Journal of Fluid Mechanics 62, 1974, 331-358

N15

Intermittent turbulence in self-similar cascades: divergence of high moments and dimension of the carrier

✦ **Abstract.** Kolmogorov's "third hypothesis" asserts that in intermittent turbulence the average $\bar{\varepsilon}$ of the dissipation ε, taken over any domain $\mathcal{D}$, is ruled by the lognormal probability distribution. This hypothesis will be shown to be logically inconsistent, save under assumptions that are extreme and unlikely. A widely used justification of lognormality due to Yaglom and based on probabilistic argument involving a self-similar cascade, will also be discussed. In this model, lognormality indeed applies strictly when $\mathcal{D}$ is "an eddy," typically a three-dimensional box embedded in a self-similar hierarchy, and may perhaps remain a reasonable approximation when $\mathcal{D}$ consists of a few such eddies. On the other hand, the experimental situation is better described by considering averages taken over essentially one-dimensional domains $\mathcal{D}$.

The first purpose of this paper is to carry out Yaglom's cascade argument, labelled as "microcanonical," for such averaging domains. The second is to replace Yaglom's model by a different, less constrained one, based upon the concept of "canonical cascade." It will be shown, both for one-dimensional domains in a microcanonical cascade, and for all domains in canonical cascades, that in every non-degenerate case the distribution of $\bar{\varepsilon}$ differs from the lognormal distribution. Depending upon various parameters, the discrepancy may be moderate, considerable, or even extreme. In the latter two cases, one finds that the moment $\left\langle \bar{\varepsilon}^q \right\rangle$ is infinite if q is high enough. This avoids various paradoxes (to be explored) that are present in Kolmogorov's and Yaglom's approaches.

The paper's third purpose is to note that high-order moments become infinite only when the number of levels of the cascade tends to infinity, as is the case when the internal scale η tends to zero. Granted the usual

value of η, this number of levels is actually small, so the limit may not be representative. This issue was investigated through computer simulation. The results bear upon the question of whether Kolmogorov's second hypothesis applies in the face of intermittency.

The paper's fourth purpose is as follows: Yaglom noted that the cascade model predicts that dissipation occurs only in a portion of space of very small total volume. In order to describe the structure of this portion of space, the concept of the "intrinsic fractional dimension" D of the carrier of intermittent turbulence will be introduced

The paper's fifth purpose is to study the relation between the parameters ruling the distribution of $\bar{\varepsilon}$, and those ruling its spectral and dimensional properties. Both conceptually and numerically, these various parameters turn out to be distinct, opening several problems for empirical study. ✦

1. INTRODUCTION AND SYNOPSIS

A striking feature of the distributions of turbulent dissipation in the oceans and the high atmosphere is that both are extremely "spotty" or "intermittent," and that their intermittency is hierarchical. In particular, both are very far from being homogeneous in the sense of the 1941 Kolmogorov-Obukhov theory, in which the rate of dissipation ε was assumed to be uniform in space and constant in time. In intermittent turbulence, ε must be considered a function of time and space. Let $\bar{\varepsilon}$ $(\mathcal{D})$ be its spatial average over a domain $\mathcal{D}$. Several approaches to intermittency view $\bar{\varepsilon}$ as lognormally distributed: in Obukhov 1962, lognormality is a pragmatic assumption; in Kolmogorov 1962, it is a basic "third hypothesis" applicable to every domain $\mathcal{D}$; in Yaglom 1966, it is derived from a self-similar cascade model. Yaglom also finds that there is equality between the parameter μ of the lognormal distribution and the exponent in the expressions ruling the correlation and spectral properties of $\bar{\varepsilon}$.

A closely analogous cascade was considered in de Wijs 1951 & 1953, by a geomorphologist concerned with the variability in the distribution of the ores of rare metals. The results in the present paper may therefore be of help outside turbulence theory. A further incidental purpose of this paper is to provide background material to discussion of instances of interplay between multiplicative perturbations and the log-normal and scaling distributions. Such interplay occurs in other fields of science where very skew probability distributions are encountered, notably in economics.

Having mentioned this broader scope of the methods to be described, I shall leave its elaborations to other more appropriate occasions.

While substantial effort is currently being devoted to testing the lognormality experimentally, the purpose of the present paper is to probe the conceptual foundations of lognormality. Like the works of Kolmogorov and Yaglom, this discussion shall be concerned with a phenomenology whose contact with physics remains remote. In particular, the central role of dissipation will not be questioned. On the other hand, greater care will be devoted to matters of internal logical consistency and to details of the assumptions. The relation between theory and experiment will be explored, and will provide a basis for further development of the theory.

Since this paper is somewhat lengthy, the mathematics that has as yet no other application in fluid mechanics will be postponed to Sections 4 and 5. The main results will be stated without proof in this section and in Section 2. Section 3 will elaborate on the important distinction between microcanonical and canonical cascades.

(a) Part of this paper is devoted to a new calculation relative to Yaglom's cascade model for Kolmogorov's hypothesis of lognormality. Let $\overline{\varepsilon}\,(\mathcal{D})$ be the average of the dissipation ε over a spatial domain $\mathcal{D}$. One form of Yaglom's model assumes that $\mathcal{D}$ is an "eddy," perhaps a three-dimensional cube embedded in a self-similar hierarchy. On the other hand, in all actually observed averages, $\mathcal{D}$ is not a cube, but is more nearly a very thin cylinder. By following up the consequences of Yaglom's model in this case, it will be shown that $\overline{\varepsilon}\,(\mathcal{D})$ is never lognormal, and that its "qualitative" behavior can fall into any one of three classes:

- In a first class, which is drastically extreme and which will be called "regular," $\overline{\varepsilon}(\mathcal{D})$ is not far from being lognormal.

- In a second class, which is equally extreme and which will be called "degenerate," all dissipation concentrates in a few huge blobs.

- In a third class, which is intermediate between the above two and which will be called "irregular," $\overline{\varepsilon}(\mathcal{D})$ is non-degenerate but is far from lognormal.

The most striking characteristic of the third class is a parameter α_1, satisfying $1 < \alpha_1 < \infty$, which rules the moments (ensemble averages) $\langle \overline{\varepsilon}^q(\mathcal{D}) \rangle$. When $q < \alpha_1$, one has $\langle \overline{\varepsilon}^q(\mathcal{D}) \rangle < \infty$ for all values of the inner scale η, but when $q > \alpha_1$ and one has $\eta = 0$, $\langle \overline{\varepsilon}^q(\mathcal{D}) \rangle = \infty$. Finally, when $q > \alpha_1$ and η is positive but small, $\langle \overline{\varepsilon}^q(\mathcal{D}) \rangle$ is huge, and its precise value is so dependent upon η as to be meaningless. The regular class can be viewed as being the limiting case $\alpha_1 = \infty$, and the degenerate class as corresponding to $\alpha_1 \leq 1$. This eliminates numerous inconsistencies that have been noted in the literature, concerning the behavior of the moments of $\bar{\varepsilon}$ under the lognormal hypothesis.

(b) Yaglom's model involves, although only implicitly, a hypothesis of rigorous local conservation of dissipation within eddies. This feature will be is said to characterize his cascade as being "microcanonical." Another part of this paper will view conservation as holding only on the average; and the resulting cascades, called "canonical," will be investigated. When a cascade is canonical, the behavior of $\bar{\varepsilon}(\mathcal{D})$ will be seen to fall under the same three classes as have been defined above under (a), when $\mathcal{D}$ is cubic eddy, except that the parameter α_1 must be replaced by a new parameter $\alpha_3 > \alpha_1$. In the same cascade, averages taken over cylinders and eddies may fall in different classes; for example, a regular $\bar{\varepsilon}(\mathcal{D})$ when $\mathcal{D}$ is an eddy is compatible with an irregular $\bar{\varepsilon}(\mathcal{D})$ when $\mathcal{D}$ is a cylinder; also, an irregular $\bar{\varepsilon}(\mathcal{D})$ when $\mathcal{D}$ is an eddy is compatible with a degenerate $\bar{\varepsilon}(\mathcal{D})$ when $\mathcal{D}$ is a cylinder.

(c) Another aspect of this paper is purely critical. It concerns Kolmogorov's second hypothesis, which asserts that the value of η does not influence $\bar{\varepsilon}(\mathcal{D})$ in the similarity range. This will indeed be confirmed when $\bar{\varepsilon}(\mathcal{D})$ is in the regular class for every domain $\mathcal{D}$, but will be disproved when all $\bar{\varepsilon}(\mathcal{D})$ are in the degenerate class. In all other cases, the hypothesis is doubtful. Thus, the domains of validity of the second and third hypothesis are related.

(d) This paper introduces, in passing, a new concept, which will be developed fully elsewhere. In the regular and irregular classes, the bulk of intermittent dissipation is shown to occur over a very small portion of space, which will be shown to be best characterized by a parameter D called the "intrinsic fractional dimension" of the carrier. The parameter D is preferable to the relative volume, because the volume is very small and too dependent upon η.

(e) Yaglom's theory introduces yet another parameter, which characterizes the spectrum of $\bar{\varepsilon}$ and is related to a correction factor to the exponent $-5/3$ of the classic Kolmogorov power law. This parameter will be denoted by Q. The parameters α_1, D and Q will be shown to be conceptually distinct. Naturally, the introduction of any additional assumption about the cascade introduces a relation among these parameters. For example, one may, under a special assumption, come close to Kolmogorov-Yaglom theory, and find that α_1, D and Q are functions of a single parameter μ. The question of whether or not the actual parameters are distinct suggests much work to the experimentalist.

(f) For the sake of numerical illustration, a variety of one-dimensional canonical cascades was simulated on a digital computer, IBM System 360/Model 91. The results, unfortunately, cannot be described in this paper. Suffice to say that they confirm the theoretical predictions concerning the limiting behavior, but throw doubt upon the rapidity of convergence to the limit.

2. BACKGROUND AND PRINCIPAL RESULTS

2.1. Background: Yaglom's postulate of independence and lognormality

The purpose of this section is to amplify items (a), (b), (c) and (f) of Section 1. To do so, we shall first describe Yaglom's cascade model in narrower and more specific form. (It is hoped that the spirit of Yaglom's approach is thereby left unaltered.)

To begin with, the skeleton of the cascade process is taken to be made of "eddies" that are prescribed form the outset and which are cubes such that each cubic eddy at a given hierarchical level includes C cubic eddies of the immediately lower level. (C is the initial of "cell number.") This expresses the fact that the grid of eddies is self-similar in the range from η to L. Obviously, $C^{1/3}$ must be assumed to be an integer and is denoted by b. The sides (edge lengths) of the largest and the smallest eddies are equal to the external scale L and internal scale η respectively.

The unit of length will be chosen such that η and L are only dimensionless powers of b. The density of turbulent dissipation at the point $\mathbf{x}$ is denoted by $\varepsilon(\mathbf{x}, L, \eta)$ and the density average over the domain $\mathcal{D}$ will be denoted by $\bar{\varepsilon}(\mathcal{D}, L, \eta)$. Units of dissipation will again be such that $\bar{\varepsilon}$ is dimensionless. When $\mathcal{D}$ is a cubic eddy of side r and center $\mathbf{x}$ (with $-\log_b r$ an integer) we write

$$\bar{\varepsilon}(\mathcal{D}, L, \eta) = \bar{\varepsilon}_r(\mathbf{x}, L, \eta).$$

It is further assumed that the distribution of dissipation over its self-similar grid is itself self-similar. This means that, whenever $\eta \ll r < rb \ll L$, the ratio $\bar{\varepsilon}_{r/b}(\mathbf{x}_s, L, \eta)/\bar{\varepsilon}_r(\mathbf{x}, L, \eta)$ is a random variable, to be denoted by Y_s, having a distribution independent of r. Here, $\{\mathbf{x}_s\}$ is a regular grid of centers of subeddies.

Next (an assumption that goes beyond self-similarity), the successive ratios $\bar{\varepsilon}_{L/b}/\bar{\varepsilon}_L$, $\bar{\varepsilon}_{L/b^2}/\bar{\varepsilon}_{L/b}$, etc. down to $\bar{\varepsilon}_r/\bar{\varepsilon}_{rb}$, are assumed independent. This makes $\log\bar{\varepsilon}_r - \log\bar{\varepsilon}_L$ the sum of $\log_b(L/r)$ independent expressions, each of which is of the form $\log Y$. Finally, assume $\langle(\log Y)^2\rangle < \infty$, a condition which implies that $\Pr\{Y=0\}=0$. This means that $\log\bar{\varepsilon}_r - \log\bar{\varepsilon}_L$ is a finite sum from a series that would, if carried out to infinity, satisfy the central limit theorem. One concludes that $\log\bar{\varepsilon}_r$ is approximately Gaussian, meaning that $\bar{\varepsilon}_r$ is approximately lognormal.

