J. Fluid Mech. (1974), *vol.* 62, *part* 2, *pp.* 331–358

Printed in Great Britain

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

By BENOIT B. MANDELBROT

Mathematical Sciences Department, IBM Thomas J. Watson Research Center, Yorktown Heights, New York 10598

(Received 1 March 1972 and in revised form 19 September 1973)

Kolmogorov's 'third hypothesis' asserts that in intermittent turbulence the average $\bar{\epsilon}$ of the dissipation ϵ, taken over any domain D, is ruled by the lognormal probability distribution. This hypothesis will be shown to be inconsistent, save under assumptions that are extreme and unlikely. Further, a widely used justification of lognormality, due to Yaglom and based on probabilistic argument involving a self-similar cascade, will be discussed. In this model, lognormality indeed applies strictly when D is 'an eddy', typically a three-dimensional box embedded in a self-similar hierarchy, and may perhaps remain a reasonable approximation when D consists of a few such eddies. On the other hand, the experimental situation is described better by considering averages taken over essentially one-dimensional domains 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 $\bar{\epsilon}$ differs from the lognormal distribution. Depending upon various parameters, the discrepancy may be either moderate, or considerable, or even extreme. In the latter two cases, high-order moments of $\bar{\epsilon}$ turn out to be infinite. This avoids various paradoxes (to be explored) that are present in Kolmogorov's and Yaglom's approaches. The third purpose is to note that high-order moments become infinite only when the number of levels of the cascade tends to infinity, meaning that the internal scale η tends to zero. Granted the usual value of η, this number of levels is actually small, so the representativity of the limit is questionable. This issue was investigated through computer simulation. The results bear upon the question of the extent to which Kolmogorov's second hypothesis applies in the face of intermittency. The fourth purpose is as follows. Yaglom noted that the cascade model predicts that dissipation only occurs in a portion of space of very small total volume. In order to describe the structure of this portion of space, it will be shown useful to introduce the concept of the 'intrinsic fractional dimension' Δ of the carrier of intermittent turbulence. The fifth purpose is to study the relations between the parameters ruling the distribu-

tion of $\bar{\epsilon}$, and those ruling its spectral and dimensional properties. Both conceptually and numerically, these various parameters turn out to be distinct, which opens up several problems for empirical study.

---

1. Introduction and synopsis of paper

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–Oboukhov theory, in which the rate of dissipation ϵ was considered as uniform in space and constant in time. In intermittent turbulence, ϵ must be considered a function of time and space. Let $\bar{\epsilon}(D)$ be its spatial average over a domain D. Several approaches to intermittency view $\bar{\epsilon}$ as lognormally distributed: in Oboukhov (1962), lognormality is a pragmatic assumption; in Kolmogorov (1962), it is a basic 'third hypothesis' applicable to every domain D; in Yaglom (1966) it is derived from a self-similar cascade model,† which also predicts that the parameter μ of the lognormal distribution and the exponent in the expressions ruling the correlation and spectral properties of $\bar{\epsilon}$ are equal.

While substantial effort is currently being devoted to testing lognormality experimentally, the purpose of the present paper is to probe its conceptual foundations. On the one hand, as were the works of Kolmogorov and Yaglom, we shall be concerned with a phenomenology whose contact with physics remains remote. In particular, the central role of dissipation will not be disputed, despite the difficulties it shows. On the other hand, greater care will be devoted to matters of internal logical consistency and to details of the assumptions, and the theory will be further developed, in particular in view of exploring the relation between theory and experiment.

Since this paper is somewhat lengthy, the bulk of the mathematics (which as yet has no other application in fluid mechanics) will be postponed to §§ 4 and 5. The main results will be stated without proof in this section and § 2, and in the captions of the figures. 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 $\bar{\epsilon}(D)$ be the average of the dissipation ϵ over a spatial domain D. One form of Yaglom's model assumes that D is an 'eddy', perhaps a three-dimensional cube embedded in a self-similar hierarchy. On the other hand, in all actually observed averages,

† It happens that a closely analogous cascade was considered by de Wijs (1951, 1953), 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 discussions of instances of interplay between multiplicative perturbations and the lognormal and Pareto 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.

D is more nearly a very thin cylinder. By following up consequences of Yaglom's model in this case, it will be shown that $\bar{\epsilon}(D)$ is never lognormal and that its 'qualitative' behaviour can fall into any of three classes. In a drastically extreme first class, called 'regular', $\bar{\epsilon}(D)$ is not far from being lognormal. In a second extreme class, called 'degenerate', all dissipation concentrates in a few huge blobs. In the intermediate class, called 'irregular', $\bar{\epsilon}(D)$ is non-degenerate but far from lognormal. Its most striking characteristic is a parameter α_1, satisfying $1 < \alpha_1 < \infty$, which rules the moments (ensemble averages) $\langle \bar{\epsilon}^h(D) \rangle$. When $h < \alpha_1$, $\langle \bar{\epsilon}^h(D) \rangle < \infty$ for all values of the inner scale η, but when $h > \alpha_1$ and $\eta = 0$, $\langle \bar{\epsilon}^h(D) \rangle = \infty$. Finally, when $h > \alpha_1$ and η is positive but small, $\langle \bar{\epsilon}^h(D) \rangle$ 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 $\alpha_1 \leqslant 1$. Numerous inconsistencies that have been noted in the literature, concerning the behaviour of the moments of $\bar{\epsilon}$ under the lognormal hypothesis, will thereby be eliminated.

(*b*) Another part of this paper proposes a change in Yaglom's model. The latter involves, though only implicitly, a hypothesis of rigorous local conservation of dissipation within eddies, a feature which will be said to characterize his cascade as being 'microcanonical'. It will be argued that it may be useful to view conservation as holding only on the average, and the resulting cascades, called 'canonical', will be investigated. When a cascade is canonical, it will be shown that any of the three classes of behaviour of $\bar{\epsilon}(D)$, as defined above under (*a*), may be encountered even when D is an eddy, save for the replacement of the parameter α_1 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 $\bar{\epsilon}(D)$ when D is an eddy is compatible with an irregular $\bar{\epsilon}(D)$ when D is a cylinder; also, an irregular $\bar{\epsilon}(D)$ when D is an eddy is compatible with a degenerate $\bar{\epsilon}(E)$ when D is a cylinder.

