram vertonen van een rechte lijn die zo goed mogelijk aansluit bij die punten. Ten behoeve van de correlatie-index berekent men de afwijkingen t.o.v. een zo goed mogelijk aangepaste kromme lijn. Ten behoeve tenslotte van de correlatieverhouding neemt men de afwijkingen t.o.v. de gemiddelden die bij elke waarde van y kunnen worden berekend voor de bijbehorende x-waarden. Steeds worden hierbij de afwijkingen in verticale richting — d.i. hier in x-richting — gemeten. De correlatie-coëfficiënt is de strengste rechter; hij is alleen 1, wanneer de punten op een rechte lijn liggen. Bij de correlatie-index heeft men het materiaal om zo te zeggen een kans meer gegeven, door de lijn naar behoefte krom te trekken; bij de correlatieverhouding past men zich nog verder aan bij het materiaal, door de lijn der gemiddelden te nemen, die eventueel betrekkelijk willekeurig heen en weer kan springen. Men kan zelfs, wanneer men de y-waarden maar fijn genoeg meet, er voor zorgen, dat elke y-waarde maar éénmaal voorkomt en daarom de gemiddelde x-waarde die daarbij behoort, gelijk is aan de (enige) x-waarde die in die klasse voorkomt.

Dan is er volledige correlatie in elk geval, waarin maar de y-waarden alle verschillen. Dit laatste is echter natuurlijk onaanvaardbaar en in 't algemeen zal men dan ook een zo fijne indeling niet nemen. Strikt genomen zal men trouwens van de correlatieverhouding geen gebruik mogen maken, wanneer het verband, dat men moet toetsen, werkelijk van te voren is beschreven door den economist.

Of men een gevonden correlatie als bevredigend wil beschouwen, is overigens een kwestie van de tolerantie (de speling) die men toestaat en daarin bestaat nogal verschil van smaak. Bij een groot aantal waarnemingen kan men bovendien met een lagere correlatie genoegen nemen dan bij een klein aantal. Men moet zichzelf hier een standaard stellen.

Is aan deze standaard voldaan, dan komt de *metende* functie van de statistiek. Het is nu van betekenis, vast te stellen welk verband er is tussen y en x. In het geval van een rechte lijn kan men deze vraag zo uitdrukken: met hoeveel stijgt x, wanneer y met een eenheid stijgt? Dit bedrag noemt men de *regressie-coëfficiënt*; ook als men met andere dan rechte lijnen te maken heeft, noemt men de lijn die het verband aangeeft de *regressielijn* (of *-curve*), de vergelijking die het verband aangeeft, de *regressievergelijking*. De regressiecoëfficiënt en in 't algemeen de constanten die in de regressievergelijking optreden zijn natuurlijk niet exact te bepalen, doch met een onzekerheidsmarge behept; de mathematische statistiek levert daaromtrent verschillende maatstaven. Hoe hoger de correlatie en hoe groter het aantal waarnemingen, hoe kleiner deze onzekerheid.

In fig. 4 B wordt een tweede voorbeeld van een enkelvoudige correlatie gegeven, over het resultaat waarvan in § 7 reeds het een en ander werd medegedeeld.

De lijn die het verband aangeeft kan, wanneer er volledige correlatie bestaat, ondubbelzinnig worden bepaald. Zodra er geen volledige correlatie bestaat, zijn er verschillende lijnen mogelijk, al naarmate men verschillende principes gebruikt om de lijn zo goed mogelijk te

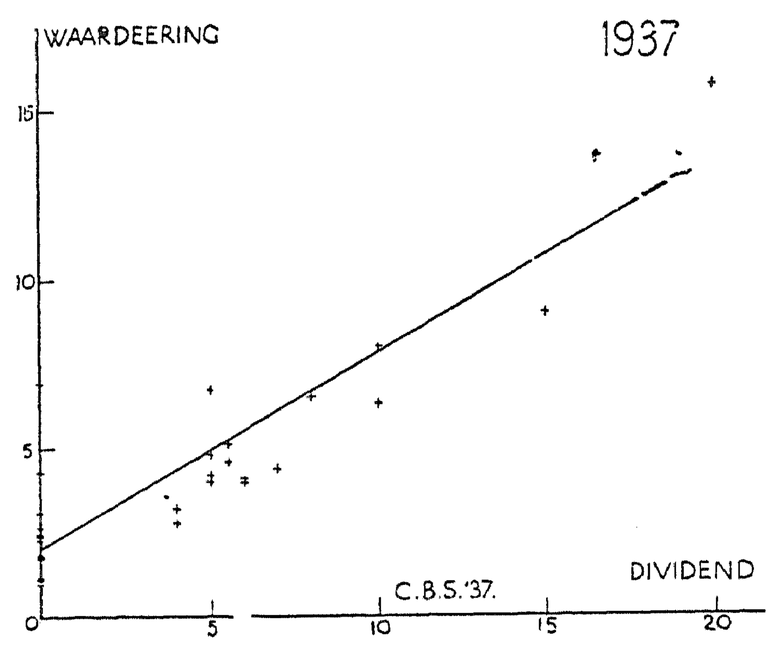

Fig. 4 B.
Voorbeeld van een enkelvoudige correlatie (ontleend aan De Nederlandse Conjunctuur): het verband tussen dividend en „waardering" van aandelen. Onder „waardering" wordt verstaan het product van de koers van het aandeel en het rendement van staatsobligaties; het geeft a.h.w. aan, met een obligatie van welk rendement men het aandeel gelijkstelt.

