Journal of Statistical Physics, Vol. 32. No. 2, 1983

Maximum Entropy Formalism, Fractals, Scaling Phenomena, and $1/f$ Noise: A Tale of Tails

Elliott W. Montroll[1,2] and Michael F. Shlesinger[1,2]

Received February 1, 1983

In this report on examples of distribution functions with long tails we (a) show that the derivation of distributions with inverse power tails from a maximum entropy formalism would be a consequence only of an unconventional auxilliary condition that involves the specification of the average value of a complicated logarithmic function, (b) review several models that yield log-normal distributions, (c) show that log normal distributions may mimic $1/f$ noise over a certain range, and (d) present an amplification model to show how log-normal personal income distributions are transformed into inverse power (Pareto) distributions in the high income range.

KEY WORDS: Clusters; self-similarity; Lévy distributions; log-normal distribution; random walks.

INTRODUCTION

In the world of the investigation of complex phenomena that requires statistical modeling and interpretation, several competing styles have been emerging, each with its own champions. The *maximum entropy formalism* people have based their style on the ideas in Boltzmann's 1877 paper that introduced the probabilistic interpretation of the thermodynamic entropy. The *fractal* people are enthusiastic about systems and processes that enjoy self-similarity characteristics that lead to statistical distributions with inverse power tails. Pareto observed an inverse power law in annual income distributions of the wealthy, and Paul Lévy developed formal statistical

[1] Institute for Physical Science and Technology, University of Maryland, College Park, Maryland 20742.
[2] La Jolla Institute, P.O. Box 1434, La Jolla, California 92038.

models that yielded such tails. The renormalization group crowd have based theories of phase transitions and other complex phenomena on scaling ideas that they hope transcend details of individual models. In their enthusiasm some of the champions of each style hope to show that theirs is the path to the solution of various unsolved problems of today's science.

The aim of this paper is to produce a few simple examples and limitations of several styles, and on the other hand to show how models may be constructed to allow one to pass from one style to another leaving a niche for each.

1. MAXIMUM ENTROPY FORMALISM[1-4]

In 1877 Boltzmann introduced the *entropy function*

$$H = -\int p(x) \log p(x)\, dx \qquad \text{with} \quad \int p(x)\, dx = 1 \tag{1}$$

In the maximum entropy formalism one seeks the distribution function $p(x)$ that maximizes the entropy subject to auxilary conditions [with $F_1(x) = c_1 = 1$]

$$\int F_i(x) p(x)\, dx = c_i \qquad \text{with} \quad i = 1, \ldots, l \tag{2}$$

By the method of Lagrange multipliers one introduces parameters $\lambda_1, \ldots, \lambda_l$ to be chosen later so that variations in a functional of $p(x)$,

$$F(p) = -\int p(x)\left[\log p(x) + \lambda_1 + \lambda_2 F_2 + \cdots + \lambda_l F_l\right] dx \tag{3a}$$

vanish as

$$\delta F(p) = -\int \left[\log p(x) + 1 + \lambda_1 + \lambda_2 F_2 + \cdots + \lambda_l F_l\right] \delta p(x)\, dx = 0 \tag{3b}$$

that is, with

$$p(x) = \exp\left[-(1 + \lambda_1 + \lambda_2 F_2 + \cdots + \lambda_l F_l)\right] \tag{4}$$

Now let us examine this process in an inverse manner. Suppose an interesting distribution $p(x)$ is known. What auxiliary conditions are required so that the chosen function maximizes the entropy for that distribution?

The simplest example is

$$p(x) = \begin{cases} \mu e^{-\mu x} & \text{with} \quad 0 \leqslant x < \infty \\ 0 & \text{for} \quad x < 0 \end{cases} \tag{5}$$

Then

$$H = -\int_0^\infty \mu e^{-\mu x}\left[\log \mu - \mu x\right] dx \tag{6}$$

λ_1 would be chosen to be $-(1 + \log \mu)$, λ_2 as μ, and the function $F_2 = x$.

Hence the decaying exponential maximizes the entropy function subject to the constraint, the mean value of x, i.e., x weighted by $p(x)$ is fixed.

The next more complicated case we might examine is the normal distribution

$$p(x) = (2\pi\sigma^2)^{-1/2}\exp(-x^2/2\sigma^2), \qquad -\infty < x < \infty \tag{7}$$

Then

$$H = \int_{-\infty}^{\infty} p(x)\left[-\tfrac{1}{2}\log(2\pi\sigma^2) - (x^2/2\sigma^2)\right] dx \tag{8}$$

Here $1 + \lambda_1$ would be chosen to cancel the $-\frac{1}{2}\log 2\pi\sigma^2$ and the function F_2 would be taken proportional to x^2. If $p(x)$ is an exponential of a polynomial of degree l, then the auxiliary conditions might be selected to be l moment conditions

$$\int_{-\infty}^{\infty} x^j p(x)\,dx = \mu_j, \qquad j = 1, 2, \ldots, l \tag{9}$$

The chosen examples (5) and (7) are basic in equilibrium statistical mechanics.[1,2]

As a further example consider the entropy of the Cauchy distribution.

$$p(x) = a/\pi(x^2 + a^2), \qquad -\infty < x < \infty \tag{10a}$$

Then

$$H = -\int_{-\infty}^{\infty} p(x)\left[\log(a/\pi) - \log(x^2 + a^2)\right] dx \tag{10b}$$

The $1 + \lambda_1$ might be chosen to cancel the $\log(a/\pi)$ but an *a priori* choice of $F_2(x)$ as being proportional to $\log(x^2 + a^2)$ would be rather unlikely. One might argue that such a function could be derived from an infinite number of moment conditions. This would not be possible because the distribution has no convergent higher even moments.

Any other distribution with long inverse power tail $p(x) \sim Ax^{-\nu}$ would suffer the same difficulty as the Cauchy distribution since for large x the auxiliary function $F(x)$ would have to behave as $\nu \log x$, a function that has not been considered a natural one for use in auxiliary conditions. The general situation is even worse since one of the most natural long-tailed inverse power distributions that is connected with some physical models is the Lévy distribution, which is generally defined only through its Fourier integral representation.

