IV. INFORMETRIC MODELS

IV.0. INTRODUCTION

It is part of human nature to try to make models of the many phenomena in this world in order to predict certain activities. Humans are likewise endeavouring to explain these models. While these facts are clear for the exact sciences, where a start was made hundreds of years ago, the start in the social sciences is of a relatively more recent age. The discipline in this area that has developed the best so far is undoubtedly econometrics. Other well-known disciplines are psychometrics and quantitative linguistics.

Even more recently one has started to model aspects in library and documentation sciences. No single paper has been written in this field before the 20^{th} century. Apart from some historic works, the development of informetric models was only started a few decades ago.

The reader is advised to read the introduction of this book for an account of recent and fluctuating terminology.

IV.1. HEURISTIC REFLECTIONS ON INFORMETRIC MODELS AND HISTORICAL EXAMPLES

IV.1.1. General approach

An approach to the application of mathematics to the empirical sciences has been propounded by Stefan Körner (1969). He suggested that three steps are needed :

1. Inexact empirical concepts have to be replaced by exact mathematical concepts.
2. Exact conclusions are then deduced from these mathematical concepts.
3. The exact mathematical conclusions are then replaced by empirical concepts.

We will adopt this approach in this part on informetric models.

IV.1.2. Information Production Processes (IPP). Sources and items

We will use the concept of '*information production process*' *(IPP)* in which there are two kinds of entities : the *sources* and the *items* produced by these sources. Exact definitions follow in subsequent sections. Let us give some examples.

1. In econometrics we can give the example of a group of workers or employees and study their productivity (Theil (1967)). Productivity can be measured in several ways : in terms of quantity (numbers of produced items), in terms of quality or in terms of profits (amount of money earned in a certain time period). In this example, the choice of the term 'production' is quite clear. More generally speaking, the examples which follow can also be considered as information production processes.

2. In demography one considers cities and villages in conjunction with their populations.

3. In linguistics one considers words (as entities or 'types', a term often used in linguistics (Herdan (1960))) and their occurrences (or 'tokens' in linguistic terms) in a given text (book, article, ...) (Zipf (1949)).

4. In bibliometrics one can study books in a library and the number of times they are loaned out, say, in a year (Burrell and Cane (1982)).

5. One can also study a group of researchers and the number of publications they produce, say in a ten-year period (Lotka (1926)).

6. In bibliometrics, one can consider a bibliography (on a specified topic), in which the contributing journals produce papers (Bradford (1934)).

7. Papers themselves can be considered as sources rather than as items in the previous case. Thus, a set of papers can be considered together with the citations they receive within a designated time period (Garfield (1979a). In this connection, there is a 'cited' relationship between papers. An

interesting point is the well-known fact that another example can be constructed when 'cited' is changed into 'citing'.

IV.1.3. Empirical laws and corresponding mathematical functions

The regularity that is the simplest to be introduced is Lotka's law.

IV.1.3.1. Lotka's law

In 1926, A.J. Lotka (1926) studied a 10-year Cumulative Index of Authors listed in Chemical Abstracts (1907-1916) and Auerbach's 'Geschichtstafeln der Physik' (1910).

He found the following regularity : if f(j) denotes the number of authors with j publications, then

$$f(j) = \frac{C}{j^{\alpha}} \quad , \qquad [IV.1.1]$$

where $\alpha \approx 2$, but not necessarily $\alpha = 2$.

If $\alpha = 2$, then, since $\sum_{j=1}^{\infty} \frac{1}{j^2} = \frac{\pi^2}{6}$ (Euler's theorem, see e.g. Patterson (1986)),

$$C = \frac{6}{\pi^2} T \approx 0.6079\ T \quad , \qquad [IV.1.2]$$

where T denotes the total number of authors. Function [IV.1.1] will be called the *Lotka* function, as it expresses Lotka's law.

The other empirical laws all relate to rankings of the IPP.

IV.1.3.2. A ranking

Henceforth, we will assume the following ranking on the sources of an IPP : the most productive source is assigned rank 1, the second most productive source rank 2, and so on. The last rank (T) is for the source with the least production; ties are broken arbitrarily (see also the next chapter for a more accurate description). All the following laws use this ranking and are therefore called 'rank-order distributions'.

IV.1.3.3. Zipf's and Mandelbrot's laws

Formulated originally in linguistics, *Zipf's law* can be expressed as (Zipf (1949)) : Order the words in a text in decreasing order of occurrence in this text. Then the product of the rank r of a word and the number of times j it is used in the text is a constant for that text :

$$r.j = E \qquad [IV.1.3]$$

or, taking $j = g(r)$:

$$g(r) = \frac{E}{r} . \qquad [IV.1.4]$$

In more general terms, one can formulate the following *general Zipf function* :

$$g(r) = \frac{F}{r^{\beta}} , \qquad [IV.1.5]$$

where F and β are constants.

Mandelbrot's law (Mandelbrot (1954,1977)) has been derived from the same context, but has an expression different from [IV.1.5] :

$$g(r) = \frac{G}{(1+Hr)^{\beta'}} , \qquad [IV.1.6]$$

where G, H and β' are constants.

IV.1.3.4. Pareto's law

This law is formulated in econometrics (Theil (1967)). It states that the number h(j) of employees with an income larger than or equal to j is

$$h(j) = \frac{L}{j^{\gamma}} , \qquad [IV.1.7]$$

where L and γ are constants. Clearly, when Subsection IV.I.3.3 and the above equation are combined (with an obvious unification of the terminology), we get :

$$r = h(j) = \frac{L}{j^{\gamma}}$$

or

$$j = \frac{L^{1/\gamma}}{r^{1/\gamma}}$$

and hence

$$g(r) = \frac{L^{1/\gamma}}{r^{1/\gamma}} . \qquad [IV.1.8]$$

In conclusion, the *Pareto function* and the Zipf function are identical though their respective laws apply to different contexts. This kind of identity is

another issue to be considered in the general context of informetrics.

IV.1.3.5. Leimkuhler's law

Consider a bibliography of papers on a specific topic, published in journals. Using the order of Subsection IV.1.3.2 and denoting by F(x) the cumulative fraction of papers in the journals of rank 1,2,...,r, where $x = \frac{r}{T}$, the cumulative fraction of the journals, we have :

$$F(x) = \frac{\log (1 + \delta x)}{\log (1 + \delta)} \quad , \qquad [IV.1.9]$$

where δ is a constant (Leimkuhler (1967)). Henceforth we shall work with the function R(r) = F(x).A (where A = the total number of papers and r = x.T). We thus have the following *Leimkuhler function* (equivalent to equation [IV.1.9]) : Let R(r) denote the cumulative number of items in the journals of rank 1,2,...,r. Then

$$R(r) = a \log (1 + br) \quad , \qquad [IV.1.10]$$

where a and b are constants.

IV.1.3.6. Bradford's law

The most intriguing of all the empirical laws is that of *Bradford* (1934), based on observations of bibliographies in Applied Geophysics, 1928-1931 (incl.) (cf. Table I.1.1) and Lubrication, 1931-June 1933.

We present it here in its original definition which is, as far as we can see, clear enough. We remark, however, that some informetrists have been confused by its formulation, giving rise to what is now known as the 'verbal' and the 'graphical' formulation of Bradford's law (which are not exactly equivalent). We present here the original 'verbal' version. It states :

> Order the journals in decreasing order of the number of papers (in this bibliography) they contain. If the journals are subdivided into p groups (according to the above order) such that each group of journals contains the same number y_0 of papers in this bibliography, then there exist r_0 and $k > 1$ such that the first group has r_0 journals, the second has $r_0 k$ journals, the third has $r_0 k^2$ journals and so on, until the last (p^{th}) group, contains $r_0 k^{p-1}$ journals.

Stated otherwise, if p is a strict positive integer (in short : $p \in \mathbb{N}$),

then there exist $r_0 \in \mathbb{N}$ and $k > 1$ (a real number) such that the first (most productive) r_0 journals produce $y_0 = \frac{A}{P}$ (where A = total number of papers) papers, the next $r_0 k$ journals again produce y_0 papers, the subsequent $r_0 k^2$ journals also produce y_0 papers, and so on, until the last (least productive) $r_0 k^{p-1}$ journals again produce y_0 papers.

One aspect of this formulation is the kind of symmetry between the journals and the papers. If we represent the bibliography and its order by a straight line (or better, an axis with the ranks of the journals as coordinates, see Fig.IV.1.1), we intuitively feel that, when going from left to right, the 'visibility' of the journals is changed into the 'visibility' of the papers.

Fig.IV.1.1 The rank axis

Later on in Chapters IV.3 and IV.4, we will delve deeper into the formal structure of IPP's and the place of the empirical laws in them. In the next section we will present a few 'explanations' of the empirical laws.

IV.2. EXPLANATIONS OF INFORMETRIC LAWS

We will present three - formally different - *explanations* of the empirical laws. We will first discuss the probabilistic theory of H.A. Simon and D. De Solla Price : the success breeds success principle (Simon (1955), Price (1976)), leading to a form of Lotka's law with exponent $\alpha = 2$. We will then deal with the analytic arguments of Bookstein (1977,1984) yielding the general Lotka law. Finally, we will give the combinatorial fractal argument of Mandelbrot (1977) that results in the law of that name.

In the fourth chapter we will show the equivalence of several informetric laws. Consequently, once one of these laws has been 'explained', the others will be too.

The reader should keep this in mind when reading this section.

IV.2.1. The success-breeds-success principle

The *success-breeds-success principle*, developed originally in Simon (1955) but simplified in Price (1976) is a probabilistic argument on the growth chances of sources in IPP's (Of course, neither Price nor Simon use the IPP terminology, but their arguments are general enough to make application to IPP's feasible; if the reader prefers a more concrete approach to IPP's, think of bibliographies).

IV.2.1.1. The success-breeds-success principle - intuitively

This principle, abbreviated SBS, states that 'the more items a source has, the greater the probability will be that this source will produce another item; still there is always a (small) probability that a source with no items will produce a first item'.

This should be intuitively clear. We will translate this principle into a mathematical formalism (cf. Price (1976), although the argument presented here is more 'rigorous').

IV.2.1.2. SBS - mathematically

Consider a set S of sources with T elements (items) from which there is a fraction $\varphi(T,j)$ in state j (i.e. a fraction $\varphi(T,j)$ has j items). Hence $\sum_{j=1}^{\infty} \varphi(T,j) = 1$ and $\mu = \sum_{j=1}^{\infty} j.\varphi(T,j)$ is the average number of items per source (at this stage of T items). Note that, in the notation of equation [IV.1.1] :

$$f(T,j) = f(j) = T\varphi(T,j) .$$

The mathematical translation of the SBS is as follows.