doen aansluiten bij het spreidingsdiagram. Gewoonlijk bepaalt men de lijn aldus dat de som der kwadraten van de residuen in de x-richting zo klein mogelijk is, d. w. z. men past het principe van de kleinste kwadraten toe op de residuen in de richting van de te verklaren variabele. De aldus verkregen lijn heet, in het bijzonder wanneer men een rechte lijn heeft genomen, de *eerste* regressielijn. Past men het principe toe in de y-richting (die van de verklarende variabele), dan heeft men de *tweede* regressielijn. Bij volledige correlatie vallen de lijnen samen. Indien de correlatiecoëfficiënt $= 0$ is, staan zij loodrecht op elkaar en vallen met de assen samen. Men kan daarom ook de afwijking tussen die twee lijnen (hetzij de hoek of een andere maatstaf) als een maatstaf aannemen van de correlatie.

Als men de eerste regressielijn neemt, veronderstelt men dat de afwijkingen tussen werkelijke en berekende waarden uitsluitend het gevolg zijn van *a*) meetfouten in de te verklaren reeks en *b*) onvolledigheden van de theorie. Als regel zal men daarbij aan *b* de grootste betekenis moeten toekennen. Als er aanleiding zou zijn ook meetfouten in de verklarende variabele aan te nemen, zou men nog andere regressielijnen kunnen bezigen.

§ 20. Toepassingsmogelijkheden op economisch gebied.

Zoals werd opgemerkt, is de toetsingsmethode der enkelvoudige correlatie alleen te gebruiken wanneer een grootheid x verklaard kan worden door één enkele andere grootheid y. Zoals men weet, is dit in de economische wetenschap echter slechts zelden het geval. In de regel wordt een grootheid x beïnvloed door vele andere grootheden. Dit is van ouds de moeilijkheid geweest bij het toetsen van economische wetten; de theoretische wetten leggen wel verband tussen telkens twee grootheden, doch onder de voorwaarde „*ceteris paribus*": dat andere factoren gelijk blijven. Wat zij meestal niet doen. Opdat de enkelvoudige correlatierekening voor het econometrisch werk betekenis heeft, moet dus de theorie kunnen doen zien, dat er gevallen zijn, waarin de ceteris-paribus-clausule ongeveer vervuld is. Wij willen nagaan in hoeverre dat blijkt het geval te zijn in het bijzonder op het gebied van de vraag- en aanbodfuncties. Daarmee onderbreken wij onze beschouwingen over de techniek der statistische toetsing dus even; wij veroorloven ons deze inbreuk op de systematiek, omdat in het levende onderzoek, zoals wij al zeiden, het heen-en-weer kaatsen tussen statisticus en econoom ook telkens plaats vind en vooral omdat er geen andere plaats is, waar dit onderwerp beter zou kunnen worden behandeld.

Wij komen hier op klassiek terrein: de statistische bepaling, zoals men het veelal uitgedrukt heeft, van vraag- en aanbodcurven, d.w.z. het verband tussen prijs en omgezette hoeveelheid. Dit kan alleen wanneer de overige factoren, dus de vraag- resp. de aanbodfactoren gelijk blijven, een geval dat wel wordt aangeduid als het Cournot-geval. Deze overige factoren zijn voor de vraagfunctie vooral het inkomen en eventueel de prijs van concurrerende artikelen en voor de aanbodfunctie de kosten van productie. Alvorens enige gevallen op te noemen die o.i. aan deze eis voldoen, moeten wij de

zojuist gegeven, gebruikelijke, opvatting iets corrigeren. Het is n.l. niet nodig dat de overige factoren — ook maar bij benadering — gelijk blijven; voldoende is dat hun invloed gering is. Dit kan daardoor gebeuren dat zij in de reactierelatie met een kleine coëfficiënt voorkomen.

Wij menen dat er inderdaad in de volgende gevallen reden is om te verwachten dat de overige factoren weinig invloed zullen hebben.

1. De vraag naar *primaire levensbehoeften* zal weinig afhankelijk zijn van het inkomen; voorzover zij dus geen sterk vervangbare of complementaire goederen zijn, zal hun vraag voornamelijk van hun eigen prijs afhangen.

2. De vraag naar goederen waarvan, door omstandigheden aan de aanbodzijde, de prijzen *extra sterk schommelen* zal relatief weinig van de overige factoren afhangen. Voorbeelden zijn goederen met zeer sterke schommelingen in oogstopbrengsten, in het bijzonder wanneer ze bederfelijk zijn, zoals vruchten en groenten. Voorts artikelen met zeer sterke schommelingen in de productie als gevolg van de slechte aanpassingsmogelijkheid van aanbod aan vraag door de lange duur van het productieproces. Voorbeelden hiervan zijn varkensvlees en koffie. Een artikel waarbij zeer sterke schommelingen in het aanbod optreden is voorts ansjovis, waar de vangst het ene jaar soms het honderdvoudige is van het andere (vgl. § 39).

3. De vraag naar goederen met zeer *snelle* schommelingen in aanbod. Daar de wijzigingen in de inkomens als regel betrekkelijk geleidelijk geschieden — voor de volkshuishouding als geheel — kan men door korte perioden te beschouwen de invloed daarvan reduceren. Voorbeelden zijn hier verschillende goederen met een sterke seizoenbeweging in het aanbod, die niet bewaard kunnen worden. Boven werden in ander verband reeds vruchten en groenten genoemd; eieren zijn een ander voorbeeld, inzonderheid voor de tijd toen er nog minder in koelhuizen werden opgeslagen.

