Statistical Models & Turbulence (La Jolla, CA, 1972)
Springer, 1972, 333-351

N14

Lognormal hypothesis and distribution of energy dissipation in intermittent turbulence

• ***Editorial changes.*** The type-offset original of this chapter is titled "Possible refinement of the lognormal hypothesis concerning the distribution of energy dissipation in intermittent turbulence." That old text repeatedly refers to Kolmogorov's third hypothesis as "appearing untenable" or as being "probably untenable." After this text had been typeset, and when it was being copy-edited, I shortened the title and took the liberty of skipping the qualifiers "appear" or "probably." Their use was *probably* motivated by the profound admiration for Kolmogorov (expressed in Chapter N2, Section 5), by my consistent reluctance to belabor error, and *probably* also by expected disapproval had I done so.

The exponent q was denoted by p in the original, and the Sections were not numbered. An earlier version of this paper had the more straightforward title, "Note on intermittency obtained through multiplicative perturbations." An excellent "advance abstract" of this chapter is provided by the following excerpts from this earlier version. •

✦ **Advance abstract.** My very modest aim is to analyze the assertion common to [several authors], that Kolmogorov's ε, and more generally the local intensity of turbulence, should be expected to follow the lognormal distribution ... Unfortunately, except for ... cases of limited interest, [the usual] derivation [of lognormality] will be shown to be seriously in error, and as a consequence the earlier [authors'] assumptions of lognormality are seriously flawed. Indeed, multiplicative perturbations predict a distribution that not only grossly differs from the lognormal, but also depends upon details in the process of multiplicative perturbations that is being

used. For those concerned with experimental check, this last aspect, naturally, is disquieting.

My remarks apply equally to [a variety of other] subject matters but I shall phrase them in terms of turbulence.

Multiplicative perturbations in general. [The authors cited], as I read them, argue that when turbulence is intermittent, its local intensity at a point in space-time (Kolmogorov's ε) can be considered the multiplicative resultant of many independent and stationary random functions, each corresponding to a different range of eddy wavelengths. The resultant of the eddies whose frequency λ satisfies $1 < \lambda < f$ is thus an approximate local intensity, which will be designated as $X'(t, f)$. The purpose of models of this kind is to list assumptions from which one can deduce the properties of $X'(t, f)$ and, much more important, the properties of local averages of $X'(t, f)$, as defined in one dimension by the integral $X(t, f) = \int_0^t X'(s, f)ds$....

The reason averages are important is that even the most refined measurements concern zones that are much larger than the range of viscosity in a fluid... The most important issue is the extent to which the properties of $X(t, f)$ depend on f when f is large, and how they behave as $f \to \infty$. Also, given that [all] models are unavoidably artificial and oversimplified, an important issue for such models is the extent to which the properties of $X(t, f)$ continue to depend on the original mechanism postulated for the perturbations.

Roughly, [the previous authors invoke], in succession, a central limit theorem, a mean value theorem and finally a law of large numbers. [But this argument] happens to be grossly incorrect, except that the lognormal law does apply to $X(t, f)$ as a rough approximation for finite f when $|t|$ is not much above $1/f$, and for arbitrary f when the effects of high frequency eddies are very weak – in a sense to be defined below. In all other cases, the correct distribution of $X(t, \infty)$ is far from lognormal; in many ways, it is a more interesting [distribution]. In particular, we have the striking result that if the high frequency eddies are sufficiently strong, – again in a sense to be defined below – [and if f is large, then] $X(t, f)$ [nearly] vanishes, save under circumstances of minute probability, under which it may be enormous. To get farther [away] from ... [asymptotic] averaging would be difficult. ✦

✦ **Abstract.** Obukhov, Kolmogorov and others argued that energy dissipation in intermittent turbulence is lognormally distributed. This hypothesis is shown to be untenable: depending upon the precise formulation chosen, it is either unverifiable or inconsistent. The paper proposes a variant of the generating model leading to the lognormal. This variant is consistent, appears tractable, and for sufficiently small values of its unique parameter μ it yields the lognormal hypothesis as a good approximation. As μ increases, the approximation worsens, and for high enough values of μ, the turbulence ends by concentrating in very few huge "blobs." Still other consistent alternative models of intermittency yield distributions that differ from the lognormal in the opposite direction; these various models in combination suggest several empirical tests. ✦

1. INTRODUCTION

A striking feature of the distributions of turbulent dissipation in the oceans and the high atmosphere is that both are extremely "spotty" or "intermittent" in a hierarchical fashion. In particular, both are very far from being homogeneous in the sense of the 1941 Kolmogorov-Obukhov theory. Nevertheless, many predictions of this classic theory have proved strikingly accurate. Self-similarity and the $k^{-5/3}$ spectrum have not only been observed, but are found to hold beyond their assumed domain of applicability. An unexpected embarrassment of riches, and a puzzle!

For many scientists, studying turbulence is synonymous with attempting to derive its properties, including those listed above, from the Navier-Stokes equations of fluid mechanics. But one can also follow a different tack and view intermittency and self-similar statistical hierarchies as autonomous phenomena.

Early examples of this approach in the literature are few in number, but they go very far back in time, and have involved several disciplines. In the field of cosmology, intermittency had already been faced in the eighteenth century, and its study underwent bursts of activity in the period 1900-1920 and today. Furthermore, concern with intermittency arose, independently and nearly simultaneously, in the fields of turbulence (including work by Obukhov, and later by Gurvich & Kolmogorov, Novikov & Stewart and Yaglom), in the study of geomorphology – especially in the study of the distribution of rare minerals (including work by deWijs and Matheron) – and finally in my own work concerning many non-thermal noises. They go under such names as "burst noise," "impulse noise," "flicker noise," and " $1/f$ noise," and may be considered forms of electromagnetic turbulence. As it happens, despite obvious differences, all the scientists working in these fields have followed the same few generic paths. What has brought these various applications together is not yet clear: it may be either their common underdevelopment or genuine kinship.

We shall be concerned with one of these generic paths, which may be designated as the "method of self-similar random multiplicative perturbations." It had two widely distinct sources, the first was a footnote remark in Landau & Lifshitz 1953 concerning the 1941 Kolmogorov theory of self-similar homogenous turbulence. This remark was taken up by Obukhov 1962, and discussed and developed by Kolmogorov 1962, Yaglom 1966 and Gurvich and Yaglom 1967. The second source lay in works by deWijs 1951, and then Mathéron 1962 and his school, on the distribution of rare minerals.

Using the vocabulary of turbulence, let η and L designate the Kolmogorov micro- and macro-scales, and let $\varepsilon(\mathrm{x}, r, \eta, L)$ be the average energy dissipation over the cube of side r and center x. Obukhov and Kolmogorov hypothesize, and de Wijs and Yaglom attempt to derive, the property that $\log[\varepsilon(\mathrm{x}, r, \eta, L)]$ is a normal random variable of variance equal to $A(\mathrm{x}, t) + \mu \log (L/r)$, where the term $A(\mathrm{x}, t)$ depends on the characteristics of the large scale motion and μ is a parameter, possibly a universal constant. The above assertion is usually called "Kolmogorov's third hypothesis."