At this point, the reader may digress to the appendices A1 and A2, which comment about lognormality.

2.2. Dissipation averaged over thin cylinders

Nevertheless, there are several reasons why, even when all of Yaglom's assumptions are accepted, the argument sketched above does not suffice to justify Kolmogorov's third hypothesis, that $\bar{\varepsilon}$ is lognormal for all $\mathcal{D}$. First (not the basic reason), Yaglom's argument is rigorous only when $\mathcal{D}$ is a cubic eddy. When $\mathcal{D}$ (while three-dimensional) is not an eddy, lognormality is at best approximate. The reason why this argument is not basic is that, for every three-dimensional $\mathcal{D}$ the moments $\langle\bar{\varepsilon}^q\rangle$ are finite for all q. A second argument is more basic and concerns the comparison of theory and experiment. Even though averages taken over three-dimensional domains, $\mathcal{D}$ may be appropriate for a theoretical characterization of turbulence (including the hoped-for linkage between the present phenomenology and actual physics), such averages cannot be measured experimentally. Actual measurements, by necessity, involve averages taken over thin cylinders in time and space. By G.I. Taylor's "frozen turbulence hypothesis," such domains can be replaced by thin cylinders through the spatial flow. When the radius of such a $\mathcal{D}$ is of the same order of magnitude as the inner scale η of the turbulence, $\mathcal{D}$ can be approximated by a one-dimensional straight segment. Thus one must raise the question of whether or not the distribution of $\bar{\varepsilon}(\mathcal{D}, L, \eta)$ remains approximately lognormal when $\mathcal{D}$ is one dimensional. Yaglom does not raise explicitly this question in his works, nor do later writers concerned with the extent to which the observed data fit the lognormal distribution. But all these authors imply that the dimensionality of the averaging domain $\mathcal{D}$ has little effect on the distribution of $\bar{\varepsilon}(\mathcal{D}, L, \eta)$.

This paper will show this implicit belief to be unwarranted. More precisely, whenever $\mathcal{D}$ is not an eddy, the distribution of $\bar{\varepsilon}(\mathcal{D}, L, \eta)$ changes depending which detailed assumptions are made about the cascade process. The assumptions, made in this paper will now be described. We shall evaluate the average of $\overline{\varepsilon}(\mathcal{D}_\eta, L, \eta)$ when $\mathcal{D}_\eta$ is a cylinder of length r and radius η, and we shall find that, except in trivial circumstances, the average is *not* lognormal.

Throughout, the hierarchy of eddies is not viewed as a physical phenomenon, but as a formal device for constructing the fully cascaded state of the medium. Every stage of the division of space will be assumed to preserve the total dissipation. Hence the average over the whole sample of the local averages over eddies. The simplest procedure is to assume nothing else about the corresponding Yaglom ratios Y. The resulting cascade will be called "microcanonical." Consider successive ratios of the form $\overline{\varepsilon}(\mathcal{D}', L, \eta)/\overline{\varepsilon}(\mathcal{D}'', L, \eta)$, where $\mathcal{D}'$ and $\mathcal{D}''$ are cylinders of identical length r but different cross-sections (with $\mathcal{D}'$ embedded in $\mathcal{D}''$). It will be found that these ratios are *not* independent. In order to formalize the limit process of Yaglom, we shall view the internal scale η as a variable tending to zero. In Sections 3 and 4 it will be proved that, except in a trivial case, the distribution of the limit $\overline{\varepsilon}(\mathcal{D}_0, L, 0)$ where $\mathcal{D}_0$ is an infinitely thin cylinder, is never lognormal. In some cases, the difference is small, but in other cases it is great, implying that the influence of the dimension of $\mathcal{D}$ over the distribution of $\overline{\varepsilon}(\mathcal{D}_0, L, 0)$ may be critical. The extent of the divergence of the distribution from the lognormal is expressed to a significant extent by the value of a parameter, denoted as α_1, which is defined as the second zero (the first being $q = 1$) of the expression

$$\overline{\tau}_1(q) = \log_b \left\langle Y^q \right\rangle - (q - 1).$$

The definition of α_1 is motivated in Section 4.3, and illustrated on Figure 1. (The latter uses the notation W instead of Y; the relationship between the two will be explained in Section 3.)

• The first class is called "regular," and includes all Y that are bounded by b. It is characterized by $\alpha_1 = \infty$. The resulting $\overline{\varepsilon}(\mathcal{D}_0, L, 0)$ differ little from lognormality. This factor is random but essentially independent of η, and all its moments are finite.

• The second class is called "degenerate," and corresponds to Y's that are extremely scattered. It is characterized by $\alpha_1 \leq 1$. The resulting $\overline{\varepsilon}(\mathcal{D}_0, L, 0)$ vanishes almost surely. In particular, $\left\langle \overline{\varepsilon}^q(\mathcal{D}_0, L, 0) \right\rangle = 0$ for every q.

Even though "physical intuition" suggests the opposite, the fact that $\overline{\varepsilon}(\mathcal{D}_0, L, 0) \equiv 0$, hence $\left\langle \overline{\varepsilon}(\mathcal{D}_0, L, 0) \right\rangle = 0$, is perfectly compatible with the combination of $\lim_{\eta \to 0} \left\langle \overline{\varepsilon}(\mathcal{D}_\eta, L, \eta) \right\rangle = 1$, $\lim_{\eta \to 0} \left\langle \overline{\varepsilon}^q(\mathcal{D}_\eta, L, \eta) \right\rangle = 0$ for $q < 1$, and

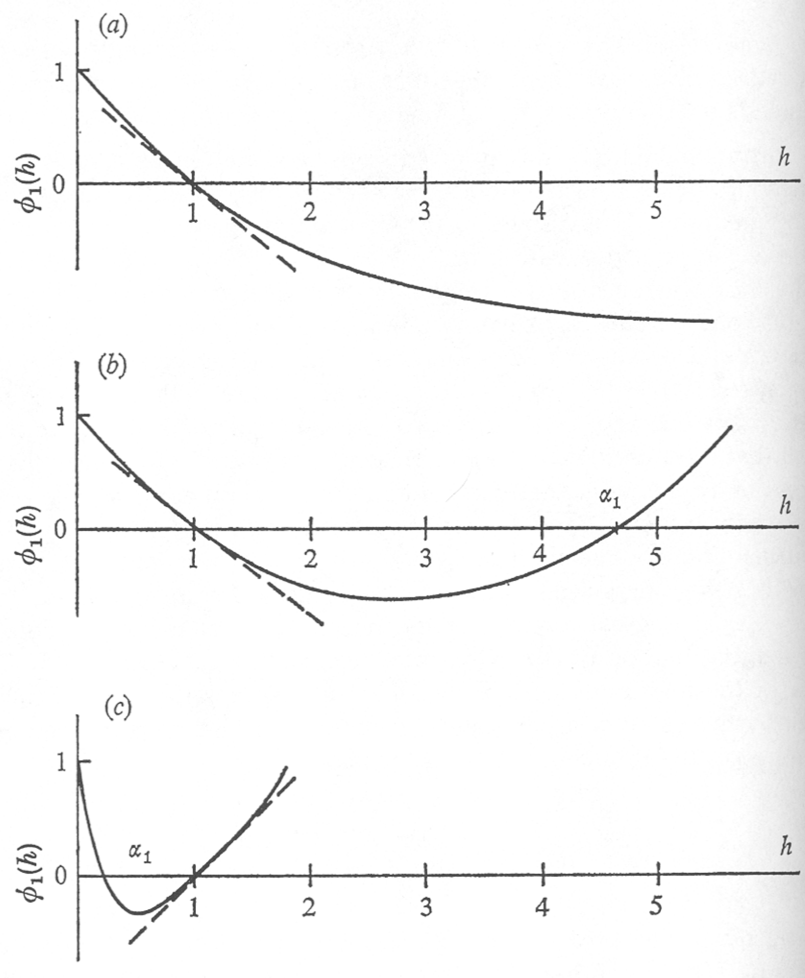

FIGURE C15-1. The distribution of the averaged dissipation $\bar{\varepsilon}$ is determined by that of the random weight W, which is roughly speaking Yaglom's ratio between the average dissipation with a subeddy and an eddy. We plot the function $\overline{\tau}_1(q) = \log_b\langle W^q \rangle - (q-1)$, which is always cup-convex. Through $\overline{\tau}_1(q)$, all presently interesting aspects of intermittency are described as follows. The spectral properties of $\bar{\varepsilon}$ (the only one to have been examined before the present study) are determined by the value of $\overline{\tau}_1(2)$. In addition, the distribution properties (finiteness of moments) depend on the value of α_1, defined as the root (other than $q=1$) of the equation $\overline{\tau}_1(q)=0$. Finally, the fractional dimension of the carrier of turbulence depends on the values of $\overline{\tau}'_1(1)$. Thus, from the viewpoint of properties of $\bar{\varepsilon}$, its distributions falls into the following three classes. (a) Regular class: $\overline{\tau}'_1(1)<0$, $\overline{\tau}_1(2)<0$ and $\alpha_1=\infty$. (b) Irregular class: $\overline{\tau}'_1(1)<0$, $\overline{\tau}'_1(2)<0$ and $1<\alpha_1<\infty$. (c) Degenerate class: $\overline{\tau}'_1(1)>0$, $\overline{\tau}_1(2)>0$ and $0<\alpha_1<1$.

$\lim_{\eta \to 0}\left\langle \overline{\varepsilon}^q(\mathcal{D}_\eta, L, \eta)\right\rangle = \infty$ for $q > 1$. Such a drastic discrepancy between the moments of the limits and the limits of the moments is ordinarily achieved by a deliberate effort to create a mathematical pathology. A classical illustration is the sequence for which $\overline{\varepsilon}(\mathcal{D}, L, \eta)$ equals $1/\eta$ with probability η, and equals zero with probability $1 - \eta$. Here, to the contrary, this discrepancy direct practical consequences. (This is a bit reminiscent of the singularity, familiar in fluid mechanics, encountered when the coefficient of viscosity tends to zero.) Thus the degenerate case suggests that, when η is non-zero but small, dissipation concentrates in a few huge blobs.

• The third class is called "irregular," and includes all Y's that are not scattered beyond reason, but can exceed b. It is characterized by $1 < \alpha_1 < \infty$. In this class, $\left\langle \overline{\varepsilon}(\mathcal{D}, L, \eta)\right\rangle$ remains identically equal to one, while higher moments $\left\langle \overline{\varepsilon}^q \right\rangle$ behave as follows: when $q < \alpha_1$, they remain finite as $\eta \to 0$, but when $q > \alpha_1$, they tend to infinity. This implies that when η is positive but small, their values are extremely large and in practice can be considered infinite.

When a probabilist knows that moments behave as stated above, with the loose additional requirement that the function $\Pr\{\overline{\varepsilon} > x\}$ is "smooth," the simplest distribution he is likely to envisage is the "scaling," defined as follows: $\min \overline{\varepsilon} = x_0 = \alpha_1/(\alpha_1 - 1) > 0$ and $\Pr\{\overline{\varepsilon} > x\} = (x/x_0)^{-\alpha_1}$. The next simplest possibility is $\Pr\{\overline{\varepsilon} > x\} = C(x)x^{-\alpha_1}$, where the prefactor $C(x)$ is a function that varies "smoothly and slowly" as $x \to \infty$. (Examples are functions with a non-trivial limit, and functions that vary like $\log x$ or $1/\log x$). Such random variables $\overline{\varepsilon}$ are called "asymptotically scaling" or "Paretian." To test for their occurrence, it is common practice to plot $\log \Pr\{\overline{\varepsilon} > x\}$ as a function of $\log x$: the tail of the resulting curve should be straight and of slope α_1. However, the more interesting prediction concerns the case when η is very small but positive. In that case, all moments of $\overline{\varepsilon}(\mathcal{D}, L, \eta)$ are very large but finite. If its distribution is again plotted in log-log coordinates, it must end on a tail that plunges down more rapidly than any straight line of finite slope. But, the behavior of the moments of $\overline{\varepsilon}$ as $\eta \to 0$ also yields a definite prediction for small η, namely: *the log-log plot of the distribution is expected to include a long "penultimate" range within which it is straight and of slope* α_1. *This is one of the principal predictions of the present work.*

This contrast between Yaglom's conclusions and mine turns out to be parallel to the contrast between two classic chapters of probability theory. (a) In the theory of sums of many nearly independent random variables, the asymptotic distribution is, under wide conditions, universal: it is Gaussian. (b) In the theory of the number of offspring in a birth-and-death process, the asymptotic distribution depends upon the distribution of a number of offspring per generation: it is not universal. Using statistical mechanics, the thermodynamics properties of matter had been reduced to theory (a) above, which is why they are largely independent of microscopic mechanical detail. What Yaglom claims in effect is that the same is true of turbulent intermittency. To the contrary, Section 4.3 shows that turbulence is closer to theory (b), the resulting absence of universality probably being intrinsic. More precisely, the theory underlying this paper is an aspect of the "theory of birth, death and random walk."