(*c*) Another aspect of this paper is purely critical, and concerns Kolmogorov's second hypothesis, which asserts that the value of η does not influence $\bar{\epsilon}(D)$ in the similarity range. Such will indeed be shown to be the case when $\bar{\epsilon}(D)$ is in the regular class for every domain D, but not when all $\bar{\epsilon}(D)$ 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, to 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, in preference to its relative volume (which is very small and too dependent upon η), a parameter Δ called the 'intrinsic fractional dimension' of the carrier.

(*e*) Yaglom's theory introduces yet another parameter, which characterizes the spectrum of $\bar{\epsilon}$ and is related to a correction factor to the exponent $-\frac{5}{3}$ of the classic Kolmogorov power law. This parameter will be denoted by Q. The parameters α_1, Δ and Q will be seen 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 deduce them all as 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 behaviour, 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 § 1. To do so, it is necessary to begin by describing Yaglom's cascade model, in a form that has been narrowed and made more specific (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 from the outset and 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 'cell number'). This expresses the fact that the grid of eddies is self-similar in the range from η to L. Obviously, $C^{\frac{1}{3}}$ must be assumed to be an integer and is denoted by Γ. The sides (edge lengths) of the largest and smallest eddy are equal to the external scale L and internal scale η respectively.

The unit of length will be assumed to be chosen so that η and L are not only dimensionless but are powers of Γ. The density of turbulent dissipation at the point $\mathbf{x}$ is denoted by $\epsilon(\mathbf{x}, L, \eta)$, and the density average over the domain D by $\bar{\epsilon}(D, L, \eta)$. (This sharpening of the previous notation $\bar{\epsilon}(D)$ is necessary because some arguments below will amount formally to varying the values of L and η.) Units of dissipation will again be assumed to be chosen so that $\bar{\epsilon}$ is dimensionless. When D is a cubic eddy of side r and centre $\mathbf{x}$ (with $-\log_\Gamma r$ an integer) we write

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

It is further assumed that the distribution of dissipation over its self-similar grid is itself self-similar, in the sense that, whenever $\eta \ll r < r\Gamma \ll L$, the ratio $\bar{\epsilon}_{r/\Gamma}(\mathbf{x}_s, L, \eta)/\bar{\epsilon}_r(\mathbf{x}, L, \eta)$ is a random variable, to be denoted by Y_s, whose distribution is independent of r. (Here, $\{\mathbf{x}_s\}$ is a regular grid of centres of subeddies.)

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

† At this point, the reader may digress to the appendix, §§ A 1 and A 2, which involve two comments 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 $\bar{\epsilon}$ is lognormal for all D. First (not the basic reason), Yaglom's argument is rigorous only when D is a cubic eddy. When 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 D, the moments $\langle \bar{\epsilon}^h \rangle$ are finite for all h. A second more basic argument has to do with the comparison of theory and experiment. Even though averages taken over three-dimensional domains D may be appropriate to 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, as frozen at a given instant in time. When the radius of such a D is of the same order of magnitude as the inner scale η of the turbulence, D can be approximated by a one-dimensional straight segment. Thus one must raise the question of whether or not, for such D's, the distribution of $\bar{\epsilon}(E, L, \eta)$ remains approximately lognormal. This question is not raised explicitly in Yaglom's work. On the other hand, since he and subsequent writers are concerned with the extent to which the observed data fit the lognormal distribution, they assume implicitly that the dimensionality of the averaging domain D has little effect on the distribution of $\bar{\epsilon}(D, L, \eta)$.

This paper will show this implicit assumption to be unwarranted. More precisely, whenever D is not an eddy, the computation of the distribution of $\bar{\epsilon}(D, L, \eta)$ will be seen to require further detailed assumptions about the cascade process. One such set of assumptions, compatible with Yaglom's, is sufficiently simple to allow detailed study. It turns out that, except under trivial circumstances, $\bar{\epsilon}(D_\eta, L, \eta)$, where D_η is a cylinder of length r and radius η, is *not* lognormal.

To introduce this simplest set of detailed assumptions about the cascade, it suffices to view the hierarchy of eddies as a more or less formal device for constructing the fully cascaded state of the medium, through division of space by imaginary lines into even smaller chunks. If so, every stage of splitting preserves the total dissipation, and hence the average of local averages. The simplest procedure is to assume nothing else about the corresponding Yaglom ratios Y. Such a cascade will be called 'microcanonical'. Here, it turns out that successive ratios of the form $\bar{\epsilon}(D', L, \eta)/\bar{\epsilon}(D'', L, \eta)$, when D' and D'' are cylinders of identical length r but different cross-sections (with D' embeded in D''), 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 §§ 3 and 4 it will be proved that, except in a trivial case, the distribution of the limit $\bar{\epsilon}(D_0, L, 0)$, where 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 D over the distribution of $\bar{\epsilon}(D_0, L, 0)$ may be critical. The extent of the divergence of the distribution from the lognormality is expressed to a significant extent by the