We add $h > 0$ new sources to the system and hence the number of sources has increased from T to T+h. Let $\mu(T)$ and $\mu(T+h)$ be the average number of items per source in respectively the populations of size T and T+h. The total increase of items is therefore

$$\mu(T+h)(h+T) - \mu(T)T \ .$$

The SBS now states that these items are to be attributed to the sources according to the number of items they already have (except for the h new sources : they keep 1 item). So $\mu(T+h)(h+T) - \mu(T)T-h$ new items are sprinkled evenly over $\mu(T)T$ (old) items (i.e. unevenly over the sources to which these items belong). Consequently, there are

$$\eta(T,h) = \frac{\mu(T+h)(h+T) - \mu(T)T-h}{\mu(T)T}$$

new items per old item and hence, in the class of sources with j items (where j is fixed, momentarily), there are

$$j\,T\varphi(T,j)\ \eta(T,h) = j\,f(T,j)\ \eta(T,h)$$

new items and hence (assuming that sources do not grow by 2 or more items in one state change $T \to T+h$), there are

$$j\,f(T,j)\ \eta(T,h)$$

transitions from state j to state j+1.
Consequently, there are $j\,f(T,j)\ \eta(T,h)$ transitions *out* of state j and $(j-1)\ r(T,j-1)\ \eta(T,h)$ transitions *into* state j $(j > 1)$ (if $j = 1$, then we have h entries, as assumed). Therefore, we have the following system :

$$f(T+h,j) - f(T,j) = -j\,f(T,j)\ \eta(T,h) + (j-1)\ f(T,j-1)\ \eta(T,h)$$

if $j > 1$, and

$$f(T+h,1) - f(T,1) = -f(T,1)\ \eta(T,h) + h$$

if $j = 1$.
Consequently, using the form of $\eta(T,h)$,

$$\frac{f(T+h,j) - f(T,j)}{h} \begin{cases} = (-j\ f(T,j) + (j-1)\ f(T,j-1)) \cdot \dfrac{\mu(T+h)(h+T) - \mu(T)T - h}{h\mu(T)T} & (j > 1), \\ = -f(T,1) \cdot \dfrac{\mu(T+h)(h+T) - \mu(T)T - h}{h\mu(T)T} + 1 & (j = 1) . \end{cases} \quad [IV.2.1]$$

Taking the limit for h going to zero ($h > 0$) gives (again using $f(T,j) = T\varphi(T,j)$) :

$$\frac{\partial[T.\varphi(T,j)]}{\partial T} \begin{cases} = (-j\ \varphi(T,j) + (j-1)\ \varphi(T,j-1)) \left(\dfrac{\mu'(T)T}{\mu(T)} + \dfrac{\mu(T) - 1}{\mu(T)}\right) & (j > 1) \\ = -\varphi(T,1) \left(\dfrac{\mu'(T)T}{\mu(T)} + \dfrac{\mu(T) - 1}{\mu(T)}\right) + 1 & (j = 1) \end{cases} \quad [IV.2.2]$$

We simplify these difference-differential equations by assuming $\mu'(T) = \frac{1}{T}$ for every T. This requirement is acceptable as a first approximation since, as T usually is high, $\mu'(T) = \frac{1}{T} \approx 0$. Then [IV.2.2] becomes :

$$\frac{\partial[T.\varphi(T,j)]}{\partial T} \begin{cases} = -j\ \varphi(T,j) + (j-1)\ \varphi(T,j-1) & (j > 1) , \\ = -\varphi(T,1) + 1 & (j = 1) . \end{cases} \quad [IV.2.3]$$

Now

$$\frac{\partial[T.\varphi(T,j)]}{\partial T} = \varphi(T,j) + T\ \frac{\partial[\varphi(T,j)]}{\partial T} \quad .$$

Hence [IV.2.3] now results in

$$T\ \frac{\partial[\varphi(T,j)]}{\partial T} \begin{cases} = -(j+1)\ \varphi(T,j) + (j-1)\ \varphi(T,j-1) & (j > 1) , \\ = -2\ \varphi(T,1) + 1 & (j = 1) . \end{cases} \quad [IV.2.4]$$

which is a system of difference-differential equations.

First approximation

To solve system [IV.2.4] in a simple case, we suppose that function φ is independent of T (i.e. that the relative fractions of sources with j items remain the same). Hence, in this case, [IV.2.4] yields (writing $\varphi(j) = \varphi(T,j)$):

$$\begin{cases} (j+1)\,\varphi(j) = (j-1)\,\varphi(j-1) & (j > 1) \ , \\ \varphi(1) = \frac{1}{2} & (j = 1) \ , \end{cases} \qquad \text{[IV.2.5]}$$

which is easy to solve recursively.
For every $j \geq 1$:

$$\varphi(j) = \frac{j-1}{j+1}\,\frac{j-2}{j}\,\frac{j-3}{j-1} \cdots \frac{1}{3} \cdot \frac{1}{2}$$

$$\varphi(j) = \frac{1}{j(j+1)} \ . \qquad \text{[IV.2.6]}$$

Although this is not exactly Lotka's inverse square law (equation [IV.1.1] for $\alpha = 2$ where $f = \varphi T$), it resembles strongly when the simplified model is taken into account. Note that

$$\sum_{j=1}^{\infty} \varphi(j) = \sum_{j=1}^{\infty} \frac{1}{j(j+1)} \ ,$$

$$= \lim_{\ell \to \infty} \sum_{j=1}^{\ell} \frac{1}{j(j+1)} \ ,$$

$$= \lim_{\ell \to \infty} \left(\sum_{j=1}^{\ell} \frac{1}{j} - \sum_{j=1}^{\ell} \frac{1}{j+1} \right) \ ,$$

$$= \lim_{\ell \to \infty} \left(1 - \frac{1}{\ell+1}\right) = 1 \ .$$

Thus

$$\sum_{j=1}^{\infty} \varphi(j) = 1 \ . \qquad \text{[IV.2.7]}$$

D. de Solla Price then continues by refining (generalising) φ :

Second approximation

In a generalisation of the above, φ is supposed to be a separable function of j and T, i.e. φ is the product of a function F of T (not of j) and a function ψ of j (not of T) :

$$\varphi(T,j) = F(T).\psi(j) \ . \qquad \text{[IV.2.8]}$$

Then, equations [IV.2.4] result in (the proof is omitted here) :

$$\varphi(T,j) = (m-1).B(j,m) \quad , \qquad [IV.2.9]$$

where m is a constant and where B denotes the classical beta function. Function [IV.2.9] is called the Cumulative Advantage Distribution and is shown to reasonably approximate Lotka's law [IV.1.1].

In the authors' opinion, the reasoning (more than the outcome) of the SBS (which must be credited to Simon), is original and interesting. Note that the SBS is a dynamic principle, even though we did not use a time parameter t. The time evolution is 'hidden' in the fact that one studies the growth of the sources from T to T+h ($h \to 0$). The resulting distribution is, however, independent of time. Even nowadays, not much has been done on the study of time-dependent IPP's. See Section IV.8.8 for a few notes about time-dependent IPP's.

IV.2.2. The function-analytic arguments of Bookstein

IV.2.2.1. Main argument

Let, as above, $\varphi(j)$ denote the fraction of the sources with j items (the variable T has been omitted). Of course, if the function φ must be applicable to all kinds of situations (e.g. bibliographies on different subjects, different time periods, etc.), adaptable parameters must occur in φ. As a simplification of this fact, *Bookstein* supposes that

$$\varphi(j) = B.\xi(j) \quad , \qquad [IV.2.10]$$

where $B > 0$ is a time-adaptable parameter and ξ is a fixed function (with respect to time) but might still contain a parameter to cope with other variable aspects (such as different subjects). By taking

$$\varphi(j) = B\ \xi(j) = B\ \xi(1)\ \frac{\xi(j)}{\xi(1)} = B_1\ \eta(j)$$

and re-using the notation B and ξ, we may assume that $\xi(1) = 1$.

We now use the following function analytic principle (replacing, in a way, the SBS principle of the previous section) : the fraction of the sources with j items in an IPP over a time period of length t is equal to the fraction of the sources with jq items in an IPP over a time period of length tq.

Adopting this principle (where B and B' denote the different time-

dependent parameters, with the function ξ being fixed), gives :

$$B\ \xi(j) = B'\xi(tj) \quad . \qquad \text{[IV.2.11]}$$

For $j = 1$, using $\xi(1) = 1$, this produces

$$B = B\ \xi(1) = B'\xi(t) \quad . \qquad \text{[IV.2.12]}$$

Combining [IV.2.11] and [IV.2.12] yields

$$B'\xi(t)\ \xi(j) = B'\xi(tj)$$

Therefore, since $B' > 0$,

$$\xi(t)\ \xi(j) = \xi(tj) \qquad \text{[IV.2.13]}$$

for every $t > 0$ (or, if you wish, $t \in \mathbb{N}$) and $j \in \mathbb{N}$.

This is a functional equation (see e.g. Kuczma (1985)) that can easily be solved by assuming ξ (and hence φ and f) to exist on $j \in [1,\infty[$ and to be differentiable on this interval (a quite natural supposition). The argument is as follows : let $h > 0$ be arbitrary. Hence [IV.2.13] implies

$$\xi(j+h) = \xi(j(1+\frac{h}{j})) \quad ,$$

$$= \xi(j)\ \xi(1 + \frac{h}{j}) \quad .$$

Hence also

$$\frac{\xi(j+h) - \xi(j)}{h} = \frac{\xi(j)\ [\xi(1 + \frac{h}{j}) - 1]}{h} \quad ,$$

$$= \frac{\xi(j)}{j} \cdot \frac{\xi(1 + \frac{h}{j}) - \xi(1)}{\frac{h}{j}} \quad .$$

Taking the limit for $h \to 0$ yields

$$\xi'(j) = \frac{\xi(j)}{j}\ \xi'(1) \quad .$$

Thus

$$\xi'(j) + D\frac{\xi(j)}{j} = 0 \quad , \qquad \text{[IV.2.14]}$$

where D is a constant.
Equation [IV.2.14] is a linear first-order homogeneous differential equation and is solved as

$$\xi(j) = Ej^{-\beta} \quad ,$$

where E is an arbitrary positive constant and $\beta = D = -\xi'(1)$.
Therefore, using [IV.2.10] and taking $\alpha = \beta$, we find equation [IV.1.1], Lotka's function $f = T\varphi$,

$$f(j) = \frac{C}{j^{\alpha}} \quad , \qquad \text{[IV.1.1]}$$

where $C = TBE\ (> 0)$.

Lotka's law [IV.1.1] was introduced earlier, in Subsection IV.1.3.1 for IPP's consisting of authors and their publications. In the past one has often been faced with the problem of *multiple authorship*. In formulating Lotka's law, should one
(I) count only the first author,
(II) count all authors,
(III) assign weights per author?
(cf. Subsections III.2.4.2 and III.7.2, where the same problem has been studied in relation to citations).

In one and the same situation, one should clearly stick to one of the above methods. But which one must be chosen? While Bookstein did not answer this question, he did show that, no matter how the authors are counted, a law of the form [IV.1.1] will always result. In other words our choice of author counting cannot destroy a Lotka-type law. This will be demonstrated below.

IV.2.2.2. Author counts

As in the previous section, we assume $j \in [1,\infty[$. Also, to simplify the problem, we again take

$$\varphi(j) = B.\xi(j) \quad , \qquad \text{[IV.2.15]}$$

where $B > 0$ is now a parameter dependent on the choice of the author weights and ξ is a function independent of the choice of the author weights (to be defined below).