2. LÉVY DISTRIBUTIONS[5–9]

Lévy distributions, as in the case of the normal distribution, made their first appearance in studies of distributions of sums in independent random variables. Consider the sum of two independent random variables x_1 and x_2, both with mean values zero:

$$X = x_1 + x_2, \qquad \text{with} \quad -\infty < x < \infty \tag{11a}$$

The characteristic function of X is defined to be

$$f(k) \equiv \langle e^{ikX} \rangle_{\text{av}} = \int\int_{-\infty}^{\infty} p_1(x_1) p_2(x_2) \left[\{ \exp[i(x_1 + x_2)k] \} \right] dx_1\, dx_2$$

$$= f_1(k) f_2(k), \qquad \text{with} \quad f_j(k) \equiv \int_{-\infty}^{\infty} p_j(x) \exp(ikx)\, dx \tag{11b}$$

The Fourier transform of the probability density $p(x)$ of X is $f(k)$ since

$$\langle e^{ikx} \rangle_{\text{av}} = \int_{-\infty}^{\infty} p(x) e^{ikx}\, dx \tag{11c}$$

Lévy posed the question, What is the most general form of $p_j(x)$ with the property that if $p_1(x)$ and $p_2(x)$ have the same form, $p(x)$ will also have that form. This is certainly the case for the normal distribution with σ_1^2 being the dispersion for $p_1(x)$ and σ_2^2 for $p_2(x)$, since

$$f(k) = \left[\exp(-k^2\sigma_1^2) \right] \left[\exp(-k^2\sigma_2^2) \right] = \exp\left[-k^2(\sigma_1^2 + \sigma_2^2) \right]$$

Lévy chose the general form to be essentially

$$f_j(k) = \exp(-a_j |k|^\alpha), \qquad 0 < \alpha \leqslant 2 \tag{12}$$

since

$$f(k) = \exp\left[-|k|^\alpha (a_1 + a_2) \right] = \exp(-|k|^\alpha a) \qquad \text{with} \quad a = a_1 + a_2$$

Then

$$p(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \exp(-ikx) \exp(-|k|^\alpha a)\, dk \tag{13}$$

The Cauchy case corresponds to $\alpha = 1$. The range of α does not exceed 2 since it can be shown that the probability density $p(x)$ for $\alpha > 2$ will be negative for some x, contrary to the requirement that $p(x)$ be everywhere nonnegative. The above discussion extends immediately to the sum of any number of random variables. The traditional central limit theorem that states that the sum of a number of independent random variables has a Gauss distribution is based upon the postulate that certain moments exist for each of the random variables. The Lévy distribution violates this requirement since no second moment exists.

The asymptotic behavior of (13) for large $|x|$ is easily shown to have the asymptotic form

$$p(x) \sim (\alpha a / \pi |x|^{\alpha+1}) \Gamma(\alpha) \sin \tfrac{1}{2}\pi\alpha \qquad \text{as} \quad |x| \to \infty, \quad \text{if} \quad 0 < \alpha < 2 \tag{14}$$

The normal distribution with $\alpha = 2$ is singular with its exponential tail. In view of (14), $p(x)$ has no integer moments of order 2 or greater. Long-tailed distributions are very different from those with Gaussian or exponential tails. Consider a man of average height (the distribution heights of adults of either sex are essentially Gaussian). No one twice the average height exists. On the other hand, consider a man of average annual income (the income distribution has a long tail). It is easy for him to find another with twice his income; his wealthier discovery could easily find another with twice his income, etc. If $p(x)$ represents the jump length distribution in a random walk, a long tail implies the existence of occasional long jumps that take the walker far away from clusters of points that have already been visited, thus starting new clusters. If the second moment of $p(x)$ is infinite (with the first being zero in a symmetric walk), then no scale exists to measure a "typical" jump length. Jumps of all sizes will occur forming a self-similar set of clusters of sites visited. Long tails may imply self-similar scaling. A Lévy flight is a random walk, in the continuum, with jump lengths governed by a $p(x)$ with a Lévy distribution. For a walker beginning at the origin, the probability for being at site x at time t is given by (3) with a replaced by t on the right-hand side. The self-similar clustering nature of a Lévy flight is more apparent when the walk is restricted to a lattice.[10]

Consider, in two dimensions, a random walk on a square lattice with the symmetric single jump probability density

$$p(\mathbf{l}) = \frac{n-1}{4n} \sum_{j=0}^{\infty} n^{-j} \left[\delta_{(l_x, l_y),(\pm b^j, 0)} + \delta_{(l_x, l_y),(0, \pm b^j)} \right] \tag{15}$$

where n and b are integers greater than unity. This $p(\mathbf{l})$ allows for jumps of all orders of magnitude, but with each succeeding order of magnitude displacement occurring with an order of magnitude less probability. This random walker makes about n jumps of a unit length forming a cluster of sites visited before a jump of length b occurs and the trajectory leaves the initial cluster. Then about n such clusters of sites are visited, each of about size n, and separated from a neighboring cluster by a distance b are formed, before a jump of length b^2 occurs, etc. Of course, in a particular simulation a wide degree of fluctuation is possible so jumps of length b^4 may precede jumps of length b^2, etc. For a modest number of steps, the set of points visited by the random walker will be clustered. The condition which ensures that this self-similar clustering persists and that Gaussian behavior is avoided as the number of jumps tends to infinity, is (as with the Lévy flights), that the mean-squared displacement per jump $\overline{l^2}$ must be infinite. Then one expects that the set of sites visited should have the fractal dimension[9] α where

$$\alpha = \ln n / \ln b \tag{16}$$

because on the average there will be n subclusters per cluster, with each subcluster scaled down by a length b relative to the cluster.

The quantity $\overline{l^2}$ is given by

$$\overline{l^2} = \sum_{\mathbf{l}} |\mathbf{l}|^2 p(\mathbf{l}) = \frac{n-1}{n} \sum_{j=0}^{\infty} (b^2/n)^j \tag{17}$$

and it is infinite when $b^2 > n$. The structure function $f(\mathbf{k})$ of the random walk is given by the Fourier transform of $p(\mathbf{l})$, which is

$$f(\mathbf{k}) = \frac{n-1}{2n} \sum_{j=0}^{\infty} n^{-j}\left[\cos\left(b^j k_x\right) + \cos\left(b^j k_y\right)\right] \tag{18}$$

For $b^2 > n$, either of the series with k_x or k_y is the celebrated example of Weierstrass' function which is continuous but nowhere differentiable. The Weierstrass function is self-similar in that it looks oscillatory on every length scale thus is not differentiable (i.e., there are no well-defined tangents to the curve). A Taylor series expansion of $f(\mathbf{k})$ will fail because the coefficient of k_x^2 or k_y^2 is related to $\overline{l^2}$ which is infinite. Instead, it can be shown using Mellin transforms that

$$f(\mathbf{k}) = 1 + \tfrac{1}{2}|k_x|^\alpha Q(k_x) + \tfrac{1}{2}|k_y|^\alpha Q(k_y) + \frac{n-1}{2n} \sum_{j=1}^{\infty} \frac{\left[(-1)^j\left(k_x^{2j} + k_y^{2j}\right)\right]}{(2j)!\left[1 - n^{-1}b^{2j}\right]} \tag{19}$$

where

$$Q(k) = \frac{n-1}{n \ln b} \sum_{j=-\infty}^{\infty} \Gamma(q)\cos\left(\frac{\pi q}{2}\right)\exp\left(\frac{-2\pi i j \ln|k|}{\ln b}\right)$$

and $q = -\ln n/\ln b + 2\pi i j/\ln b$.