4. De *aanbod*functie van goederen met betrekkelijk

stabiele kosten doch grote schommelingen in vraag, zoals *minerale* grondstoffen, kan eveneens genoemd worden.

In een aantal andere gevallen is aan de ceteris-paribus-clausule niet voldaan, zodat x verklaard moet worden uit meerdere variabelen y, z, enz. Daarbij is het echter vaak toch mogelijk om het meervoudig verband dat hierbij optreedt, te herleiden tot een enkelvoudig, dat dan door enkelvoudige correlatierekening op de proef kan worden gesteld. Veelal voert men dan te voren bepaalde „*correcties*" aan het materiaal uit of gaat men over op andere variabelen. Wij willen een aantal van de meest bekende opnoemen.

1. Als bij een vraagfunctie de voornaamste andere factor die zich heeft gewijzigd de *omvang van de bevolking* is, beschouwt men de gevraagde hoeveelheid *per hoofd* der bevolking, d.w.z. men voert als correctie uit een deling van de hoeveelheden door de bevolking.

2. Als de voornaamste andere factor die zich heeft gewijzigd het *prijsniveau van alle andere goederen* is — waarbij men alle andere goederen als concurrerend met het beschouwde goed aanneemt — deelt men de prijs van het onderzochte goed door het algemene prijsniveau, men werkt met, zoals de Amerikanen zeggen, „*deflated prices*". Hierin ligt overigens een hypothese opgesloten, die even vermeld dient te worden, n.l. dat de invloed van het algemene prijsniveau juist zodanig is, dat in de vraagfunctie juist het quotiënt van de prijs en het algemene prijsniveau optreedt. Het is denkbaar dat de functie anders is.

3. Als — bij een *aanbodfunctie* — de andere factoren die beschouwd moeten worden, de kosten van productie van de eenheid zijn, trekt men deze van de prijs af en voert als nieuwe variabele in de *winstmarge*. Als deze kosten zelf nog weer opgebouwd gedacht worden uit grondstof- en loonkosten, zorgt men er voor, dat een index van de grondstofprijzen en een loonindex zodanig tot een gemiddelde worden herleid, dat hun gewichten beantwoorden aan de samenstelling van de productiekosten.

4. Als — eveneens bij een aanbodfunctie — rekening gehouden moet worden met wijzigingen in de *capaciteit* der productie, drukt men de aangeboden hoeveelheden uit in procenten van de capaciteit.

5. Als — in een vraagfunctie — rekening gehouden moet worden met de invloed van *prijzen van concurrerende goederen*, kan men aannemen dat de verhouding van de prijs tot die van het concurrerende goed in die vraagfunctie optreedt. Hier past overigens eenzelfde opmerking als onder 2 werd gemaakt. Men kan nog een stapje verder gaan en ook de hoeveelheid vervangen door de verhouding der hoeveelheden van de twee concurrerende goederen. Als dus van het beschouwde goed een hoeveelheid x_1 wordt gevraagd bij een prijs p_1 en van het concurrerende goed een hoeveelheid x_2 bij een prijs p_2, dan neemt men aan dat

$$\frac{x_1}{x_2} = f\left(\frac{p_1}{p_2}\right)$$

Een soortgelijke gedachte ligt ten grondslag aan een vraagfunctie voor investeringsgoederen, door sommige schrijvers verdedigd, waarbij men deze vraag afhankelijk veronderstelt van het quotiënt

$$\frac{z}{m\,q},$$

waarbij z de verwachte winst,
q de prijs van het investeringsgoed en
m de rentestand is. Dit quotiënt kan worden opgevat als een verhouding van „winstpercentages", n.l. $\frac{z}{q}$ op het te investeren goed en m op het geldkapitaal dat men er voor opoffert.

6. Als — in een vraagfunctie — het *inkomen* de factor is waarmee men, behalve de prijs, ook nog rekening moet houden, kan men de invloed daarvan *a priori* schatten op grond van de budgetstatistiek en de waargenomen

hoeveelheden vervangen door „voor inkomenswijzigingen gecorrigeerde hoeveelheden".

7. Als de overige factoren alle geacht kunnen worden slechts zeer geleidelijk te veranderen — b.v. grootte van de bevolking, verandering van smaak en gewoonten, eventueel ook inkomensveranderingen in zeer rustige tijden — kan men aannemen, dat deze slechts invloed hebben op de trend van het verbruik. Door zowel van hoeveelheden als van prijzen de trendafwijkingen te nemen, zal men dan twee variabelen hebben, welker samenhang niet gestoord is. Men zou hier verkeerd doen door alleen bij het verbruik de trend uit te schakelen en niet bij de prijs; want de trend van de prijs zal ook op de trend van het verbruik invloed oefenen; de schommelingen van het verbruik kunnen echter slechts worden verklaard door de schommelingen in de prijs, onder schommelingen hier trendafwijkingen te verstaan.

Indien een dezer aprioristische herleidingen tot enkelvoudig verband niet mogelijk is, dient men gebruik te maken van de multipele correlatierekening.

§ 21. Meervoudige correlatie.