The ease of verifying this prediction increases as the slope α_1 becomes less steep. In an approximation discussed in Section 4.8, the value of α_1 can be inferred from the spectral exponent $Q=\mu$, as being equal to about $2/\mu \sim 4$. This suggests that moments should misbehave for $q \geq 4$. Further discussion of the validity of this prediction must be postponed until more data are available.

2.3. Validity of the microcanonical assumption and reasons for introducing the canonical cascades

The second purpose of this paper is to probe Yaglom's assumption that the ratios of the form $\bar{\varepsilon}(\mathcal{D}_s, L, \eta)/\bar{\varepsilon}(\mathcal{D}, L, \eta)$, relative to a subeddy $\mathcal{D}_s$ and to an eddy $\mathcal{D}$ containing $\mathcal{D}_s$, are independent. We noted that this is satisfied by the microcanonical model, in which the cascade is merely a way of splitting up space. But less formalistic interpretations are conceivable.

For example, one may keep to the approximation that a cascade divides an eddy exactly into subeddies, but combine splitting with some kind of diffusion, in such a way that conservation of dissipation only holds on the average. (Physically, the "dissipation"invoked in this proposal would correspond to the energy transfer between eddy sizes, rather than to the ultimate conversion of eddy kinetic energy into heat.)

The resulting model is to be called "canonical." It is interesting because (a) when $\mathcal{D}$ is a cylinder the results it yields are essentially the same as in the microcanonical model, and (b) Yaglom's ratios turn out to be so strongly interdependent that $\bar{\varepsilon}(\mathcal{D})$ fails to be lognormal even when $\mathcal{D}$ is an eddy. The theory of the canonical $\bar{\varepsilon}(\mathcal{D})$ with an eddy $\mathcal{D}$ follows the same pattern as the theory of the microcanonical $\bar{\varepsilon}(\mathcal{D})$, with $\mathcal{D}$ a cylinder. Thus, it can fall into any of the three classes noted in Section 2.2, with the change that one must replace α_1 by a new parameter α_3. Ordinarily, $\alpha_3 > \alpha_1$.

Since in some cases the predictions of the canonical and the microcanonical models are very different, the degree of validity of Yaglom's model depends on the solidity of the foundations of the microcanonical assumption. It would be nice if either kind of cascade turned out to have a more precise relationship with the physical breakdown of eddies but, so far, no connection has been established. As a matter of fact, the accepted role dissipation plays in the current phenomenological approach to turbulence should perhaps be downgraded, and the canonical model of a cascade be rephrased in terms of energy transfer between different scale sizes. Nevertheless, attempting this task would go beyond the purpose of the present work, and we shall stick to the logical analysis of the cascades. The relative advantages and disadvantages of the two main models are as follows.

Yaglom's argument. In the case of cylinders, this argument requires amplification that may lead to substantially non-lognormal results; in the case of cubes, it is disputable.

The canonical alternative. In the case of cylinders, this alternative appears to be a nearly inevitable approximation, and in the case of cubes, it may well be an improvement.

2.4. Kolmogorov's "second hypothesis of similarity"

The possibility of peculiar behavior of the moments leads us to probe Kolmogorov's second hypothesis, which was stated originally (1941) for homogeneous turbulence and was generalized in Kolmogorov 1962 to intermittent turbulence. Intuitively, if $\mathcal{D}$ is a domain of characteristic scale $\gg \eta$, the second hypothesis states that the distribution of $\bar{\varepsilon}(\mathcal{D}, L, \eta)$ is nearly independent of η. To restate it rigorously, let us make η into a parameter and let it tend to zero. Kolmogorov's second hypothesis may merely state that $\lim_{\eta \to 0} \bar{\varepsilon}(\mathcal{D}, L, \eta)$ exists. If it does, and if (as is usual in mathematics) the concept of a limit is interpreted through "convergence of probability distributions," then for both the canonical and the microcanonical cascades the second hypothesis will indeed be satisfied. But mathematical convergence need not be intuitively satisfactory, and the second hypothesis ought perhaps to be interpreted in stronger terms.

When $\mathcal{D}$ is an eddy of the microcanonical model, and for other D's models leading to the regular class, we have $\bar{\varepsilon}(\mathcal{D}, L, \eta) \to \bar{\varepsilon}(\mathcal{D}, L, 0)$ mathematically and, for all $q > 0$, $\left\langle \bar{\varepsilon}^q(\mathcal{D}, L, \eta) \right\rangle \to \left\langle \bar{\varepsilon}^q(\mathcal{D}, L, 0) \right\rangle$. In this case, as long as η is small, the "error term" $\bar{\varepsilon}(\mathcal{D}, L, \eta) - \bar{\varepsilon}(\mathcal{D}, L, 0)$ should be expected to be small. Kolmogorov's intuitive second hypothesis holds uncontroversially. When convergence is regular for both eddies and cylinders, the Kolmogorov-Yaglom lognormal approximation is (up to a fixed correction factor) workable.

In the degenerate convergence class, on the contrary, $\Pr\{\bar{\varepsilon}(\mathcal{D}, L, 0) = 0\} = 1$. For small but positive η, $\bar{\varepsilon}(\mathcal{D}, L, \eta)$ and $\bar{\varepsilon}(\mathcal{D}, L, 0)$ may be mathematically close, but are intuitively very different. The actual behavior of $\bar{\varepsilon}$ when convergence is degenerate appears to resemble the illustrative example given above. As a result, Kolmogorov's second hypothesis is not really applicable to this class.

When convergence is irregular and η is small, the error term $\bar{\varepsilon}(\mathcal{D}, L, \eta) - \bar{\varepsilon}(\mathcal{D}, L, 0)$ is extremely likely to be small. But in cases when it is not small, it may be very large, and its own moments of high order may be infinite. In this case, the Kolmogorov second hypothesis is controversial, but its degree of validity improves as α_1 increases.

2.5. Relationship between the canonical cascade and the "limiting lognormal" model in M 1972j{N14}

Both the canonical and the microcanonical variants allow the distribution of dissipation between neighboring subeddies to be highly discontinuous. However, M 1972j{N14}, has investigated yet another alternative cascade model, using a "limiting lognormal process." Its principal characteristic is that it generates its own eddies of different shapes, and that the distribution of dissipation within eddies is continuous. This feature will appear especially attractive when the study of the geometry of the carrier of turbulence is pushed beyond the concept of fractional dimension, to include matters of connectedness. The limiting lognormal model can be viewed, though it was developed first, as an improvement upon a canonical cascade with a lognormal weight *W*. Section A4 will describe its main characteristics.

2.6. Study by computer simulation of the rapidity of convergence in the canonical cascade process

As always in the application of probability theory, limit cascades (involving infinitely many stages) are of practical interest primarily because the formulae relative to actual cascades (in which the number of stages is large but finite) are unmanageable. The present paper goes a step further, by including "qualitative" arguments about the nature of error terms for finite cascades. In addition, I have arranged numerous computer simulations. The very tentative conclusions are (a) that many of the involved discrepancies from lognormality should manifest themselves only in a relatively small number of exceptionally large observations, and (b) that they depend greatly upon high-order moments of the Yaglom ratio Y, which express comparatively minute characteristics of the cascade model.

If these inference are confirmed, then lognormality may combine the worst of two worlds: it could prove fairly reasonable qualitatively, while its use for any calculation that involves moments could not be trusted. If so, even Obukhov, who did evaluate moments, would prove less pragmatic than he thought. However, having expressed those fears, I hasten to say that I do not share them, and that I believe the study of intermittency to be very enlightening as to the nature of turbulence.

3. INTRODUCTION TO CANONICAL AND MICROCANONICAL CASCADES

3.1. A detailed cascade model

To be able to make a prediction about $\bar{\varepsilon}(\mathcal{D}, L, \eta)$ when $\mathcal{D}$ is a cylinder, one must make assumptions about the local distribution of ε within eddies. We shall build a model by making η smaller and smaller.

Initially, $\eta = L$ and the original dissipation $\bar{\varepsilon}(\mathbf{x}, L, L)$ is uniformly distributed in space. At the beginning of each successive stage of the cascade, one assumes that dissipation density is uniform within each eddy of side r. This is also the initial distribution one observes if $\eta = r$; it can therefore be denoted by $\varepsilon(\mathbf{x}, L, r)$. At the end of each stage of the cascade, the dissipation density is uniform in each subeddy of side r/b. When the center of an eddy of side r is denoted by $\mathbf{x}$ the centers of the immediately smaller subeddies will be denoted by $\mathbf{x}_s$, with $0 \leq s \leq C-1$; they form a regular lattice. The corresponding densities will be denoted by $\varepsilon(\mathbf{x}_s, L, r/b)$. Next, designate the random variable $\varepsilon(\mathbf{x}_s, L, r/b)/\varepsilon(\mathbf{x}, L, r)$ by W_s. The ratio W and Yaglom's ratio Y differ by the fact that W involves local densities, while Y involves averages, but in the microcanonical model the concepts of W and Y will merge. Homogeneity suggests that, at each cascade stage, the s random variables of the form W_s have the same distribution. Self-similarity and Kolmogorov's second hypothesis suggest in addition that the distribution is the same for all values of s, r, L and η. The final stage ends with eddies of side η, and with density $\bar{\varepsilon}(\mathbf{x}, L, \eta)$.

The low- and high-frequency multiplicative factors of $\bar{\varepsilon}(\Omega, L, \eta)$. The random variable $\varepsilon(\mathbf{x}, L, \eta)$ resulting from the above cascade has a single parameter: L/η. Moreover, since the actions of eddies of sides above and below r are quite separate, $\varepsilon_r(\mathbf{x}, L, \eta)$ can be written as the product of two statistically independent factors, which can be studied separately. These factors are, respectively, $\bar{\varepsilon}_r(\mathbf{x}, L, r)$ and $\bar{\varepsilon}_r(\mathbf{x}, r, \eta)$. The former is independent of η and has r/L as its sole parameter; it is a "low frequency factor." The latter is independent of L and has r/η as the sole parameter; it is a "high frequency factor." More generally, when $\mathcal{D}$ is not an eddy but is included in an eddy of side r,

$$\bar{\varepsilon}(\mathcal{D}, L, \eta) = \bar{\varepsilon}_r(\mathbf{x}, L, r)\bar{\varepsilon}(\mathcal{D}, r, \eta).$$

3.2. The approximate lognormality of the low frequency factor $\bar{\varepsilon}_r(\mathbf{x}, L, r)$ and the question of whether or not W can take the value zero

To study the low frequency factor, it suffices to follow Yaglom, as in Section 2.1. One notes that $\log \bar{\varepsilon}_r(\mathbf{x}, L, r)$ is the sum of $\log_b(L/r)$ random factors of the form W. Assuming $\langle (\log W)^2 \rangle < \infty$, the Gaussian central limit theorem seems to suggest that $\log \bar{\varepsilon}_r \mathbf{x}, L, r$ is approximately normally distributed, it would follow that $\bar{\varepsilon}_r(\mathbf{x}, L, r)$ is approximately lognormal.

A finite $\langle (\log W)^2 \rangle$ implies in particular that $W = 0$ has zero probability. On the other hand, there is a model by Novikov & Stewart 1964 which assumes that $W = 0$ has a positive probability. In that case, $\bar{\varepsilon}_r(\mathbf{x}, L, r)$ is usually a mixture: with some positive probability, it vanishes, and with the remaining probability, it is lognormal. In the present paper, to allow $W = 0$ will not cause any complication, and in fact will allow consideration of useful simple examples.

3.3. The high frequency factor; limit behavior for $\eta \to 0$

This limiting behavior is ruled by the following theorem (stated at the intermediate level of generality at which the proof is simplest, a level more general than is required and less general than is possible).

Theorem. Let the domain $\mathcal{D}$ be simple, meaning that $\mathcal{D}$ is the sum of a finite number of eddies when $E = 3$, and the sum of a finite number of eddy edges when $E = 1$. Consider $\bar{\varepsilon}(\mathcal{D}, L, \eta)$ (for fixed $\mathcal{D}$ and L) as a random function of η. Assume $\langle W \rangle = 1$ and let $\eta \to 0$. Then, with probability equal to 1, $\bar{\varepsilon}$ tends to a finite limit random variable.