Any of the above situations (I), (II) or (III) can be described by the following unique formalism : let the total number of articles (items) be A and let $v_i \in [0,1]$ be the weight of a given author in article $i \in \{1,\ldots,A\}$. Hence the 'number of articles' published by this author is

$$j = \sum_{i=1}^{A} v_i = \sum_{i=1}^{A} \frac{v_i}{A} A =: rA \quad . \qquad [IV.2.16]$$

Consider next a second weighting system for authors, expressed by weights $v_i' \in [0,1]$ as above. This same author now receives a 'number of articles'

$$j' = \sum_{i=1}^{A} v_i' = \sum_{i=1}^{A} \frac{v_i'}{A} A =: r'A \quad . \qquad [IV.2.17]$$

One can, of course, assume that $r' \geq r$ (or otherwise, interchange the v_i and v_i') . Consequently

$$A = \frac{j}{r} = \frac{j'}{r'}$$

Hence

$$j' = \frac{r'}{r} j \quad .$$

If we write $\theta = \frac{r'}{r}$, then

$$j' = \theta j \quad , \qquad [IV.2.18]$$

where $j' \in [1,+\infty[$, since $\theta \geq 1$.
Of course θ is dependent on the author. Let η be the (unknown) distribution of θ over the authors (we assume here that θ is a continuous variable, being an approximation in the case of a large group of authors).

Based on [IV.2.15], let $B\,\xi(j)$ be the fraction of authors with j articles (measured according to the first method) and let $B'\xi(j)$ be the fraction of authors with j articles (measured according to the second method).

Any 'production' j' measured according to the second method is obtained, per θ, of a production $\frac{j'}{\theta}$ measured according to the first method (indeed, consider equation [IV.2.18]).Hence, for every $j' \in [1,\infty[$:

$$B'\xi(j') = \int B\,\xi\left(\frac{j'}{\theta}\right) \eta(\theta)\, d\theta \quad . \qquad [IV.2.19]$$

Taking $j' = 1$, we see

$$B' = B'\xi(1) = \int B\ \xi(\tfrac{1}{\theta})\ \eta(\theta)\ d\theta \quad .$$

Substituting this in [IV.2.19], we find :

$$\int B\ \xi(\tfrac{1}{\theta})\ \eta(\theta)\ \xi(j')\ d\theta = \int B\ \xi(\tfrac{j'}{\theta})\ \eta(\theta)\ d\theta \quad . \qquad [IV.2.20]$$

This must be true for every η (as it is the expression of 'every weight assignment method'). A result states in classical Lebesgue integration theory that (see e.g. Apostol (1974)) one must have

$$B\ \xi(\tfrac{1}{\theta})\ \xi(j') = \xi(\tfrac{j'}{\theta}) \qquad [IV.2.21]$$

for every $j' \in [1,\infty[$ and θ (varying in an interval). (For the mathematicians (the others can skip this argument), the exact Lebesgue-theory states that : if h, h' are integrable and λ is a Lebesgue measure such that

$$\int_D h\ d\lambda = \int_D h'd\lambda \qquad [IV.2.22]$$

for every integrable set D, then $h = h'$, λ - a.e. (where a.e. means 'almost everywhere'). Hence $h - h' = 0$, λ - a.e.. If $h - h'$ is also continuous, then $h - h' = 0$ everywhere, and hence $h = h'$. Condition [IV.2.22] is satisfied here by taking η above equal to χ_D, the characteristic function of D. The condition that $h - h'$ be continuous is satisfied here since ξ is assumed to be continuous.)

Of course, as in Subsection IV.2.2.1, it follows that ξ must be a power function and hence that f is of the form [IV.1.1].

This is a strong result, implying that only Lotka-type frequency functions [IV.1.1] can be used if one wishes to keep the same type of frequency function for different weight assignment methods. This is an important fact in favour of Lotka's functions [IV.1.1]. For further remarks on this topic, see Bookstein (1977,1984).

IV.2.3. Mandelbrot's combinatorial-fractal argument

This explanation of the informetric laws is restricted to IPP's in linguistics, namely the case of texts consisting of words (types) and their occurrences (tokens) in the text. Such IPP's can be regarded as special IPP's since words consist of letters and a few other signs. We will first present the derivation of Mandelbrot's law and indicate then how it can be linked to

fractal theory (see Feder (1988), a basic book on fractal theory).

IV.2.3.1. Derivation of Mandelbrot's law

Let our alphabet (including all necessary numbers or signs) consist of N 'letters'. Consider a text as consisting of letters and blanks. Every letter or blank fills up the spaces in the text. We assume (greatly simplified situation) that every letter has an equal chance of being used. Let ρ be this probability. Therefore, since there are also blanks in the text :

$$\rho = P(\text{letter}) < \frac{1}{N} \quad .$$

The probability of having a word consisting of k letters is then (this is also an approximation of reality)

$$P(k) = (1 - N\rho)\ \rho^k$$

for every $k = 0,1,2,\ldots$. Hence, we also have

$$P(k) = P_0\ \rho^k \quad . \qquad [IV.2.23]$$

Let r be the rank of such a word if we arrange the words in decreasing order of use. This results in

$$1 + N + N^2 + \ldots + N^{k-1} < r \leq 1 + N + N^2 + \ldots + N^k \ .$$

Thus, with $1 + N + \ldots + N^s = \frac{N^{s+1} - 1}{N - 1}$, where $s = k-1$ and $s = k$,

$$N^k < r(N-1) + 1 \leq N^{k+1} \quad . \qquad [IV.2.24]$$

Now [IV.2.23] implies that

$$k = \frac{\log\left(\frac{P}{P_0}\right)}{\log \rho} \ . \qquad [IV.2.25]$$

Hence [IV.2.24] becomes

$$N^{\frac{\log (P/P_0)}{\log \rho}} < r(N-1) + 1 \leq N^{\frac{\log (P/P_0)}{\log \rho} + 1} \quad .$$

Consequently,

$$e^{\frac{\log N}{\log \rho} \log (P/P_0)} < r(N-1) + 1 \leq e^{\frac{\log N}{\log \rho} \log (P/P_0) + \log N} . \quad [IV.2.26]$$

We denote (for reasons to be explained later on) the following :

$$D_S = - \frac{\log N}{\log \rho} . \quad [IV.2.27]$$

Then [IV.2.26] becomes :

$$(\frac{P}{P_0})^{-D_S} < r(N-1) + 1 \leq N(\frac{P}{P_0})^{-D_S} . \quad [IV.2.28]$$

The last approximation makes r(N-1) + 1 equal to the average of the left and the right part in [IV.2.28] (in the light of continuous IPP's - see below - this approximation is acceptable). Hence :

$$r(N-1) + 1 = \frac{N+1}{2} (\frac{P}{P_0})^{-D_S}$$

or

$$P = \frac{P_0 (\frac{N+1}{2})^{1/D_S}}{(1 + r(N-1))^{1/D_S}} .$$

When this is multiplied by T, the total number of words in the text,indeed results in the function g as defined in [IV.1.6]. Note that

$$\beta' = \frac{1}{D_S} , \quad [IV.2.29]$$

a fact that will be used later on (note 3 in Section IV.4.2).

In the case in which N = 2 (e.g. a text consisting of binary codes) we see that

$$P = \frac{P_0 (\frac{3}{2})^{1/D_S}}{(1 + r)^{1/D_S}}$$

or, in terms of general constants and the function g :

$$g(r) = \frac{F}{(1+r)^{\beta'}} \quad . \qquad \text{[IV.2.30]}$$

This law is called 'Zipf's law' (cf. [IV.1.5], where the ranks are one lower) and is therefore a special case of Mandelbrot's law (cf. Zipf (1949)). Interpreting [IV.2.30] for continuous r, we also find Pareto's law. The special place of all these laws (as well as several others) will be examined in Chapter IV.4.

The fact that, more or less the same laws are appearing in different disciplines has lead Egghe (1989a,b,c) to model the mechanisms of general IPP's such as bibliographies, texts, economic production processes, social systems etc. This will be executed in Chapter IV.3.

We close this section by offering an interpretation of the parameter D_S in the above reasoning. This interpretation is implicit in Mandelbrot (1977), but is not very clearly formulated. Therefore we will try to enhance its clarity.

IV.2.3.2. IPP's (more specifically : texts) as fractals

IV.2.3.2.1. Introduction to fractal theory

Fractal theory was introduced by Mandelbrot (see e.g. Mandelbrot (1954, 1977)), but an interesting reader is also referred to Feder (1988) for a more recent and structured text on fractal theory. This is not the place to give a detailed account of fractals; the reader will benefit more from an intuitive account of this important and (relatively) new area in mathematics.

Fractal theory arose from the fact that - in the real word - it is not easy to measure distances. In fact, if we do not indicate the distance from the measurerto the object, it is not well defined. For the purposes of illustration, let us look at a map of Australia (see Fig.IV.2.1). Given the scale of the map we can estimate the length of the coastline.

Of course, a map with a larger scale will show more detail, so that one might conclude that a more accurate measurement is possible. This is not really so. What happens is that the closer we look at the coastline, the longer its length becomes and, when graphing the measured length of the coastline as a function of the distance d of the measurer, we obtain a graph that has an infinite limit for d going to zero. Indeed, at very close distances, one must measure the length of all small fractions in the coastline, including river cuts, irregularities in the stones, etc. One could even go to the micro-level of molecules, but this is not necessary in order to see that coast length goes to infinity when d goes to zero.

Fig.IV.2.1 Map of Australia

An equivalent method of expressing the distance of the measurers to the coastline is to cover the coastline by squares having a fixed side length δ and to count the number of squares $N(\delta)$ needed to cover the coastline, as a function of δ, for δ going to zero. As an example, we reproduce the graph of $N(\delta)$ for the coast of Norway (Fig.IV.2.2), based on Feder (1988).

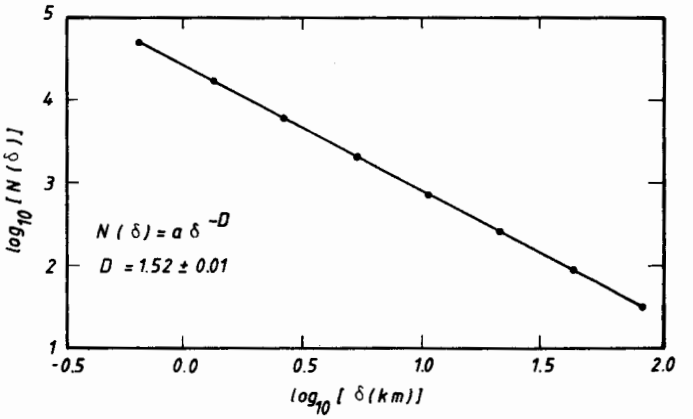

Fig.IV.2.2 The number of squares of size δ needed to cover the coastline of Norway, as a function of δ

Here one fits through linear regression on a log-log scale $N(\delta) = a\delta^{-D}$, where $D \approx 1.52$ and a is a constant. The attentive reader can check that, for a straight line or a simple curve, one necessarily gets $D = 1$, while for a

rectangle one finds D = 2; a being here the length of the straight line, the area of the rectangle respectively. The exact definition of D is omitted here (see Feder (1988), p.11 ff.). D is called the '*Hausdorff-Besicovitch' dimension* of the fractal and represents - intuitively - the degree of higher dimensionality of the line.