In addition, the expectation of $\log \varepsilon$ is ordinarily assumed equal to $-(\mu/2)\log(L/r) - A(\mathrm{x}, t)/2$. Finally, the averages of $\varepsilon(\mathrm{x}, r, \eta, L)$ corresponding to cubes whose scale equals the micro-scale of turbulence, are assumed to have a certain correlation function of the form required by self-similarity;

$$E\,[\varepsilon(\mathrm{x} + \Delta\mathrm{x}, \eta, \eta, L)\,\varepsilon(\mathrm{x}, \eta, \eta, L)] \sim (L/|\Delta\mathrm{x}|)^{\mu} \text{ for large } \Delta x.$$

I will call this last expression the "Gurvich-Zubkovskii correlation." Observe that neither this correlation nor Kolmogorov's "third" hypothesis involve η explicitly, which expresses that they obey Kolmogorov's "second" hypothesis of 1941, which he had maintained unchanged in 1962.

The purpose of the present paper is, first, to show that the above "third hypothesis" raises serious conceptual difficulties which make it untenable; secondly, to propose an improved alternative. The practical relevance of my criticism has not yet been established. It depends upon the value of μ, and each field of application will have to investigate it specifically.

{P.S.1998 Following the probabilists' custom, "random variable" will be shortened to "r.v.."}

2. CRITIQUES OF VARIOUS FORMS OF LOGNORMALITY

Allow me to make the historical background more precise: Obukhov introduces lognormality as an "approximate hypothesis:" on the ground that the lognormal "represents any essentially positive characteristic." Kolmogorov treats lognormality as a "third hypothesis" to be derived from other assumptions. And deWijs and Yaglom derive lognormality from a "cascade" argument. Each approach requires a separate reexamination and critique.

Obukhov's approximate hypothesis. Because it is approximate it can only be examined on pragmatic grounds. Its weakness is that it cannot support the elaborate calculations of moments which have been built on it, because the population moments of the random variable (r.v.) $\exp(Y)$ are extremely sensitive to small deviations of Y from normality.

For example, consider a normal random variable G, a Poisson r.v., P, and a Bernoulli r.v., B, obtained as the sum of a large number H of binomial r.v.'s B_H. When their respective means and variances are equal and large, those three r.v.'s are indeed considered by probabilists as being "nearly identical." But this concept of "near identity" tells little about higher moments of the same order of G, P, and B. A fortiori, the moments of e^G, e^P and e^B of all orders are so influenced by the tails of the various distributions that their values may be very different. For example, suppose they all have the same mean δ and variance δ, and denoted the possible values and probabilities of the binomials B_h contributing to B by B', B'', π' and π'', with $\pi' B' H \ll \delta$, $\pi'' B'' B < \delta$, and $\pi' B' > \pi'' B''$. We then have:

$$E(e^G)^q = \exp(q\delta + \delta q^2/2) = \exp[\delta(q + q^2/2)]$$

$$E(e^P)^q = \exp(-\delta + \delta e^q) = \exp[\delta(e^q - 1)]$$

$$E(e^B)^q = \exp E(e^{\Sigma B} h)^q = (Ee^{qB} h)^H = (e^{q\pi' B'} + e^{q\pi'' B''})^H.$$

For large δ, $E(e^B)^q$ increases like $e^{qH\pi' B'}$, that is, less rapidly than $e^{q\delta}$; $E(e^G)^q$ increases more rapidly, and $E(e^P)^q$ even more rapidly.

Even for $q = 1$, one finds different values, respectively

$$E(e^G) = \exp(1.5\delta) \text{ and } E(e^P) = \exp(1.7\delta).$$

The coefficients of variation are

$$\frac{E[(e^G)^2]}{[E(e^G)]^2} = e^\delta \quad \text{and} \quad \frac{E[(e^P)^2]}{[E(e^P)]^2} = e^{(e-1)^{2\delta}} \sim e^{3\delta}.$$

They differ even more, and higher order moments differ strikingly. In short, as soon as one takes exponentials, B and P *cease* to be good approximations to the normal G. It follows that the significance of moment calculations under Obukhov's approximate hypothesis of lognormality is entirely unclear.

This last finding must be reviewed from the viewpoint of the observation, due to Orszag 1970, that the moments of the lognormal increase too fast to satisfy the so-called Carleman criterion. Consequently, lognormal y intermittent turbulence is *not* determined by its moments. The moments of Poisson intermittent turbulence increase even more rapidly, while those of the binomial do satisfy the criterion. However, we have noted that this property is sensitive to minor deviations from normality, so I hesitate to consider this question solved.

Kolmogorov's suggestion that the lognormal hypothesis be considered as strictly valid. This suggestion encounters a different kind of difficulty. Indeed, let us show that the assumption that the r.v.'s $\log \varepsilon(\mathrm{x}, r, \eta, L)$ are normal for every x and every r is incompatible with the assumption that the correlation of ε follows the self-similar (Gurvich-Zubkovskii) form.

Indeed, if $A(\mathrm{x}, t)$ is replaced by the constant term exp $\bar{\varepsilon}$, the lognormal hypothesis yields:

$$E[r^3\varepsilon(\mathrm{x}, r, \eta, L)]^q = \bar{\varepsilon}^q r^{3q - q(q-1)\mu/2} L^{q(q-1)\mu/2}.$$

When r reaches it maximum value, which is $r = L$, all these moments reduce to $\bar{\varepsilon}^q r^{3q}$, as they should. But we must examine more closely *how* they tend to this limit. Suppose $\mu > 3$ and focus on the second moment ($q = 2$). The exponent of r in the above expression takes the value $6 - \mu < 3$. This inequality expresses that, when r is doubled, $E[r^3\varepsilon]^2$ is multiplied by a factor that is *smaller* than 8. On the other hand, the fact that the Gurvich-Zubkovskii correlation is positive implies that the factor in question must be *greater* than 8. This is a contradiction, as previously announced.

When $\mu < 3$, the contradiction moves up to higher moments, namely to moments such that q satisfies

$$3q - q(q-1)\mu/2 < 3, \quad \text{i.e.} \quad q/3 > 2/\mu.$$

This last criterion will be encountered repeatedly in the rest of the paper.

There is another internal contradiction. Consider the variables log ε corresponding to 8 neighboring small cubes obtained by subdividing a bigger cube. When they are lognormal, consistency also requires the variables log ε corresponding to the big cube to be lognormal. However, sums (and hence averages) of independent lognormal variables are themselves *not* lognormal, which suggests that when the eight small cubes's variables are nearly statistically independent, the above requirement is violated. In particular, when μ is very large, the correlation between variables over neighboring small cubes is very small, which suggests that the dependence is small and that the said requirement is violated.