Proof. This proof is written as a digression addressed to readers having an elementary knowledge of the theory of "martingales." This theory is the next most obvious mathematical generalization of the theory of products of independent random variables of unit expectation, such as Yaglom ratios. In order to conform to the usual presentation of martingales, let us view the actual value of the inner scale η as the "present value," values $\eta' < \eta$ and $\eta' > \eta$ being viewed respectively as "future" and "past." A martingale is a random function such that the expectation of a "future" value, conditioned by knowing the present value and any number of past values, is equal to the present value. Here "time" is discrete, being equal to $-\log_b \eta$. Assume that $\mathcal{D}$ is an eddy of side r; a similar argument applies to other simple $\mathcal{D}$'s. Denote its subeddies of side r/b by $\mathcal{D}_s$. We know that

$$\bar{\varepsilon}(\mathcal{D}, L, \eta/b) = \frac{1}{C} \sum_{s=0}^{C-1} W_s \bar{\varepsilon}(\mathcal{D}_s, L, \eta).$$

Designate by E_C the conditional expectation, given the present and any one of past values of $\bar{\varepsilon}(\mathcal{D}, L, \eta)$. Since $\langle W \rangle = 1$, we have

$$E_C \bar{\varepsilon}(\mathcal{D}, L, \eta/b) = \frac{1}{C} \sum_{s=0}^{C-1} E_C \bar{\varepsilon}(\mathcal{D}_s, L, \eta) = E_C \bar{\varepsilon}(\mathcal{D}, L, \eta) = \bar{\varepsilon}(\mathcal{D}, L, \eta).$$

This proves that $\bar{\varepsilon}(\mathcal{D}, L, \eta)$ is a martingale. Being non-negative, $\bar{\varepsilon}$ obeys a convergence theorem (Doob 1953, p. 319): as $\eta \to 0$, $\bar{\varepsilon}(\mathcal{D}, L, \eta)$ has a limit random variable to be denoted by $\bar{\varepsilon}(\mathcal{D}, L, 0)$.

Corollary. In the case of cubic eddies, $\bar{\varepsilon}_r(\mathbf{x}, r, \eta)$ converges to a limit $\bar{\varepsilon}_r(\mathbf{x}, r, 0)$. By self-similarity, the limit is independent of r, so it can be denoted by $\bar{\varepsilon}_1(\mathbf{x}, 1, 0)$.

Remark. The above theorem means that, when $r/\eta \gg 1$, one knows $\bar{\varepsilon}(\mathcal{D}, L, \eta)$. "approximately" without knowing the exact value of η. However, any more detailed information about the quality of approximation involves the character of the convergence of $\bar{\varepsilon}(\mathcal{D}, L, \eta)$ to $\bar{\varepsilon}(\mathcal{D}, L, 0)$ (regular, irregular or degenerate), and in turn requires more detailed assumptions about the model (e.g., about the set of random variables W).

3.4. The microcanonical cascade

Definition. *A cascade will be called microcanonical if the sum* $\sum_{s=0}^{C-1} W_s$ *of the weights* W_s *corresponding to all the subeddies of any eddy is precisely equal to* C. {P.S. My more recent papers use the more self-explanatory term, *conservative.*.}

As a corollary, $\langle W \rangle = 1$ and $W < C$.

The microcanonical condition expresses that, at each cascade stage, the total dissipation $r^3\varepsilon(\mathbf{x}, L, r)$ within an original eddy is replaced by an *equal dissipation* distributed among its C subeddies of centers $\mathbf{x}_s$, namely

$$\sum_{s=0}^{C-1} \frac{r^3}{C}\,\varepsilon(\mathbf{x}_s, L, r/b) = \sum_{s=0}^{C-1} \frac{W_s}{C}\,[r^3\varepsilon(\mathbf{x}, L, r)]$$

Hence, as long as $\eta < r$, one has $\bar{\varepsilon}_r(\mathbf{x}, L, \eta) = \varepsilon(\mathbf{x}, L, \eta)$. This result is independent of η, and shows that the high frequency factor $\bar{\varepsilon}_r(\mathbf{x}, L, \eta)$ is identically equal to 1, which makes it independent of η. Consequently, Yaglom's ratio Y_s coincides with W_s and his postulate of independence is satisfied. Thus, *the theory of microcanonical averages taken over three-dimensional eddies is seen to coincide with Yaglom's theory*.

The converse, that Yaglom's theory is identical to the microcanonical theory, is also true, under certain additional constraints, but there is no need to digress for the proof.

Notice that the microcanonical weights W_s are statistically dependent. In particular, if $s \neq t$,

$$\langle W_s W_t \rangle = \langle W_s E_C(W_t \mid \text{ knowing } W_s) \rangle = \left\langle \frac{W_s(C - W_s)}{(C-1)} \right\rangle$$

$$= 1 - (\langle W^2 \rangle - 1)/(C-1) < 1.$$

Therefore, $\langle W_s W_\varepsilon \rangle < \langle W_s \rangle \langle W_t \rangle$. This inequality expresses that any two weights are negatively correlated (see Section A5). There are analogous results for higher cross-moments; for example, $\langle W_s^2 W_t \rangle < 1$.

3.5. The canonical cascade

Definition. *A cascade will be called canonical if the weights W_s are statistically independent and satisfy* $\langle W \rangle = 1$, meaning that the sum of the weights is *equal to C on the average*. In order to obtain features that go beyond the microcanonical case, it is critically important to allow W to exceed the ceiling $W = C$.

The canonical variant as an approximation for cylinder averages in a microcanonical cascade. Consider a cylinder of length r constituted by a string of elementary eddies of side η hugging one edge (to be called the "marked edge") of a cubic eddy of side r. The dissipation in this cylinder can be obtained through a sequence of two different subcascades. The first subcascade, applicable until an eddy of side r has been reached, follows the mechanism described in Section 3.3, typically ending up with a lognormal $\bar{\varepsilon}_r(\mathbf{x}, L, r)$. The second subcascade is ruled by a different mechanism. The first difference is that each stage only picks those subeddies placed along the marked edge, but we know their number is not C but $b = C^{1/3}$. The second difference is that the conditions imposed on the corresponding weights are

$$\text{(a):}\ \ W_s < C,\ \text{(b)}\ \sum_{s=0}^{b-1} W_s \le C\ \ \text{and (c)}\ \ \langle W_s \rangle = 1.$$

By contrast, if the second subcascade had been microcanonical with b subeddies per eddy, the weights would have obeyed the conditions

$$(a'):\ \ W_s < b\ \ \text{and}\ \ (b'):\ \ \sum_{s=0}^{b-1} W_s = b,$$

which are much stronger. As $C \to \infty$ and $b/C \to 0$, conditions (a) and (b) above become increasingly less demanding in comparison with (a′) and (b′).

This observation gives us a choice between two procedures. The line sections can be studied directly and rigorously. But there is a more attractive alternative: *the second subcascade generating a line average can be approximated by a canonical cascade*. In a canonical cascade, the condition $W < C$ may, in a first approximation, be waived. One may even approximate W by a lognormal random variable, despite the fact that the lognormal is unbounded.

Hence, even if the cascade ruling the cubic eddies is microcanonical, the theory of canonical cascades turns out to be a useful approximation. Incidentally, its most striking result, divergence of high moments, is confirmed by direct argument.

The effect of the condition W < C on the difference between the results of the microcanonical and the canonical models. When three-dimensional canonical eddies are regular, the values of $\bar{\varepsilon}$ given by the canonical and microcanonical theories are identical except for a random prefactor. In this case, the necessary and sufficient condition for regularity is $W < C$. Under the more demanding condition $W < b$, one-dimensional averages are also regular but when $b < \max W < C$, three dimensional and one-dimensional averages belong to different classes and may differ significantly.

4. CLASSIFICATION OF CASCADES ACCORDING TO THE BEHAVIOR OF THE MOMENTS OF $\bar{\varepsilon}$

4.1. A basic recurrence relation for $\bar{\varepsilon}(D, L, \eta)$

Let $\mathcal{D}$ be an eddy of dimension $E = 3$. The definition of Section 3.3 yields, irrespective of the rule of dependence between the W's,

$$\bar{\varepsilon}_{br}(\mathbf{x}, L, \eta) = C^{-1} \sum_{s=0}^{C-1} \bar{\varepsilon}_r(\mathbf{x}_s, L, \eta),$$

where $\{\mathbf{x}_s\}$ is a regular grid of centers of subeddies. Factor the ε on both sides into products of low and high frequency components as follows:

$$\bar{\varepsilon}_{br}(\mathbf{x}, L, br) \bullet \bar{\varepsilon}_{br}(\mathbf{x}, br, \eta) = C^{-1} \sum_{s=0}^{C-1} \bar{\varepsilon}_r(\mathbf{x}_s, L, r) \bullet \bar{\varepsilon}_r(\mathbf{x}_s, r, \eta).$$

Next replace $\bar{\varepsilon}_r(\mathbf{x}_s, L, r)$ by $W_s \bar{\varepsilon}_{br}(\mathbf{x}, L, br)$ and divide both sides by $\bar{\varepsilon}_{br}(\mathbf{x}, L, br)$. We obtain

$$\bar{\varepsilon}_{br}(\mathbf{x}, br, \eta) = C^{-1} \sum_{s=0}^{C-1} W_s \bar{\varepsilon}_r(\mathbf{x}_s, r, \eta).$$

Finally, taking account of self-similarity, we obtain the following basic recurrence relation:

$$\bar{\varepsilon}_1(\mathbf{x}, 1, \eta / br) = C^{-1} \sum_{s=0}^{C-1} W_s \bar{\varepsilon}_1(\mathbf{x}_s, 1, \eta / r).$$

When $E = 1$, so that $\mathcal{D}$ and $\mathcal{D}_s$ are straight intervals of length r, one has the very similar relation

$$\bar{\varepsilon}(\mathcal{D}, 1, \eta/br) = b^{-1} \sum_{s=0}^{b-1} W_s \bar{\varepsilon}(\mathcal{D}_s, 1, \eta/r).$$

Derivation of the moments of eddy averages from the basic recurrence relation. For $q = 1$, it suffices to check that the relation $\langle \bar{\varepsilon}_1(\mathbf{x}, 1, \eta) \rangle = 1$ and the above recurrence relation are compatible. For $q > 1$, the recurrence relation for $\bar{\varepsilon}_1$ can be used to deduce a recurrence relation for the sequence of the moments $\langle \bar{\varepsilon}_1^q(\mathbf{x}, 1, b^{-k}) \rangle$. The form of the latter depends on the rule of dependence between the W's. Throughout, we shall set $r = 1$, which will simplify the notation.

The microcanonical case. We know that $\bar{\varepsilon}_1(\mathbf{x}, 1, \eta) = 1$, but we want to verify that $\langle \bar{\varepsilon}_1^q(\mathbf{x}, 1, \eta) \rangle = 1$. Indeed, for $q = 2$, we have

$$\langle \bar{\varepsilon}_1^2(\mathbf{x}, 1, \eta/b) \rangle = C \left\langle \left(\frac{W}{C} \right)^2 \right\rangle \langle \bar{\varepsilon}_1^2(\mathbf{x}, 1, \eta) \rangle + C(C-1) \left\langle \frac{W_s}{C} \frac{W_t}{C} \right\rangle [\langle \bar{\varepsilon}_1(\mathbf{x}, 1, \eta) \rangle]^2$$

$$= \frac{\langle W^2 \rangle}{C} \langle \bar{\varepsilon}_1^2(\mathbf{x}, 1, \eta) \rangle + \frac{C-1}{C} \left[1 - \frac{\langle W^2 \rangle - 1}{C-1} \right].$$

Starting from $\langle \bar{\varepsilon}_1^2(\mathbf{x}, 1, 1) \rangle = 1$, we obtain

$$\langle \bar{\varepsilon}_1^2(\mathbf{x}, 1, 1/b) \rangle = \langle W^2 \rangle / C + 1 - C^{-1} - \langle W^2 \rangle / C + C^{-1} = 1.$$

The recurrence relation reduces to the identity $1 = 1$, as it should. The recurrence relations for $q > 2$ also reduce to identities.

The canonical case. Now, the recurrence relation for the moments takes the form

$$\langle \bar{\varepsilon}_1^2(\mathbf{x}, 1, \eta/b) \rangle = C\langle (W/C)^2 \rangle \langle \bar{\varepsilon}_1^2(\mathbf{x}, 1, \eta) \rangle + 2[(1/2)C(C-1)][\langle (W/C) \rangle \langle \varepsilon_1 \rangle]^2$$

$$= (\langle W^2 \rangle /C)\langle \bar{\varepsilon}_1^2(\mathbf{x}, 1, \eta) \rangle + (C-1)/C.$$

This is no longer an identity, but rather it establishes that the necessary and sufficient condition for $\lim_{\eta \to 0} \langle \bar{\varepsilon}_1^2(\mathbf{x}, 1, \eta) \rangle < \infty$ is $\langle W^2 \rangle /C < 1$. Similarly, we have the following important property

$$\lim_{\eta \to 0} \langle \bar{\varepsilon}_1^q(\mathbf{x}, 1, \eta) \rangle < \infty \text{ if and only if } \frac{\langle W^q \rangle}{C^{q-1}} < 1.$$

Conclusion. For eddy averages, the asymptotic behavior of the moments depends on the nature of the cascade.