The fractal dimension α appears naturally and is equal to $\ln n/\ln b$. When $\overline{l^2}$ is infinite, then α is less than 2. The function $Q(k_x)$ is oscillatory, periodic in $\ln k_x$ with period $\ln b$, i.e., $Q(k_x) = Q(bk_x)$. It can be shown that in the continuum limit a Lévy flight of dimension α is recovered.[10]

The self-similarity of the random walk is reflected in the following scaling equation, which is derived directly from Eq. (18):

$$f(k) = n^{-1} f(b\mathbf{k}) + \frac{n-1}{2n}(\cos k_x + \cos k_y). \tag{20}$$

The same argument can be put into a temporal context.[10] If $p(x)$ represents a distribution of waiting times between a sequence of events, some events might occur after short intervals, but occasionally with small probability long pauses (intermittances) would exist that would be followed by a burst. A record of such events then becomes an alternation of bursts and intermittancies. When the mean time between bursts becomes infinite, no characteristic time exists and a self-similar set of bursts will occur. Then, the time between events is governed by a probability density $p(x)$ with an inverse power tail similar to (14) with $0 < \alpha < 1$.

It is difficult to imagine that anyone in an *a priori* manner would introduce a set of auxiliary conditions that could yield the logarithmic term that appears in the entropy function associated with the Lévy distribution:

$$H = -\int_{-\infty}^{\infty} p(x) \log\left(\frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-ixk} e^{-|k|^{\alpha}} dk\right) dx \tag{21}$$

Hence the wonderful world of clusters and intermittancies and bursts that is associated with Lévy distributions would be hidden from us if we depended on a maximum entropy formalism that employed simple traditional auxiliary conditions.

3. THE LOG-NORMAL DISTRIBUTION[11-14]

The log-normal distribution has the distinction of having both a long tail and moments, being essentially at the edge of this possibility, with most distributions with longer tails having no moments.

The introduction of the log-normal distribution into statistics was motivated by an observation by Francis Galton[11] that certain classes of events are better classified through geometric means than through arithmetic means. Galton followed Quitelat's program of introducing the normal distribution into social science and statistics of human measure and behavior. Galton was an imaginative 19th century English eccentric[12] who inherited a small fortune, traveled extensively, wrote travel guides, and became one of the founding fathers of the once popular "science" of eugenics (now considered racist by some). He established the Eugenics Institute at University College in London. Karl Pearson was appointed its first Galton professor. Following Galton, D. McAlister[13] in 1879 introduced the log-normal distribution in a paper entitled The Law of the Geometric Mean. A definitive review of the log-normal distribution is given by Atchison and Brown.[14] We were first introduced to the Galton and McAlister papers by this reference. Let us now review some mechanisms that lead to the log-normal distribution.

First consider a complex task whose successful completion requires the successful completion of n independent subtasks. The probability, P, of success for the primary task in a unit time is (p_j being the probability of success of the jth subtask)

$$P = p_1 p_2 p_3 \cdots p_n \tag{22a}$$

so that

$$\log P = \log p_1 + \log p_2 + \log p_3 + \cdots + \log p_n \tag{22b}$$

Since the p_i are independent random variables so are the $\log p_i$. If the appropriate moments exist for $\log p_i$ and n is "large," then the central limit theorem is applicable and $\log P$ has a Gauss distribution. This argument was first given by W. Shockley[15] to interpret the observation that the productivity of scientists publishing research papers has a log-normal distribution. Shockley's data were taken from publication records of scientists at the Brookhaven National Laboratories. In that case the p_j's were probabilities of attributes necessary for the publication of a paper; p_1 being the probability of having an idea for investigation, p_2 that of having the competence to pursue the investigation, p_3 that of obtaining some interesting results, etc.

Events that lead to log-normal distributions are analogous to circuits that are described by an "and" logic (series networks). For a signal to pass through the circuit all the "and" gates must be open. Events that lead to a Gauss distribution are the analogs of an "or" logic (or parallel networks). Pulses may proceed through any open gates, the total current being the sum of the current through each gate (the central limit theorem being related to the sum of independent variables).

Log-normal distributions have been observed in many diverse fields such as income distributions, body weights, sound measurements (in decibels), rainfall, etc. Modern scaling theories have led to a log-normal distribution of electrical resistance in materials with random scatterers which cause the localization of electrons.[16-19]

An older example is the size distribution of crushed ore; Kolmogorov[20] "deduced" the crushed ore distribution by a mechanism first proposed by the Dutch astronomer Kapteyn. Suppose that a random variable initially has a value X_0 and that through a succession of breaks achieves the values $X_1, X_2, X_3, \ldots, X_N$ where the difference $X_n - X_{n-1}$ is a random portion of X_{n-1} so that

$$X_n - X_{n-1} = R_n X_{n-1}$$

with R_n being a random variable that ranges between -1 and 0 (since $X_{n-1} > X_n$). Then

$$\sum_{n=1}^{N} (X_n - X_{n-1})/X_{n-1} = \sum_{n=1}^{N} R_n$$

If we change the sum to an integral letting $dx = (X_n - X_{n-1})$, then

$$\int_{X_0}^{X_N} dx/x = \log X_N - \log X_0 = R_1 + \cdots + R_N$$

When the individual R_n distributions have appropriate moments an application of the central limit theorem implies that $\log(X_N/X_0)$ has a normal distribution, or the size distribution, X_N/X_0, is log normal. Notice that the

Shockley model is recovered by writing

$$X_N = \prod_{n=1}^{N} (1 + R_n) X_0$$

and identifying Shockley's p_n with $(1 + R_n)$ and his P with X_N / X_0.

4. FROM THE LOG-NORMAL TO THE $1/f$ DISTRIBUTION[21]

Another commonly observed distribution is the $1/f$ distribution. If the range of f is to be $0 < f < \infty$, then it must be part of some other distribution since the normalization integral

$$\int_0^\infty df/f$$

diverges in the large and small f regime. We now show that over some range of the appropriate variable the log normal distribution mimics a $1/f$ distribution.

If $\log x$ has a normal distribution then the variable x has the following distribution:

$$\begin{aligned} F\left(\log \frac{x}{\bar{x}}\right) d\left(\log \frac{x}{\bar{x}}\right) &= \frac{\exp\left\{ -\left[\log(x/\bar{x})\right]^2 / 2\sigma^2 \right\}}{(2\pi\sigma^2)^{1/2}} \, d \log x/\bar{x} \\ &= \frac{\exp\left\{ -\left[\log(x/\bar{x})\right]^2 / 2\sigma^2 \right\}}{(2\pi\sigma^2)^{1/2}} \, \frac{dx/\bar{x}}{x/\bar{x}} = g\left(\frac{x}{\bar{x}}\right) \frac{dx}{\bar{x}} \end{aligned} \tag{23a}$$

Generally, a $1/f$ distribution is demonstrated graphically on log–log graph paper by plotting log g as a function of $\log x$. For this purpose, we write

$$\log\left[g(x/\bar{x}) \right] = -\log(x/\bar{x}) - \left\{ \left[\log(x/\bar{x})\right]^2 / 2\sigma^2 \right\} - \tfrac{1}{2}\log(2\pi\sigma^2) \tag{23b}$$

the last term being a constant. Let us measure the variable x in multiples f of its mean value $\bar{x}$, with $x = f\bar{x}$. Then Eq. (23b) becomes

$$\log\left[g(f) \right] = -\log f - \tfrac{1}{2}\left[(\log f)/\sigma\right]^2 - \tfrac{1}{2}\log(2\pi\sigma^2) \tag{24}$$