To sum up, for moderately large values of μ, the lognormal hypothesis could only be consistent with some special rule of dependence for which the correlation function is not positive. I can't imagine any such rule, and circumstantial evidence to be described below makes me doubt such a rule exists. This suggests that *Kolmogorov's strict hypothesis is untenable.*

Lognormality obtained as the conclusion of the deWijs-Yaglom (WY) cascade arguments. One may expect the third form of lognormality flawed. Let us review WY step by step. First step: pave space with a regular grid of eddies: the elementary eddies are cubes of side η, eddies of the next stage are cubes of side 2η, so each contains 8 elementary eddies, etc. Second step: assume that $r = \eta 2^n$ for some integer n while $L = \eta 2^N$ for some integer N, and rewrite $\varepsilon(\mathrm{x}, r, \eta, L)$ as the product:

$$\frac{\varepsilon(\mathrm{x}, r, \eta, L)}{\varepsilon(\mathrm{x}, 2r, \eta, L)} \ \frac{\varepsilon(\mathrm{x}, 2r, \eta, L)}{\varepsilon(\mathrm{x}, 4r, \eta, L)} \ \frac{\varepsilon(\mathrm{x}, 2^{N-1}r, \eta, L)}{\varepsilon(\mathrm{x}, 2^{N}r, \eta, L)} \ \varepsilon(\mathrm{x}, L, \eta, L).$$

Third step: identify the last term as $\bar{\varepsilon}$, and assume the ratios in the above expression to be independent identically distributed r.v.'s. Fourth step: one applies the central limit theorem to the sum of the logarithms of the above rations. Conclusion: when the cube of center x and side r is one of the above eddies, the distribution of the corresponding $\varepsilon(\mathbf{x}, r, \eta, L)$ is lognormal.

Our criticisms of Obukhov's and Kolmogorov's approaches extend to the WY generative model. In addition, WY predictions concern the eddies themselves, so that direct verification is impossible. On the other hand, when our cube of center **x** and radius r is *not* an eddy, $\varepsilon(x, r, \eta, L)$ is *not* lognormal. For example, when r is large and one cube overlaps several big eddies, $\varepsilon(x, r, \eta, L)$ is the average of several independent lognormal variables; as we have seen this implies it is not lognormal. To establish the distribution of ε over an arbitrary cube, one would have to average the distribution corresponding to cubes having the same r and overlapping various numbers of eddies.

3. AN ALTERNATIVE TO LOGNORMALITY: LIMIT LOGNORMAL RANDOM PROCESSES

The basic difficulty with the WY cascade argument is, I think, due to the fact that it imposes *local conservation of dissipation.* This is expressed by the fact that various random ratios of the form $\varepsilon(x, r/2, \eta, L)/\varepsilon(x, r, \eta, L)$ corresponding to different parts of an eddy are required to have an average of one. Especially when μ is large, this requirement implies that such ratios are strongly negatively correlated, a feature which is foreign to the Gurvich-Zubkovskii correlation, but (as we saw) is needed in the Kolmogorov argument.

Conservation on the average. By way of contrast, the variant of the model of multiplicative perturbations proposed in the present paper can be characterized by the feature that *conservation of dissipation is assumed*, not on the local, but *only on the global level.* That is, this model visualizes the cascade process as being combined with powerful mixing motion, and with exchanges of energy that disperse dissipation and free the above ratios from having to average to one. Moreover, in order to better satisfy self-similarity, the hierarchy of eddy breakdowns is taken as continuous rather than discrete. Under these conditions, one can relate ε to a sequence of random functions (r.f.'s) $F'(x, \lambda, L)$ such that the $\log F'(x, \lambda, L)$ are Gaussian, with the variance $\mu \log(L/\lambda)$, the expectation $-(\mu/2)\log(L/\lambda)$ and a spectral density equal to $\mu/2k$ for $1/L < k < 1/\lambda$, and to 0 elsewhere. Consequently, the covariance $C(s, \lambda)$ of $\log F'(x, \lambda, L)$ will be assumed to satisfy $C(s) = \lim_{\lambda \to \infty} C(s, \lambda) = -\mu \log(2\pi e^{\gamma} s/L)$. (Here, γ is the Euler constant, whose value is about 0.577.)

For fixed x and L, $F'(x, \lambda, L)$ is clearly a sequence of lognormal r.v.'s whose expectation is identically 1, while their variance, and hence their skewness and their kurtosis, all increase without bound as $\lambda \to \infty$. The quantity $F(\mathrm{x}, r, \lambda, L)$ will then be defined as the integral of F' over a cube of center x and side r (which need not be any specific cube designated as "eddy"), and will be viewed as $r^3\varepsilon(\mathrm{x}, r, \lambda, L)$, namely as the approximate total dissipation that only takes account of perturbations whose wavelength lies between λ and L.

Note that, in contrast to the WY model, there is no specific grid of eddies in the present model. One resemblance to WY is that when $\eta < \lambda \ll r_1 < r_2 < r_3 \ll L$ and $r_2/r_1 = r_3/r_2$, the ratios

$$\frac{F(\mathrm{x}, r_2, \lambda, L)}{F(\mathrm{x}, r_3\lambda, L)} \quad \text{and} \quad \frac{F(\mathrm{x}, r_3, \lambda, L)}{F(\mathrm{x}, r_2, \lambda, L)}$$

have identical distributions. One difference is that those ratios need not be independent. A second difference is that WY assume that randomness in $F(\mathrm{x}, L, \lambda, L)$ lies entirely beyond the model, while in the present variant the "A (x, t) " is in part due to eddy action.

Our task is to derive the distribution of $F(\mathrm{x}, r, \lambda, L)$. In particular, the smallest value of λ is η, and we must check whether or not the distribution of $F(\mathrm{x}, r, \eta, L)$, for $r > \eta$, is independent of η. If it is, then Kolmogorov's second hypothesis, unchanged from 1941 to 1962, is satisfied.

A delicate passage to the limit. Our procedure will be to keep x and r fixed and view $F(\mathrm{x}, r, \lambda, L)$ as a r.f. of λ. From the mathematical viewpoint, this r.f. happens to be a "martingale" and, of course, $F \geq 0$. Doob's classical "convergence theorem for positive martingales" (Doob 1953, p. 319) states that $\lim_{\lambda \to 0} F(\mathrm{x}, r, \lambda, L) = F(\mathrm{x}, r, 0, L)$ exists. This result suggests it may be legitimate for small but positive λ say for $\lambda = \eta$, to view $F(\mathrm{x}, r, \lambda, L)$ as differing from $F(\mathrm{x}, r, 0, L)$ by a "perturbation term."