For most fractals, there is an easier way of calculating D. If we have so-called *self-similar fractals*, the same graphs can be duplicated on a smaller scale.

Example :

We take the example of the triadic Koch curve (see Fig.IV.2.3).

Continuing this process in the limit yields the fractal we are examining here. Every time the level is increased by 1, we need N = 4 times the previous graph with a scale of $r = \frac{1}{3}$. The similarity dimension is defined as

$$D_S = - \frac{\log N}{\log r} , \qquad [IV.2.31]$$

which in this case is $D_S = \frac{\log 4}{\log 3} > 1$ (a proper fractal!). $D = D_S$ for self-similar fractals and is also called the '*fractal dimension*'. For fractals in a plane one has $1 \leq D \leq 2$ and for fractals in space $2 \leq D \leq 3$, etc. $D = D_S$ replaces measuring lengths since the latter is impossible.

IV.2.3.2.2. Application to texts and to general informetrics

Suppose we have a text consisting of words and blanks, where the words are composed of letters from an N-letter-alphabet. Assume, as in Subsection IV.2.3.1, that

$$\rho = P(\text{letter}) < \frac{1}{N} .$$

The different stages (comparable to the triadic Koch curve) in this case are :

$n = 0$: empty text ,

$n = 1$: the N letters = the N words, say A,B,C,... ,

$n = 2$: the N^2 words AA,AB,...
BA,BB,...
.....

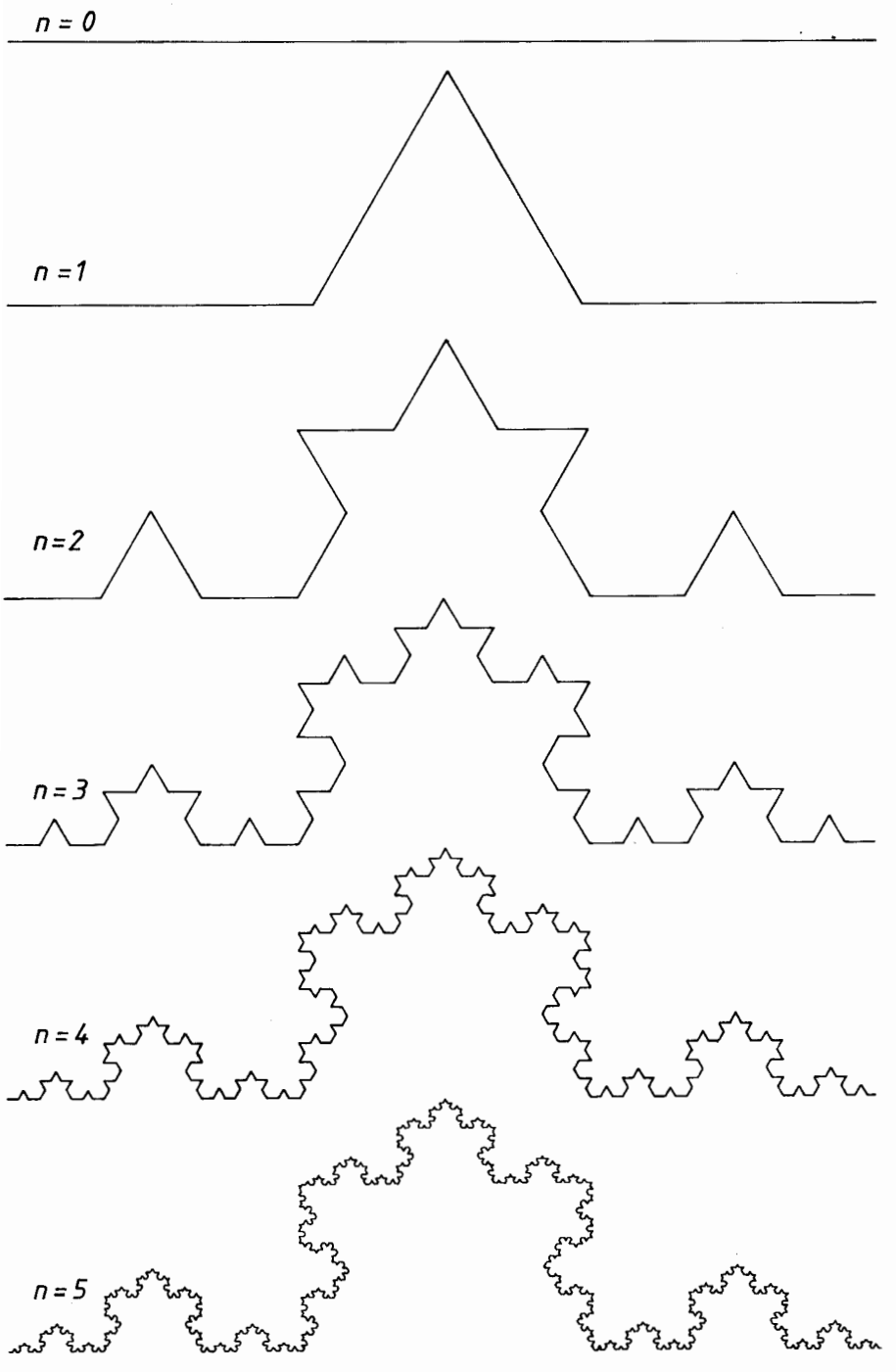

Fig.IV.2.3 Construction of the triadic Koch curve

etc. At each level n we have

$$P(\text{word}) = \rho^n$$

and N^n words, so that the change from one level to a higher level gives a multiplication of the self-similar objects by N, together with a scale factor of ρ. Hence, according to [IV.2.31] :

$$D_S = - \frac{\log N}{\log \rho} \quad .$$

We return to equation [IV.2.27]. We can therefore interpret, in Mandelbrot's law,

$$g(r) = \frac{G}{(1 + Hr)^{\beta'}} \qquad \text{[IV.1.6]}$$

$\frac{1}{\beta'}$ to be the fractal dimension of the IPP (see [IV.2.29]).

Problem :

Mandelbrot's arguments only apply to the case of IPP's being texts, and hence to IPP's in linguistics. Prove this interpretation once again for general IPP's.

This problem was raised by B.C. Brookes (oral communication).

IV.3. THE FORMAL THEORY OF IPP's, THEIR MECHANISMS AND DUALITY

Instead of trying to explain or to construct informetric laws, one can also look at the mechanism of the object under study : an information production process (IPP), see Section 1. We present our model as developed in Egghe (1989a,b,c,d) and restrict our discussion to continuous IPP's. Continuous IPP's are closed models for large IPP's as they are encountered in practice. The powerful theory of infinitesimal analysis allows *continuous* IPP's to be modelled more easily. We henceforth drop the adjective 'continuous'.

Based on the heuristic ideas in Chapter IV.1 of the interaction of sources and items in IPP's, we formally introduce :

IV.3.1. Definition of Information Production Processes (IPP)

An *IPP* is a triplet of the form

$$(S,I,V) , \qquad [IV.3.1]$$

where $S = [0,T]$ (the closed interval starting in 0 and ending in T), $I = [0,A]$ and V is a strictly increasing differentiable function

$$V : S \to I \qquad [IV.3.2]$$

such that $V(0) = 0$ and $V(T) = A$.

The elements of S are called *sources*; the elements of I are called *items*. From now on we will consider $V(r)$ (for every $r \in S \setminus \{0\}$) to be the cumulative number of items in all sources $s \in [T-r,T]$ (taking $[T-r,T]$ rather than $[0,r]$ for technical reasons, which will be made clear below).

IV.3.2. Duality in IPP's

Let

$$(S,I,V) = ([0,T],[0,A],V)$$

be an arbitrary IPP.

The *dual IPP* of the IPP (S,I,V) is defined to be the IPP

$$(I,S,U) = ([0,A],[0,T],U) , \qquad [IV.3.3]$$

where

$$U(i) = T - V^{-1}(A-i) \qquad [IV.3.4]$$

(here V^{-1} denotes the inverse function of V). It can easily be seen that the

dual IPP of the IPP (I,S,U) is again the IPP (S,I,V).

We also define

$$\sigma(i) = U'(i) \qquad [IV.3.5]$$

for every $i \in I$ and

$$\rho(r) = V'(r) \qquad [IV.3.6]$$

for every $r \in S$.
(Here U' and V' denote the derivative of U and V respectively).

Note that, since $V(0) = 0$, $V(r) = \int_0^r \rho(r')dr'$ (by means of [IV.3.6]).
From [IV.3.4] it also follows that $U(0) = 0$; hence by using [IV.3.5],
$U(i) = \int_0^i \sigma(i')di'$, for every $r \in [0,T]$ and $i \in [0,A]$.

When expressed as a function of i in the IPP (S,I,V) (hence $i = V(r)$), $\rho(r)$ becomes :

$$\rho(i) = V'(V^{-1}(i)) \ . \qquad [IV.3.7]$$

We have the following results :

Lemma IV.3.2.1 (Egghe (1989a,b,c))

$$\sigma(i) = \frac{1}{\rho(A-i)} \ , \qquad [IV.3.8]$$

for every $i \in I$.

Proof :
For every $i \in I$, we have, using [IV.3.4] :

$$U'(i) = \frac{dU}{di}(i) \ ,$$

$$= \frac{1}{V'(V^{-1}(A-i))} \ ,$$

$$= \frac{1}{\rho(A-i)} \ ,$$

via [IV.3.7]. Hence [IV.3.5] gives :

$$\sigma(i) = \frac{1}{\rho(A-i)} \quad ,$$

for every $i \in I$. Note that $\rho \neq 0$ everywhere since V is strictly increasing. □

The functions σ and ρ each introduce a coordinate system on $[0,A] \times [0,T]$, different for σ and ρ. So whenever we use a coordinate (i,r) we have to specify whether it belongs to the σ-system (IPP (I,S,U) : U(i) = r) or to the ρ-system (IPP (S,I,V) : V(r) = i). Thus (i,r) is a coordinate in the σ-system if and only if $(A-i,T-r)$ is a coordinate in the ρ-system (this is seen via [IV.3.4]).

Corollary IV.3.2.2

In the IPP (I,S,U) we have :

1. *$\rho(i)$ is the density function of the items (with respect to the sources), in the point $A-i \in I$.*
2. *$\sigma(i)$ is the density function of the sources (with respect to the items), in the point $i \in I$.*

Proof :

This follows readily from [IV.3.7] and [IV.3.8] respectively, the definition of U and V and the previous remarks. □

Alternatively (and equivalently), the functions ρ and σ could have been used as defining functions of an IPP and its dual.