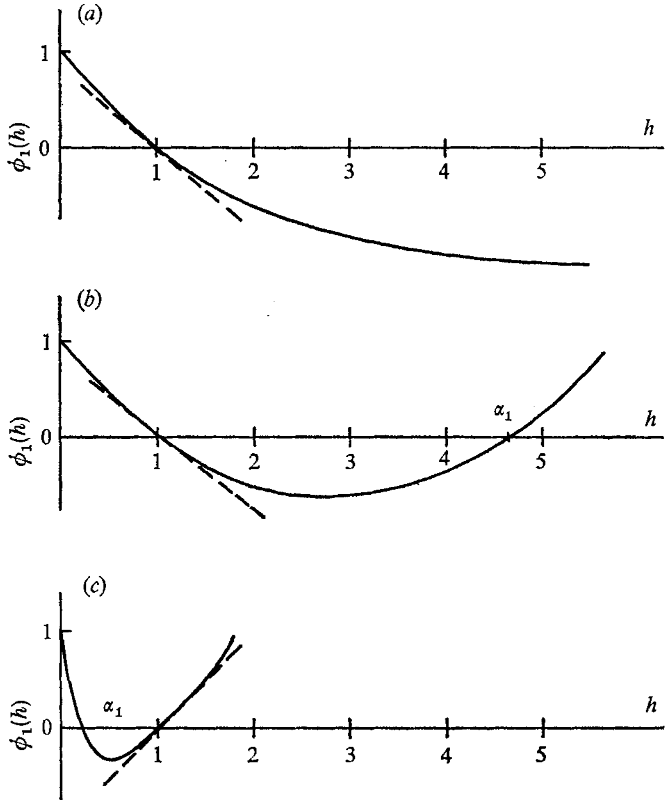

FIGURE 1. The distribution of the averaged dissipation $\bar{\epsilon}$ is determined by that of the random weight W, which is roughly speaking Yaglom's ratio between the average dissipation with a subeddy and an eddy. We plot the function $\phi_1(h) = \log_\Gamma\langle W^h\rangle - (h-1)$, which is always convex. Through $\phi_1(h)$, all presently interesting aspects of the said dependence are described as follows. The spectral properties of $\bar{\epsilon}$ (the only one to have been examined before the present study) are determined by the value of $\phi_1(2)$. In addition, the distribution properties (finiteness of moments) depend on the value of α_1, defined as the root (other than $h = 1$) of the equation $\phi_1(h) = 0$. Finally, the fractional dimension of the carrier of turbulence depends on the values of $\phi_1'(1)$. Thus, from the viewpoint of properties of $\bar{\epsilon}$ of present interest, its distributions can fall into the following three classes. (a) Regular class: $\phi_1'(1) < 0$, $\phi_1(2) < 0$ and $\alpha_1 = \infty$. (b) Irregular class: $\phi_1'(1) < 0$, $\phi_1(2) < 0$ and $1 < \alpha_1 < \infty$. (c) Degenerate class: $\phi_1'(1) > 0$, $\phi_1(2) > 0$ and $0 < \alpha_1 < 1$.

value of a parameter, denoted as α_1, which is defined as the second zero (the first being $h = 1$) of the equation

$$\phi_1(h) = \log_\Gamma\langle Y^h\rangle - (h-1).$$

The definition of α_1 is motivated in § 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 § 3).

The first class, called 'regular', includes all Y that are bounded by Γ. Here, one has $\alpha_1 = \infty$. The resulting $\bar{\epsilon}(D_0, L, 0)$ only differs from lognormality by a factor that is random, but is essentially independent of η, and has, like the lognormal distribution, finite moments to every order.

The second class, called 'degenerate', corresponds to Y's that are 'sufficiently scattered' (in fact more scattered than is likely in nature; nevertheless, the case must be fully understood). Here one has $\alpha_1 \leqslant 1$. The resulting $\bar{\epsilon}(D_0, L, 0)$ vanishes almost surely. In particular, one has, for every h, $\langle \bar{\epsilon}^h(D_0, L, 0)\rangle = 0$. Perhaps against 'physical intuition', $\bar{\epsilon}(D_0, L, 0) \equiv 0$, and hence $\langle \bar{\epsilon}(D_0, L, 0)\rangle = 0$ is compatible with the combination of $\lim_{\eta\to 0} \langle \bar{\epsilon}(D_\eta, L, \eta)\rangle = 1$, $\lim_{\eta\to 0} \langle \bar{\epsilon}^h(D_\eta, L, \eta)\rangle = 0$ for $h < 1$, and $\lim_{\eta\to 0} \langle \bar{\epsilon}^h(D_\eta, L, \eta)\rangle = \infty$ for $h > 1$. Thus, one is dealing with a comparatively rare instance when the possible discrepancy between the moments of the limits and the limits of the moments is not a mathematical pathology, but has possible direct practical consequences.† A classical illustration of the possibility of obtaining this discrepancy is the sequence for which $\bar{\epsilon}(D, L, \eta)$ equals $1/\eta$ with probability η, and equals zero with probability $1-\eta$. Thus the degenerate case suggests that, when η is non-zero but small, dissipation concentrates in a few huge blobs.

The third class, called 'irregular', includes all Y's that are not too scattered, but nevertheless can exceed Γ. Here, one has $1 < \alpha_1 < \infty$. Then, $\langle \bar{\epsilon}(D, L, \eta)\rangle$ remains identically equal to one, while higher moments $\langle \bar{\epsilon}^h\rangle$ behave as follows: they remain finite when $h < \alpha_1$, but when $h > \alpha_1$, they tend to infinity as $\eta \to 0$, which 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 $\Pr(\bar{\epsilon} > x)$ is 'smooth', the simplest distribution he is likely to envisage is the 'hyperbolic', defined as follows: $\min \bar{\epsilon} = x_0 = \alpha_1/(\alpha_1 - 1) > 0$ and $\Pr(\bar{\epsilon} > x) = (x/x_0)^{-\alpha_1}$. The next simplest possibility is $\Pr(\bar{\epsilon} > x) = C(x)\, x^{-\alpha_1}$, where $C(s)$ varies 'smoothly and slowly' as $x \to \infty$ (for example, has a non-trivial limit, or perhaps varies like $\log x$ or $1/\log x$). Such random variables $\bar{\epsilon}$ are called 'asymptotically hyperbolic' or 'Paretian'. To test for their occurrence, it is common practice to plot $\log \Pr(\bar{\epsilon} > x)$ 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. If so, all moments of $\bar{\epsilon}(D, L, \eta)$ are very large but finite. If its distribution is again plotted in log–log co-ordinates, it must end on a tail that plunges down more rapidly than any straight line of finite slope. However, the behaviour of the moments of $\bar{\epsilon}$ as $\eta \to 0$ also yields a definite prediction for small η, namely the following: *the log–log plot of the distribution is expected to include a long*

† This is a bit reminiscent of the singularity, familiar in fluid mechanics, encountered when the coefficient of viscosity tends to zero.

‡ *Footnote added in* 1972, *during revision.* Two papers by Novikov (1969, 1971), of which a referee has made me aware, help to put these results in focus. On p. 236 of his 1971 paper, Novikov observes that the moments of $\bar{\epsilon}$ do not tend towards those of the lognormal distribution. On the other hand, he claims that "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 the 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 by not being discussed again.

'penultimate' range within which it is straight and of slope α_1. This is one of the main predictions of the present work.†

The ease of verifying this prediction increases as the slope α_1 becomes less steep. The value of α_1 can (in an approximation discussed in § 4.8) be inferred from the spectral exponent $Q = \mu$ as equal to about $2/\mu \sim 4$. This suggests that moments should misbehave for $h \geqslant 4$. Further discussion of empirical results is better postponed until more data are available.