However, the convergence theorem allows two possibilities. {P.S. 1998: This is a major but unavoidable complication.} The limit may be either non-degenerate, that is, have a positive probability of being finite and positive, or degenerate, that is, almost surely reduced to 0 {P.S. 1998: In the former case, the expectation of the limit is 1. In the latter case, the expectation of the limit is 0, despite the fact that the limit of the expectations is 1.}

When $F(\mathrm{x}, r, 0, L)$ is nondegenerate, then for small but positive values of λ, such as $\lambda = \eta$, a perturbation term dependent on λ is required, but one may consider that Kolmogorov's second hypothesis *holds.*

But when $F(\underline{x}, r, 0, L)$ is degenerate, then for small λ, $F(\mathrm{x}, r, \lambda, L)$ either nearly vanishes, with probability nearly 1, or is extraordinarily large, with a very small probability. This last probability tends to zero with λ, but it is finite if $\lambda > 0$, which explains why the normalizing constraint $EF(\mathrm{x}, r, \lambda, L) = 1$ could be imposed without contradiction. Nevertheless, the perturbation term is non-negligible, so $F(\mathrm{x}, r, 0, L)$ is a bad approximation and Kolmogorov's second hypothesis *fails.* The preceding alternative shows the importance of determining which of the above alternates holds for given μ. More precisely, we shall seek when and to which extent the lognormal approximation to $F(\mathrm{x}, r, L)$ is reasonable.

First main result. $F(\mathrm{x}, r, L)$ is *degenerate* when $\mu > 6$, and *nondegenerate* when $\mu < 6$.

Second main result. In the nondegenerate case $\mu < 6$, the moment $EF^q(\mathrm{x}, r, 0, l)$ is finite when $\mu < 6/q$ and infinite when $\mu > 6/q$.

This behavior suggests that, for large values of u,

$$\Pr\{F(\mathrm{x}, r, 0, L) > u\} \sim C(r, L)u^{-6/\mu}.$$

Thus, as $\mu \to 0, F(\mathrm{x}, r, 0, L)$ acquires an increasing number of finite moments, which are shown to converge towards those of the lognormal. This result constructively establishes that for small μ, the cascade scheme of deWijs and Obukhov can be modified so as to avoid the difficulties that have been listed above without significantly changing the prediction. For large μ, on the other hand, the required changes are significant.

The transition criterion $\mu = 6/q$ was already encountered in the discussion of the inconsistency of Kolmogorov's strict hypothesis. Those same high moments that seemed to behave inconsistently no longer do so here. The reason is that in the present model they are infinite throughout. Proofs of the above assertions will be given in the following two sections; each uses specific mathematical tools appropriate to its goal.

➡ ➡ ➡

FIGURE C14-1.

The graphs found on the next three pages are computer simulated approximations to one-dimensional self-similar limit lognormal r.f.'s $F(x, r, \lambda, L, \mu)$. The are plotted for successive values of x, each a multiple of r.

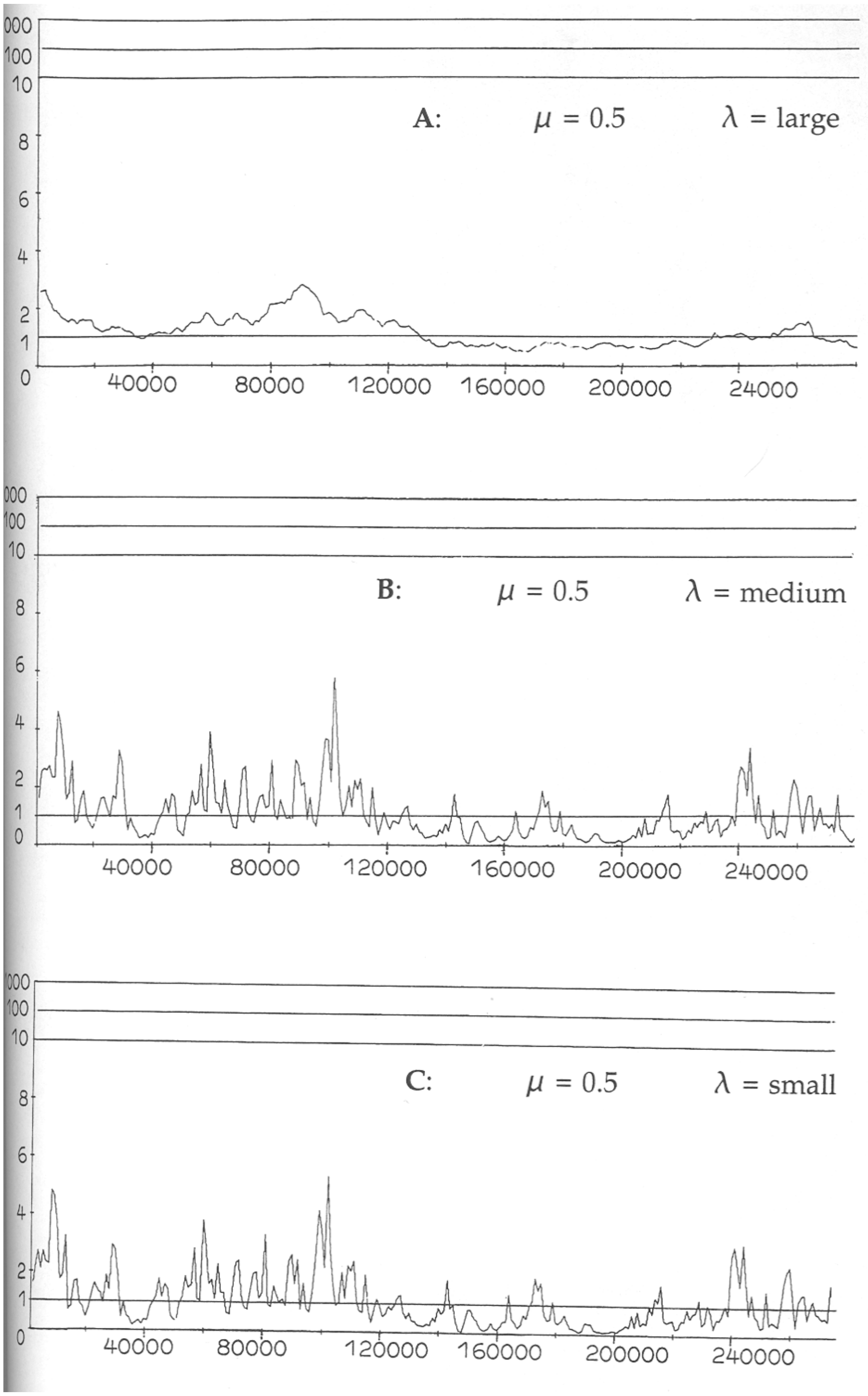
Method of construction: Define $F'(x, \lambda, L = 10^7, \mu = 2)$ as a Gaussian r.f. for discrete x with $1 \le x \le 560{,}000$, with a spectral density that is approximately equal to $1/k$ for $1/L < k < 1/\lambda$. This function was simulated on the IBM System 360/Model 91 for selected values of μ and λ, and $F(x, r, \lambda, L, \mu)$ was computed for x multiple of 1000, using the formula $\Sigma_{u=1}^{1000}[F'(u, \lambda, 10^7, 2)]^{\mu/2}$. The ordinate is the ratio $R(x, \lambda, \mu)$ between F and the median of the values of F along the sample. For each μ, the output of the program is a Calcomp tracing across a broad strip of paper. All the programs were written by Hirsh Lewitan using the fast fractional Gaussian noise algorithm described in M 1971f.