From now on we consider only IPP's with an increasing function $\rho > 0$ (so $\sigma > 0$ is also increasing, according to Lemma IV.3.2.1). This supposition is natural and does not obstruct our general approach; it reduces to an ordering of the set S in the same way as introduced in Subsection IV.1.3.2 (but now for the continuous setting). We further also assume ρ (hence also σ) to be continuous functions.

The functions ρ and σ are the central tools in our duality approach to IPP's. We can also say that ρ plays the same role in (S,I,V) as σ does in (I,S,U).

IV.3.3. The property of pure duality and classical informetrics

The following definition is logical :

Definition IV.3.3.1 :

Given an IPP, we say that we have the property of *pure duality* if there exists a constant $C > 0$ such that, for every $i \in I$:

$$\sigma(i) = C.\rho(i) \quad . \qquad [IV.3.9]$$

Stated otherwise, we have pure duality when the dual functions are proportionally the same.

Which IPP's satisfy the property of pure duality?

Theorem IV.3.3.2 (Egghe (1989a,b)) :
Let (S,I,V) be any IPP. This IPP then satisfies the pure duality property, i.e. there exists a constant C > 0 such that

$$\sigma(i) = C.\rho(i)$$

for every $i \in I = [0,A]$, *if and only if*

$$\sigma(i)\,\sigma(A-i) = C$$

for every $i \in I$.

Proof :
This follows from [IV.3.9] and lemma IV.3.2.1, equation [IV.3.8]. □

This result and Subsection IV.1.3.6 on Bradford's law lead us to a new definition, which will prove to be very useful in what follows : the group-free Bradford law for IPP's (and corresponding Bradford function).

Definition IV.3.3.3 (Egghe (1989a,b,d)) :
Let (S,I,V) be any IPP. We say that this IPP satisfies the *group-free law of Bradford* if, for every $i \in I$,

$$\sigma(i) = M.K^i \quad , \qquad [IV.3.10]$$

where $M > 0$ and $K > 1$ are constants.

Equation [IV.3.10] is called the 'group-free Bradford function'.

The number K is called the 'group-free Bradford factor' and, of course, is independent of p in Subsection IV.1.3.6 (p does not exist here!). This definition allows us to recognise Bradford's law as a function just like the other informetric laws discussed in Chapter IV.1. We furthermore have the following result :

Theorem IV.3.3.4 (Egghe (1989a,b)) :
If the IPP satisfies the group-free law of Bradford, then this IPP satisfies the pure duality property, i.e. there is a constant C > 0 such that

$$\sigma(i) = C.\rho(i)$$

for every $i \in I$.

Proof :
Indeed

$$\sigma(i) = M.K^{i}$$

and hence

$$\sigma(A-i) = M.K^{A-i}$$

for every $i \in I$. Consequently

$$\sigma(i)\ \sigma(A-i) = M^{2}K^{A}$$

for every $i \in I$. By using theorem IV.3.3.2, this IPP satisfies the pure duality property. □

Note 1 :

The following consequence is interesting. Suppose that we have an IPP (S,I,V) (in practice, a large discrete one). If this IPP satisfies Bradford's law, then the informetric 'calculus' ρ in (S,I,V) is the same as the informetric 'calculus' σ in the dual IPP (I,S,U). This means, for instance, that if (S,I,V) is a Bradfordian set of citation data (e.g. $S \rightarrow I$, where $\rightarrow$ is the relation 'citing') then the 'cited' set (I,S,U) satisfies the *same* informetric laws with the same proportional parameters!

Note 2 :

In Chapter IV.4 we will prove that Bradford's group-free law is equivalent to Bradford's classical law, where the number p of groups is arbitrary (in $\mathbb{N}$).

IV.3.4. General duality properties and applications to Lotka's laws

In the previous section we proved an initial result on duality in IPP's, namely pure duality.

This section deals with the more general aspects of duality, valid for general IPP's. We then apply these aspects to Lotka type laws (to be introduced below), to find conditions on the types of Lotka laws.

IV.3.4.1. Basic equations for σ and ρ, in general IPP's

Let (S,I,V) be any IPP with dual functions σ and ρ.

We introduce the following function :

$$f : [\rho(0),\rho(A)] \to \mathbb{R}^+ \text{ (the positive real numbers)}$$

$$j \to f(j) ,$$

where f(j) is defined as the density function (with respect to the IPP (S,I,V)) of the number of sources as a function of j. Hence, by definition, for every $i \in I$,

$$\int_{\rho(0)}^{\rho(i)} f(j)\, dj$$

denotes the cumulative number of sources for which $j \in [\rho(0),\rho(i)]$, equivalently on the coordinates (in I)

$$i' = \rho^{-1}(j) \in [0,i] .$$

This is, because of Corollary IV.3.2.2, equal to :

$$\int_0^i \sigma(A-i')\, di' .$$

Therefore we have (to be used alternatively as the defining relation for f) :

Source - relationship

$$\int_0^i \sigma(A-i')\, di' = \int_{\rho(0)}^{\rho(i)} f(j)\, dj \qquad [IV.3.11]$$

for every $i \in I$.

The integral equation [IV.3.11] is difficult to handle because it is inversely retarded. Luckily, equation [IV.3.11] is equivalent to the following easy integral equation :

Item - relationship

$$\int_{\rho(0)}^{\rho(i)} f(j)j\, dj = i \qquad [IV.3.12]$$

for every $i \in I$.

The simple proof can be found in Egghe (1989a). From now on we will also assume $\rho(0) = 1$. Although this is not really necessary, it is convenient and

is always true in practice.
Hence we have the system :

$$\begin{cases} \rho(i) = \dfrac{1}{\sigma(A-i)} \\ \\ \displaystyle\int_1^{\rho(i)} f(j)j\, dj = i \end{cases} \qquad \text{[IV.3.13]}$$

for every $i \in I = [0,A]$.

We now turn to an initial application of this dual formalism.

IV.3.4.2. Exclusion of certain Lotka functions

The underlying theorem is a result for general functions f (as defined in the previous section) that are continuous and strictly positive on the interval $[1,\infty[$. Considering f on the interval $[1,\infty[$ does not mean that we have sources with an unlimited number of items. We merely assume the existence of the continuous function as an extension of the original function. The function f is then, in practice, restricted to the interval $[1,\rho(A)]$.

Theorem IV.3.4.2.1 (Egghe (1989a,c)) :
If f (restricted to $[1,\rho(A)]$) is the density function of the number of sources in $j \in [1,\rho(A)]$ in a general IPP, and if f is continuous and strictly positive on $[1,\infty[$, then

$$A < \int_1^{\infty} f(j)j\, dj \; . \qquad \text{[IV.3.14]}$$

Proof :
From [IV.3.13] we find that

$$A = \int_1^{\rho(A)} f(j)j\, dj \; . \qquad \text{[IV.3.15]}$$

Suppose that

$$\int_{\rho(A)}^{\infty} f(j)j\, dj = 0 \; .$$

Then the function $j \to f(j)j$ is zero almost everywhere on $[\rho(A),\infty[$ in the Lebesgue sense. But f is continuous. Hence the function $j \to f(j)j$ is identically zero on $[\rho(A),\infty[$ (Apostol (1974)). Thus f(j) is zero on $[\rho(A),\infty[$, a contradiction (cf. also the argument in Subsection IV.2.2.2). Consequently

$$\int_{\rho(A)}^{\infty} f(j)j\,dj > 0 \ . \qquad \text{[IV.3.16]}$$

Equations [IV.3.15] and [IV.3.16] together yield [IV.3.14]. □

This result has an unexpected effect on the Lotka functions :

Corollary IV.3.4.2.2 (Egghe (1989a,c)) :
Suppose that (S,I,V) is an IPP with function f (we define this function to be the general Lotka function, cf. Section IV.1.3.1) :

$$f(j) = \frac{C}{j^{\alpha}} \qquad \text{[IV.3.17]}$$

for every $j \in [1,\infty[$, *where* $\alpha > 1$. *Then*

$$\alpha < \frac{C}{A} + 2 \ . \qquad \text{[IV.3.18]}$$

Proof :
From the previous theorem we see that

$$A < \int_{1}^{\infty} f(j)j\,dj \ . \qquad \text{[IV.3.19]}$$

Hence, upon integrating function [IV.3.17] (which obviously satisfies the requirements of the above theorem), we have :

a. If $\alpha \leq 2$, then [IV.3.18] is automatically satisfied.
b. If $\alpha > 2$, then

$$\int_{1}^{\infty} f(j)j\,dj = \frac{C}{\alpha - 2} \ . \qquad \text{[IV.3.20]}$$

Therefore [IV.3.19] and [IV.3.20] yield

$$A < \frac{C}{\alpha - 2} \ ,$$

and hence [IV.3.18]. □

This, in turn creates a further surprise.

Corollary IV.3.4.2.3 (Egghe (1989a,c)) :
If (S,I,V) is as in the previous corollary, then $\alpha \geq 3$ *implies :*

$$f(1) = C \geq A \ . \qquad \text{[IV.3.21]}$$

Proof :
Indeed, corollary IV.3.4.2.2 produces

$$\alpha < \frac{C}{A} + 2 \ .$$

Hence $\alpha \geq 3$ implies

$$f(1) = C \geq A \ . \qquad \square$$

Note :

Although it is theoretically possible to have [IV.3.21] (since f is a density function), the case in which $\alpha \geq 3$ is very likely to be excluded if the Lotka function [IV.3.17] must fit a practical IPP. Indeed, in practical, discrete IPP's, $C = f(1)$ denotes the number of sources with one item and hence $C < A$.

In the literature we indeed find examples where $\alpha \geq 3$ (see e.g. Pao (1986)). They do not contradict the above remarks since the fittings are statistical and are therefore not based on a mathematical theory. In addition, practical data can differ from Lotka's function (random fluctuations). Furthermore, in most cases we do not know the complete IPP (some of the least productive sources are usually missing) or we do not use the complete IPP (as in Pao (1986)) : in this case A is lower than in reality and hence, according to corollary IV.3.4.2.2, $\alpha \geq 3$ is possible.

We can, however, conclude that, *in general, $\alpha < 3$ will be encountered more often than $\alpha \geq 3$*. The above theory is an initial theoretical basis for this.

IV.4. THE LAWS THAT ARE EQUIVALENT TO LOTKA'S LAW $f(j) = \frac{C}{j^{\alpha}}$, $j \in [1,\rho(A)]$, $\alpha > 1$

IV.4.1. The case $\alpha = 2$

Basically this case has been known for several years (see e.g. Egghe (1985) or Rousseau (1990)), but it will be presented here based on the duality system [IV.3.13] (cf. Egghe (1989a)).

Theorem IV.4.1.1 :

Let (S,I,V) be any IPP. Then the following assertions are equivalent :

(i) The IPP satisfies Lotka's function [IV.3.17] with $\alpha = 2$, on $j \in [1,\rho(A)]$.

(ii) The IPP satisfies Mandelbrot's function for $\beta' = 1$ (cf. Subsection IV.1.3.3) : Consider the IPP (I,S,U). Let $g(r)$ denote the density of the number of items in $r \in [0,T]$. Then

$$g(r) = \frac{G}{(1+Hr)^{\beta'}} \quad , \qquad [IV.4.1]$$

where G, H and β' are constants and $r \in [0,T]$.
Note that $g(r) = \rho(T-r)$ for every $r \in [0,T]$.