A necessary and sufficient condition. In order for the inequality $\langle W^q \rangle /C^{q-1} < 1$ to hold for all q, it is necessary and sufficient that $W < C$.

Proof of necessity. The inequality $\langle W^q \rangle /C^{q-1} < 1$, i.e., $\langle (W/C)^q \rangle < 1/C$, implies that

$$\max(W/C) = \lim_{q \to \infty} [\langle (W/C)^q \rangle]^{1/q} < \lim_{q \to \infty} C^{-1/q} = 1.$$

Proof of sufficiency. Knowing that $\langle W \rangle = 1$ and $W < C$, $\langle (W/C)^q \rangle$ is maximized by setting $\Pr\{W = C\} = 1/C$ and $\Pr\{W = 0\} = 1 - VC$. In this extreme case, $\langle (W/C)^q \rangle = 1/C$, so in all other cases $\langle W^q \rangle /C^{q-1} < 1$.

Derivation of the moments of line averages from the basic recurrence: the microcanonical case. The recurrence relation for moments is now replaced by

$$\left\langle \bar{\varepsilon}^2(\mathcal{D}, 1, \eta/b)\right\rangle = b\left\langle (W/C')^2\right\rangle\left\langle \bar{\varepsilon}^2(\mathcal{D}_s, 1, \eta)\right\rangle + (b-1)b^{-1}\left\langle W_s W_t\right\rangle$$

$$= \frac{\left\langle W^2\right\rangle}{b}\left\langle \bar{\varepsilon}^2(\mathcal{D}_s, 1, \eta)\right\rangle + \frac{b-1}{b}\left[1 - \frac{\left\langle W^2\right\rangle - 1}{(C-1)}\right].$$

This is no longer an identity: the necessary and sufficient condition for

$$\lim_{\eta\to 0}\left\langle \bar{\varepsilon}^2(\mathcal{D}, 1, \eta)\right\rangle < \infty$$

has become $\left\langle W^2\right\rangle/b < 1$. Similarly, we have the following important property:

$$\lim_{\eta\to 0}\left\langle \bar{\varepsilon}_q^2(\mathcal{D}, 1, \eta)\right\rangle < \infty \text{ if and only if } \frac{\left\langle W^q\right\rangle}{b^{q-1}} < 1.$$

The canonical case. The recurrence relation is unchanged when the dimension changes from $E=3$ to $E=1$, except for the replacement of C by b. Therefore, we fall back on the condition $\left\langle W^q\right\rangle/b^{q-1} < 1$ of the preceding paragraph.

Conclusion. For line averages, the finiteness of the limiting moments is *not* dependent on the nature of the cascade. On the other hand, the value of the limiting moment, when finite, is smaller when the cascade is microcanonical; for example, for $q=2$, it is smaller by the factor $1-(\left\langle W^2\right\rangle - 1)/(C-1)$.

4.2. The determining functions Ψ(*q*)

In order to apply the above results to classify cascades, and in order to carry the theory further, we form the expression

$$\Psi(q) = \log_C\left\langle W^q\right\rangle,$$

which will be called the "determining function." More specifically, when $\mathcal{D}$ is E -dimensional, we shall need the quantities

$$\overline{\tau}_E(q) = (3/E)\Psi(q) - (q-1).$$

To define various parameters of dissipation, different features of these functions must be examined. First of all, $\langle W^q \rangle/(C^{E/3})^{q-1} < 1$ is synonymous with $\overline{\tau}_E(q) < 0$, and so the values of the zeros of $\overline{\tau}_E(q)$ are of interest.

For all q, a general theorem of probability theory shows that $\Psi(q)$ is a convex function of q (see Feller 1971, p. 155), and so are all the functions $\overline{\tau}_E$. Hence, $\Psi(1) = \overline{\tau}_E(1) = 0$, and $\overline{\tau}_E(q)$ has, at most, one root other than 1. This root it will be designated by α_E. The conditions $\overline{\tau}_1(q) < 0$ and $\overline{\tau}_2(q) < 0$ are both at least as demanding as $\overline{\tau}_3(q) < 0$, so when $\alpha_1 > 1$, the α_E satisfy $\alpha_1 \leq \alpha_2 \leq \alpha_3$.

A further investigation of the $\overline{\tau}_E$ involves their slopes for $q = 1$, more specifically the expressions

$$D_E = -E\overline{\tau}'_E(1) = -E\left\langle W \log_{C^{E/3}}(W/C)^{E/3}\right\rangle$$

$$= -3\left\langle W \log_C W\right\rangle + E.$$

Writing $D_3 = D$, we have $D_2 = D - 1$ and $D_1 = D - 2$. The value of D_E will be useful, because an E -dimensional average in a canonical cascade is degenerate when $D_E < 0$ and non-degenerate when $D_E > 0$. In particular, when $W < C^{E/3}$, $D_E > 0$. (The transition case $D_1 = 0$ deserves the attention of the mathematicians, but is too complicated to be tackled in this paper.) More precisely, in the degenerate case $D_1 < 0$, one has $\alpha_E < 1$, and the value of α_E plays no special role. But in the non-degenerate case $D_E > 0$, α_E satisfies $\alpha_E > 1$, and its value serves to determine whether the cascade is regular ($\alpha_E = \infty$) or irregular ($\alpha_E < \infty$).

We know that for all $q > 1$, $\alpha_E = \infty$ is equivalent to $\langle W^q \rangle/(C^{E/3})^{q-1} < 1$. It follows that the necessary and sufficient condition for $\alpha_E = \infty$ is $W < C^{E/3}$, an inequality already featured in Section 2.2.

When $D_E > 0$, the quantity D_E plays an independent role as the intrinsic dimension of the support of $\overline{\varepsilon}$ within an E -dimensional $\mathcal{D}$ (see Section 4.8).

Since $W \log W$ is concave, $\langle W \log W \rangle > \langle W \rangle \log \langle W \rangle = 0$; therefore, the intrinsic dimension D_E never exceeds the embedding dimension E.

Finally, the "second-order" dependence properties of the dissipation, namely its correlation and its spectrum, depend on the value of $\Psi(2)$. Yaglom showed the correlation between the averages taken over small domains $\mathcal{D}$, separated by the distances d, to be proportional to d^{-Q}, with

$$Q = 3 \log_C \langle W^2 \rangle = 3\Psi(2) = 3[\overline{\tau}_3(2) + 1].$$

To obtain a lower bound on Q, note that from $\langle W \rangle = 1$ it follows that $\langle W^2 \rangle > 1$ and hence that $Q > 0$. More precisely,

$$Q > 3(1 + \overline{\tau}'_3(1)) > 3 - D.$$

As for the upper bounds, under the constraints $\langle W \rangle = 1$ and $W < C$, we know that the maximum of $\langle W^2 \rangle$ occurs when $\Pr\{W = C\} = 1/C$ and $\Pr\{W = 0\} = 1 - 1/C$, in which case $\langle W^2 \rangle = C$ and so $Q = 3$. More generally, $W < C^{E/3}$ implies $Q < E$. When $W/C^{E/3} > 1$ on the contrary, it may happen that $Q > E$.

Studies involving correlations of higher order q depend similarly on values of f up to the argument q. Since we shall stop at the second order, our classification of canonical cascades will depend solely on the values of Q, D and α_E. These parameters are conceptually distinct, and their numerical values are only related by the conditions of compatibility $Q > 3 - D$ and $(\alpha_E - 1)D_E > 0$. The question of whether or not their actual values are related should be investigated experimentally.

The relationship between W, $\Psi(q)$ and the various other parameters deserves additional mathematical investigations. A knowledge of C and of the distribution of W determines $\Psi(q)$ for all q, and thus determines all the parameters. On the other hand, a knowledge of C and of the values of $\Psi(q)$ for integer values of q an integer need not determine W uniquely. A sufficient condition is that the moments satisfy the Carleman criterion (see Section A2). This technicality is important because this criterion fails in the case of a lognormal W.

4.3. Examples of determining functions

Rectilinear determining functions. Ψ and $\overline{\tau}_E$ are linear functions of q if W is binomial, i.e., $\Pr\{W = 1/p\} = p$ and $\Pr\{W = 0\} = 1 - 1/p$. If so,

$$\left\langle W^q \right\rangle = pp^{-q} = p^{1-q}, \text{ thus } \overline{\tau}_3(q) = (1-q)\log_C(pC),$$

which is a degenerate form of convex function.

Digression, This example reduces to the classical theory of birth-and-death processes (see Harris 1963). After K "generations," each elementary subeddy either is empty or includes a non-random mass of turbulence equal to p^{-K}. Discarding this last factor, the mass in an eddy and the number of its non-empty elementary subeddies are equal. Their probability distribution is readily determined: between each generation and the next, non-empty elementary eddy can be interpreted as acquiring random "offspring" made of M lower order elementary eddies, with M following a binomial distribution of expectation $C^{E/3}p$. When $M = 0$, the eddy "dies out." When $M > 1$, new eddies are born. Classical results on birth-and-death processes show that the number of offspring after the kth generation is ruled by the following rule. When $p \leq 1/C^{E/3}$, so that $D_E \leq 0$, it is almost certain that the offspring will eventually die out. When $p > 1/C^{E/3}$, so that $D_E > 0$, one forms the ratio of the number of offspring to its expected value $(C^{E/3}p)^K = b^{D_E K}$. One finds that this ratio tends asymptotically towards a non-degenerate limiting random variable that has finite moments of every positive order.

Asymptotically rectilinear determining functions. Now let us suppose only that W is bounded. Designate its greatest attainable value by $\max W$. This means that $\Pr\{W > \max W\} = 0$, but $\Pr\{W > \max W - \theta\} > 0$ for all $\theta > 0$. (A more correct mathematical idiom for $\max W$ is "almost sure supremum.") It follows that $\lim_{q \to \infty}(\log_C \left\langle W^q \right\rangle / q) = \log_C \max W$, which implies that $\Psi(q)$ has an asymptotic direction of finite slope $\log_C \max W$. Conversely, in order for this asymptotic slope to be finite, it is necessary that $W < \max W < \infty$. Also, $\overline{\tau}_E(q)$ has an asymptotic direction of slope $-1 + \log \max W / \log C^{E/3}$. When $\max W < C^{E/3}$, this slope is negative and $\overline{\tau}_E(q) = 0$ has no root other than 1; in other words, $\alpha_E = \infty$. When $\max W > C^{E/3}$, and particularly when $\max W = \infty$, one has $\alpha_E < \infty$. This confirms our assertion that, in general (except for some inequalities), the values of D, Q and the α_E are independent. See the caption of Figure 2.

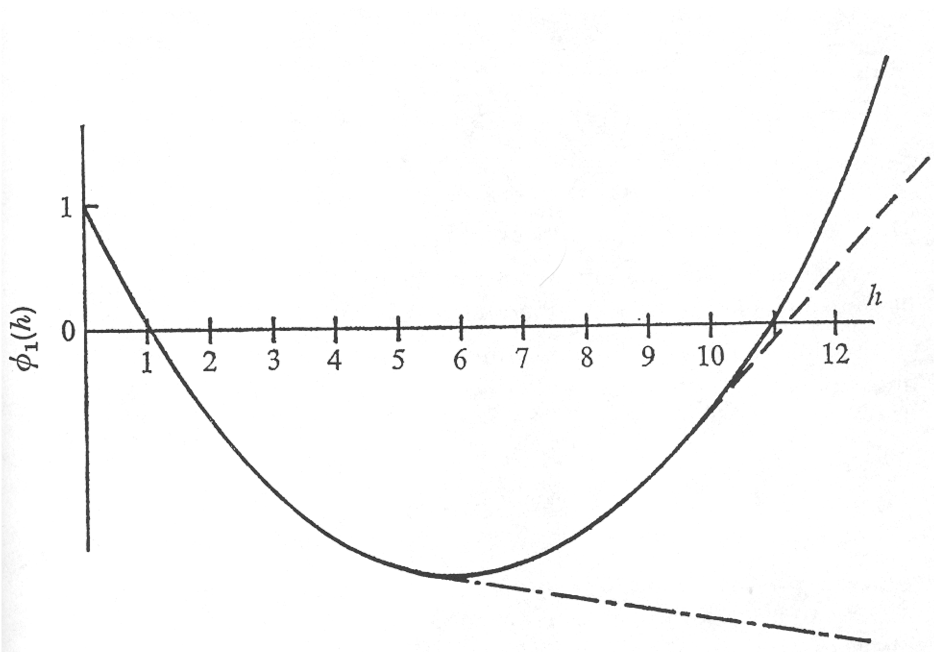

FIGURE C15-2. Characterization of the distribution of $\bar{\varepsilon}(\mathcal{D})$ when (a) $\log W$ is normal; $\overline{\tau}_1(q)$ is then a parabola and $\overline{\tau}_1(q)=0$ has two finite roots (solid line); (b) $\log W$ is a sum of sufficiently many terms to be a good approximation to the normal distribution; $\overline{\tau}_1(q)$ is then nearly parabolic for $q<\alpha_1$ (dashed line); (c) $\log W$ is the sum of comparatively few terms; even when the quality of approximation to the normal distribution is good by other standards, it may be poor from the viewpoint of $\bar{e}$; in the zone of interest, $\overline{\tau}_1(q)$ is far from parabolic and $\overline{\tau}_1(q)=0$ may have a single finite root, i.e. $\alpha_1=\infty$ (dash-dot line). Thus the degree of sensitivity of various properties of $\bar{\varepsilon}$ are very different. On the one hand, the moment properties of the distribution of $\bar{\varepsilon}$ depend greatly upon fine details, namely the tail of the distribution of W: a lognormal W never fails in the regular class, but a "nearly log normal" W may do so. On the other hand, the value of $\overline{\tau}_1(2)$, hence of the spectral properties of $\bar{\varepsilon}$, and even more the value of $\overline{\tau}'_1(1)$, hence of the fractional dimension, will be essentially the same for the three cases as drawn.