2.3. *Validity of the microcanonical assumption and introduction of canonical cascades*

The second purpose of this paper is to probe Yaglom's assumption that the ratios of the form $\bar{\epsilon}(D_s, L, \eta)/\bar{\epsilon}(D, L, \eta)$, relative to a subeddy D_s and to an eddy D containing 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 (while keeping the approximation that in a cascade an eddy divides exactly into subeddies) one may view cascading (still a way of building up the fully cascaded state) as combining splitting with some kind of diffusion, in such a way that conservation of dissipation only holds on the average.‡ The resulting model, to be called 'canonical', is interesting because (*a*) when 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 $\bar{\epsilon}(D)$ fails to be lognormal even when D is an eddy. The theory of the canonical $\bar{\epsilon}(D)$ with D an eddy follows the same pattern as the theory of the microcanonical $\bar{\epsilon}(D)$ with D a cylinder. Thus, it can fall into either of the three classes noted in § 2.2, with the change that one must replace α_1 by a new parameter α_3. Ordinarily, one has $\alpha_3 > \alpha_1$.

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 contact 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

† 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, namely, Gaussian. (*b*) In the theory of the number of offspring in a birth-and-death process, the asymptotic distribution depends upon the distribution of the number of offspring per generation, so it is no longer universal. Using statistical mechanics, the thermodynamic 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; I have found it, on the contrary (see § 4.3), to be closer to theory (*b*), the resulting absence of universality being probably intrinsic. More precisely, the theory underlying this paper is an aspect of the 'theory of birth, death and random walk'.

‡ The 'dissipation' here invoked corresponds physically to the rate of energy transfer between eddy sizes, rather than to the ultimate rate of conversion of eddy kinetic energy into heat.

canonical model of a cascade be rephrased in terms of energy transfer between different scale sizes. Nevertheless, attempting this would go beyond the purpose of the present work, and we shall stick to the logical analysis of the cascades. The relative advantages of the two main models are as follows.

Yaglom's argument. In the case of cylinders, it 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, it 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 behaviour of the moments makes it useful to probe Kolmogorov's second hypothesis, as stated originally (1941) for homogeneous turbulence and as generalized in Kolmogorov (1962) to intermittency. Intuitively, if D is a domain of characteristic scale $\gg \eta$, the hypothesis is that the distribution of $\bar{\epsilon}(D, L, \eta)$ is nearly independent of η. To state this rigorously, make η into a parameter and let it tend to zero. Kolmogorov's second hypothesis might merely express that $\lim_{\eta\to 0} \bar{\epsilon}(D, L, \eta)$ should exist. If so, 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 D is an eddy of the microcanonical model, and for other D's and models leading to the regular class above, we have $\bar{\epsilon}(D, L, \eta) \to \bar{\epsilon}(D, L, 0)$ mathematically and, for all $h > 0$, $\langle \bar{\epsilon}^h(D, L, \eta)\rangle \to \langle \bar{\epsilon}^h(D, L, 0)\rangle$. In this case, as long as η is small, it is permissible to consider the 'error term' $\bar{\epsilon}(D, L, \eta) - \bar{\epsilon}(D, L, 0)$ as being itself 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, $\Pr\{\bar{\epsilon}(D, L, 0) = 0\} = 1$. For $\eta > 0$ but small, $\bar{\epsilon}(D, L, \eta)$ and $\bar{\epsilon}(D, L, 0)$ may be mathematically close, but are intuitively very different. The actual behaviour of $\bar{\epsilon}$ 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, for small η the error term $\bar{\epsilon}(D, L, \eta) - \bar{\epsilon}(D, L, 0)$ is extremely likely to be small, but in cases when it is not small, it may be very large, so that its own moments of high order are infinite. In this case, the Kolmogorov second hypothesis is controversial, its degree of validity improving as α_1 increases.

2.5. *Relationship with the 'limiting lognormal' model*

Every cascade model involves eddies. The present paper follows Yaglom in assuming them to be prescribed in advance. In addition, both the canonical and the microcanonical variants allow the distribution of dissipation between neighbouring subeddies to be highly discontinuous. However, an earlier study, Mandelbrot (1972), has investigated still 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 A 4 will describe its main characteristics.

2.6. *Simulation study 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 (i) that many of the involved discrepancies from lognormality should manifest themselves only in a relatively small number of largest observations, and (ii) that they depend greatly upon high-order moments of the Yaglom ratio Y, which express comparatively minute characteristics of the cascade model. If this is 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 Oboukhov, who did evaluate moments, would prove less pragmatic than he thought. Nevertheless, 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 $\bar{\epsilon}(D, L, \eta)$ when D is a cylinder, one needs assumptions about the local distribution of ϵ within eddies. We shall build a model formally, by making η smaller and smaller.

Initially, $\eta = L$ and the dissipation, equal to $\bar{\epsilon}(\mathbf{x}, L, L)$, is uniformly distributed in space. Each successive stage of the cascade begins with dissipation whose density is uniform in each eddy of side r. Such an initial distribution is the same as if one had $\eta = r$; it can therefore be denoted by $\epsilon(\mathbf{x}, L, r)$. The stage ends with dissipation whose density is uniform in each subeddy of side r/Γ. When the centre of an eddy of side r is denoted by $\mathbf{x}$, the centres of the immediately smaller subeddies will be denoted by $\mathbf{x}_s$, with $0 \leqslant s \leqslant C-1$; they form a regular lattice. The corresponding densities can be denoted by $\epsilon(\mathbf{x}_s, L, r/\Gamma)$. Next, designate the random variable $\epsilon(\mathbf{x}_s, L, r/\Gamma)/\epsilon(\mathbf{x}, L, r)$ by W_s. The ratio W and

Yaglom's ratio Y differ by the fact that W involves local densities and Y involves averages, but in one model (the microcanonical) the concepts of Y and W 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 $\bar{\epsilon}(\mathbf{x}, L, \eta)$.

Factoring out of $\bar{\epsilon}(D, L, \eta)$. The random variable $\epsilon(\mathbf{x}, L, \eta)$ 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, one can write $\epsilon_r(\mathbf{x}, L, \eta)$ as the product of two statistically independent factors, which can be studied separately, and identify them respectively with $\bar{\epsilon}_r(\mathbf{x}, L, r)$ and $\bar{\epsilon}_r(\mathbf{x}, r, \eta)$. The former is a 'low frequency factor', being independent of η and having r/L as the sole parameter; the latter is a 'high frequency factor', being independent of L and having r/η as the sole parameter. More generally, when D is not an eddy but is included in an eddy of side r, one has

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

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

For the low frequency factor, it suffices to follow Yaglom, as in § 2.1,

$$\log \bar{\epsilon}_r(\mathbf{x}, L, r)$$

being the sum of $\log_\Gamma (L/r)$ random factors of the form W. Assuming

$$\langle (\log W)^2 \rangle < \infty,$$

$\log \bar{\epsilon}_r$ is approximately normally distributed, and $\bar{\epsilon}_r(\mathbf{x}, L, r)$ 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, $\bar{\epsilon}_r(\mathbf{x}, L, r)$ 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 be any complication, and in fact will allow consideration of examples useful because of their simplicity.