If the distribution $g(f)$ is to be $1/f$, then only the linear term in log $g(f)$ and the constant term should remain in Eq. (24), as would be the case as $\sigma \to \infty$. Let σ be large but finite and let f be expressed as a power n of e: $f = \exp n$. Then

$$\log\left[g(f) \right] = -n - \tfrac{1}{2}(n/\sigma)^2 - \tfrac{1}{2}\log(2\pi\sigma^2) \tag{25}$$

When σ is large, we can estimate the largest integer value of n that allows Eq. (25) to be regarded as linear in n to within a prescribed precision (always omitting the constant term in the precision estimation). If the middle term on the right-hand side of Eq. (25) is to be less than a fraction θ of the first term, then

$$\tfrac{1}{2}(n/\sigma)^2 \leqslant \theta|n| \quad \text{or} \quad |n| \leqslant 2\theta\sigma^2 \tag{26}$$

Suppose that $\theta = 0.1$ and $\sigma = 5$. Then, for any $|n| \leqslant 5$, $g(f)$ mimics a $1/f$ distribution to within 10%. This corresponds to 11-integer n values or, from Eq. (26), 11 e-folds, which is equivalent to four orders of magnitude. Generally, the function $g(f)$ mimics a $1/f$ distribution for $(4\theta\sigma^2 + 1)$ e-folds to within a relative error θ. Clearly, the larger σ, the more orders of magnitude the mimicking persists.

Now consider a task whose successful conclusion follows the completion of N subtasks. Then, the sum similar to (22b) is composed of N terms. The sum of the squares, σ_j^2, of the component random variables, σ^2 should be the order of N with $\sigma^2 = N\bar{\sigma}^2$ where

$$\bar{\sigma}^2 = \frac{1}{N} \sum_{j=1}^{N} \sigma_j^2 \tag{27}$$

The greater N the greater the number of e-folds or decades over which the distribution function for primary task ("grand task") successes mimics a $1/f$ distribution.

5. ON INCOME DISTRIBUTIONS

It is commonly observed that over a large range of an independent variable, distributions might be of a standard type such as normal or log-normal but then suffer a transition in the last few percentiles of a population into an inverse power law. This transition is analyzed here through a special example, the U.S. annual income distribution. That distribution is plotted in Fig. 1 for the period 1935–1936 on log-normal graph paper.[22] On such graph paper a cumulative log normal distribution would be a straight line. That is the case for the first 98–99 percentile; however, afterwards a transition to a Pareto inverse power distribution occurs. One of the earliest observers of the log-normal distribution of incomes was R. Gibret.[23] More recent critical examination of the fitting of the log-normal distribution to data is given in Refs. 24 and 25. Badger[22] has given a useful summary of the application of various statistical distribution functions to income data.

We now indicate how the log-normal distribution might be interpreted in terms of a maximum entropy strategy. Then we suggest a model to "describe" the transition to the Pareto form.

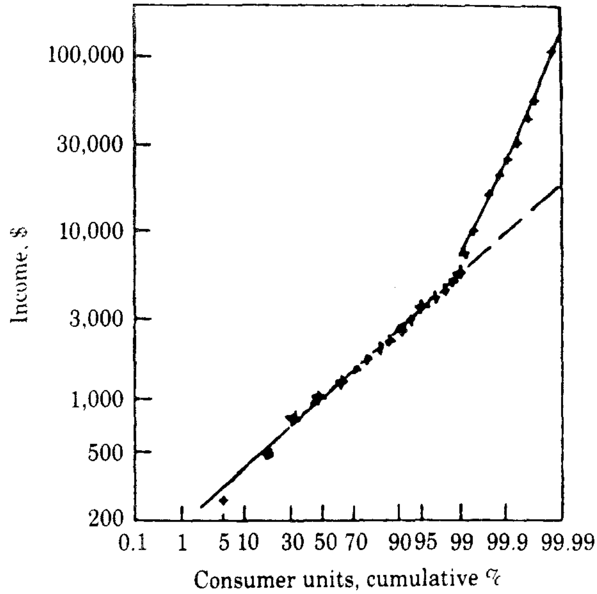

Fig. 1. Distribution of families and single individuals by income level, 1935/1936. Data are from Ref. 35. Most of the data follow a log-normal distribution, while the last 1% is governed by a Pareto tail.

Through various transactions, money is transferred from individual to individual in a manner analogous to the transfer of energy from gas molecule to gas molecule through collisions. By transfer of goods or services (or welfare or charity), every family has someone with an annual income. One might argue that the many transactions cause money to become randomly distributed but, through various constraints due to training, motivation, risk-taking, inheritance, luck, intimidation, skill, etc., some people obtain larger annual incomes than others. We will still apply the entropy principle, but at first without any clear understanding of the constraints. Then we introduce our backward method of choosing the constraint that implies the observed distribution.

Let us suppose that the distribution of annual incomes is log normal, as indicated in Fig. 1. Then the probability that one's annual income lies between x and $x + dx$ is

$$(2\pi\sigma^2)^{-1/2}\exp\{-\log[x/\bar{x}]^2/2\sigma^2\}\,dx/x = p(x)\,dx \tag{28}$$

The factor dx/x is exactly the variation of the Bernoulli utility function $U(x)$ defined so that[26]

$$dU = dx/x \tag{29}$$

The classical significance to this form is that a process involving a transfer of money dx has a different meaning to persons of different levels of income. Transactions made by persons of different income levels might be more equivalent if they involved the same fraction of the income of the participants. Hence, according to Daniel Bernoulli, the basic function which determines one's course of action is the utility function

$$U(x) = \log(x/\bar{x}) \tag{30}$$

Notice that with U considered to be the basic function of our process the normal distribution of U would follow from the maximization of an entropy function[27]

$$H = -\int p(U) \log p(U)\, dU \tag{31}$$

under the auxiliary condition $p(U)$ being normalized and

$$\langle U^2 \rangle = \int U^2 p(U)\, dU = \text{const} \tag{32}$$

That the integral of U^2 is essentially constant over a long time interval is apparent from the data in Table I. We may write

$$\int U^2 p(U)\, dU = \int [\log(x/\bar{x})]^2 p[\log(x/\bar{x})](\bar{x}/x)\, d(x/\bar{x}) \tag{33}$$

From the table the fraction of the national family income in a given population quintile remained almost constant over the period of 18 years of the selected data. The mean income shifted, generally going to a higher level, but relative to the mean the distribution in a given interval remained invariant. Hence in the transition from one year to another incomes would have suffered an annual inflation factor (or deflation factor α) so that

$$x \to \alpha x \quad \text{and} \quad \bar{x} \to \alpha \bar{x} \quad \text{and} \quad x/\bar{x} \to \alpha x/\alpha \bar{x} = x/\bar{x}$$

but yet (33) would have remained invariant. This is a consequence of there being no basic scale in the process.