Hierbij gaat het erom de veranderingen in een reeks x te verklaren uit de veranderingen in enige andere reeksen y, z, ... Ook hier is *grafische verkenning* zeer gewenst. De historische grafieken laten zich zonder moeilijkheden ook hier gebruiken. Alleen moet nu de vraag niet luiden of de schommelingen van x en y dezelfde zijn — eventueel na correcties voor trend of lags —, doch of de schommelingen van x kunnen worden opgebouwd uit een combinatie van de schommelingen in y, z... enz. Ofschoon algemenere methoden mogelijk zijn, zullen wij ons hier in hoofdzaak bepalen tot relaties die *additief* zijn en de vraag stellen of er een functie te vinden is:

$$x' = a y + b z + \ldots + a' y^2 + \ldots \qquad (21.1)$$

met constanten a, b, a^1, ... zodanig, dat de afwijkingen

tussen x en x' zo klein mogelijk zijn. Hier zijn nog termen van hogere graad opgenomen. Meestal zal men beginnen met zelfs die weg te laten en te proberen of een *lineaire* functie voldoende is:

$$x' = a y + b z + \ldots . \qquad (21.2)$$

Met behulp van de methode der historische grafieken kan men vaak reeds oppervlakkig zien of er kans is op een dergelijk verband. Dit toch wil zeggen (vgl. fig. 5) dat men de reeks x goed moet kunnen benaderen door een som van y, z, enz., waarbij deze laatste variabelen eventueel eerst nog in een bepaalde — vrij te kiezen — verhouding vergroot of verkleind zijn. Men moet dus de bijzonderheden in het verloop van x terugvinden, hetzij in y of in z of in beide.

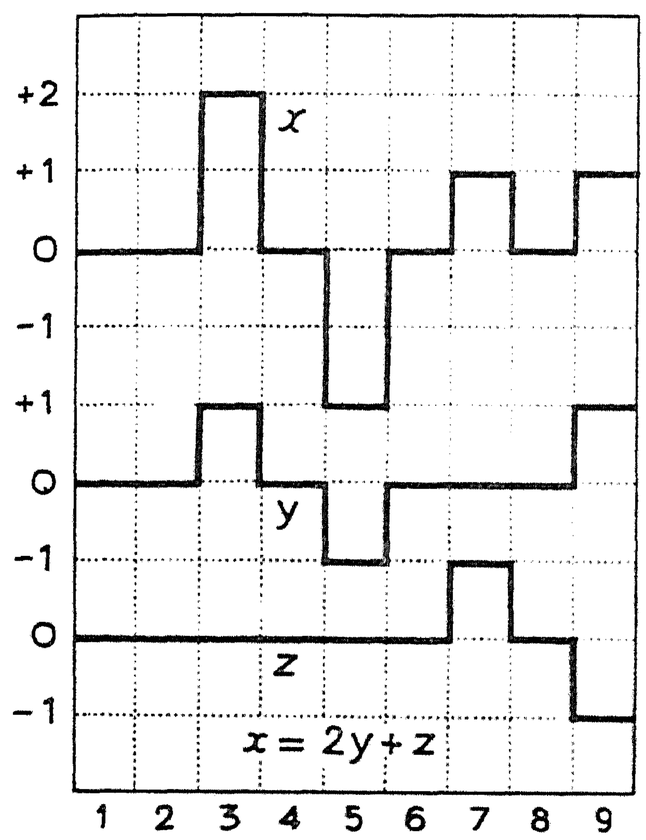

Fig. 5
Theoretisch voorbeeld van multipele correlatie. Men ziet dat de lijn x kan worden „opgebouwd" uit de lijnen y en z, wanneer men maar neemt $x = 2y + z$.

De methode van het spreidingsdiagram zou hier eigenlijk in de ruimte moeten worden toegepast. Als er een nauw verband is tussen de variabelen, moeten de punten in een oppervlak, eventueel een plat vlak, liggen. Het werken met een ruimtelijk diagram, dat overigens slechts zou kunnen worden gebruikt, wanneer er maar twee verklarende variabelen zijn, biedt echter practische moeilijkheden en men moet zich dus op andere wijze behelpen, zoals b.v. met de methode van Ezekiel. Wij verwijzen daartoe naar Ezekiel's boek en willen hier niet ingaan op d methoden tot het vinden van het beste verband la

deze weg, doch slechts op het uitbeelden daarvan, nadat men het langs de weg der berekening heeft bepaald.

Deze berekening geschiedt opnieuw met de methode van de kleinste kwadraten. Men bepaalt a, b, ... (en eventueel a' ...) zodanig, dat de som der kwadraten van de residuen

$$x - x'$$

een minimum is. Daarmee is dan weer de vergelijking van de *eerste regressielijn* verkregen. Men kan ook hier weer andere regressielijnen onderscheiden, nl. de 2e, 3e... tot en met n^{de}, wanneer er in totaal n variabelen en dus $n - 1$ verklarende variabelen zijn. De tweede wordt verkregen door een relatie

$$y' = c\,x + e\,z + \ldots \qquad (21.3)$$

te veronderstellen en de voorwaarde te stellen, dat de som der kwadraten van $y - y'$ minimaal is. (In de wiskundige literatuur over deze materie heeft men een systematische wijze van noteren van al de coëfficiënten die bij deze verschillende vergelijkingen optreden. Eenvoudigheidshalve en in overeenstemming met de opzet van dit werkje gaan wij hieraan nu voorbij). Wederom zal, wanneer de correlatie volledig is — d.w.z. de punten in een plat vlak liggen, in het geval van een lineaire regressievergelijking — het verband tussen x en de overige variabelen dat men uit de 2e regressievergelijking vindt, identiek zijn met dat uit de 1e.