(iii) The IPP satisfies Leimkuhler's function (cf. Subsection IV.1.3.5) : In the IPP (I,S,U) : Let $R(r)$ denote the cumulative number of items in the sources $s \in [0,r]$, for every $r \in [0,T]$. Then

$$R(r) = a \log (1 + br) \quad , \qquad [IV.4.2]$$

where a and b are constants, and $r \in [0,T]$. Note that $R = U^{-1}$.

(iv) The IPP satisfies Bradford's law for every $p \in \mathbb{N}$, $p \geq 3$ (cf. Subsection IV.1.3.6) : We can divide the set I into p equal parts, each of length y_0 such that the division in S corresponding to V has length

$$r_0, r_0k, r_0k^2, \ldots, r_0k^{p-1} \qquad [IV.4.3]$$

for a certain r_0 and $k > 1$. This k is, of course, p-dependent, $k = k(p)$.

(v) The IPP satisfies the group-free Bradford function (cf. Subsection IV.3.3.3) :

$$\sigma(i) = M.K^i \quad , \qquad [IV.4.4]$$

where M and K are constants, $K > 1$, and $i \in I = [0,A]$.

Assuming the validity of these equivalent statements, we have the following relations between the parameters :

$$a = \frac{y_0}{\log k} = \frac{1}{\log K} \quad , \qquad \text{[IV.4.5]}$$

$$b = \frac{k-1}{r_0} = \frac{\log K}{M} \quad . \qquad \text{[IV.4.6]}$$

Here y_0, k and r_0 form a valid Bradford triple as in (iv) above (depending on p) but a and b are independent of p (as are M and K) :

$$G = \rho(A) = ab \quad , \qquad \text{[IV.4.7]}$$

$$H = \frac{\rho(A)}{C} = b \quad , \qquad \text{[IV.4.8]}$$

$$K = k(p)^{\frac{p}{A}} \quad , \qquad \text{[IV.4.9]}$$

for every $p \in \mathbb{N}$. Here we write $k = k(p)$.
Consequently, one also has

$$C = a \quad , \qquad \text{[IV.4.10]}$$

$$y_0 = C \log k \quad , \qquad \text{[IV.4.11]}$$

$$r_0 = \frac{C}{\rho(A)} (k-1) \quad . \qquad \text{[IV.4.12]}$$

Proof :
We will show that (i) ⇒ (ii) ⇒ (iii) ⟺ (iv) ⇒ (v) ⇒ (i), which is sufficient for the proof of equivalence. For the direct proof of some other implications, see Egghe (1989a)).

(i) ⇒ (ii)

The proof is based on the general relation for $j \in [1,\rho(A)]$:

$$g^{-1}(j) = r(j) = \int_j^{\rho(A)} f(j')\, dj' \quad . \qquad \text{[IV.4.13]}$$

Although this is intuitively clear, we will offer an exact proof. Consider a valid triple (r,i,j) in (I,S,U), i.e. $i \in [0,A]$, $r = U(i) \in [0,T]$ and $j \in [1,\rho(A)]$. From corollary IV.3.2.2 we have $j = \rho(A-i)$.

Using [IV.3.5] and the notes following it, we get

$$r = U(i) = \int_0^i \sigma(i')\, di' \ .$$

Hence, using the transformation $i'' = A - i'$:

$$r = -\int_{i''=A}^{i''=A-i} \sigma(A-i'')\, di'' = \int_{A-i}^{A} \sigma(A-i'')\, di''$$

$$= \int_0^A \sigma(A-i'')\, di'' - \int_0^{A-i} \sigma(A-i'')\, di'' \ .$$

Using [IV.3.11] twice implies ($\rho(0) = 1$) :

$$r = \int_1^{\rho(A)} f(j)\, dj - \int_1^{\rho(A-i)} f(j)\, dj$$

$$r = \int_{\rho(A-i)}^{\rho(A)} f(j)\, dj \ .$$

Therefore, since $j = \rho(A-i)$, we find

$$r = \int_j^{\rho(A)} f(j')\, dj' \ .$$

Because, by definition, $j = g(r)$ we also have $g^{-1}(j) = r$. Hence [IV.4.13] is proved.

Since, assuming [IV.3.17] with $\alpha = 2$,

$$g^{-1}(j) = r(j) = \int_j^{\rho(A)} \frac{C}{j'^2}\, dj'$$

$$= C\left(\frac{1}{j} - \frac{1}{\rho(A)}\right) ,$$

we also have, again taking $j = g(r)$ and $r = r(j)$:

$$g(r) = \frac{\rho(A)}{1 + \frac{\rho(A)}{C}\, r} \ . \qquad \text{[IV.4.14]}$$

Thus, this implication has been demonstrated, together with the first half of the equalities in [IV.4.7] and [IV.4.8].

(ii) ⇒ (iii)

This proof is based on the general defining relation (definition of R and g) :

$$R(r) = \int_0^r g(r')\, dr' \,. \qquad [IV.4.15]$$

Equations [IV.4.8] and [IV.4.1] (for $\beta' = 1$) give :

$$R(r) = \frac{G}{H} \log (1 + Hr) \,, \qquad [IV.4.16]$$

for every $r \in [0,T]$, yielding Leimkuhler's law. This yields also the second half of the equalities in [IV.4.7] and [IV.4.8].

(iii) ⇒ (iv)

Let $p \in \mathbb{N}$ be fixed but arbitrary. Let $y_0 = \frac{A}{p}$ and r_0 be such that $R(r_0) = y_0$. Define $k > 1$ such that $R(r_0 + r_0 k) = 2\, y_0$. Using [IV.4.2] we see that, if $r = r_0 + r_0 k + \ldots + r_0 k^{i-1}$ $(i = 2,\ldots,p)$, then

$$R(r) = i\, y_0 \,. \qquad [IV.4.17]$$

Indeed, from $R(r_0 + r_0 k) = 2\, y_0 = 2\, R(r_0)$ we find $k = 1 + br_0$. So

$$r = r_0 + r_0 k + \ldots + r_0 k^{i-1} \,,$$

$$= r_0 \frac{k^i - 1}{k - 1} \,,$$

$$= \frac{(1 + br_0)^i - 1}{b} \,.$$

Hence

$$R(r) = i\, R(r_0) = i\, y_0 \,,$$

for every $i = 2,\ldots,p$.

This relation is equivalent to Bradford's law for p groups. A similar argument could be demonstrated for every $p \in \mathbb{N}$. Hence (iv) is proved.

Since we need the equivalence of (iii) and (iv) in the proof of (iv) ⇒ (v), we will also show that (iv) ⇒ (iii) (this is also the direct proof that Bradford's classical law implies Leimkuhler's law and is presented here for the first time in complete detail) :

(iv) ⇒ (iii)

This proof is not trivial and requires several steps, which are given below.

A. If we show that functions R and R^{-1} are differentiable, then they also must be continuous. The fact that they are differentiable follows from [IV.4.15]. Now $g(r) = \rho(T-r)$ for every $r \in [0,T]$ as follows from the definition of g. Hence [IV.4.15] becomes :

$$R(r) = \int_0^r \rho(T-r')\, dr'$$

for every $r \in [0,T]$. Thus $R'(r) = \rho(T-r)$ for every $r \in [0,T]$ and since

$$(R^{-1})'(i) = \frac{1}{R'(R^{-1}(i))} \quad ,$$

$$= \frac{1}{\rho(T - R^{-1}(i))}$$

the proof is complete, since $\rho \geq 1$.

B. Denote by A the set

$$A = \{r_0 \frac{k^i - 1}{k - 1} \,||\, (r_0,k,p^\ell) \text{ is a valid triple in Bradford's law, where } p \text{ is fixed (take e.g. } p = 3) \; \ell \in \mathbb{N}, \text{ and } i = \frac{q}{p^\ell} A, \text{ where } q = 1,2,\ldots,p^\ell.\}.$$

This set A is *dense* in $[0,T]$. (A set $X \subset \mathbb{R}$ is said to be dense in a set $Y \subset \mathbb{R}$ if every element of Y can be written as the limit of a sequence of elements of X.)

Proof :

Since Bradford's law is valid for every $p \in \mathbb{N}$, we can consider Bradford situations for several numbers of groups : $p, p^2, p^3, \ldots$. In each case p^ℓ the item set $[0,A]$ is divided at points

$$A_\ell = \{\frac{A}{p^\ell}, \frac{2A}{p^\ell}, \ldots, A\} ,$$

which is a subset of the divisions in the case $p^{\ell+1}$:

$$A_{\ell+1} = \{\frac{A}{p^{\ell+1}}, \frac{2A}{p^{\ell+1}}, \ldots, \frac{pA}{p^{\ell+1}} = \frac{A}{p^\ell}, \ldots, A\} .$$

By taking $\ell \in \mathbb{N}$ high enough we can make the length between the consecutive division points as small as we wish. From the form of A_ℓ we see that $\bigcup_{\ell \in \mathbb{N}} A_\ell$ is dense in [0,A]. As R^{-1} is continuous, we see that

$$[0,T] = R^{-1}([0,A])$$

$$= R^{-1}\left(\overline{\bigcup_{\ell \in \mathbb{N}} A_\ell}\right) ,$$

where $\overline{\bigcup_{\ell \in \mathbb{N}} A_\ell}$ denotes the closure of the set $\bigcup_{\ell \in \mathbb{N}} A_\ell$, i.e. the set $\bigcup_{\ell \in \mathbb{N}} A_\ell$ and all limits of sequences of elements in $\bigcup_{\ell \in \mathbb{N}} A_\ell$.

Thus $[0,T] \subset \overline{R^{-1}\left(\bigcup_{\ell \in \mathbb{N}} A_\ell\right)}$, by part A.

But, as given by (iv),

$$R^{-1}\left(\bigcup_{\ell \in \mathbb{N}} A_\ell\right) = A ,$$

since, for every i : $r_0 + r_0 k + \ldots + r_0 k^{i-1} = r_0 \frac{k^i - 1}{k - 1}$.

Hence $A \subset [0,T] \subset \overline{A}$, or A is dense in [0,T].

C. Fix $p \in \mathbb{N}$ arbitrarily. We apply Bradford's law.

We have $R(r) = i\, y_0$ for

$$r = r_0 + r_0 k + \ldots + r_0 k^{i-1} ,$$

$$r = r_0 \frac{k^i - 1}{k - 1} ,$$

where $i = 1,2,\ldots,p$.