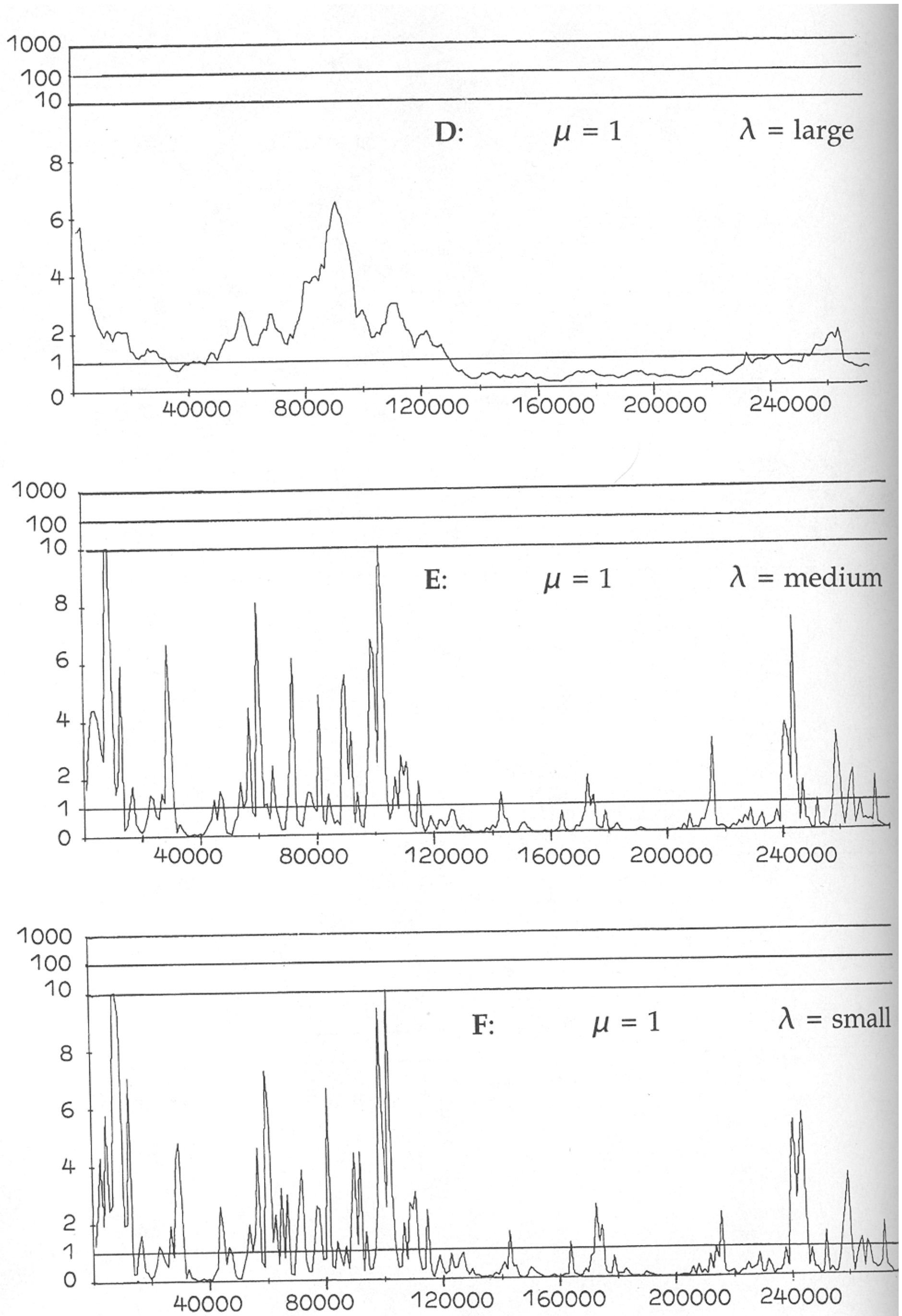
On the next three pages, portions of these graphs are shown for $\mu = 0.5$, $\mu = 1$, and $\mu = 4$, respectively, with λ decreasing down the page.

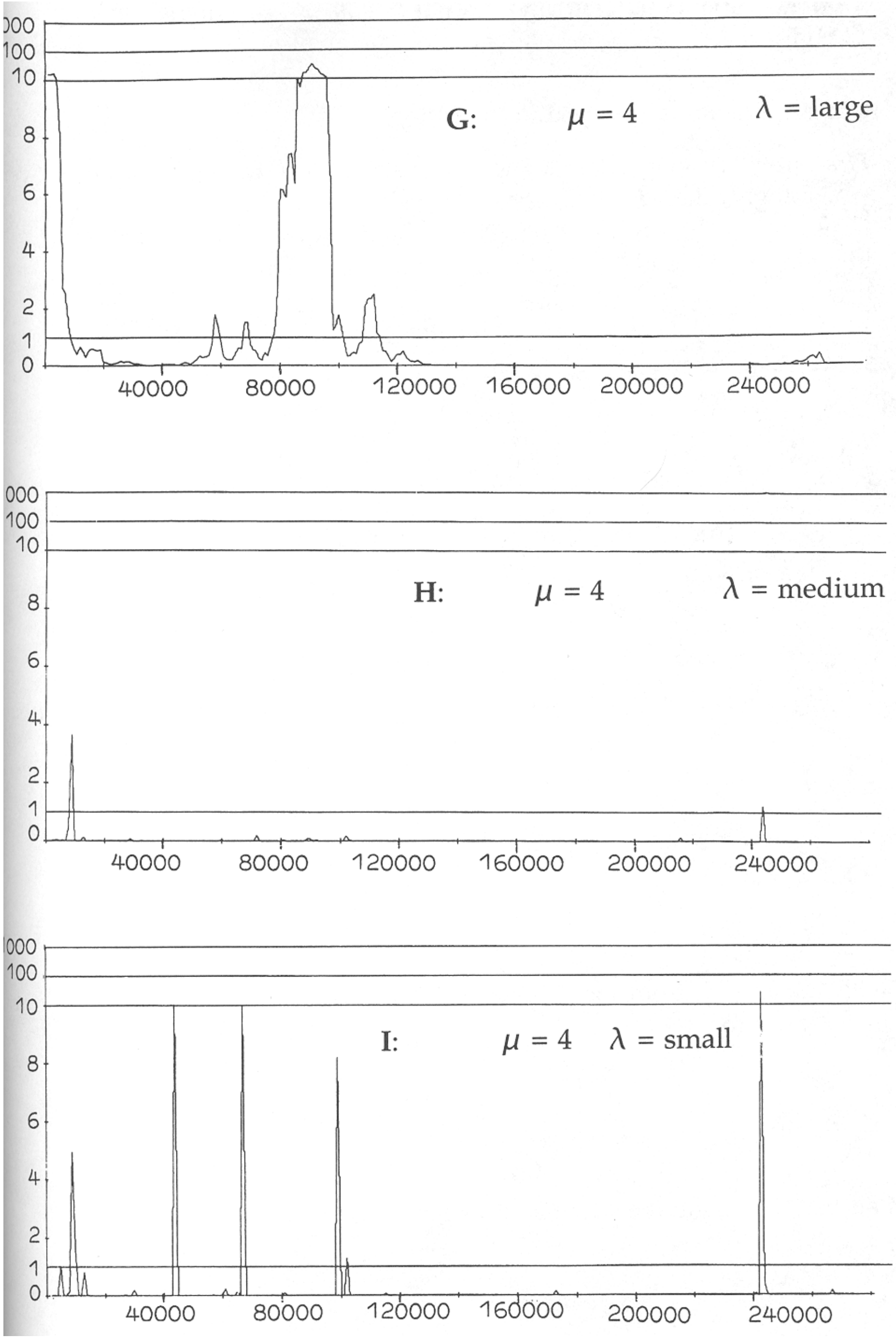
Analysis of the results. The theory predicts that when μ is small (graphs A, B, and C), the ratio R (x, λ, μ) converges to a limit. The simulations clearly confirm that R soon ceases to vary. The ostensible limit is clearly non-Gaussian, but not extremely so.