In de vergelijking (21.2), die als regel zal worden gebruikt, omdat men gewoonlijk mag aannemen dat de afwijkingen vooral het gevolg zijn van een onvolledige theorie, noemen wij de term $a\,y$ „de invloed van y", $b\,z$ de invloed van z, enz. Deze term is onafhankelijk van de eenheden, waarin y is gemeten, omdat, wanneer y door een kleiner getal wordt weergegeven (ten gevolge van grotere eenheden), a automatisch in dezelfde verhouding groter wordt. Wanneer wij, om de gedachten te

bepalen, aannemen, dat het onderzoek betrekking heeft op een vraagcurve, waarbij dus
x de gevraagde hoeveelheid en
y de prijs is, terwijl er voorts nog voorkomen
z en u, twee vraagfactoren, dan zullen wij

$$x - bz - cu \tag{21.4}$$

de „voor de invloed der vraagfactoren gecorrigeerde waarde van x" noemen. Deze grootheid moet nu een enkelvoudige correlatie vertonen met de prijs y. Wij kunnen de aard van het verband grafisch weergeven door een spreidingsdiagram te maken, waarbij als coördinaten deze twee grootheden zijn gebezigd en spreken dan van een *partiëel spreidingsdiagram*. Soortgelijke spreidingsdiagrammen kan men maken voor elk der verklarende variabelen, dus ook b.v. tussen z en $x - ay - cu$, de voor y en u gecorrigeerde waarden van x.

Men kan van de vergelijking

$$x' = ay + bz + cu \tag{21.5}$$

ook nog een interpretatie geven, die de aansluiting bij de theorie der vraagcurven doet uitkomen. De termen $bz + cu$ kan men nl. de „shift" noemen. Deze naam sluit aan bij de gebruikelijke grafische voorstelling van de vraagcurve, waarbij men aan de „overige factoren", in dit geval z en u, constante waarden toekent en dan een verband tussen x en y overhoudt, dat grafisch wordt voorgesteld. Wanneer nu z en u andere waarden aannemen, betekent dat, dat de vraagcurve met zichzelf evenwijdig *verschoven* wordt, een „shifting" vertoont. Men kan deze voorstellingswijze ook nog in verband brengen met de voorstellingswijze in de ruimte, waarvan wij al eerder spraken. Gemakshalve beperken wij ons dan tot maar één vraagfactor z. Het verband $x = ay + bz$ wordt, zoals gezegd, door een hellend plat vlak voorgesteld. Geeft men nu aan z een vaste waarde, dan beschouwt men alleen de punten op het platte vlak die op één rechte lijn liggen. Neemt men voor z een andere waarde, dan

krijgt men de punten van een andere rechte lijn in dat vlak. De projecties van deze rechte lijnen op het xy-vlak zijn dan de vraaglijnen voor de verschillende waarden van z, de „shifting demand curves".

De bepaling van de regressie-coëfficiënten is in een principiëel belangrijk geval onmogelijk: nl. wanneer er een lineair verband bestaat tussen twee of meer der verklarende variabelen onderling. Het eenvoudigste geval is dat, waarbij de schommelingen in twee dier variabelen, b.v. y en z, evenredig zijn. In dat geval kan men voor z een zeker aantal keren y substitueren. Zou men een regressievergelijking hebben gevonden, dan zouden zich daaruit dus onmiddellijk willekeurig veel andere laten afleiden, die een precies even goed resultaat zouden opleveren. Het resultaat is dus niet ondubbelzinnig. Anders uitgedrukt, de invloeden van y en z zijn niet te scheiden.

Er zijn ook ingewikkelder gevallen mogelijk. Het kan zijn, dat y niet met z evenredig is, maar dat

$$y = pz + qu, \tag{21.6}$$

waarin p en q constanten zijn. Dan kan men y geheel of ten dele vervangen door het rechterlid dezer uitdrukking en daardoor het aantal mogelijke resultaten vergroten.

Men moet niet denken dat in dergelijke gevallen, waarin men zegt, dat er *multicollineariteit* bestaat, uit de berekening niets is te leren. Indien y en z evenredig zijn is weliswaar hun invloed niet te scheiden, doch die van u is wel te bepalen. Als (21.6) geldt, kan men weliswaar de invloed van geen der drie verklarende variabelen afzonderlijk vaststellen, doch bepaalde functies van die invloed weer wel. Het hangt van het uiteindelijke doel der berekeningen af, of men het resultaat kan gebruiken of niet.

§ 22. Beoordeling van de onzekerheid in de resultaten.

Uit twee oorzaken zijn de resultaten van een multipele

correlatieberekening niet volkomen nauwkeurig. Ten eerste zijn er toevallige afwijkingen, zich uitend in de spreiding van de punten en veroorzaakt, hetzij door meetfouten, hetzij door het verwaarlozen — in de theorie der onderzochte relatie — van een deel der factoren. Ten tweede is er het verschijnsel van de multicollineariteit of van een toestand die deze benadert. De twee oorzaken kunnen geheel los van elkaar optreden, doch zullen somtijds beide aanwezig zijn.