Table I. Percent Distribution of Family Personal Income by Quintiles and Top 5% of Consumer Units for Selected Years[22]

Quintiles	1944	1947	1950	1951	1954	1956	1959	1962
Lowest	4.9	5.0	4.8	5.0	4.8	4.8	4.6	4.6
Second	10.9	11.0	10.9	11.3	11.1	11.3	10.9	10.9
Third	16.2	16.0	16.1	16.5	16.4	16.3	16.3	16.3
Fourth	22.2	22.0	22.1	22.3	22.5	22.3	22.6	22.7
Highest	45.8	46.0	46.1	44.9	45.2	45.3	45.6	45.5
Total	100	100	100	100	100	100	100	100
Top 5%	20.7	20.9	21.4	20.7	20.3	20.3	20.3	19.6

The above analysis is very attractive; however, it gives us no insight into the appearance of the Pareto inverse power tail beyond the 99 percentile in Fig. 1. No one would dispute the fact that the wealthy differ from the lower 99% in the manner that they accumulate income. While most people are paid by the hour, the wealthy frequently accumulate their extra wealth by some amplification process; that process varying from case to case. At the height of the Beatles' popularity any new recording by them was purchased by millions of fans. The leverage people in the investment business have their style of amplification. During certain periods of prosperity easy money becomes available for investment, sometimes in stock, sometimes in real estate or perhaps in silver or Rembrandts. A common characteristic of such time is that the daring may exploit the easy money to acquire some speculative commodity through a small margin payment, say, 10% with a promise to pay the remainder later. If the commodity doubles in price a 10% margin payment is amplified into a ninefold profit. J. P. Morgan was given his first million by his father. He invested a considerable fraction of that in the manner described above, reinvesting the profit, etc., to become much richer than he would have, had he accepted the offer of a privatdozentship in mathematics at Göttingen University offered to him by Felix Klein. Perhaps one of the most common lower level modes of amplification is for an individual to organize an operation with others working for him so that his income is amplified through the efforts of others (a modest-sized business, for example).

We now introduce a model to indicate how Pareto–Lévy tails may be derived from a log-normal distribution (or indeed from any one of a broad class of distributions with second moments) by accounting for the process of amplification, by the amplification of amplifications, etc.[21] Let $g(x/\bar{x})$ denote the basic distribution written in terms of the dimensionless quantity $x/\bar{x}$, $\bar{x}$ being the mean value of the observed x if the tail of the distribution is neglected. With a small probability, λ, suppose that in the new amplifier class one has the same distribution function g that is natural for the process but that $\bar{x}$ is amplified to $N\bar{x}$. Then the basic quantity $g(x/\bar{x})\,dx/\bar{x}$ is converted to $g(x/N\bar{x})\,dx/N\bar{x}$. In the second stage of amplification, which we postulate to occur with a probability λ^2, the mean value $\bar{x}$ becomes $N^2\bar{x}$. The new distribution $G(y)$ (with $y \equiv x/\bar{x}$) that allows for the possibility of continuing levels of amplification is

$$G(y) = (1-\lambda)\left[g(y) + \frac{\lambda}{N}\, g\left(\frac{y}{N}\right) + \frac{\lambda^2}{N^2}\, g\left(\frac{y}{N^2}\right) + \cdots \right] \tag{34}$$

where λ is a parameter that determines the range of the initial distribution $g(y)$. The factor

$$\beta = (1-\lambda)$$

is introduced to ensure the proper normalization of $G(y)$. It is easy to see that by replacing y by y/N in (34) that

$$G(y) = \frac{\lambda}{N} G(y/N) + (1-\lambda) g(y) \tag{35}$$

The determination of the complete solution of our inhomogeneous scaling formula (35) is rather complex but it is easy to obtain our desired asymptotic properties of $G(y)$. First suppose $\lambda \to 0$. Then there is no amplifier class in the population and $G(y)$ becomes the same as $g(y)$. If λ is small, say, 0.01, and N is about 10, then $G(y)$ is still close to $g(y)$ since the first term in (35) may be neglected. However, when y becomes large $g(y) \to 0$. Let us suppose this decay is faster than that of $G(y)$. Then the asymptotic form of $G(y)$ is determined by the simpler scaling formula

$$G(y) = (\lambda/N) G(y/N)$$

If we suppose that $G(y) = Ay^{-1-\mu}$, then direct substitution yields

$$\mu = \left[\log(1/\lambda)\right]/\log N \tag{36}$$

Thus the Pareto exponent appears as a fractal dimension. The evaluation of A requires a more subtle analysis since in general it may be periodic in $\log\lambda$ with period $\log N$.

The best value of μ to fit the tail of the 1935–1936 data was found by Badger to be 1.63. If we put the probability of being in the special amplifier class as $\lambda = 0.01$ the average amplification factor N would be about 16.8. This number is not surprising since one of the most common modes of significant income amplification is to organize a modest-sized business with the order of 15–20 employees.

6. ON $1/f$ NOISE

The noise spectrum of a "purely random" process is one associated with an autocorrelation function of the form $c(t,\tau) = e^{-t/\tau} = c(t/\tau)$, τ being the "relaxation time" of the process. The power spectrum of such a random process is, at frequency f,

$$S(f,\tau) = 4\,\mathrm{Re}\int_0^\infty c(t,\tau) e^{2\pi i f t}\, dt \tag{37a}$$

$$= 4\,\mathrm{Re}\int_0^\infty e^{-t/\tau} e^{2\pi i f t}\, dt$$

$$= 4\tau/\left[1 + (2\pi f \tau)^2\right] \tag{37b}$$

A commonly observed noise spectrum has $1/f$ behavior over a broad frequency range. To proceed from (37a) to such a spectrum it is presumed that in a complex system there is not a single τ but rather a distribution (or ensemble) of relaxation times $\rho(\tau)$ with normalization

$$\int_0^\infty \rho(\tau)\,d\tau = 1 \tag{38}$$

Then the power spectrum becomes

$$S(f) = \int_0^\infty \frac{4\tau\rho(\tau)\,d\tau}{\left[1 + (2\pi f\tau)^2\right]} \tag{39}$$

The structure of the distribution function $\rho(\tau)$ depends upon the character of the noisy dynamical system of interest. Here we shall postulate it to be determined by a process characterized by Eq. (22a). We assume that the events determining the $\rho(\tau)$ may be described as the successful completion of "grand processes" whose progress is the consequence of the successful completion of a number of independent subprocesses. If we interpret the P of Eq. (22a) as the probability per unit time that a "grand success" occurs, then $\tau = 1/P$ is the time required for a single grand success. The distribution function of τ, weighting the many ways that a success might be achieved is obtained by rewriting (22a) as

$$\log\tau = \log(1/P) = \sum_{i=1}^{N} \log(1/p_i) \tag{40}$$

Since each p_i is a random variable, so is $\log 1/p_i$. Then, if N becomes large and appropriate moments exist, the central limit theorem implies that $\log\tau$ has a normal distribution or τ a log-normal one. We let $\bar{\tau}$ be the mean value of τ. In a system or network whose noise is generated by charge carrier mobility, the sequence defining the p_i might be a successful succession of the surmountings of a sequence of energy barriers by the carrier.