3.3. *The high frequency factor: limiting behaviour for* $\eta \to 0$

This limiting behaviour 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 D be simple, meaning for $i = 3$ that D is the sum of a finite number of eddies, and for $i = 1$ that D is the sum of a finite number of eddy edges. Consider $\bar{\epsilon}(D, L, \eta)$ (for fixed 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, $\bar{\epsilon}$ tends to a finite limiting random variable.

Proof. This proof is written as a digression addressed to readers having an elementary knowledge of the theory of 'martingales' (which 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 respectively viewed 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. 'Time' is here discrete, being equal to $-\log_\Gamma \eta$. Assume that D is an eddy of side r; a similar argument applies to other simple D's. Denote its subeddies of side r/Γ by D_s. We know that

$$\bar{\epsilon}(D, L, \eta/\Gamma) = \frac{1}{C}\sum_{s=0}^{C-1} W_s \bar{\epsilon}(D_s, L, \eta).$$

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

$$E_C \bar{\epsilon}(D, L, \eta/\Gamma) = \frac{1}{C}\sum_{s=0}^{C-1} E_C \bar{\epsilon}(D_s, L, \eta) = E_C \bar{\epsilon}(D, L, \eta) = \bar{\epsilon}(D, L, \eta).$$

This proves that $\bar{\epsilon}(D, L, \eta)$ is a martingale. Moreover, $\bar{\epsilon}$ is non-negative. Hence, it obeys a convergence theorem (Doob 1953, p. 319) which asserts that, as $\eta \to 0$, $\bar{\epsilon}(D, L, \eta)$ converges to some limiting random variable, to be denoted as $\bar{\epsilon}(D, L, 0)$.

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

Remark. The above theorem means that, when $r/\eta \gg 1$, one knows $\bar{\epsilon}(D, L, \eta)$ '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 $\bar{\epsilon}(D, L, \eta)$ to $\bar{\epsilon}(D, L, 0)$ (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*

A cascade will be called microcanonical if the sum $\sum_{s=0}^{C-1} W_s$ *of the weights* W_s *that correspond to all the subeddies of any eddy is precisely equal to* C. As a corollary, $\langle W \rangle = 1$ and $W < C$. The microcanonical condition expresses that, at each cascade stage, the total dissipation $r^3 \epsilon(\mathbf{x}, L, r)$ within an original eddy is replaced by an *equal dissipation* distributed among its C subeddies of centres $\mathbf{x}_s$, namely

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

Hence, as long as $\eta < r$, $\bar{\epsilon}_r(\mathbf{x}, L, \eta) = \epsilon(\mathbf{x}, L, r)$, independently of η. This shows the high frequency factor $\bar{\epsilon}_r(\mathbf{x}, r, \eta)$ to be identically equal to 1, and in particular 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 $s \neq t$

$$\begin{aligned}\langle W_s W_t \rangle &= \langle W_s E_C(W_t, \text{knowing } W_s)\rangle = \langle W_s(C-W_s)/(C-1)\rangle \\ &= 1-(\langle W^2\rangle -1)(C-1)^{-1} < 1 = \langle W_s\rangle\langle W_t\rangle.\end{aligned}$$

This inequality expresses that any two weights are correlated negatively (see § A 5). Higher cross-moments, e.g. $\langle W_s^2 W_t\rangle$, are also < 1.

3.5. *The canonical cascade*

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.* It will become critical to allow W to exceed the ceiling $W = C$.

The canonical variant as an approximation for cylinder averages in a microcanonical cascade. Consider the 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 § 3.3, ending up typically with a lognormal $\bar{\epsilon}_r(\mathbf{x}, L, r)$. 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, and we know their number is not C but $\Gamma = C^{\frac{1}{3}}$. The second difference is that the conditions imposed on the corresponding weights are (*a*) $W_s < C$, (*b*) $\sum_{s=0}^{\Gamma-1} W_s \leqslant C$ and (*c*) $\langle W_s\rangle = 1$. By contrast, if the second subcascade had been microcanonical with Γ subeddies per eddy, the weights would have obeyed the conditions (*a*′) $W_s < \Gamma$ and (*b*′) $\sum_{s=0}^{\Gamma-1} W_s = \Gamma$, which are much stronger. As $C \to \infty$ and $\Gamma/C \to 0$, conditions (*a*) and (*b*) above become increasingly less demanding in comparison with (*a*′) and (*b*′).

This observation gives us a choice between two procedures. One can study the line sections directly and rigorously. Alternatively, *the second subcascade generating a line average can be approximated by a canonical cascade.* In this case, the condition $W < C$ may, in a first approximation, be waived; W may even be approximated by a lognormal random variable, though the latter is unbounded.

Hence, even if the cascade ruling the cubic eddies is trusted to be 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 the direct argument.

The effect of the condition $W < C$ upon the difference between the results of the microcanonical and the canonical models. In order for three-dimensional canonical eddies to be regular, meaning that averages yielded by the canonical and microcanonical theories only differ by a factor that is random but rather innocuous, the necessary and sufficient condition is $W < C$. When $W < \Gamma$, the same applies to one-dimensional averages. But when $\Gamma < \max W < C$, the two kinds of averages belong to different classes and thus may differ significantly.