Hence

$$r = r_0 \left(\frac{k^{\frac{R(r)}{y_0}} - 1}{k-1}\right) ,$$

yielding

$$R(r) = \frac{y_0}{\log k} \log \left(1 + \left(\frac{k-1}{r_0}\right) r\right) , \qquad \text{[IV.4.18]}$$

which is Leimkuhler's function for

$$r = r_0 \left(\frac{k^i - 1}{k - 1}\right) , \qquad \text{[IV.4.19]}$$

$i = 1,2,\ldots,p$. Here we see that

$$a = \frac{y_0}{\log k} ,$$

$$b = \frac{k-1}{r_0} .$$

D. Let a_i and b_i respectively be the above values when there are p^i divisions ($i = 1,2,3,\ldots$) ($p \in \mathbb{N}$ is fixed, take e.g. $p = 3$). Then

$$\begin{cases} a_i = a_{i+1} , \\ b_i = b_{i+1} , \end{cases} \qquad \text{[IV.4.20]}$$

for every $i = 1,2,\ldots$.

1st proof :

Indeed, for every $i = 1,2,\ldots$ the divisions with p^{i+1} groups are a refinement of the divisions with p^i groups. So we have p^i common points. Select any two of them : r_1 and $r_2 \in S$, $r_1 \neq r_2$. Then

$$\begin{cases} R(r_1) = a_i \log (1 + b_i r_1) = a_{i+1} \log (1 + b_{i+1} r_1) , \\ R(r_2) = a_i \log (1 + b_i r_2) = a_{i+1} \log (1 + b_{i+1} r_2) . \end{cases}$$

This system has only one solution :

$$\begin{cases} a_i = a_{i+1} , \\ b_i = b_{i+1} . \end{cases} \qquad \text{[IV.4.21]}$$

2nd proof :

We show this for (a_1,b_1) and (a_2,b_2) respectively. Equation [IV.4.21] then follows by induction.

Let (r_0,y_0,k,p) and (r_0',y_0',k',p^2) be the respective Bradford parameters. Then

$$\begin{cases} a_1 = \dfrac{y_0}{\log k}\,, & a_2 = \dfrac{y_0'}{\log k'}\,, \\ b_1 = \dfrac{k-1}{r_0}\,, & b_2 = \dfrac{k'-1}{r_0'}\,, \end{cases} \qquad \text{[IV.4.22]}$$

according to C. But clearly

$$y_0 = p\, y_0' \,, \qquad \text{[IV.4.23]}$$

so that

$$r_0 = r_0' + r_0'k' + \ldots + r_0'k'^{p-1}\,,$$

$$r_0 = r_0' \left(\frac{k'^p - 1}{k' - 1}\right) . \qquad \text{[IV.4.24]}$$

Also,

$$r_0' = \frac{T(k'-1)}{k'^{p^2} - 1}\,,$$

so that

$$r_0 = T\,\frac{k'^p - 1}{k'^{p^2} - 1} . \qquad \text{[IV.4.25]}$$

But, as

$$r_0 = T\,\frac{k-1}{k^p - 1}\,, \qquad \text{[IV.4.26]}$$

we see from [IV.4.23] and [IV.4.24] that (since the function $f(x) = 1 + x + \ldots + x^{p-1}$ is strictly increasing and is hence injective)

$$k = k'^p \,. \qquad \text{[IV.4.27]}$$

Next, [IV.4.23] and [IV.4.27] produce

$$a_1 = \frac{y_0}{\log k} = \frac{py_0'}{\log k'^p} = \frac{y_0'}{\log k'} = a_2 \quad , \qquad \text{[IV.4.28]}$$

and via [IV.4.24] and [IV.4.27], we get

$$b_1 = \frac{k-1}{r_0} = \frac{k'^p - 1}{r_0} = \frac{k'-1}{r_0'} = b_2 \quad . \qquad \text{[IV.5.29]}$$

As all a's and all b's are equal, we have verified the validity of

$$R(r) = a \log (1 + br) \qquad \text{[IV.4.30]}$$

in the points $r \in A$. This indeed follows from C and D.

E. Since R is continuous, by virtue of A being dense in [0,T] and since the function

$$r \to a \log (1 + br)$$

is already a continuous extension of R to [0,T], we can conclude that

$$R(r) = a \log (1 + br)$$

for every $r \in [0,T]$, where a and b are constants.
We have also shown the first half of the equalities in [IV.4.5] and [IV.4.6], where a and b are independent of p.
Hence with these equations and with [IV.4.7] and [IV.4.8] equations [IV.4.10], [IV.4.11] and [IV.4.12] are also shown.

Note :

From [IV.4.5] it follows that

$$k(p)^p = \text{constant} \quad , \qquad \text{[IV.4.31]}$$

independent of p.

Proof :

Let k_1 correspond to a Bradford division into p_1 groups and k_2 correspond to a Bradford division into p_2 groups. Then, according to [IV.4.5] and the above reasoning of '(iv) implies (iii)', we have

$$a = \frac{\frac{A}{p_1}}{\log k_1} = \frac{\frac{A}{p_2}}{\log k_2} \quad . \qquad [IV.4.32]$$

Hence

$$\log k_1^{p_1} = \log k_2^{p_2} = \frac{A}{a} \quad . \qquad [IV.4.33]$$

Equation [IV.4.33] now yields

$$k_1^{p_1} = k_2^{p_2}$$

and thus

$$k(p)^p = \text{constant} \; .$$

(iv) ⇒ (v)

Since (iii) ⇔ (iv) it suffices to prove (iii) ⇒ (v), which is easier. Since, by definition,

$$R(r) = U^{-1}(r) \qquad [IV.4.34]$$

for every $r \in [0,T]$, we get

$$U(i) = R^{-1}(i) = \frac{e^{i/a} - 1}{b} \qquad [IV.4.35]$$

for every $i \in [0,A]$. Hence

$$\sigma(i) = U'(i) = \frac{1}{ab} e^{\frac{i}{a}} \quad ,$$

$$\sigma(i) = M.K^i \quad ,$$

for all $i \in [0,A]$, where

$$M = \frac{1}{ab} \; ,$$

$$K = e^{\frac{1}{a}} \quad . \qquad [IV.4.36]$$

This also demonstrates the second half of the equalities [IV.4.5] and [IV.4.6]. From these equalities it also follows that

$$\log K = \frac{\log k}{y_0} \quad . \qquad \text{[IV.4.37]}$$

Since $y_0 = \frac{A}{p}$ (by virtue of the fact that we have Bradford's law for every $p \in \mathbb{N}$ because (iii) $\Leftrightarrow$ (iv)), we find

$$K = k(p)^{\frac{p}{A}} \quad ,$$

where we write $k = k(p)$, for clarity. Hence [IV.4.9] is also proved. Consequently, all the equations are proved.

(v) $\Rightarrow$ (i)

Since

$$\sigma(i) = M.K^i$$

for every $i \in [0,A]$ and since differentiating [IV.3.11] gives

$$\sigma(A-i) = f(\rho(i))\ \rho'(i) \qquad \text{[IV.4.38]}$$

for every $i \in [0,A]$, we see (using [IV.3.8] as well) that

$$f(\rho(i)) = \frac{1/\log K}{\rho^2(i)} \quad . \qquad \text{[IV.4.39]}$$

Taking $C = \frac{1}{\log K}$, we see that

$$f(j) = \frac{C}{j^2}$$

for every $j \in [1,\rho(A)] = [\rho(0),\rho(A)]$. □

Note :

The implication (iv) $\Rightarrow$ (v) can be proved directly, but even then we need the equivalence of (iii) and (iv), or at least the fact that $k(p)^p$ is constant. For the proof see Egghe (1989d) .

Corollary IV.4.1.2 :

If the IPP satisfies Bradford's law for p groups ($p \in \mathbb{N}$), then the Bradford factor $k = k(p)$ has the value

$$k = \rho(A)^{\frac{1}{p}} \quad . \qquad \text{[IV.4.40]}$$

Proof :
Using [IV.4.11] (valid for a fixed but arbitrary $p \in \mathbb{N}$), we see that (since y_0 is obviously $\frac{A}{p}$)

$$k = e^{\frac{A}{pC}} \quad , \qquad \text{[IV.4.41]}$$

But, using [IV.3.12], we have

$$A = \int_1^{\rho(A)} j\, f(j)\, dj \quad ,$$

$$A = \int_1^{\rho(A)} \frac{C}{j}\, dj \quad ,$$

$$A = C \log \rho(A) \quad . \qquad \text{[IV.4.42]}$$

Equations [IV.4.41] and [IV.4.42] now yield

$$k = \rho(A)^{\frac{1}{p}} \quad . \qquad \square$$

This equation will have to be slightly adapted when discrete practical bibliographies are fitted to Bradford's law (equation [IV.5.3]).

Corollary IV.4.1.3 :
If the IPP satisfies Bradford's group-free function, then the continuous Bradford factor K has the value

$$K = \rho(A)^{\frac{1}{A}} \quad . \qquad \text{[IV.4.43]}$$

Proof :
This follows readily from equations [IV.4.9] and [IV.4.40]; equation [IV.4.40] can be used since, in the above theorem, (v) implies (iv). $\square$

IV.4.2. The general case : $\alpha \neq 2$

The following theorem can be found in Egghe (1989a) or Egghe (1989c).

Theorem IV.4.2.1 :
Let (S,I,V) be any IPP. Then the following assertions are equivalent :

(i) The IPP satisfies Lotka's function [IV.3.1] on $j \in [1,\rho(A)]$ (general $\alpha \neq 2$).

(ii) The IPP satisfies Mandelbrot's function [IV.4.1] (general β').

(iii) The IPP satisfies the general Leimkuhler function : In the IPP (I,S,U), let $R(r)$ denote cumulative number of items in the sources $s \in [0,r]$ for every $r \in [0,T]$. Then

$$R(r) = \frac{C}{2-\alpha}\left[\rho(A)^{2-\alpha} - \left(\rho(A)^{1-\alpha} - \frac{1-\alpha}{C}\, r\right)^{\frac{2-\alpha}{1-\alpha}}\right], \qquad \text{[IV.4.44]}$$

where

$$\rho(A) = \left(\frac{A(2-\alpha)}{C} + 1\right)^{\frac{1}{2-\alpha}} . \qquad \text{[IV.4.45]}$$

(iv) The IPP satisfies Bradford's general group-free law :

$$\sigma(i) = \left(\left(\frac{A(2-\alpha)}{C} + 1\right) - i\,\frac{2-\alpha}{C}\right)^{\frac{1}{\alpha-2}} , \qquad \text{[IV.4.46]}$$

for every $i \in [0,A]$.

Proof :

We will demonstrate the implications (i) ⇒ (iv) ⇒ (iii) ⇒ (ii) ⇒ (i).

(i) ⇒ (iv)

From [IV.3.17] and [IV.3.12] we have

$$\int_1^{\rho(i)} \frac{C}{j^{\alpha-1}}\, dj = i$$

for every $i \in I$. Hence

$$\frac{C}{2-\alpha}\left(\rho(i)^{2-\alpha} - 1\right) = i .$$

Consequently

$$\rho(i) = \left(\frac{i(2-\alpha)}{C} + 1\right)^{\frac{1}{2-\alpha}} \qquad \text{[IV.4.47]}$$

under the condition that

$$\frac{i(2-\alpha)}{C} + 1 > 0 \qquad \text{[IV.4.48]}$$

for every $i \in [0,A]$. To prove this, invoke corollary IV.3.4.2.2, yielding, if $\alpha > 2$:

$$\frac{A}{C}(2-\alpha) + 1 > 0 \quad . \qquad \text{[IV.4.49]}$$

a) If $\alpha < 2$ then

$$\frac{i(2-\alpha)}{C} + 1 > 0$$

always.

b) If $\alpha > 2$ then

$$\frac{A(2-\alpha)}{C} + 1 = \min_{i\in[0,A]} \left(\frac{i(2-\alpha)}{C} + 1\right) . \qquad \text{[IV.4.50]}$$

So, [IV.4.49] and [IV.4.50] imply

$$\frac{i(2-\alpha)}{C} + 1 > 0$$

for every $i \in [0,A]$.