Our log-normal distribution has the form

$$\rho\left(\frac{\tau}{\bar{\tau}}\right)d\tau = \frac{\exp\left\{-\left[\log(\tau/\bar{\tau})\right]^2/2\sigma^2\right\}}{\left(2\pi\sigma^2\right)^{1/2}}\;\frac{d\tau/\bar{\tau}}{\tau/\bar{\tau}} \tag{41}$$

Then the power spectrum becomes

$$S(\omega,\sigma) = \bar{\tau}\int_0^\infty \frac{4(\tau/\bar{\tau})\exp\left(-\left[\log(\tau/\bar{\tau})\right]^2/2\sigma^2\right)d(\tau/\bar{\tau})}{\left(2\pi\sigma^2\right)^{1/2}(\tau/\bar{\tau})\left[1 + \omega^2(\tau/\bar{\tau})^2\right]} \tag{42}$$

with ω being the dimensionless frequency

$$\omega = 2\pi f\bar{\tau} \tag{43}$$

Then, defining z to be $z = \tau/\bar{\tau}$,

$$S(\omega,\sigma)/4\bar{\tau} = \int_0^\infty \frac{\exp\left[-(\log z)^2/2\sigma^2\right]dz}{(2\pi\sigma^2)^{1/2}\left[1+\omega^2 z^2\right]} \tag{44}$$

Most of our analysis is made after introducing the transformation

$$y = \log z \tag{45}$$

Then

$$S(\omega,\sigma)/4\bar{\tau} = \int_{-\infty}^\infty \frac{\exp\left[-(y^2/2\sigma^2)+y\right]dy}{(2\pi\sigma^2)^{1/2}(1+\omega^2 e^{2y})}$$

$$= \frac{1}{\omega(2\pi\sigma^2)^{1/2}}\int_{-\infty}^\infty \frac{\exp(-y^2/2\sigma^2)\,dy}{\omega e^y + (\omega e^y)^{-1}} \tag{46}$$

Several asymptotic results are immediately apparent. We note from (46) that in the large σ limit $(\sigma \to \infty) S(\omega,\sigma)$ mimics $1/\omega = 1/f\bar{\tau}$ noise:

$$S(\omega,\tau) \sim (2\pi\sigma^2)^{-1/2}\int_0^\infty \frac{dz}{1+\omega^2 z^2} = \frac{1}{2\omega\sigma}\left(\frac{\pi}{2}\right)^{1/2} \tag{47}$$

Most of the remainder of this section will be devoted to finding, for a given σ, the range of frequencies for which $S(\omega,\sigma)$ mimics a $1/\omega$ spectrum. However before proceeding to this main theme we record two other asymptotic results:

(a) as $\omega \to 0$, (46) becomes

$$\frac{S(\omega,\sigma)}{4\tau} \sim (2\pi\sigma^2)^{-1/2}\int_{-\infty}^\infty \exp\left[-\left(\frac{y^2}{2\sigma^2}\right)+y\right]dy$$

$$= \frac{\exp(\frac{1}{2}\sigma^2)}{(2\pi)^{1/2}}\int_{-\infty}^\infty \exp\left[-\frac{1}{2}\left(\frac{y}{\sigma}-\sigma\right)^2\right]d\left(\frac{y}{\sigma}\right) = \exp\left(\frac{1}{2}\sigma^2\right) \tag{48}$$

(b) as $\omega \to \infty$, (46) becomes

$$S(\omega,\sigma)/4\bar{\tau} \sim \omega^{-2}(2\pi\sigma^2)^{-1/2}\int_{-\infty}^\infty \exp\left[-(y^2/2\sigma^2)+y\right]dy$$

$$= \omega^{-2}\exp\left(\tfrac{1}{2}\sigma^2\right) \tag{49}$$

The systematic expansion of $S(\omega, \sigma)$ is obtained from (46) by noting that

$$\left[\omega e^{y} + (\omega e^{y})^{-1}\right]^{-1} = \begin{cases} \omega e^{y}\left[1 - (\omega e^{y})^{2} + (\omega e^{y})^{4} \cdots\right] & \text{if } \omega e^{y} < 1 \\ (\omega e^{y})^{-1}\left[1 - (\omega e^{y})^{-2} + (\omega e^{y})^{-4} \cdots\right] & \text{if } \omega e^{y} > 1 \end{cases} \tag{50}$$

Since at $\omega e^{y} = 1$, $y = \log(1/\omega)$,

$$\begin{aligned} &(2\pi\sigma^2)^{1/2} S(\omega, \sigma)/4\bar{\tau} \\ &\quad = \omega^{-1}\int_{-\infty}^{-\log\omega} \omega e^{y}\left\{1 - (\omega e^{y})^{2} + (\omega e^{y})^{4} \cdots\right\} e^{-y^2/2\sigma^2}\,dy \\ &\quad + \omega^{-1}\int_{-\log\omega}^{\infty} (\omega e^{y})^{-1}\left\{1 - (\omega e^{y})^{-2} + (\omega e^{y})^{-4} \cdots\right\} e^{-y^2/2\sigma^2}\,dy \end{aligned} \tag{51}$$

The integration in terms of the error function is carried out in the Appendix. When (A10) is substituted into (46) it is found that

$$\begin{aligned} S(\omega, \tau)/4\bar{\tau} = \frac{1}{2\sigma\omega}\left(\frac{\pi}{2}\right)^{1/2} \exp\left[-\frac{1}{2}\left(\frac{\log\omega}{\sigma}\right)^2\right] \\ \times \Bigg\{\left[1 - \frac{1}{2}\left(\frac{\pi}{2\sigma}\right)^2 + \frac{5}{8}\left(\frac{\pi}{2\sigma}\right)^4 + \cdots\right] \\ + \frac{\pi^2}{2}\left(\frac{\log\omega}{2\sigma^2}\right)^2\left[1 - \frac{5}{2}\left(\frac{\pi}{2\sigma}\right)^2 + \frac{61}{8}\left(\frac{\pi}{2\sigma}\right)^4 + \cdots\right] \\ + \frac{5\pi^4}{24}\left(\frac{\log\omega}{2\sigma^2}\right)^4\left[1 - \frac{61}{10}\left(\frac{\pi}{2\sigma}\right)^2 + \frac{277}{8}\left(\frac{\pi}{2\sigma}\right)^4 + \cdots\right] \\ + \cdots\Bigg\} \end{aligned}$$

If $|\sigma^{-1}\log\omega| < 2\sigma$ the powers of $(\log\omega)$ may be neglected from the terms in the brackets and $S(\omega, \tau)$ has a log-normal distribution analogous to $\rho(x/\bar{x})$ in (41). Then we may proceed in the same manner as we did in the discussion of Eqs. (23)–(26) to determine the number of decades of $\omega = f\bar{\tau}$ the noise spectrum seems to have a $1/\omega$ character. Again in our process characterized by (22a), σ^2 becomes proportional to n, the number of hurdles that must be surmounted on the road to a "grand success."

Similarly, Nelkin and Harrison[28] have shown that a regime of $1/f$ noise results in the specific context of charges moving, under the influence of an electric field, in a material containing traps with a log-normal distribution of release rates.