De onzekerheid in de resultaten, in het bijzonder in de regressiecoëfficiënten en ook in de correlatiecoëfficiënt, als gevolg van de aanwezigheid van toevallige afwijkingen, pleegt men uit te drukken door de grootte van de *standaarddeviatie* dezer grootheden. Deze geeft aan de standaarddeviatie die men — rekenend volgens de gewone definitie dezer uitdrukking — zou verkrijgen indien men niet één complex van waarnemingen had, maar een groot aantal soortgelijke complexen, waarbij telkens de toevallige afwijkingen andere zouden zijn geweest. Men stelt zich die complexen, evenals het ene in de werkelijkheid opgetreden complex, voor als *steekproeven* uit een groter denkbeeldig complex dat aangeeft „wat er even goed had kunnen gebeuren". Dit grotere complex blijft natuurlijk enigszins hypothetisch. Fisher, die de theorie heeft gegeven, op grond waarvan deze standaarddeviaties gewoonlijk worden berekend, heeft aangenomen dat de verklarende variabelen nauwkeurig bekend zijn, doch de te verklaren variabele een toevallige component bezit, die verdeeld is volgens de wet van Gauss. De gebruikelijke formules zijn uit deze premissen afgeleid.

Deze formules leveren uitkomsten, die zelf weer in hoge mate onzeker kunnen worden, wanneer het complex waarnemingen bij benadering multicollineair is, waarop gewezen is door Frisch. Deze heeft daarom getracht op andere wijze te komen tot een schatting van de onzekerheid in de resultaten. Om althans enkele hoofdzaken van zijn werkwijze, die bestaat uit de bepaling van zgn. *bunch maps* of *regressiediagrammen*, te begrijpen, gaan

wij uit van de stelling dat, wanneer er volkomen correlatie en geen multicollineariteit is, alle regressievergelijkingen tot hetzelfde resultaat leiden. Dat wil zeggen, dat de 2e regressievergelijking (21.3) bij oplossing naar x:

$$x = \frac{y}{c} - \frac{e}{c} z \ldots \qquad (22.1)$$

coëfficiënten oplevert, die dezelfde zijn als in de eerste:

$$x = a y + b z + \ldots$$

Dus: $$a = \frac{1}{c}, \quad b = -\frac{e}{c}, \text{ enz.}$$

Wanneer er daarentegen bij benadering multicollineariteit heerst, zal men grote afwijkingen vinden tussen de op de ene en op de andere manier bepaalde beïnvloedingscoëfficiënten. Regressiediagrammen (zie fig. 6)

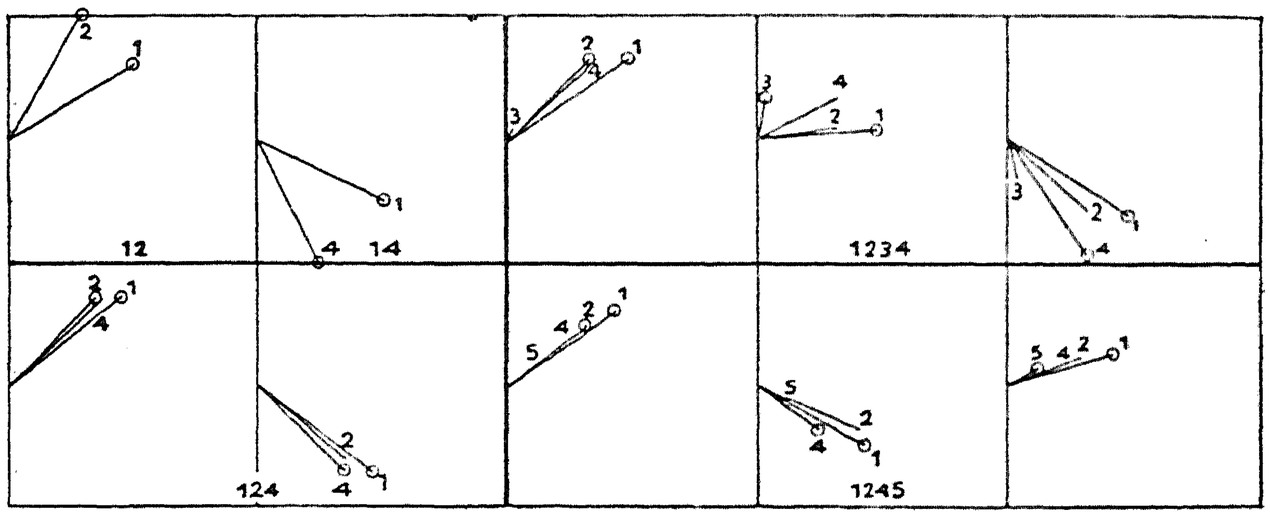

Fig. 6.
Voorbeeld van een regressiediagram (bunch map) (voor de betekenis vergelijke men de tekst).

geven nu een overzicht van de verschillende waarden, die vóor eenzelfde beïnvloedingscoëfficiënt worden verkregen. Deze waarden zijn daarin aangegeven door de helling van lijnen uit de oorsprong of „stralen". Deze stralen dragen een nummer. Elk nummer geeft een variabele aan. Twee der stralen zijn altijd op bijzondere wijze aangegeven; het zijn de hoofdstralen. De twee nummers

die daarbij staan geven de variabelen aan, waartussen de beïnvloedingscoëfficiënt bepaald is; het onderste nummer de als gevolg beschouwde, het bovenste nummer de als oorzaak beschouwde variabele. De nummers geven voorts aan het nummer van de regressievergelijking, waaruit de bij elke straal behorende waarde van de beïnvloedingscoëfficiënt is berekend, of de richting, waarin de residuen tot een minimum zijn gemaakt. Het aantal stralen is gelijk aan het aantal variabelen, dat in de beschouwing is opgenomen.