4. Classification of cascades according to the behaviour of the moments of $\bar{\epsilon}$

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

Let D be an eddy of dimension $i = 3$. The definition of § 3.3 yields, irrespective of the rule of dependence between the W's,

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

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

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

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

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

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

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

When $i = 1$, D and D_s being straight intervals of length r, one has, similarly,

$$\bar{\epsilon}(D, 1, \eta/\Gamma r) = \Gamma^{-1} \sum_{s=0}^{\Gamma-1} W_s \bar{\epsilon}(D_s, 1, \eta/r).$$

Derivation of the moments of eddy averages from the basic recurrence relation. For $h = 1$, it suffices to check that the relation $\langle \bar{\epsilon}_1(\mathbf{x}, 1, \eta)\rangle = 1$ and the above recurrence relation are compatible. For $h > 1$, the recurrence relation for $\bar{\epsilon}_1$ can be used to deduce a recurrence relation for the sequence of the moments $\langle \bar{\epsilon}_1^h(\mathbf{x}, 1, \Gamma^{-k})\rangle$. 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 $\bar{\epsilon}_1(\mathbf{x}, 1, \eta) = 1$, but we want to verify that $\langle \bar{\epsilon}_1^h(\mathbf{x}, 1, \eta)\rangle = 1$. Indeed, for $h = 2$,

$$\begin{aligned}\langle \bar{\epsilon}_1^2(\mathbf{x}, 1, \eta/\Gamma)\rangle &= C\langle (W/C)^2\rangle \langle \bar{\epsilon}_1^2(\mathbf{x}, 1, \eta)\rangle + C(C-1)\langle (W_s/C)\,(W_t/C)\rangle [\langle \bar{\epsilon}_1(\mathbf{x}, 1, \eta)\rangle]^2 \\ &= (\langle W^2\rangle/C)\langle \bar{\epsilon}_1^2(\mathbf{x}, 1, \eta)\rangle + (C-1)\,C^{-1}[1 - (\langle W^2\rangle - 1)/(C-1)].\end{aligned}$$

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

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

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

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

$$\begin{aligned}\langle\bar\epsilon_1^2(\mathbf{x},1,\eta/\Gamma)\rangle &= C\langle(W/C)^2\rangle\langle\bar\epsilon_1^2(\mathbf{x},1,\eta)\rangle+2[\tfrac{1}{2}C(C-1)][\langle(W/C)\rangle\langle\epsilon_1\rangle]^2\\ &= (\langle W^2\rangle/C)\langle\bar\epsilon_1^2(\mathbf{x},1,\eta)\rangle+(C-1)/C.\end{aligned}$$

This is no longer an identity, but rather it establishes that the necessary and sufficient condition for $\lim_{\eta\to 0}\langle\bar\epsilon_1^2(\mathbf{x},1,\eta)\rangle<\infty$ is $\langle W^2\rangle/C<1$. Similarly,

$$\lim_{\eta\to 0}\langle\bar\epsilon_1^h(\mathbf{x},1,\eta)\rangle<\infty \quad \text{if and only if} \quad \langle W^h\rangle/C^{h-1}<1.$$

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

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

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

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

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

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

$$\begin{aligned}\langle\bar\epsilon^2(D,1,\eta/\Gamma)\rangle &= \Gamma\langle(W/C')^2\rangle\langle\bar\epsilon^2(D_s,1,\eta)\rangle+(\Gamma-1)\,\Gamma^{-1}\langle(W_sW_t)\rangle\\ &= (\langle W^2\rangle/\Gamma)\langle\bar\epsilon^2(D_s,1,\eta)\rangle+(\Gamma-1)\,\Gamma^{-1}[1-(\langle W^2\rangle-1)(C-1)^{-1}].\end{aligned}$$

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

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

has become $\langle W^2\rangle/\Gamma<1$. Similarly

$$\lim_{\eta\to 0}\langle\bar\epsilon^h(D,1,\eta)\rangle<\infty \quad \text{if and only if} \quad \langle W^h\rangle/\Gamma^{h-1}<1.$$

The canonical case. The recurrence relation is unchanged when the dimension changes from $i=3$ to $i=1$, except for the replacement of C by Γ. Therefore, we fall back on the condition $\langle W^h\rangle/\Gamma^{h-1}<1$ 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; e.g. for $h=2$, it is smaller by the factor $1-(\langle W^2\rangle-1)/(C-1)$.

4.2. *The determining functions*

In order to apply the above results to classify cascades, and in order to carry the theory further, it is convenient to consider the expression

$$f(h)=\log_C\langle W^h\rangle,$$

to be called the 'determining function'; more specifically, when D is i-dimensional, we shall need

$$\phi_i(h)=(3/i)f(h)-(h-1).$$

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

For all h, by a general theorem of probability theory, $f(h)$ is a convex function of h (see Feller 1971, p. 155), and so are all the ϕ_i. One has $f(1) = \phi_i(1) = 0$, and so $\phi_i(h)$ has at most one root other than 1; it will be designated α_i. The conditions $\phi_1(h) < 0$ and $\phi_2(h) < 0$ are both at least as demanding as $\phi_3(h) < 0$, so when $\alpha_1 > 1$, the α_i satisfy $\alpha_1 \leqslant \alpha_2 \leqslant \alpha_3$.

A further investigation of the ϕ_i involves their slopes for $h = 1$, more specifically the expressions

$$\Delta_i = -i\phi_i'(1) = -i\langle W \log_{C^{\frac{1}{3}i}}(W/C)^{\frac{1}{3}i})\rangle$$

$$= -3\langle W \log_C W\rangle + i.$$

Writing $\Delta_3 = \Delta$, we have $\Delta_2 = \Delta - 1$ and $\Delta_1 = \Delta - 2$. The value of Δ_i will be useful, because an i-dimensional average in a canonical cascade is degenerate when $\Delta_i < 0$ and non-degenerate when $\Delta_i > 0$; in particular, when $W < C^{\frac{1}{3}i}$, $\Delta_i > 0$. (The transition case $\Delta_1 = 0$ deserves the attention of the mathematicians, but is too complicated to be tackled in this paper.) More precisely, in the degenerate case $\Delta_1 < 0$, α_i satisfies $\alpha_i < 1$, and its value will play no special role. But in the non-degenerate case $\Delta_i > 0$, α_i satisfies $\alpha_i > 1$ and serves to determine whether the cascade is regular ($\alpha_i = \infty$) or irregular ($\alpha_i < \infty$).

We know that $\alpha_i = \infty$ is equivalent to $\langle W^h\rangle/(C^{\frac{1}{3}i})^{h-1} < 1$ for all $h > 1$, and so we know that the necessary and sufficient condition for $\alpha_i = \infty$ is $W < C^{\frac{1}{3}i}$, already featured in § 2.2.