As μ increases (graphs D, E, and F), the point of ostensible convergence moves towards decreasing values of λ, and the non-Gaussian character of the ostensible limit of R becomes increasingly apparent. In particular, an increasing proportion of the cumulated F becomes due to a decreasing number of sharp peaks and blobs. (The peaks are truncated at 10 for the sake of legibility, but it was decided *not* to plot log R instead of R.)

Finally, $\mu = 2$ (graphs G, H, and I) is the critical value of μ in one dimension (as $\mu = 6$ was in three dimensions). Fron this value on, R ceases to converge to a limit. This lack of convergence is clearly seen on this simulation.


A: μ = 0.5 λ = large
B: μ = 0.5 λ = medium
C: μ = 0.5 λ = small
40000
80000
120000
160000
200000
24000
240000


D: $\mu = 1$ λ = large
1000
100
10
8
6
4
2
1
0
40000 80000 120000 160000 200000 240000
E: $\mu = 1$ λ = medium
1000
100
10
8
6
4
2
1
0
40000 80000 120000 160000 200000 240000
F: $\mu = 1$ λ = small
1000
100
10
8
6
4
2
1
0
40000 80000 120000 160000 200000 240000


G: μ = 4 λ = large
H: μ = 4 λ = medium
I: μ = 4 λ = small
40000
80000
120000
160000
200000
240000

4. THE LIMIT LOGNORMAL MODEL FOR $\mu > 6$: DEGENERATE LIMIT

Our argument will proceed in three steps.

First step: In very skew lognormal distributions, the expectation is overwhelmingly due to occasional large values. Therefore, let $\log F'$ be Gaussian with variance $\mu \log(L/\lambda)$ and expectation $-(\mu/2)\log(L/\lambda)$, implying $F' \equiv 1$, and let $N(\lambda)$ be a function such that

$$\lim_{\lambda \to 0} N(\lambda) = \infty, \text{ while } \lim_{\lambda \to 0} N(\lambda)/\sqrt{\log(L/\lambda)} = 0.$$

Also define three functions, Threshold (λ, L), F'_+, and F'_- as follows:

$$\text{Threshold } (\lambda, L) = (L/\lambda)^{\mu/2} \exp[-N(\lambda)\sqrt{\mu \log(L/\lambda)}\,];$$

$$\text{if } F'(x, \lambda, L) \geq \text{ Threshold } (\lambda, L), \text{ then } F'_+ = F' \text{ and } F'_- = 0$$

$$\text{if } F'(x, \lambda, L) < \text{ Threshold } (\lambda, L), \text{ then } F'_+ = 0 \text{ and } F'_- = F'.$$

Finally, let F_+ and F_- be the integrals of F'_+ and F'_-.

The motivation of the above definitions lies in the value of the expectation

$$EF_+(x, r, \lambda, L) = \frac{r^3}{\sqrt{2\pi\mu \log(L/\lambda)}} \int \exp\left\{x - \frac{[x + (\mu/2)\log(L/\lambda)]^2}{4\mu \log(L/\lambda)}\right\} dx,$$

with integration from log{ Threshold (λ, L)} to infinity. Transforming the integrand into

$$\exp\left\{-\frac{[x - (\mu/2)\log(L/\lambda)]^2}{4\mu \log(L/\lambda)}\right\},$$

and then changing the variable of integration, we obtain

$$EF_{+}(\mathrm{x}, r, \lambda, F) = \frac{r^3}{\sqrt{2\pi}} \int_{-N(\lambda)}^{\infty} f \exp(-z^2/2)\, dz.$$

Conclusion. This form of EF_{+} shows that the contribution of F_{-} to F is asymptotically negligible for $\lambda \to 0$, and that the above choice of $N(\lambda)$ has been appropriate to insure that F is arbitrarily closely approximated by F_{+}. Moreover, for $\lambda > 0$, the function F' is a.s. continuous, so the variation of F is a.s. concentrated on those intervals where $F' = F'_{+}$.

Second step: Over any cube of side λ, F' and therefore F'_{+}, is near constant. Hence

$$F(\mathrm{x}, r, \lambda, L) \sim \lambda^3 \sum F' + (\mathrm{x}, \lambda, L)$$

with summation carried out over those points of a regular lattice of side λ for which $F' > \mathrm{Threshold}(\lambda, L)$.

Third step: We suspect that there exist circumstances under which $\lim_{\lambda \to 0} F(\mathrm{x}, r, \lambda, L) = 0$. At least some of these circumstances may also fulfill the stronger sufficient condition that in a cube of side L, the random number of lattice sites for which $F'_{+} > 0$ tends to 0 *almost surely.* A sufficient condition for the latter property is that the number in question should tend to 0 *on the average.* This expected number equals

$$(L/\lambda)^{-3} \Pr \{ F'(\mathrm{x}, \lambda, L) > \mathrm{Threshold}(\lambda, L) \}.$$

In terms of the r.v.

$$\frac{\log F' + (\mu/2) \log(L/\lambda)}{\sqrt{\mu \log(L/\lambda)}},$$

which is a reduced Gaussian G, the Pr in the equation before last becomes

$$\Pr \{G > \sqrt{\mu \log(L/\lambda)} - N(\lambda)\}.$$

Using the standard tail approximation of G, the expected number in question is about

$$\frac{(L/\lambda)^3 \exp[-(\mu/2)\log(L/\lambda)]}{\sqrt{2\pi\mu \log(L/\lambda)}} = \frac{(L/\lambda)^{3-\mu/2}}{\sqrt{2\pi\mu \log(L/\lambda)}} .$$

Note that this last approximation is independent of $N(\lambda)$ for $\lambda \to 0$.

For this expression to tend to 0 with λ, *a sufficient* condition is $\mu > 6$. (It is also a necessary condition, but this is besides the point; see below.) It follows that, when $\mu > 6$, $\lim_{\lambda \to 0} F(\mathrm{x}, r, \lambda, L) = 0$ almost surely. Obviously, the limit is far from being distributed lognormally.