De methode van Frisch bestaat nu in het bijzonder daarin, dat hij dat aantal variabelen geleidelijk uitbreidt. Wanneer een nieuwe variabele toegevoegd wordt, kan men dus in zijn regressiediagrammen waarnemen, welke invloed dat heeft op de stralenbundels. Voegt men een variabele toe, die voor de verklaring vanuit statistisch oogpunt essentieel is, d.w.z. ten naaste bij de nog onverklaarde residuen verklaart, dan zal men zien, dat de stralenbundels zich sluiten. Is de variabele die men toevoegt echter bijna een lineaire combinatie der reeds aanwezige variabelen — of hij nu essentiëel vanuit economisch oogpunt is of niet — dan zullen de bundels „exploderen": de typische onbepaaldheid in de regressiecoëfficiënten treedt op. Men kan dus waarnemen welke nieuw toegevoegde variabelen „nuttig", welke „schadelijk" zijn. Er is een tussenklasse van „overbodige" variabelen, die weinig invloed hebben. De statisticus kan alleen „instaan" voor regressiecoëfficiënten, waarvan de bundel een gesloten beeld toont. Frisch laat zien, dat onder bepaalde veronderstellingen de onzekerheidsmarge kan worden gemeten door de afstand der uiterste aldus berekende waarden, d.w.z. de hellingen der uiterste stralen van de bundel.

De regressiediagrammen kunnen ons nog meer leren, doch dat zou ons in wiskundige details voeren.

Voor één verkeerde wijze van gebruik der bunch maps moet worden gewaarschuwd. Soms denkt men, dat het exploderen van de stralenbundel bij het toevoegen van

een nieuwe variabele een teken is, dat deze variabele niet moet worden opgenomen. Daarover kan echter slechts de economische theorie, niet de regressie-analyse beslissen. Het kan dan zeer goed, dat volgens de economische theorie de betrokken variabele essentiëel is. De conclusie moet dan luiden: jammer genoeg kunnen wij niet alle beïnvloedingscoëfficiënten bepalen, omdat — toevallig of niet — enige der essentiële variabelen onderling gecorreleerd waren. Zoals reeds gezegd, hangt het dan van het probleem in kwestie af, of men toch nog nuttige conclusies kan trekken.

T. Koopmans heeft de elementen van de theorieën van Fisher en van Frisch verenigd tot één theorie, die echter een gecompliceerd wiskundig karakter heeft en zich moeilijk laat parafraseren. Daarom zij naar de literatuur verwezen.

§ 23. Het gebruik van de regressie-analyse in de econometrie.

De wijze waarop de econometrist van de in het voorgaande beschreven methoden gebruik maakt is op systematische wijze door Koopmans ongeveer aldus beschreven. [1]) De statisticus kan bij de toetsing van een door de economische theorie opgestelde relatie onvoorwaardelijke en voorwaardelijke (van zekere voorwaarden afhankelijke) conclusies trekken. De *onvoorwaardelijke* zijn echter slechts *negatief*. Deze conclusies kunnen bij de controlerende functie optreden. Indien nl. de correlatie blijkt te laag te zijn (in vergelijking tot de gestelde standaard), dan luidt de conclusie: de relatie is met de feiten in strijd. Anders gezegd: de theorie is onjuist, of althans onvolledig. De positieve conclusies die men kan trekken komen bij het controleren der theoretische relatie niet voor: men kan niet statistisch *bewijzen*, dat een theorie juist is. Al is de correlatie hoog, andere theorieën leiden

[1]) The Logic of Econometric Business Cycle Research, Journal of Political Economy 1940.

misschien ook tot een hoge correlatie. Positieve conclusies komen wel voor bij de metende functie, maar dan zijn ze *voorwaardelijk*; b.v. aldus: als de aangenomen relatie juist is, wat betreft de keuze der variabelen en (eventueel) de lags, dan is de meest waarschijnlijke waarde van de beïnvloedingscoëfficiënten zo- en zoveel.

Wil men dus tot positieve conclusies komen, dan is *additionele informatie* nodig om te komen tot de voorwaarden zojuist genoemd; deze is afkomstig van de economische theorie. Het is redelijk dat deze additionele informatie aan de volgende principes is onderworpen:

(a) het principe van *statistische censuur*: zij mag niet in strijd zijn met de gegevens;
(b) het principe van *wetenschappelijke efficiëntie*: zij moet geen gevolg zijn van de reeds aangenomen stellingen en de gegevens tezamen;
(c) het principe van de *solide basis*: zij moet zo plausibel mogelijk zijn; en
(d) het principe van de *voldoende basis*: zij moet zo zijn, dat men positieve conclusies kan trekken. Tussen deze laatste twee principes kan veelal een zekere spanning bestaan en men moet trachten of aan beide kan voldaan worden.