Van der Ziel[29] as early as 1950 in his work on noise in semiconductors appreciated the importance of the $1/\tau$ factor in $\rho(\tau)$ to cancel the τ in (39) thus leading to $S(f)$ being proportional to $1/f$. He proposed an activated hopping process, where

(i) $\tau \propto \exp(E/kT)$ (activated hopping)

(ii) $\rho(\tau) = \tau^{-1}$ for $\tau_{\min} < \tau < \tau_{\max}$

This implied that the distribution of activation energies $f(E)$ is uniform over the range $(E_{\min}, E_{\max})$ because $d\tau/\tau = dE/kT = f(E)\,dE$. This model was neglected[30] because a peaked (at the order of an eV) rather than a constant $f(E)$ is expected. However, Dutta and Horn[31] have shown that a $1/f^{\alpha}$ spectrum (with α near 1) results from a peaked $f(E)$ if the width of this distribution is broad compared to kT. The price one pays is that α now becomes specifically temperature dependent. This has been experimentally verified for silver films (31). Similar temperature behavior, but with systematic quantitative deviations from the Dutta–Horn activated energy model, has been found for silicon-on-sapphire wafers.[32]

Machlup[34] stressed that the scale invariant $\rho(\tau)d\tau = d\tau/\tau$ which leads to $1/f$ noise is a reflection that nature " . . . is sufficiently chaotic to possess . . . a large ensemble of mechanisms with no prejudice about scale . . . ." In this view the semiconductor noise is just a particular example of the basic message that nature is scale invariant. The log-normal distribution has the required scale-invariant characteristic. We have listed several models that lead to the log-normal distribution. Machlup in the 1950s while at the Bell Laboratories was assigned the $1/f$ noise problem by Shockley. We have used Shockley's 1957 model of scientific productivity (as applied more generally to multiplicative random processes) to derive the scale-invariant distribution of relaxation times sought by Machlup.

A broad variety of systems that display $1/f$ noise is discussed in Ref. 33.

This research was partially supported by the Advanced Research Project Agency of the U.S. Department of Defense.

APPENDIX: ON THE EVALUATION OF THE POWER SPECTRUM

In (46) for the power spectrum we encounter the integral

$$I(\omega, \sigma) = \int_{-\infty}^{\infty} \frac{\exp(-y^2/2\sigma^2)}{\omega e^y + (\omega e^y)^{-1}}\,dy$$

We first note that

$$\left[\omega e^{y} + (\omega e^{y})^{-1}\right]^{-1} = \begin{cases} \omega e^{y}\left[1 - (\omega e^{y})^{2} + (\omega e^{y})^{4} \cdots\right] & \text{if } \omega e^{y} < 1 \\ (\omega e^{y})^{-1}\left[1 - (\omega e^{y})^{-2} + (\omega e^{y})^{-4} \cdots\right], & \text{if } \omega e^{y} > 1 \end{cases} \tag{A.1}$$

Since at $\omega e^{y} = 1$, $y = \log(1/\omega)$

$$\begin{aligned} I(\omega, \sigma) = &\int_{-\infty}^{\log(1/\omega)} \omega e^{y} \left\{1 - (\omega e^{y})^{2} + (\omega e^{y})^{4} \cdots\right\} \exp(-y^{2}/2\sigma^{2})\, dy \\ &+ \int_{\log(1/\omega)}^{\infty} (\omega e^{y})^{-1}\left\{1 - (\omega e^{y})^{-2} + (\omega e^{y})^{-4} \cdots\right\} \\ &\times \exp\left[-(y^{2}/2\sigma^{2})\right] dy \end{aligned} \tag{A.2}$$

Then, if we replace y by $-y$ in the first integral

$$\begin{aligned} I(\omega, \sigma) = &\int_{\log \omega}^{\infty} \omega e^{-y} \left\{1 - (\omega e^{-y})^{2} + (\omega e^{-y})^{4} \cdots\right\} \exp\left[-(y^{2}/2\sigma^{2})\right] dy \\ &+ \int_{\log(1/\omega)}^{\infty} (\omega e^{-y})^{-1}\left\{1 - (\omega e^{y})^{-2} + (\omega e^{y})^{-4} \cdots\right\} \\ &\times \exp\left[-(y^{2}/2\sigma^{2})\right] dy \end{aligned} \tag{A.3}$$

A typical integral on the right-hand side of this equation has the form

$$\begin{aligned} &\frac{\sqrt{\pi}}{2} \int_{a}^{\infty} \exp\left[-(2n+1)y\right] \exp(-y^{2}/2\sigma^{2})\, dy \\ &\quad = (2\sigma^{2})^{1/2} \exp\left(\tfrac{1}{2}\sigma^{2}\right)(2n+1)^{2} \operatorname{Erfc}\left[a + (2n+1)\sigma^{2}\right]/(2\sigma^{2})^{1/2} \end{aligned}$$

Hence

$$\begin{aligned} \frac{\sqrt{\pi}}{2} I(\omega, \sigma) = &(2\sigma^{2})^{1/2} \sum_{n=0}^{\infty} (-1)^{n} \omega^{2n+1} \exp\left(\tfrac{1}{2}\sigma^{2}\right)(2n+1)^{2} \\ &\times \operatorname{Erfc}\left[\log \omega + (2n+1)\sigma^{2}\right]/(2\sigma^{2})^{1/2} \\ &+ (2\sigma^{2})^{1/2} \sum_{n=0}^{\infty} (-1)^{n} \omega^{-2n-1} \exp\left(\tfrac{1}{2}\sigma^{2}\right)(2n+1)^{2} \\ &\times \operatorname{Erfc}\left[\log(1/\omega) + (2n+1)\sigma^{2}\right]/(2\sigma^{2})^{1/2} \end{aligned} \tag{A.4}$$

For either large or small σ the argument of the Erfc function is large hence we may use the asymptotic formula of Laplace

$$\frac{\sqrt{\pi}}{2}\operatorname{Erfc} x \sim \frac{1}{2} e^{-x^2}\left(\frac{1}{x} - \frac{1}{2x^3} + \frac{3}{4x^5} - \frac{3\cdot 5}{8x^7} + \cdots\right) \tag{A.5}$$

We first find the contribution of the factor $\frac{1}{2}\exp(-x^2)$ to the terms in (A.4). Then we calculate influence of the factor in the parentheses of (A.5) to (A.4). Since

$$\begin{aligned} &\tfrac{1}{2}\exp - \left[\log\omega + (2n+1)\sigma^2\right]^2/2\sigma^2 \\ &\quad = \tfrac{1}{2}\left\{\exp\left[-(\log\omega)^2/2\sigma^2\right]\right\}\exp\left\{-\log\omega^{2n+1}\right\}\exp - \tfrac{1}{2}\sigma^2(2n+1)^2 \end{aligned} \tag{A.6}$$

$$I(\omega,\sigma) = \tfrac{1}{2}(2\sigma^2)^{1/2}\exp\left[-(\log\omega)^2/2\sigma^2\right]\sum_{n=0}^{\infty}(-1)^n\left\{c_n(\omega) + c_n(1/\omega)\right\}$$

where

$$\begin{aligned} c_n(\omega) = {} & \frac{(2\sigma^2)^{1/2}}{\log\omega + (2n+1)\sigma^2} - \frac{(2\sigma^2)^{3/2}}{2\left[\log\omega + (2n+1)\sigma^2\right]^3} \\ & + \frac{3(2\sigma^2)^{5/2}}{4\left[\log\omega + (2n+1)\sigma^2\right]^5} - \cdots \end{aligned} \tag{A.7}$$