Parabolic determining functions. Suppose that $\log W$ is Gaussian of mean and variance $\langle \log W \rangle$ and $\sigma^2 \log W = \langle (\log W)^2 \rangle - \langle \log W \rangle^2$. Then W is lognormal, and

$$\langle W^q \rangle = \exp(q \langle \log W \rangle + q^2 2^{-1} \sigma^2 \log W).$$

That is,

$$\Psi(q) = \log_C \langle W^q \rangle = (q \langle \log W \rangle + q^2 2^{-1} \sigma^2 \log W) \log_C e.$$

This function $\Psi(q)$, and the functions $\overline{\tau}_E(q)$ are represented by parabola.

Digression. The lognormal distribution not being fully determined by its moments (see Section 2), other weights W may lead to the same $\Psi(q)$).

To insure that $\langle W \rangle = 1$, we must have $-\langle \log W \rangle = (1/2)\sigma^2 \log W$, a quantity to be denoted (in order to fit Kolmogorov's notation) by $(\mu/6) \log C$. It follows that

$$\Psi(q) = \frac{1}{6}(q-1)q\mu.$$

Hence, $D_E = -3\Psi'(1) + E = E - \mu/2$, $\alpha_E = 2E/\mu$ and $Q = \mu$. Here, the values of D, Q and the α's are all functions of μ, and are strongly interdependent; this is an exceptional circumstance (see Section 4.9).

The determining function when $\log W$ *is a sum of many uniform random variables* (Figure 2). Suppose that $\log W$ is bounded, for example a sum of many uniformly distributed random variables. Such a sum is near-Gaussian according to the customary definition of nearness. When q is small, the resulting $\overline{\tau}_E(q)$ will nearly coincide with the parabola in the preceding paragraph, but asymptotically it will be a straight line. The graph of $\overline{\tau}_E(q)$ will be a portion of parabola continued by a straight tangent. If we add many uniform components, this tangent will have positive slope, therefore $\log W$ and its Gaussian approximation will yield about the same value for α. If, on the other hand, we add few uniform random variables, the asymptotic tangent will have negative slope and the lognormal approximation will be entirely worthless.

As an illustration, let us describe one of our early computer simulations of a cascade. We thought that W was ostensibly lognormal and we expected $\alpha_E < \infty$. But the actual results were completely at variance with the expectations. They remained mysterious until it was recalled that to generate the Gaussian random variable, our computer added 12 random variables uniformly distributed on $[-0.5, +0.5]$. When the same program was run again, this time adding 48 and then 192 uniform random variables, the results changed to full conformity with the expectations. (See also Sections 4.3 and 4.9.)

4.4. Regular classes

A classification of cascades can be based either on a single value of E, or on two or three values, typically $E = 1$ and $E = 3$.

***The regular class for fixed** E*. Here $\bar{\varepsilon}(\mathcal{D}_E, L, 0)$ is, by definition, a nondegenerate random variable with all moments finite. In a canonical cascade, the necessary and sufficient condition for $\lim_{\eta \to 0} \langle \bar{\varepsilon}_1^q(\mathbf{x}, L, \eta) \rangle < \infty$ can be written as $\bar{\tau}(q) < 0$ for all $q > 1$, in $\alpha_E = \infty$, or, as $W < C^{E/3}$. As a corollary, $Q < 3$.

A formal argument would consist in replacing the limits of the moments with the moments of the limit, and would suggest that $\langle \bar{\varepsilon}_1^q(\mathbf{x}, L, 0) \rangle < \infty$ holds if and only if $\bar{\tau}_E(k) < 0$ for all $q > 1$. This formal argument is justified by a theorem in Doob 1953, p.319. Ordinarily, physicists do not feel that such justifications deserve attention but Section 4.5 will show that in this context they must be taken seriously.

A microcanonical cascade is always regular from the viewpoint of eddy averages. In addition, the condition $W < C^{E/3}$ is necessary and sufficient to insure that W be admissible as weight in a microcanonical cascade with the same E. Consider, then, both the microcanonical and the canonical cascades corresponding to a weight W of the regular class. The only effect of changing the definition of the cascade is to change the high frequency term of $\bar{\varepsilon}_r(\mathbf{x}, L, 0)$ from 1 to some random variable having finite moments of all orders. In other words, the only difference between the full canonical random variable $\bar{\varepsilon}_r$ and its lognormal low frequency term lies in a numerical factor whose values are about the same when $\eta = 0$ and when η is small but non-zero. Such a factor is comparatively innocuous.

As a specific example, if W is binomial with

$$\Pr\{W = 1/p\} = p \text{ and } \Pr\{W = 0\} = 1 - p,$$

$\bar{\tau}_E(q) < 0$ if and only if $pC^{E/3} > 1$, i.e., $p > 1/C^{E/3}$.

The uniformly regular class defined by $\alpha_1 = \infty$***, i.e.,*** $D_1 > 0$**.** In this case, it is true for all values of E that E's, and $\bar{\varepsilon}(\mathcal{D}, L, 0)$ is a non-degenerate random variable with finite moments. In this class, $\varepsilon(\mathcal{D}, L, \eta)$ may be said to be "approximately lognormal." Yaglom has implicitly assumed that this situation prevails in practice. This may, but need not, be so. Only experiments can tell.

4.5. Degenerate classes

The degenerate class for fixed E. This class is defined by $\Pr\{\bar{\varepsilon}(\mathcal{D}), L, 0) = 0\} = 1$. A sufficient condition is $D < 0$, from which it follows that $\alpha \leq 1$.

The proof (see Section A3) consists of showing that the number of elementary subeddies of side r contributing to the bulk of $\bar{\varepsilon}_r(\mathbf{x}, L, 0)$ is roughly equal to $(L/r)^{D_E}$. From $D_E < 0$ it follows that, as $\eta \rightarrow 0$, this number tends to zero, and so does $\bar{\varepsilon}(\mathcal{D}, L, \eta)$. Two examples come to mind.

First example. $\Pr\{W = 1/p\} = p$ and $\Pr\{W = 0\} = 1 - p$ with $p < 1/C$.

Second example. Since a lognormal distribution is unbounded, *a lognormal cascade is never regular*. Because $D_E = E - \mu/2$, the cascade is degenerate when $D_E < 0$, i.e., when $\mu > 2E$. In particular, $\bar{\varepsilon}_r(\mathbf{x}, L, \eta)$ is lognormal only when $r = \eta$.

The uniformly degenerate class. When $\bar{\varepsilon}(\mathcal{D}_E, L, 0)$ is degenerate for $E = 3$, i.e., when $D < 0$, $\bar{\varepsilon}$ is also degenerate for $E = 2$ and $E = 1$.

4.6. Irregular classes

The irregular class for fixed E. This class is defined by $\Pr\{\bar{\varepsilon}(\mathcal{D}_E, L, 0) = 0\} > 0$, with $\left\langle \bar{\varepsilon}^q(\mathcal{D}_E, L, 0) \right\rangle < \infty$ for small enough $q > 1$, but $\left\langle \bar{\varepsilon}^q(\mathcal{D}_E, L, 0) \right\rangle = \infty$ for large finite q. The class is characterized by $1 < \alpha_E < \infty$, and the cut-off between finite and infinite moments is $q = \alpha_E$ (see M 1974c{N16} and Kahane 1973).

The uniformly irregular class. When $\bar{\varepsilon}(\mathcal{D}_E, L, 0)$ is non-degenerate for $E = 3$ and is irregular for $E = 1$, then it is irregular for all E.

4.7. Mixed classes

Since $\alpha_1 < \alpha_2 < \alpha_3$, it is possible that a cascade ($E = 3$) and its cross-sections ($E = 1$ and $E = 2$) belong to different classes. Neglecting the behavior for $E = 2$, three possibilities are open. We shall give one example of each.

An example from the mixed regular-degenerate class. $\Pr\{W = 1/p\} = p$ and $\Pr\{W = 0\} = 1 - p$ with $C^{-1} < p < C^{-1/3}$. Here $\overline{\tau}_3(q) < 0$ but $\overline{\tau}_1(q) > 0$ for all $q > 1$. That is, the cascade is regular for $E = 3$ but degenerate for $E = 1$.

Comments. I doubt that this mixed class is ever encountered in practice, because it implies that the spatial distribution of dissipation is extraordinarily sparse, sparser than anything I would consider as likely.

An example from the mixed regular-irregular class: lognormal W. Take $C = 27$, so that $C^{1/3} = b = 3$, with the random variable W satisfying $\Pr\{W = 3.7\} = 0.1$ and $\Pr\{W = 0.7\} = 0.9$. Since $W < C$, the resulting three-dimensional cascade is regular, and it may correspond to a canonical approximation of a microcanonical cascade. On the other hand, it is not true that $W < b$, although it is true that $\overline{\tau}'_1(1) < 0$. As a result, a one-dimensional cascade corresponding to this W is irregular.

Comments. I consider this last situation to be a very strong possibility. If and when it occurs in practice, the distribution of one-dimensional averages is not at all lognormal. *One task for the experimental study of turbulence should be to check whether or not such a mixture ever occurs. It may be that different circumstances yield either this mixture or the uniformly regular class; if so, those circumstances should be classified according to the class to which they lead.*

An example from the mixed irregular-degenerate class: lognormal W. If $2 < \mu < 6$, the full three-dimensional pattern is irregular, while one-dimensional cross-sections are degenerate.

4.8. Digression: D_E as a fractional intrinsic dimension

Select two arbitrary small thresholds ω and γ. When $D_E > 0$, it can be shown (see Section A3) that the eddies of side r can be divided into two groups. The eddies of the first group contain a proportion greater than $1 - p$ of the whole dissipation. However, their number lies between $(L/r)^{D_E - \gamma}$ and $(L/r)^{D_E + \gamma}$, which makes the group relatively very small. As a result, almost all eddies belong to the second group. But, the total dissipation they contain is at most equal to ω, which makes it negligible.

It is convenient to call D_E an intrinsic dimension; alternatively (because it need not be an integer) it can be called a fractional dimension.

The notion that a geometric figure can have a fractional dimension was conceived in 1919 by a pure mathematician, Felix Hausdorff. This concept is closely related to the Cantor set, and both have the reputation of lacking any conceivable application, and of "turning off" every natural scientist. I believe that this reputation is no longer deserved. I hope to show (elsewhere) that fractional dimension is in fact something very concrete and that different aspects of it are useful measurable physical characteristics. Examples are the degree of the wiggliness of coastlines, described in M 1967s, the degree of clustering of galaxies, and the intensity of the intermittency of turbulence.