The preceding argument is heuristic, but it is the best I can do in three dimensions. The one dimensional version of the limit lognormal process is easier to study, and the heuristics can be made rigorous by using the Rice formula for the extreme values of a random function. (This is one more reason why it would be desirable to generalize the Rice formula to higher dimensions.)

Extended to the case $\mu < 6$, the preceding argument suggests that the bulk of the variation of F concentrates in approximately $(L/\lambda)^{3-\mu/2}$ cubes of side λ. As $\lambda \to 0$, each cube either is eliminated or becomes subdivided into numerous subcubes. This conclusion is correct, but the above heuristic proof is mathematically very incomplete. The reason is that condition $\lim_{\lambda \to 0} X(\lambda) < \infty$ does *not* exclude the possibility that $\lim_{\lambda \to 0} X(\lambda) = 0$ almost surely. Mathematical concerns of such nature are usually dismissed by physicists, but in the present instance the misbehavior of F for $\mu > 6$ suggests that extreme care is necessary, and different tools are needed to tackle $\mu < 6$.

5. THE LIMIT LOGNORMAL MODEL FOR μ<6: NON-DEGENERATE LIMIT

In this Section, the moments $EF^q(\mathrm{x}, r, 0, L)$ will be evaluated for integer q, and then compared with the moments $E[rF'(\mathrm{x}, r, L)]$. This last lognormal r.v. provides some kind of link with the WY model. Indeed, it is tempting to reason as follows: $F'(\mathrm{x}, r, L)$ varies little over a cube of side r, while the ratio $F'(\mathrm{x}, \lambda, L)/F'(\mathrm{x}, r, L)$, which equals $F'(\mathrm{x}, \lambda, r)$, varies rapidly. Averaged over the cube of side r, this last ratio is bound to be close to its expectation, which is equal to one. This would imply that $F'(\mathrm{x}, r, \lambda, L)$ is approximated reasonably by $rF'(\mathrm{x}, r, L)$. We shall now check whether or not this is really the case.

The case $\boldsymbol{q=2}$**.** Integrating over the domain where all coordinates of **u** and **v** lie between 0 and r, we have

$$EF^2(\mathrm{x}, r, \lambda, L) = E\int\int\int\int\int\int \exp[\log F'(\mathbf{u}, \lambda, L) + \log F'(\mathbf{v}, \lambda, L)]d\mathbf{u}\, d\mathbf{v}$$

$$= \int\int\int\int\int\int E \exp[\log F'(\mathbf{u}, \lambda, L) + \log F'(\mathbf{v}, \lambda, L)]d\mathbf{u}\, d\mathbf{v}.$$

The expression in the exponential is a Gaussian r.v. of expectation $-\mu \log(L/\lambda)$ and variance $2\mu \log(L/\lambda) + 2C(|\mathbf{u}-\mathbf{v}|, \lambda)$. As a result,

$$EF^2(\mathrm{x}, r, \lambda, L) = \int\int\int\int\int\int \exp[C(\mathbf{u}-\mathbf{v}, \lambda)]d\mathbf{u}\, d\mathbf{v}.$$

Keep r fixed, with $r \ll L$, and let $\lambda \to 0$. The preceding integral continues to converge if and only if $\mu/2 < 3/2 = 3/q$, in which case its limit for $\lambda \to 0$ equals

$$(2\pi e^{\gamma}/L)^{-\mu}\int\int\int\int\int\int |\mathbf{u}-\mathbf{v}|^{-\mu} d\mathbf{u}\, d\mathbf{v}.$$

Alternatively, carry the integration over the variables, $\mathbf{u}' = \mathbf{u}/r$ and $\mathbf{v}' = \mathbf{v}/r$, whose values vary from 0 to 1. Then the above second moment converges to

$$r^{6-\mu}L^{\mu}[(2\pi e^{\gamma})^{-\mu}\int\int\int\int\int\int |\mathbf{u}'-\mathbf{v}'|^{-\mu} d\mathbf{u}'\, d\mathbf{v}'].$$

By way of contrast, the would-be approximating lognormal $rF'(\mathrm{x}, r, L)$ has a second moment equal to $r^{6-\mu}L^{\mu}$. The ratio between the limit and the approximate moment is the quantity in brackets. As $\mu \to 0$, its integrand and it prefactor $[2\pi e^{\gamma}]^{-\mu}$ both tend to 1, and so does the ratio itself.

Suppose that it is true that $\lim_{\lambda \to 0} EF^q(\mathrm{x}, r, \lambda, L) = EF^q(\mathrm{x}, r, 0, L)$, which is unfortunately not established by the preceding formal calculation. If this were true, it would follow that as $\mu \to 0$, $rF'(\mathrm{x}, r, L)$ becomes a good second order approximation to $F(\mathrm{x}, r, 0, L)$.

The case $q \geq 3$. By a completely similar calculation, we find that

$$\text{iff } \mu/2 < 3/q, \ \lim_{\lambda \to 0} EF^q(\mathrm{x}, r, \lambda, L) < \infty.$$

This suggests that $rF'(\mathrm{x}, r, L)$ is a good approximation to $F(\mathrm{x}, r, 0, L)$ up to the order $6/\mu$. When μ is very small, F^q has very many finite moments and its low order moments lie near those of the lognormal; one is tempted to describe F itself as being near lognormal.

Now we must tackle the mathematical difficulty concerning the agreement or discrepancy between $\lim_{\lambda \to 0} EF^q(\mathrm{x}, r, \lambda, L)$ and $EF^q(\mathrm{x}, r, 0, L)$. I am able to give only an incomplete answer to this question. Let $\tilde{p}(\mu)$ be the largest integer satisfying $\mu/2 < 3/\tilde{p}(\mu)$. When $\tilde{p}(\mu) \geq 3$, which implies $\mu < 2$, a standard theorem on martingales (Doob 1953, p.319, Theorem VII, 4.1, clause iii) suffices to establish that $\mu < 2$ is a sufficient condition for $F(\mathrm{x}, r, 0, L)$ to be nondegenerate, meaning that $\Pr\{F(\mathrm{x}, r, 0, L) > 0\} > 0$. In addition, this theorem establishes that $EF^q(\mathrm{x}, r, 0, L) = \lim_{\lambda \to 0} EF^q(\mathrm{x}, r, \lambda, L)$. In particular, since $2 < \tilde{p}(\mu)$, the above obtained $\lim_{\lambda \to 0} EF^2(\mathrm{x}, r, \lambda, L)$ is indeed the second moment of $F(\mathrm{x}, r, 0, L)$.

A bit of additional manipulation establishes that, for all values of q

$$EF^q(\mathrm{x}, r, 0, L) = (r/L)^{q(q-1)\mu/2} E(\mu, q),$$

where $0 < E(\mu, q) < \infty$ if $r \ll L$ and $q < \tilde{p}(\mu)$, and $E(\mu, 1) \equiv 1$.

6. MISCELLANEOUS REMARKS

Different forms of correlation. The preceding theory of distribution and of correlation concerns *cubes* of side r or η. Experimental measurements, on the contrary, generally concern averages of ε along thin cylinders of fixed uniform cross section and varying length r. Appropriate changes must be made to extend our results to this case.