When σ is large the first term in $c_n(\omega)$ may be expanded as

$$\begin{aligned} \frac{\sqrt{2}}{\sigma(2n+1)} & \left\{1 - \frac{\log\omega}{(2n+1)\sigma^2} + \frac{\log^2\omega}{(2n+1)^2\sigma^4}\right. \\ & \left. - \frac{\log^3\omega}{(2n+1)^3\sigma^6} + \frac{\log^4\omega}{(2n+1)^4\sigma^8} \cdots\right\} \end{aligned} \tag{A.8}$$

Since $c_n(1/\omega)$ has the same form except with $\log\omega$ replaced by $\log(1/\omega) = -\log\omega$ the odd powers of $\log\omega$ cancel each other from $c_n(\omega)$ and $c_n(1/\omega)$. Since

$$\sum_{n=0}^{\infty}\frac{(-1)^n}{(2n+1)^{2l+1}} = \frac{\pi^{2l+1}|E_{2l}|}{2^{2l+2}(2l)!} \tag{A.9a}$$

with E_l being the Euler numbers

$$E_2 = -1,\quad E_4 = 5,\quad E_6 = -61,\quad E_8 = 1385,\quad E_{10} = -50521, \ldots \tag{A.9b}$$

and

$$\sum_{n=0}^{\infty} \frac{(-1)^n}{2n+1} = \frac{\pi}{4} \tag{A.9c}$$

the summations over the reciprocals of powers of the odd integers are easily performed in (A.6) using the expansions of the various terms in (A.7) (following the scheme of (A.8)). We finally obtain

$$\begin{aligned} I(\omega,\sigma) = \frac{1}{2}\pi \Bigg\{ & \left[1 - \frac{1}{2}\left(\frac{\pi}{2\sigma}\right)^2 + \frac{5}{8}\left(\frac{\pi}{2\sigma}\right)^4 \cdots\right] \\ & + \frac{\pi^2}{2}\left(\frac{\log\omega}{2\sigma^2}\right)^2\left[1 - \frac{5}{2}\left(\frac{\pi}{2\sigma}\right)^2 + \frac{61}{8}\left(\frac{\pi}{2\sigma}\right)^4 \cdots\right] \\ & + \frac{5\pi^4}{24}\left(\frac{\log\omega}{2\sigma^2}\right)^4\left[1 - \frac{61}{10}\left(\frac{\pi}{2\sigma}\right)^2 + \frac{277}{8}\left(\frac{\pi}{2\sigma}\right)^4 \cdots\right] \\ & + \cdots \Bigg\} \end{aligned} \tag{A.10}$$

REFERENCES

1. L. Boltzmann, Vienna Academy 1877: Paper number 39 in Gesammelte werke.
2. J. C. Slater, *Introduction to Chemical Physics* (McGraw-Hill, New York, 1938).
3. E. T. Jaynes, *Phys. Rev.* **106**:620 (1957).
4. R. D. Levine and M. Tribus, eds., *The Maximum Entropy Formalism* (MIT Press, Cambridge, Massachusetts, 1979).
5. P. Lévy, *Théorie de L'addition des variables aléatoires* (Gauthier-Villars, Paris, 1937).
6. B. V. Gnedenko and A. N. Kolmogorov, *Limit Distributions for Sums of Independent Random Variables* (Addison-Wesley, Cambridge, Massachusetts, 1954).
7. W. Feller, *An Introduction to Probability Theory and its Applications*, Vol. 2. (Wiley, New York, 1966).
8. E. W. Montroll and B. J. West, in *Fluctuation Phenomenon*, J. L. Lebowitz and E. W. Montroll, eds. (North-Holland Publishing Co., Amsterdam, 1979), p. 62.
9. B. B. Mandelbrot, *The Fractal Geometry of Nature* (W. H. Freeman, San Francisco, 1982).
10. B. D. Hughes, M. F. Shlesinger, and E. W. Montroll, *Proc. Natl. Acad. Sci. (USA)* **78**:3287 (1981); M. F. Shlesinger and B. D. Hughes, *Physica* **A109**:597 (1981).
11. F. Galton, *Proc. R. Soc.* **29**:365 (1879).
12. K. Pearson, *Life, Letters and Labors of Francis Galton* (4 vols., 1914–1930, London).
13. D. McAlister, *Proc. R. Soc.* **29**:367 (1879).
14. J. Atchison and J. A. C. Brown, *The Lognormal Distribution*, (Cambridge University Press, Cambridge, 1963).
15. W. Shockley, *Proc. IRE* **45**:279 (1957).
16. P. W. Anderson, D. J. Thouless, E. Abrahams, and D. C. Fisher, *Phys. Rev. B* **22**:3519 (1980).

17. P. W. Anderson, *Phys. Rev.* **B23**:4828 (1981).
18. A. Douglas Stone, J. D. Joannopoulos, and D. J. Chadi, *Phys. Rev.* **B24**:5583 (1981).
19. M. Ya Azbel, *Phys. Rev.* **B26**:4735 (1982).
20. A. N. Kolmogorov, *C. R. Acad. Sci. USSR* **31**:99 (1941).
21. E. W. Montroll and M. F. Shlesinger, *Proc. Natl. Acad. Sci.* **79**:3380 (1982).
22. W. Badger, in *Mathematical Models as a Tool for the Social Sciences*, B. J. West, ed. (Gordon and Breach, New York, 1980), p. 87.
23. R. Gibret, *Les inégalites economiques* (Paris, 1931).
24. L. R. Klein, *An Introduction to Econometrics* (Prentice-Hall, Englewood Cliffs, New Jersey, 1962).
25. I. B. Kravis, *The Structures of Income, Some Quantitative Essays* (University of Pennsylvania Press, Philadelphia, 1962).
26. H. T. Davis, *Theory of Econometrics* (Principia Press, Bloomington, Indiana, 1941).
27. E. W. Montroll and W. Badger, *Introduction to Quantitative Aspects of Social Phenomena* (Gordon and Breach, New York, 1974).
28. M. Nelkin and A. K. Harrison, *Phys. Rev.* **B26**:6696 (1982).
29. A. van der Ziel, *Physica* **6**, 359 (1950).
30. D. A. Bell, Electrical Noise (Van Nostrand, London, 1960), p. 228.
31. P. Dutta and P. M. Horn, *Rev. Mod. Phys.* **53**, 497 (1981).
32. M. B. Weissman, R. D. Black, P. J. Restle, and T. Ray, *Phys. Rev. B*, **27**:1428 (1983).
33. Sixth International Conference on Noise in Physical Systems, P. H. E. Meijer, R. D. Mountain, and R. J. Souler, Jr., eds. (National Bureau of Standards, Washington, D.C., Special publication No. 614, 1981).
34. S. Machlup, in Ref. 33, p. 157.
35. U.S. National Resources Committee, *Consumer Incomes in the U.S.* 1935–1936 (U.S. Government Printing Office, Washington, D.C., 1939).