In de meeste econometrische onderzoekingen met behulp der multipele correlatie luidt de additionele informatie ongeveer aldus:

(i) De niet in de relatie opgenomen factoren geven slechts aanleiding tot kleine toevallige afwijkingen of tot
(ii) een trendinvloed, dan wel tot
(iii) afwijkingen in individuele jaren, die buiten beschouwing worden gelaten.
(iv) De wel in de relatie opgenomen factoren (verklarende variabelen) werken op een gegeven wiskundige wijze (vaak lineair of slechts additief), terwijl
(v) de lags gegeven zijn of althans positief moeten zijn en

(vi) bepaalde regressiecoëfficiënten ook een voorgeschreven teken hebben; soms is de waarde dier coëfficiënten aan bepaalde grenzen gebonden.

De laatste drie punten vereisen nog enige nadere toelichting. Zoals men weet, is de beperking in (iv) belichaamd niet ernstig, wanneer de afwijkingen die de variabelen van hun gemiddelde vertonen niet groot zijn. De lags behoeven niet van tevoren gegeven te zijn; men kan ze ook onbepaald laten en door de eis van zo goed mogelijke aansluiting tussen x en x' laten bepalen. Dit kan door eenvoudig proberen of ook wel door in een relatie met als verklarende variabelen y en z niet alleen deze variabelen voor hetzelfde tijdstip op te nemen, doch ook voor andere, b.v.

$$x' = a y + a' y_{-1} + b z + b' z_{-1}.$$

De relatieve grootte der regressiecoëfficiënten a en a' leert dan, hoe groot men (bij benadering) de lag moet nemen tussen x en y. Is b.v. $a = 2$ en $a' = 1$, dan is met grote benadering — wanneer althans de tijdseenheid klein is t.o.v. de periode der schommelingen in y —:

$$x' = (a + a')\, y_{-\frac{1}{3}} + \ldots$$

Punt (vi) is het belangrijkst. Bepaalde regressiecoëfficiënten moeten in verband met hun economische betekenis een bepaald teken of een bepaalde waarde hebben. Zo zal de beïnvloedingscoëfficiënt van de prijs in een vraagvergelijking negatief, in een aanbodvergelijking positief moeten zijn. Een regressiecoëfficiënt die de betekenis heeft van b.v. de loonquote in de productie of van de marginale „propensity to consume" zal beneden 1 moeten liggen.

Veelal zal men beginnen zulke eisen niet te stellen, doch de regressievergelijking te berekenen.

Zoals reeds vroeger werd medegedeeld, zijn de volgende uitkomsten mogelijk: I) de correlatie is lager dan de gestelde standaard; de theorie wordt verworpen; II) de correlatie is goed en de regressiecoëfficiënten hebben een kleine onzekerheidsmarge; dan is meting mogelijk met

betrekkelijk weinig voorwaarden; III) de correlatie is goed, maar de regressiecoëfficiënten of lags zijn betrekkelijk onzeker. In dit geval kan men veelal door verdere voorwaarden van de soort bij (v) en (vi) genoemd het gebied van de onzekerheid beperken. Men kan de waarde van een of meer der lags of regressiecoëfficiënten vastleggen of gedeeltelijk vastleggen.

Het meeste succes is natuurlijk verkregen, als dergelijke additionele informatie niet nodig is. Men kan dan soms toch nog betrekkelijk onverwachte conclusies bereiken, b.v. dat de invloed van een der factoren klein is. Uit de onderzoekingen van schrijver dezes, door Koopmans besproken, is b.v. gebleken dat bij bepaalde later te noemen veronderstellingen (zie § 36) de invloed van de rentestand op de investeringsbedrijvigheid klein is; eveneens de invloed van het z.g.n. acceleratieprincipe.

Als men succes heeft, kan men trachten om een der voorwaarden te verzachten; b.v. de wiskundige vorm generaliseren, of meer factoren in de relatie opnemen. Men kan de premissen waarvan oorspronkelijk was uitgegaan, verdelen in *aanvaardbare* en *twijfelachtige* premissen en nu alleen de aanvaardbare in de berekening opnemen. Ten aanzien van de twijfelachtige kan men nu tot de volgende conclusies komen:

A) deze zijn fout;
B) ze zijn goed, gegeven de aanvaardbare premissen;
C) het is niet uit te maken of ze goed zijn of niet, maar als men ze weglaat, is meting onmogelijk of
D) idem, maar ze zijn irrelevant voor de meting.

Conclusies van de soort B zijn van een type, dat bij de eerste opzet niet aanwezig was.

Als voorbeeld van een twijfelachtige premisse kan men de veronderstelling nemen dat een der regressiecoëfficiënten moet liggen binnen een gegeven interval.

A) doet zich nu voor, als er geen goede correlatie is, waarbij die coëfficiënt in dat interval ligt;
B) als er alleen goede correlaties zijn, waarbij dat zo is;

C) en D) als er goede correlaties voorkomen zowel met waarden binnen als buiten dat interval.

Een ander voorbeeld is: aangenomen wordt, dat y invloed heeft op x, maar twijfelachtig is of ook z invloed heeft.

A) doet zich voor, als de regressiecoëfficiënt van z ongeveer nul is;
B) als de correlatie met z goed en zonder z slecht is;
C) en D) wanneer de onzekerheidsmarges in de coëfficiënt voor z groot zijn.

Andere situaties, voor positieve uitspraken van den statisticus gunstiger, doen zich voor, wanneer de economische theorie leidt tot de opstelling van alternatieve hypotheses voor de twijfelachtige; dus wanneer de econoom zegt: het is niet zeker of z invloed heeft, maar als z geen invloed heeft, dan heeft u invloed.