In these applications, it is best to use a semi-formal variant of dimension called the "similarity dimension," which is of more limited validity than Hausdorff's concept, but incomparably simpler. It is rooted in some elementary features of the usual concept of dimension, as applied to segments of a straight line, to rectangles and to parallelepipeds. A line has dimension $D=1$, and for every positive integer C, the segment where $0 \leq x < X$ exactly subdivides into C non-overlapping segments of the form $(n-1)X/C \leq x < nX/C$, where $1 \leq n \leq C$. To obtain one of these parts from the whole one performs a similarity of ratio $\rho(C) = C^{-1}$. In the same way, a plane has dimension $D=2$, and for every integer $\sqrt{C}$, the rectangle where $0 \leq x < X$ and $0 \leq y < Y$ subdivides exactly into C non-overlapping rectangles of the form

$$\frac{(k-1)X}{\sqrt{C}} \leq x < \frac{kX}{\sqrt{C}} \quad \text{and} \quad \frac{(h-1)Y}{\sqrt{C}} \leq y < h\frac{Y}{\sqrt{C}},$$

where $1 \leq k \leq C$ and $1 \leq h \leq C$. To obtain one of these parts from the whole, one performs a similarity of ratio $\rho(C) = 1/C^{1/2} = 1/b$. More generally, a D-dimensional rectangular parallelepiped can, for every integer $C^{1/D} = 1/b$, be decomposed into C parallelepipeds. To obtain one of these parts from the whole, one performs a similarity of ratio $\rho(C) = 1/C^{1/D}$.

For each of the above figures, the dimension D satisfies the relation

$$D = \frac{-\log C}{\log \rho(C)}.$$

This suggests that the concept of dimension can be generalized to the set on which the bulk of intermittent turbulence is concentrated. Here $1/\rho(C) = L/r$, and for every $\gamma > 0$,

$$(L/r)^{D_E - \gamma} < C < (L/r)^{D_E + \gamma}, \text{ i.e., } D_E - \gamma < \log C / \log \rho(C) < D_E + \gamma.$$

In other words the dimension of that set is D_E. The intuitive notion that turbulence concentrates on an extremely sparse set is expressed numerically by the inequality $D_E < E$. By choosing W appropriately, the dimension D_E can take any value between 0 and E. Note also that, when $D < 3$, $D_3/3 = D/3$ is greater than $D_2/2 = (D-1)/2$, which in turn is greater than $D_1 = (D-2)$. The inequality $D < 3$ expresses that a figure does not fill the space dimensionally, and the inequality $D/3 > (D-2)$ expresses that the intersections of such a figure by straight lines are dimensionally even less filling.

I have great faith in the practical usefulness of fractional dimension and hope it will be explored further. In particular, it opens up the issue of the degree of connectedness of the volume where the dissipation concentrates. However, neither the microcanonical nor the canonical models appear to provide a satisfactory framework, because both allow the dissipation to be divided very discontinuously. Therefore connectedness should be studied in some other context, say, that of the limiting lognormal model.

4.9. Further comments on the lognormal approximation to W, and on parabolic approximations to $\Psi(q)$

Suppose that W is non-lognormal and bounded, and satisfies $\sigma^2 \log W < \infty$, and let $\overline{W}$ be its lognormal approximation and let the corresponding determining functions be $\Psi(q)$ and $\overline{\Psi}(q)$, with the obvious definitions for $\overline{\tau}_E(q)$. We have already noted that for $q \to \infty$, $\Psi(q)$ has a finite asymptotic slope $\log_C \max W$, while $\overline{\Psi}(q)$ is parabolic. Therefore, their asymptotic behaviors differ qualitatively. On the other hand, the behavior of $\Psi(q)$ and the $\overline{\tau}_E(q)$ for small q depends only on $\langle \log W \rangle$ and $\sigma^2 \log W$, therefore it remains unchanged when $\log W$ is replaced by its normal approximation. Hence the following consequences.

Since the moment of order $q = 2/3$ is likely to be covered by this approximation, the conclusions Kolmogorov and Yaglom had obtained by applying the "two thirds-law," may well be essentially unchanged.

For $q = 1$, $\overline{\Psi}(1)$ need not equal 1. Also, $\overline{\tau}'_E(1)$ need not equal $\overline{\tau}_E'(1)$. In extreme instances (see the end of Section 4.2) they may have different signs. It may happen that the "real" $\overline{\tau}'_E(1)$ is negative, meaning that the cascade is non-degenerate, while $\overline{\tau}_E'(1)$ is positive, suggesting that the cascade is degenerate. The approximations of the values of Q and of the α_E may be even poorer.

A different lognormal approximation to W, to be called $\overline{W}$, is achieved by approximating $\Psi(q)$ by a parabola $\overline{\Psi}(q)$ satisfying $\overline{\Psi}(0) = \overline{\Psi}(1) = 0$ and having the correct slope $\overline{\Psi}'(1)$. The mean and variance of $\log W$ are determined by the properties of $\overline{\Psi}(q)$ near $q=0$. From the practical viewpoint, all that is of direct interest is the portion of $\overline{\tau}_E(q)$ that lies between $q=1$ and $q=\alpha_E$; it follows that, when $\Psi(q)$ is smooth and the original α_E is small, the error introduced by the lognormal approximation $\overline{W}$ may well be acceptable. Whenever such is the case, the various properties of $\overline{\varepsilon}(\mathcal{D}, L, 0)$ linked to D_E, Q and α_E turn out to be related, after all. Given the inaccuracy inherent in experimental work, this implies that it may be necessary to return to the situation that used to prevail when the single characteristic parameter Q was believed sufficient. To the contrary, when the original α_E is large, and especially when it is infinite, the error in using $\widetilde{W}$ is very large; the process of approximation changes the class to which such a cascade belongs.

Digression. (This is another occurrence of a phenomenon also encountered in Section A.1: the moments of exp V are very sensitive to apparently slight deviation of V from normality.)

APPENDIX

A1. Approximate versus strict lognormality; the differences are deep, hence the use of approximate lognormality to calculate moments is unsafe

Let us consider the following random variables: a normal (Gaussian) random variable G, a Poisson random variable P and a bounded random variable B obtained as the sum of a large number K of random variables $B_k = \log R_k$, each of which bounded by the same $\beta < \infty$. We want G, P and B to be nearly identical. Since the mean and the variance are equal in the case of P, they must be assumed equal for G and B. Finally, we want G, P and B to be near identical. Therefore, the value δ common to mean and their variance must be large. It follows that

$$\left\langle (e^G)^q \right\rangle = \exp(q\delta + \delta q^2/2) = \exp[\delta(q + q^2/2)],$$

$$\left\langle (e^P)^q \right\rangle = \exp(-\delta + \delta e^q) = \exp[\delta(e^q - 1)],$$

$$\left\langle (e^B)^q \right\rangle \le \exp(qK\beta).$$

Thus, $\langle (e^B)^q \rangle$ increases at most exponentially with q, $\langle (e^G)^q \rangle$ increases more rapidly than any exponential, and $\langle (e^P)^q \rangle$ more rapidly still. The expectations, equal respectively to $\langle e^G \rangle = \exp(1.5\delta)$ and $\langle e^P \rangle = \exp(1.7\delta)$, are already very different. The coefficients of variation, defined as

$$\frac{\langle (e^G)^2 \rangle}{\langle e^G \rangle^2} = e^{\delta} \text{ and } \frac{\langle (e^P)^2 \rangle}{\langle e^P \rangle^2} = e^{(e-1)2\delta} \sim e^{3\delta},$$

differ even more, and higher order moments differ strikingly. In short, it may be that B and P are nearly normal from the usual viewpoint (which is that of the so-called "weak topology"). But from the present viewpoint they provide poor approximations to the normal distribution. However, when $q < 1$ (for example, $q = 2/3$, as in the calculation of spectra) the discrepancy is smaller.

{PS 1998. *A fourth example of a manageable sequence of moments is the log-gamma distribution.* The reduced gamma r.v. of parameter δ, to be called Gamma (δ), is defined as having the density $u^{\delta-1}e^{-u}/\Gamma(\delta)$. It is well known that the sum of two independent gamma r.v. of parameters δ' and in exponent δ'' is a gamma r.v. of parameter $\delta' + \delta''$. Let the reduced log-gamma r.v. be defined as $e^{-\sigma \text{Gamma}(\delta)} / \langle e^{-\sigma \text{Gamma}(\delta)} \rangle$. The q-th moment of the log-gamma is finite if and only if $q > -1/\sigma$; when finite, its value is

$$\langle (e^{-\sigma \text{Gamma}(\delta)})^q \rangle = \int_0^\infty u^{\delta-1} e^{-u} e^{-u\sigma q} du / \Gamma(\delta) = (1 + \sigma q)^{-\delta}.$$

In particular, irrespective of δ, $\sigma \leq -1$ yields $\langle e^{-\sigma \text{Gamma}(\delta)} \rangle = \infty$. That is $\langle W \rangle = 1$ cannot be insured unless $\sigma > -1$. To my knowledge, the four families listed in this paper and this PS are the only ones that are endowed with explicit analytic expressions, have some finite moments and possess the property of being closed under addition.}

A2. On Orszag's remark concerning the determination of turbulence by its moments

Homogeneous turbulence is presumed to be determined by its moments, and the bulk of the theory based on the Navier-Stokes equations is devoted to efforts to determine these moments theoretically. Is intermittent turbulence also so determined? To answer, note that a random variable can have the same moments as the lognormal distribution, without itself being lognormal. Feller 1951 vol. 2, p.227 of the 2nd edition, credits this example to Heyde 1963. {P.S. 1998. In fact, the lognormal was the very first example to be mentioned in the original (1894) paper on the moment problem; see Stieltjes 1914.}

The reason for this indeterminacy is that the moments of a lognormal e^G increase so fast that $\Sigma[\langle \exp(2qG)\rangle]^{-1/(2q)} < \infty$, which expresses that the lognormal distribution fails to satisfy a necessary condition due to Carleman. Orszag 1970 has observed that a corollary of this indeterminacy is that, if intermittent turbulence were indeed lognormal, it would not be determined by its moments. On the other hand, suppose that Yaglom's $\bar{\varepsilon}_r$ is the product of independent bounded factors. In that case, the moments of intermittent turbulence do satisfy the Carleman criterion; therefore, the indeterminacy noted by Orszag vanishes. (*Note added during revision*. Novikov 1971, p.235, contains a remark to the same effect.)

A3. The dimension exponent, as introduced through the number of eddies of side *r* within which dissipation is concentrated

The purpose of this section of the appendix is to show that among subeddies of side r, most of the dissipation is concentrated in a subset of about $(L/r)^{D_E}$ subeddies.

Preliminary example: binomial weights. Let

$$\Pr\{W=0\} = 1-p \text{ and } \Pr\{W=1/p\} = p,$$

so that $D = -\log_C p$. Then $\bar{\varepsilon}_L(\mathbf{x}, L, \eta) = 1$ factors into two terms: (a) the contents of a non-empty eddy, namely

$$(p^{-1})^{\log_C(L/\eta)} = (L/\eta)^{\log_C p} = (L/\eta)^{-D},$$

and (b) the number of non-empty eddies of side η contained in a big eddy of side L. Since $E\,\bar{\varepsilon}_L(\mathbf{x}, L, \eta) = 1$, the expectation of this last number must be $(L/\eta)^D$.

Second example: lognormal weights and cubic eddies. Let us begin with the low frequency factor $\bar{\varepsilon}_r(\mathbf{x}, L, r)$. With W lognormal as in Section 4.3, $\log \bar{\varepsilon}_r$ is Gaussian with variance $\mu \log(L/r)$ and expectation $(\mu/2)\log(L/r)$. To simplify the notation, we shall denote $\bar{\varepsilon}_r$ by V. When $L/r \gg 1$, this lognormal factor has the feature that its expectation is due almost exclusively to occasional large values. As a result, one can select a function threshold (L/r) in such a way that values of V below threshold (L/r) are negligible. Specifically, if one defines $\tilde{V}$ by

$$\tilde{V} = \begin{cases} V & \text{when } V > \text{ threshold } (L/r), \\ 0 & \text{otherwise,} \end{cases}$$

then $E\tilde{V}$ is arbitrarily close to 1. Let us prove that such a result is achieved when C is a function subjected to the sole requirement $\lim_{r \to 0} C(L/r)/[\log(L/r)]^{1/2} = 0$, and when the "threshold" function is chosen to satisfy

$$\text{threshold } (L/r) = (L/r)^{\mu/2} \exp\{-C(L/r)\sqrt{\mu \log(L/r)}\}.$$

Indeed,

$$\langle V \rangle = \frac{0.1}{\sqrt{2r\mu \log(L/r)}} \int \exp\left\{ x - \frac{[x + (\mu/2)\log(L/r)]^2}{4\mu \log(L/r)} \right\} dx,$$

with an integration range from log[threshold (r, L)] to ∞. The expression in braces transforms into

$$\frac{-[x - (\mu/2)\log(L/r)]^2}{4\mu \log(L/r)},$$

and by changing the variable of integration to

$$z = [x - (\mu/2)\log(L/r)]\,[2\pi \log(L/r)]^{-1/2}$$

we obtain

$$\langle \tilde{V} \rangle = (2\pi)^{-1/2} \int \exp(-z^2/2) dz,$$

with an integration range from $-C(L/r)$ to infinity. As $L/r \to \infty$, $\langle V \rangle \to 1$, which shows that the contribution of other values of V to $\bar{\varepsilon}_r$ is asymptotically negligible, and that the above choice of N was indeed appropriate to make V arbitrarily closely approximated by $\tilde{V}$. From now on, one can consider the $(L/r)^3$ cells of side r that lie within a cube of side L, and divide them into those for which $V >$ threshold (L/r), and those for which $V <$ threshold (L/r).