When $\Delta_i > 0$, the quantity Δ_i also plays an independent role as the intrinsic dimension of the support of $\bar{\epsilon}$ within an i-dimensional domain D; see § 4.8. $W \log W$ is concave, and hence $\langle W \log W\rangle > \langle W\rangle \log \langle W\rangle = 0$; therefore, the intrinsic dimension Δ_i never exceeds the corresponding physical dimension i.

Finally, the second-order dependence properties of the dissipation, namely its correlation and its spectrum, depend on the value of $f(2)$. Indeed, the correlation between averages taken over domains D that are small in comparison with the distance d between them was shown by Yaglom to be proportional to d^{-Q}, with

$$Q = 3\log_C\langle W^2\rangle = 3f(2) = 3[\phi_3(2)+1].$$

To obtain a lower bound on Q, note that from $\langle W\rangle = 1$ it follows that $\langle W^2\rangle > 1$ and hence $Q > 0$; more precisely $Q > 3(1+\phi_3'(1)) > 3-\Delta$. As for the upper bounds, when $W < C$, we know that the maximum of $\langle W^2\rangle$ (constrained by $\langle W\rangle = 1$) 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, $W < C^{\frac{1}{3}i}$ implies $Q < i$. When $W/C^{\frac{1}{3}i}$ may exceed 1, on the contrary, it is possible that $Q > i$.

Studies involving correlations of higher order h similarly depend on values of f up to the argument h. Since we shall stop at the second order, our classification of canonical cascades will depend solely on the values of Q, Δ and α_i. These parameters are conceptually distinct, and since their numerical values are only related by the conditions of compatibility $Q > 3-D$ and $(\alpha_i - 1)\Delta_i > 0$, they can depend differently upon the specific model chosen. The question of whether or not their actual values are related might stimulate experimental investigations.

The relationship between W, $f(h)$ and the different parameters deserves amplification from the viewpoint of mathematical determination. A knowledge of C and of the distribution of W determines $f(h)$ for all h, and thus determines all the parameters. On the other hand, a knowledge of C and of $f(h)$ for h an integer need not determine W uniquely. A sufficient condition is that the moments satisfy the Carleman criterion (see § A 2). This technicality is important because this criterion fails in the case of a lognormal W.

4.3. *Examples of determining functions*

Rectilinear determining functions. The functions f and ϕ_i are linear functions of h, if W is binomial with $\Pr(W = 1/p) = p$ and $\Pr(W = 0) = 1 - 1/p$. One has $\langle W^h \rangle = pp^{-h} = p^{1-h}$, so that $\phi_3(h) = (1-h)\log_C(pC)$, which is a degenerate form of convex function.

In a somewhat digressive but brief paragraph, we shall simplify this example, and through it the canonical Novikov–Stewart (1964) theory; because it can be discussed fully with the help of the classical probabilistic theory of birth-and-death processes (Harris 1963). After K stages, each elementary subeddy either is empty or includes a non-random mass of turbulence equal to p^{-K}. Save for this factor, the mass of turbulence in an eddy and the number of its non-empty elementary subeddies are equal. Their distribution is readily determined, because every time K increases by unity, each non-empty elementary eddy can be viewed as having acquired a random 'offspring' made of M lower order elementary eddies, with M following a binomial distribution of expectation $C^{\frac{1}{3}i}p$. 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 alternative. When $C^{\frac{1}{3}i}p \leqslant 1$, i.e. $\Delta_i \leqslant 0$, it is almost certain that the offspring will eventually die out. When $p > 1/C^{\frac{1}{3}i}$ and $\Delta_i > 0$, on the contrary, the number of offspring, if normalized by being divided by $(C^{\frac{1}{3}i}p)^K = C^{\frac{1}{3}\Delta_i K}$, tends asymptotically towards a non-degenerate limiting random variable that has finite hth moments for all h.

Asymptotically rectilinear determining functions. Now let us only suppose that W is bounded. Designate its greatest attainable value by $\max W$. This means that $\Pr(W > \max W) = 0$ and $\Pr(W > \max W - \theta) > 0$ for all $\theta > 0$. (A more correct mathematical idiom for $\max W$ is 'almost sure supremum'.) It follows that $\lim_{h\to\infty}(\log_C \langle W^h \rangle / h) = \log_C \max W$, which implies that $f(h)$ has an asymptotic direction of finite slope $\log_C \max W$. Conversely, for this asymptotic slope to be finite, the condition $W < \max W < \infty$ can be shown to be necessary. Also, $\phi_i(h)$ has an asymptotic direction of slope $-1 + \log\max W/\log C^{\frac{1}{3}i}$. When $\max W < C^{\frac{1}{3}i}$, this slope is negative and $\phi_i(h) = 0$ has no root other than 1 (i.e. $\alpha_i = \infty$). When $\max W > C^{\frac{1}{3}i}$, and particularly when $\max W = \infty$, $\alpha_i < \infty$. Our assertion that, in general (except for some inequalities), the values of Δ, Q and the α_i are independent, is in this case very apparent. See also the caption of figure 2.

Parabolic determining functions. Suppose that W is lognormal, namely $\log W$ is

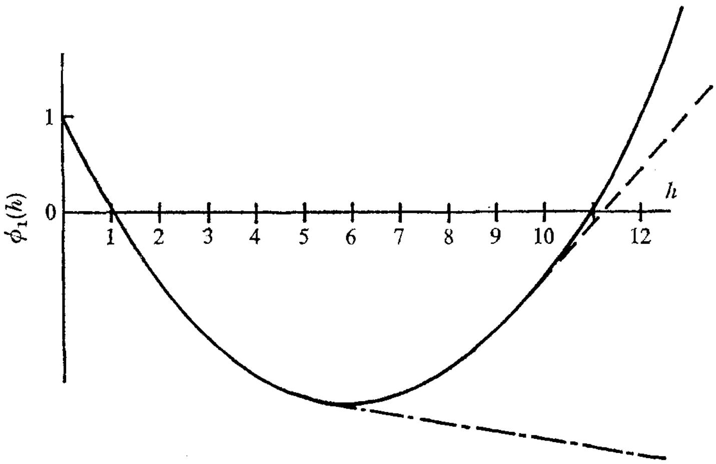

FIGURE 2. Characterization of the distribution of $\bar{\epsilon}(D)$ when, respectively, (*a*) log W is normal; $\phi_1(h)$ is then a parabola and $\phi_1(h) = 0$ has two finite roots (solid line); (*b*) log W is a sum of sufficiently many terms to be a good approximation to the normal distribution; $\phi_1(h)$ is then nearly parabolic for $h < \alpha_1$ (dashed line); (*c*) log W is a 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 $\bar{\epsilon}$; in the zone of interest, $\phi_1(h)$ is far from parabolic and $\phi_1(h) = 0$ may have a single finite root, i.e. $\alpha_1 = \infty$ (dash–dot line). Thus the degrees of sensitivity of various properties of $\bar{\epsilon}$ are very different. On the one hand, the moment properties of the distribution of $\bar{\epsilon}$ depend greatly upon fine details, namely the tail of the distribution of W: a lognormal W never falls in the regular class, but a 'nearly lognormal' W may do so. On the other hand, the value of $\phi_1(2)$, hence of the spectral properties of $\bar{\epsilon}$, and even more the value of $\phi_1'(1)$, hence of the fractional dimension, will be essentially the same for the three cases as drawn.