In conclusion, [IV.4.48] is satisfied for every $i \in [0,A]$ and every $\alpha \neq 2$; hence [IV.4.47] is also satisfied.

Next, [IV.4.46] follows from [IV.4.47] by virtue of [IV.3.8].

(iv) ⇒ (iii)

We have

$$r = \int_0^i \sigma(i')\, di'$$

if $R(r) = i$, by definition of R.

Applying [IV.4.46], we get

$$r = \int_0^i \sigma(i')\, di' = \frac{(A_1 + iA_2)^{1+A_3} - A_1^{1+A_3}}{A_2(1+A_3)} ,$$

if we take

$$\begin{cases} A_1 = \dfrac{A(2-\alpha)}{C} + 1 \ , \\ A_2 = \dfrac{\alpha - 2}{C} \ , \\ A_3 = \dfrac{1}{\alpha - 2} \ . \end{cases} \qquad \text{[IV.4.51]}$$

Hence (using $R(r) = i$), we have

$$R(r) = \frac{1}{A_2} \left[(A_1^{1+A_3} + A_2(1+A_3)r)^{\frac{1}{1+A_3}} - A_1 \right] .$$

We now interpret A_1, A_2 and A_3 by means of [IV.4.51] and thus obtain for every $r \in [0,T]$:

$$R(r) = \frac{C}{2-\alpha} \left[\left(\frac{A(2-\alpha)}{C} + 1\right) - \left(\left(\frac{A(2-\alpha)}{C} + 1\right)^{\frac{1-\alpha}{2-\alpha}} - \frac{1-\alpha}{C} r\right)^{\frac{2-\alpha}{1-\alpha}} \right] . \qquad \text{[IV.4.52]}$$

It follows from [IV.4.46], by virtue of [IV.3.8], that

$$\rho(i) = \left(\frac{i(2-\alpha)}{C} + 1\right)^{\frac{1}{2-\alpha}} .$$

Substituting this form of $\rho(A)$ in [IV.4.52] yields [IV.4.44].

<u>(iii) ⇒ (ii)</u>

Since

$$R(r) = \int_0^r g(r')\, dr'$$

(cf. general relation [IV.4.15]), we also have

$$g(r) = R'(r)$$

for every $r \in [0,T]$. From [IV.4.44] we derive the form of R

$$R(r) = B_1(B_2 - (B_3 + B_4 r)^{B_5}) \ ,$$

where

$$\begin{cases} B_1 = \dfrac{C}{2-\alpha} , \\ B_2 = \rho(A)^{2-\alpha} , \\ B_3 = \rho(A)^{1-\alpha} , \\ B_4 = \dfrac{\alpha - 1}{C} , \\ B_5 = \dfrac{2-\alpha}{1-\alpha} . \end{cases} \qquad \text{[IV.4.53]}$$

Hence

$$g(r) = \frac{-B_1 B_4 B_5}{B_3^{1-B_5} (1 + \frac{B_4}{B_3} r)^{1-B_5}} ,$$

which is Mandelbrot's law with exponent

$$\beta' = 1 - B_5 = \frac{1}{\alpha - 1} . \qquad \text{[IV.4.54]}$$

(ii) ⇒ (i)

We again use the general relation [IV.4.13] :

$$g^{-1}(j) = r(j) = \int_j^{\rho(A)} f(j')\, dj' .$$

Hence

$$f(j) = -r'(j) = -(g^{-1})'(j) .$$

From

$$j = g(r) = \frac{G}{(1 + Hr)^{\beta'}}$$

we derive

$$f(j) = -\frac{1}{\beta' H} \left(\frac{G^{\frac{1}{\beta'}}}{j}\right)^{\frac{1+\beta}{\beta}} ,$$

which is Lotka's law. □

Note 1 :

The form of function [IV.4.43] was first derived by Rousseau (1988c) by other methods. The method presented here as well as the functions [IV.4.46] and [IV.4.47] are due to Egghe.

Note 2 :

From formula [IV.4.46] it follows that

$$\lim_{\alpha \to 2} \sigma(i)$$

is an exponential function of the form [IV.4.4] and is therefore the function $\sigma(i)$ for $\alpha = 2$ (Bradford's group-free version). Hence our theory for $\alpha \neq 2$ gives the classical Bradford function ($\alpha = 2$) as a limiting case (as it should).

This is the first time that Bradford's law for the general Lotka law [IV.3.17] has been proved.

Note 3 :

As explained in Section IV.2.3, the Mandelbrot exponent β' equals (cf. [IV.2.29]) :

$$\beta' = \frac{1}{D_S} ,$$

where D_S is the Hausdorff (or similarity) dimension of the text (the IPP was indeed supposed to be a text, cf. the problem at the end of Subsection IV.2.3.2.2). Since [IV.4.54] implies that

$$D_S = \alpha - 1 , \qquad [IV.4.55]$$

we can draw some interesting conclusions when [IV.4.55] is combined with the results found in Subsection IV.3.4.2. Indeed, when the IPP is a text, it follows that (see [IV.3.18]) :

$$D_S < \frac{C}{A} + 1 \qquad [IV.4.56]$$

and, most commonly (see the note after Corollary IV.3.4.2.3)

$$D_S < 2 .$$

Furthermore we always assume that $\alpha > 1$. Hence we find the conclusion

$$0 < D_2 < \frac{C}{A} + 1 \qquad [IV.4.57]$$

and most commonly

$$0 < D_2 < 2 \ . \qquad \text{[IV.4.58]}$$

This looks quite natural since our models explain the mechanism of duality (i.e. 2-dimensionality) in IPP's. Still, we are left with the problem of proving [IV.4.55] and [IV.4.57] for general IPP's. A possible relationship between D_S and α has also been conjectured, independently, by Tabah and Saber (1989).

Note 4 :

Equation [IV.4.1] for H = 1 is referred to as *Zipf's (or Pareto's) function* (cf. equations [IV.1.5] and [IV.1.8]).

IV.4.3. Corollary

If the IPP satisfies [IV.3.17], then Bradford's corresponding function σ *satisfies :*

$\frac{\sigma'(i)}{\sigma(i)}$ *increases with* i *if* $\alpha < 2$,

$\frac{\sigma'(i)}{\sigma(i)}$ *decreases with* i *if* $\alpha > 2$,

$\frac{\sigma'(i)}{\sigma(i)}$ *is constant if* $\alpha = 2$.

Proof :

Suppose that $\alpha \neq 2$. Equations [IV.4.46] and [IV.4.51] yield

$$\sigma'(i) = A_2 A_3 (A_1 + iA_2)^{A_3 - 1}$$

Hence

$$\frac{\sigma'(i)}{\sigma(i)} = \frac{A_2 A_3}{A_1 + iA_2} \ .$$

Substituting the values of A_1, A_2 and A_3 in function of A, C and α gives

$$\frac{\sigma'(i)}{\sigma(i)} = \frac{1}{A(2-\alpha) + C - i(2-\alpha)} \ .$$

This is an increasing function if $\alpha < 2$ and a decreasing function if $\alpha > 2$. If $\alpha = 2$, the result is well known : equation [IV.4.4] yields

$$\frac{\sigma'(i)}{\sigma(i)} = \log K \quad ,$$

a constant. □

As is shown in Rousseau (1988c), the graph of function R (equation [IV.4.44]), drawn on a semilogarithmic (log r, R(r)) scale, shows an inflection point. Note that in informetrics this is called a '*Groos droop*', since Groos was the first to find such a 'deviation' from the log form (cf. Groos (1967) and see Subsection IV.6.3.1) for $\alpha < 2$. There is no inflection point for $\alpha > 2$ (as predicted also in Egghe (1985)). If there is a Groos droop ($\alpha < 2$), the inflection point is given by :

$$r_d = \frac{C}{2-\alpha} \left(\frac{A(2-\alpha)}{C} + 1\right)^{\frac{1-\alpha}{2-\alpha}} \quad . \qquad \text{[IV.4.59]}$$

IV.5. INFORMETRIC APPROXIMATIONS

Numerous informetric papers use *approximations* in order to apply certain mathematical properties. Stating 'axiomatically' what is allowed and what is not with respect to approximations in informetrics is not easy.

Basically, approximations are needed to cope with the fact that the above theory is being applied to IPP's (and thus to continuous models), while practical bibliographies are not : they are discrete but large. We therefore adopt the following acceptable principles for discrete IPP's in practice :

(A_1) *We may use discrete sums wherever we have used integrals in the continuous theory above. We have not done this from the beginning to preserve the theoretical elegance and for technical reasons. Some results would even have been impossible to prove in a discrete setting.*

(A_2) *$\rho(A)$, the maximal density of items, can be taken equal to the number of items in the most productive source, provided there is only one such source. This quantity will henceforth be denoted by*

$$y_m = \rho(A) \ . \qquad [IV.5.1]$$

(A_3) *y_m is large, in the absolute sense (i.e. when not in combination with other parameters).*

These principles agree with all practical (i.e. not too small) bibliographies.

We leave it to the reader to change the above equations in which $\rho(A)$ appears, into equations containing y_m, using [IV.5.1]. We do provide one example :

$$R(r) = \frac{C}{2-\alpha}\left[y_m^{2-\alpha} - \left(y_m^{1-\alpha} + \frac{\alpha-1}{C}\, r\right)^{\frac{2-\alpha}{1-\alpha}}\right] \ , \qquad [IV.5.2]$$

which is now precisely the function found by Rousseau (1988c).

We state without proof the following 'discrete' analogue of equation [IV.4.40] :

$$k = \left(e^{\gamma}\, y_m\right)^{\frac{1}{p}} \ , \qquad [IV.5.3]$$

where γ denotes Euler's number; $\gamma \approx 0.5772$.
We refer the reader to Egghe (1986a or 1989a,d) for a complete proof. This equation also implies

$$y_m = \frac{k^p}{e^\gamma} .$$

Suppose we have a discrete, practical bibliography for which Bradford's law with p groups holds. One might wonder what the value of m(i) (i = 1,2,...,p) is, where m(i) denotes the number of items in the most productive source in the i^{th} group (counted from the least productive source on). Since $y_m = m(p)$, the above equation suggests that

$$m(i) = \frac{k^i}{e^\gamma} , \qquad [IV.5.4]$$

which is intuitively clear when the last p-i groups are cut off (i.e. the p-i groups containing the most prolific sources). Equation [IV.5.4] is indeed correct, but the proof is not trivial. See Egghe (1986a or 1989a) for the complete proof.