For the former, the expectation of their total number is

$$(L/r)^3 \Pr\{V > \text{threshold } (L/r)\}.$$

In terms of the reduced Gaussian random variable

$$[\log V + (\mu/2)\log(L/r)]\sqrt{\mu \log(L/r)} = G,$$

the above probability becomes

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

Using a well-known tail approximation of G, the expected number in question is approximately equal to

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

(Note that this last approximation is independent of C.)

With cubic eddies replaced by straight segments, the only change in the above formulae is that the factor $(L/r)^3$ is replaced by L/r and hence $3-\mu/2$ by $1-\mu/2 = D_1$.

As for the cells in which $V <$ threshold (L/r), we want to show that their total contribution is negligible. The proof involves the high frequency factors $\bar{\varepsilon}_r(\mathbf{x}, r, \eta)$ and an application of the ergodic theorem. Details need not be given here.

General weights W. The assertion is that the number of eddies that are not nearly empty is about $(L/r)^D$. The proof cannot be given, but its principle can be indicated. The quantity $\langle W \log W \rangle$ is related to Shannon's concept of entropy-information, and it enters here because our problem can be restated in terms of information-theoretical asymptotical equiprobability; see Billingsley 1967.

A4. Introduction to a model of intermittency based on the limiting lognormal processes, in which eddies are randomly generated and the partition of dissipation is continuous

My earlier paper on intermittency M 1972j{N14}, involved a departure from the assumption of Section 2.1: the grid itself was made random, and was generated by the same model as the distribution of dissipation. The purpose of this section is to provide a transition from this to the earlier work.

Our point of departure consists in a prescribed grid of eddies, and a canonical cascade with lognormal weight W: $\log \varepsilon(\mathbf{x}, L, \eta)$ is Gaussian with variance $\mu \log(L/\eta)$ and expectation $-(\mu/2)\log(L/\eta)$. We know that the correlation of $\bar{\varepsilon}$ is approximately proportional to d^{-Q}, with $Q=\mu$. Because the eddies were prescribed, the random function $\varepsilon(\mathbf{x}, L, \eta)$ is non-stationary and discontinuous: between an eddy and its neighbors, there may be very large discontinuities. Both non-stationarity and discontinuity are of course quite unrealistic. One may instead demand that $\log \varepsilon(\mathbf{x}, L, \eta)$ be Gaussian and stationary, with the added restriction that it should be continuous and vary little over spans of order shorter than η. This will ensure that $\bar{\varepsilon}(\mathbf{x}, L, \eta)$ is nearly identical to $\varepsilon(\mathbf{x}, L, r)$. It remains to ensure that $\varepsilon(\mathbf{x}, L, \eta)$ has a correlation proportional to d^{-Q}. The simplest way to achieve this aim is to require $\varepsilon(\mathbf{x}, L, \eta)$ to have a truncated self-similar spectral density, namely a spectral density equal to $\eta/2\omega$ when $1/L < \omega < 1/\eta$, and equal to zero elsewhere. The resulting model may be viewed, as combining self-similarity with the maximum retrievable portion of the Kolmogorov third hypothesis.

The properties of $\bar{\varepsilon}(\mathcal{D}, L, \eta)$ relative to this model can be summarized as follows. The dimensions continue to be $D_E = E - \mu/2$ and the cascades are never regular: for $\mu < 2/E$, they are irregular with $\alpha_E = 2/\mu$, while for $\mu > 2/E$, they are degenerate. Compared with a canonical cascade with a lognormal W, the main differences involve the values of certain numerical constants.

A5. Remarks on Kolmogorov's third hypothesis of lognormality

This hypothesis states that, for every cube of center $\mathbf{x}$ and side $r > \eta$, $\bar{\varepsilon}_r(\mathbf{x}, L, \eta)$ follows the lognormal distribution, the variance of $\log \bar{\varepsilon}_r$ being equal to $\mu \log(L/r)$. Within Yaglom's context of prescribed eddies, either canonical or microcanonical, it will be shown that this hypothesis cannot hold. Then, it will be shown that in a wider context this hypothesis is tenable only if one makes unlikely additional conditions.

In the context of prescribed microcanonical eddies, the difficulty is the following. To say that $\log \bar{\varepsilon}_r$ is normal is to say that a finite number of the independent random variables $\log W_k$ add to a Gaussian distribution, and it follows, by a classical theorem (Lévy-Cramèr), that the $\log W_k$ must themselves be Gaussian, i.e., unbounded. On the other hand, we know that microcanonical weights must be bounded. Thus Kolmogorov's hypothesis cannot apply strictly for any fixed value of r, not even for $r = \eta$.

In the context of prescribed canonical eddies, the source of the difficulty is different. One can show that, in a canonical context, the correlation of every pair of ε_r is positive. (To be more precise, Yaglom's rough derivation suggests that the correlation is also positive in a microcanonical context. But a careful investigation, too long to be worth reporting, shows that for some value of d it must be negative. This follows from the fact that $\langle W_s W_t \rangle < 1$; see Section 3.6.) To the contrary, we shall see momentarily that the Kolmogorov hypotheses imply that at least some of those correlations are negative. Thus, the Kolmogorov hypotheses might conceivably hold for $r = \eta$, but the hypothesis relative to several values of r are incompatible, meaning that overall the hypotheses are internally inconsistent.

More generally, the joint assumptions that the random variables $\log \bar{\varepsilon}_r(\mathbf{x}, L, \eta)$ are normal for every r, with $\sigma^2 \log \bar{\varepsilon}_r = \mu \log(L/r)$ and $\langle \log \bar{\varepsilon}_r \rangle = -(\mu/2) \log(L/r)$, are incompatible with any model that leads to a positive reduced covariance for $\bar{\varepsilon}_\eta$. Indeed it would follow from these assumptions that $r^3\bar{\varepsilon}_r$, the mass of turbulence in a cube, satisfies

$$\left\langle [r^3 \bar{\varepsilon}_r(\mathbf{x}, L, \eta)]^q \right\rangle = r^{3q - q/2(q-1)\mu} L^{q/2(q-1)\mu}.$$

When r reaches its maximum value L, the above moment will reduce to r^{3q}, for all q. This is as it should. But the nature of convergence to this limit must be examined more closely. For example, let us subdivide our cube into (say) 2^3 portions. When $q/3 > 2\mu$, the exponent of r in the above expression is negative, implying that the value of the ratio $\langle (r^3\bar{\varepsilon}_r)^q \rangle / \langle [(r/2)^3 \bar{\varepsilon}_r]^q \rangle$ is less than 2^3. From an elementary result of probability, this means that at least two of our subcubes must have a negative reduced correlation. This conclusion, and hence Kolmogorov's form of the lognormal hypothesis, is inconsistent with the assumed positive covariance of $\bar{\varepsilon}_\eta(\mathbf{x}, L, \eta)$. (*Note added during revision*. This inconsistency is observed in Novikov 1971, p. 236. The author notes the contradictory behavior of the quantities he designates as μ_ρ; however, having noted the contradiction, he does not resolve it.)

This is the moment to point out that in my limit lognormal model (M 1972j{N14} and Section A4) the above inconsistency is avoided, because every formerly misbehaving moment turns out to be infinite.

A6. Footnote added in 1972, during revision

This text appeared as footnote to Section 2.2 of the original M 1974f, but deserves being changed into an appendix and discussed in the Annotations that follow.

A referee made me aware of Novikov 1969 and 1971 which helped put [my] results in focus. Novikov 1971 p. 236 observes that the moments of $\bar{\varepsilon}$ do not tend towards those of the lognormal distribution. Yet, "in the same manner as in [a textbook by] Gnedenko, it may be shown that the limit distribution is lognormal." The puzzling discrepancy between these results appears to be due to use of conflicting approximations. Earlier, Novikov 1969, p. 105 states that "all moments (if they exist) must have a power law character." The phrase in parentheses raises the possibility that moments may not exist, but this possibility is regrettably dismissed and is not discussed again.

ACKNOWLEDGMENT

I benefited greatly from conversations with Jay M. Berger of IBM, Erik Mollo-Christensen of MIT, Robert W. Stewart, then of the University of British Columbia, and Akiva Yaglom of the Academy of Sciences of the USSR. Computer simulations were carried out in part by Gerald B. Lichtenberger, then of Yale University, and in part by Hirsh Lewitan of IBM. During the last of the revisions of this paper for publication, Robert Kraichnan helped me to sharpen the distinctions between different forms of the cascades. Several early forms of portions of this paper have been circulated privately; excerpts and abstracts have been published in the *Proceedings of the Summer Study Program in Geophysical Fluid Mechanics*, Woods Hole, Massachusetts, 1965; *Proceedings of the IUGG-IUTAM International Symposium on Boundary Layers and Turbulence including Geophysical Applications*, Kyoto, Japan, 1966 (A Supplement to *Phys. Fluids* 10, (suppl.), 1967, p. S302); and *Proceedings of the Polytechnic Institute of Brooklyn Symposium on Turbulence of Fluids and Plasmas*, New York, 1968 (John Wiley, 1969), and in M 1966k{N}, and M 1969b{N}.

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

Editorial comments. The notation was brought close to the current usage. The letter Γ denoting a base was replaced by b; the letter h for the exponent in a moment was replaced by q; the expression $f(h)$ by $\overline{\tau}(q)$ (my old function $f(h)$ is now ordinarily denoted by $-\tau(q)$ and the notation $\overline{\tau}(q) = -\tau(q)$ avoids having to change signs throughout this paper); the D identifying a domain was replaced by $\mathfrak{D}$ the letter i for the embedding dimension was replaced by E; the letter Δ for the intrinsic fractional dimension was replaced by D; the expression V^* was replaced by $\tilde{V}$, and the integer N in Section 4.8 was replaced by C. A footnote to Section 2.2 in the original was made into Appendix A.6. The other footnotes were incorporated in the text. A few lengthy ones are printed in smaller size letters.

Comment on Appendix A6. This Appendix acknowledges that E.A. Novikov and I simultaneously showed that the q-th moment of the average dissipation $\bar{\varepsilon}$ within a volume is the volume's size raised to some power $\tau(q)$. However, the derivation of this $\tau(q)$ is only the first of three stages of the theory of multifractals.

Elaborating on M 1969b {N13}, this chapter goes on to a second step and concludes that the central limit theorem is far from being sufficient from the viewpoint of the study of multifractals, that is, the distribution of $\bar{\varepsilon}$ is *not* approximately lognormal to a significant degree. M 1974c{N16} takes a further third step. It uses the Cramer large deviations theory, hence the Legendre transform, to obtain the multifractal function $f(\alpha)$ via the probability distribution of $\bar{\varepsilon}$ plotted on suitable log-log coordinates.

To the contrary, Novikov stopped at the first step. He did observe the difficulties $\tau(q)$ may present and the fact that its non-universality contradicts lognormality. But he did not resolve these difficulties. In terms of the function $f(\alpha)$ (which he did not consider), his text implied that the graph of $f(\alpha)$ is in every case a parabola.

Other scientific comments. Additional comments on this chapter are combined with comments at the end of Chapter N16.

How did M 1974f come to be written and published. This paper was hard to write, and it is hard to read because it is far too detailed for a first formal work on a new topic. That counterproductive complication was largely the unintended result of a very long refereeing process. The annotation "received 1 March 1972" refers the date of the final draft. An editor of the *Journal of Fluid Mechanics*, Keith Moffatt, invited a paper after hearing me sketch M 1967k{N12} at Kyoto in 1966. Thus, M 1974f was preceded by at least two very distinct versions ranging back to 1968. Among successive referees the *Journal* called upon, several professed utter bafflement about the problem and the solution. Other referees made nice noises, but asked not to be called again. One referee who professed competence picked endlessly at insignificant issues, and chastised me for incompleteness. (The next significant step in the theory of multifractals did not come until the mid 1980s.) All this encouraged endless rewriting that became increasingly "defensive" and counter-productive.

Fortunately, Keith Moffatt showed great foresight and kindness, but everything that sounded like "philosophy" had to be removed. As a result, the text that *J. Fluid Mech.* received on March 1, 1972 left a remainder. After revision, this became a typescript of 90 tightly packed pages, titled *The Geometry of Turbulence.* My copy indicates it was received by *J. Fluid Mech.* on August 27, 1974, but I changed my mind. One part became M1975F and a second was largely incorporated in M 1977F and M 1982F{*FGN*}. Its summary, slightly abbreviated, is reproduced as Section N4.2.