Experimental verification of the probability distribution predicted for $F(x, r, 0, L)$. One question must be addressed: are the above results specific to the lognormal model, or do they apply more generally? It has been noted that in the scheme of multiplicative perturbations, the set on which the bulk of variation of $X(t,f)$ occurs is greatly influenced by the tails of the distribution of $\log X'(t,f)$. The central limit theorem gives no information about those tails. More generally, different models of multiplicative perturbations may seem to differ by inconsequential details, yet yield different predictions for the distribution of Kolmogorov's ε. In addition, the alternative models of intermittency belonging to the second broad class mentioned in the introduction, namely the models of Novikov & Stewart 1964 and M 1965c and 1967b{N7, N10}, lead to still different concentration sets, and to probability distributions that are *less scattered than the lognormal.* In other words, the multiplicative model is extremely sensitive to its inputs, and appropriately selected variants could account for distributions that are more scattered or less scattered than the lognormal. In truth, the theory in its present stage offers few predictions that the experimentalist can verify.

Generative models of the scaling law. The interplay we have observed between multiplicative perturbations and the lognormal and scaling distributions has incidental applications in other fields of science where very skew probability distributions are encountered. Notable examples occur in economics, e.g., in the study of the distribution of income. Having mentioned that fact, I leave its elaboration to a more appropriate occasion.

&&&&&&&&&&&& ANNOTATIONS &&&&&&&&&&&&

Technical comment on the last paragraph. The last paragraph of M1972j alludes to the following circumstances. M. & Taylor 1967 {E21} had pointed out that the stable processes can be represented (using today's words) as Wiener Brownian motions followed in fractal time. This, my first paper on multifractals, instantly suggested that replacing fractal by multifractal time would yield a new and more general mathematical process showing promise in empirical investigations. In addition, Wiener Brownian motion could be replaced by the fractional Brownian motion introduced in M 1965h{H}.

However, the "more appropriate occasion" called for in the last words of this paragraph did not materialize until after a 25 year delay. Details and references are given in Chapter E6 of M 1997E, the *Selecta* volume devoted to Finance.

Technical comment on multifractals considered as 1/f noises. A side result of this chapter is that the limit lognormal measure has a positive correlation function proportional to t^{-Q}. In loose current terminology, this Q is referred to as a "correlation dimension." In other words, multifractal measures provide an example of f^{-B} noise. This topic is discussed in greater detail in Section N2.2 and the annotations of next chapter.

Roots of lognormality and multiplicative effects in economics. In turbulent dissipation much of the total dissipation is due to deviations that are large, but the very largest peaks are too few to have a significant total contribution. I was prepared to scrutinize the claims for lognormality in Kolmogorov 1962, because I had encountered their counterparts in the totally different context of economics. (See Part III of M 1997E.) That is, I was sensitive to a number of very serious difficulties that were described much later in Chapter E9 of M1997E, unambiguously titled *A case against the lognormal distribution.*

Since most readers of this book are unfamiliar with statistics, it is good to insert at this point some background concerning multiplicative perturbations and lognormality. The standard reference when I was dealing with these matters in the context of economics was Aitchison & Brown 1957. This reference states that the lognormal distribution was first considered in 1879 by a student of Francis Galton. It was rediscovered independently many times.

A great increase in the popularity of multiplicative effects and the lognormal distribution occurred with Gibrat 1931. Robert Gibrat (a French engineer, manager and economist) focussed on economic inequalities such as those in the distribution of personal income. He found that the middle incomes are distributed lognormally, and never faced the fact that the high income tail is definitely *not* lognormal. Pareto's law asserts that this tail follows the scaling distribution. Thus, deviations from lognormality were familiar to experimentalists in fields far removed from the study of turbulence. But those tails were disregarded and in the 1930s Gibrat convinced many statisticians and scientists that the lognormal distribution is, in some way, a basic building block of randomness in nature. That is, many authors feel that no specific justification is needed when randomness is either Gaussian or lognormal, while a specific justification is required for other distributions.

It can be revealed that I have long been dubious about Gibrat's theoretical argument. Contrary to Kolmogorov, Gibrat did not put lognormality as an absolute hypothesis. Instead, very much as Obukhov 1962 was to do for turbulence, Gibrat postulated that log (income) is the sum of many factors, and applied the central limit theorem. No one seemed to be concerned by the fact, that a) one rarely deals with sums of many factors, and b) the central limit theorem says nothing about the tails, while – in economics just like for multifractals – the tails are the interesting portions of the distribution both empirically and theoretically.

Multiplicative effects with a reflecting boundary; critique of their widely advocated use in economics and finance. It is easy to modify Gibrat's argument so that, instead of the lognormal, it leads to the scaling distribution called for by Pareto's law. It suffices to set up a reflecting lower boundary. This is an ancient idea that keeps being resurrected by investigators who wish to use the methods of physics in studies of the social sciences. Strong reservations on those matters are described in Chapter E10 of M 1997E.

The difficult rigorous theory of the limit lognormal multifractal measures. While I had full confidence in the validity of this paper's heuristic results, I was eager to see a rigorous mathematical treatment argument to buttress them. This is why I approached Jean-Pierre Kahane again, sometime in 1972 or 1973, showing him my conjectures. Years before, my heuristics of turbulence inspired Kahane & M 1965{N11}, opening new and interesting "natural" developments in harmonic analysis. That fascinating discipline started with Newton's *spectrum* of light, and with the decomposition of sound into its *harmonics*. With Wiener, it became powerfully affected by the analysis of electrical noises. Authors like Zygmund remained aware of old concrete problems. But others (like my uncle), preferred to "purify" harmonic analysis by forgetting its bright and loud roots. Constant low-key irritation against this attitude makes me welcome every opportunity to demonstrate the continuing power of the "applications" to inspire "pure" mathematics.

In addition to the random multiplicative measures described in this paper, the examples shown to Kahane in 1972 or 1973 included the measures that were later described in M 1974f{N15} and M 1974c{N16}. The latter have a well-defined integer base b, hence can be characterized as as *base-bound*. By contrast, the limit lognormal measures described in this paper can be characterized *base-free*.

As argued in Chapter N1, integer bases are not part of nature, only a mathematical convenience. (I did not know then that physicists use it heavily in renormalization theory; see Chapter N3.) Therefore, I have always strongly favored the base-free measures. Unfortunately, Kahane could not handle them rigorously, as of 1972-3. The base-bound measures, to the contrary, were tackled immediately in Kahane & Peyrière 1976{N17}. It is a well-known fact that when physics and mathematics tackle the same problem, no relation need exist between the levels of technical difficulty that they encounter.

The difficulties Kahane encountered in 1972-3 proved serious, and spurred him to develop delicate new mathematical tools to tackle my construction. He confirmed my base-free conjectures and provided a generalized formal restatement of the limit lognormal multifractals, with new results, as exemplified in Kahane 1987a,b, 1989, 1991a.