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 $\langle W^h \rangle = \exp(h \langle \log W \rangle + h^2 2^{-1} \sigma^2 \log W)$, that is,

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

This $f(h)$ is represented by a parabola, and so are the $\phi_i(h)$.† To ensure that $\langle W \rangle = 1$, we must have $\langle \log W \rangle = -2^{-1}\sigma^2 \log W$, a quantity to be denoted (in order to fit Kolmogorov's notation) by $-\frac{1}{6}\mu \log C$. It follows that

$$f(h) = \tfrac{1}{6}(h-1)\, h\mu.$$

Hence, $\Delta_i = -3f'(1) + i = i - \frac{1}{2}\mu$, $\alpha_i = 2i/\mu$ and $Q = \mu$. Here, the values of Δ, Q and the α are all functions of μ, and are strongly interdependent; this is an exceptional circumstance (see §4.9).

The determining function when log W is a sum of many uniform random variables (figure 2). Suppose that log W is bounded, but near-Gaussian according to the customary definition of 'nearness', for example, that log W is a sum of many uniform random variables. Then, for low enough values of h, $\phi_i(h)$ will nearly coincide with the parabola in the preceding paragraph, but asymptotically it will be a straight line. Roughly, $\phi_i(h)$ will look like a portion of parabola con-

† Curiously, the converse is not true; since the lognormal distribution is not determined by its moments, see §2, other weights W may lead to the same $f(h)$.

tinued by a straight tangent. If we have added many uniform components, and if μ is large, this tangent will have positive slope and the two values of α, corresponding to the Gaussian log W and to its approximation, will be about the same. If, on the other hand, we add few uniform random variables, and μ is small, the tangent will have negative slope and the lognormal approximation will be entirely worthless. As an illustration, one of our computer simulations of a cascade, in which W was ostensibly lognormal and $\alpha_i < \infty$, yielded results that were completely at variance with the expectations, and remained such until it was recalled that in fact our random variable log W was not Gaussian, but rather was given by an algorithm that added 12 uniformly distributed random variables. When the same program was run again, with sums of 48 and then of 192 such random variables, the results were changed to conform with the expectations. See also §§ 4.3 and 4.9.

4.4. *Regular classes*

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

The regular class for fixed i. Here $\bar{\epsilon}(D_i, L, 0)$ is, by definition, a non-degenerate random variable with all moments finite. In a canonical cascade, the necessary and sufficient condition for $\lim_{\eta \to 0} \langle \bar{\epsilon}_1^h(\mathbf{x}, L, \eta)\rangle < \infty$ is that $\phi(h) < 0$ for all $h > 1$, i.e. $\alpha_i = \infty$, or, alternatively, $W < C^{\frac{1}{3}i}$. As a corollary, $Q < 3$. A formal argument, replacing the limits of the moments with the moments of the limit, suggests that $\langle \bar{\epsilon}_1^h(\mathbf{x}, L, 0)\rangle < \infty$ if and only if $\phi_i(k) < 0$ for all $h > 1$.

To justify this formal argument, it suffices to use a theorem of Doob (1953, p. 319). In physics, such technicalities are ordinarily disregarded, but § 4.5 shows the issue to be significant in the degenerate case, which warns us to take the matter seriously.

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

As a specific example, if W is binomial with

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

$\phi_i(h) < 0$ if and only if $pC^{\frac{1}{3}i} > 1$, i.e. $p > 1/C^{\frac{1}{3}i}$.

The uniformly regular class. When $\alpha_1 = \infty$, alternatively when $\Delta_1 > 0$, then, for all i, $\bar{\epsilon}(D, L, 0)$ is a non-degenerate random variable with finite moments. In this class, $\epsilon(D, L, \eta)$ may be said to be 'approximately lognormal'. Yaglom has implicitely assumed that this situation prevails in practice. This may, but need not, be so. Only experiment may tell.

4.5. *Degenerate classes*

The degenerate class for fixed i. This is defined by $\bar{\epsilon}(D, L, 0)$ being zero almost surely. A sufficient condition is $\Delta < 0$ (from which it follows that $\alpha \leqslant 1$).

The proof (see § A 3) consists of showing that the number of elementary sub-eddies of side r contributing to the bulk of $\bar{\epsilon}_r(\mathbf{x}, L, 0)$ is roughly equal to $(L/r)^{\Delta_i}$. From $\Delta_i < 0$ it follows that, as $\eta \to 0$, this number tends to zero, and so does $\bar{\epsilon}(D, L, \eta)$.

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.* Since $\Delta_i = i - \frac{1}{2}\mu$, the cascade is degenerate whenever $\mu > 2i$. In particular, $\bar{\epsilon}_r(\mathbf{x}, L, \eta)$ is lognormal only when $r = \eta$.

The uniformly degenerate class. When $\bar{\epsilon}(D_i, L, 0)$ is degenerate for $i = 3$, i.e. when $\Delta < 0$, $\bar{\epsilon}$ is also degenerate for $i = 2$ and $i = 1$.

4.6. *Irregular classes*

The irregular class for fixed i. This is defined by $\epsilon(D_i, L, 0)$ being non-degenerate, with $\langle \bar{\epsilon}^h(D_i, L, 0) \rangle < \infty$ for small enough $h > 1$, but $\langle \bar{\epsilon}^h(D_i, L, 0) \rangle = \infty$ for large finite h. The class is characterized by $1 < \alpha_i < \infty$, and the cut-off between finite and infinite moments is $h = \alpha_i$; see Mandelbrot (1973) and Kahane (1973).

The uniformly irregular class. When $\bar{\epsilon}(D_i, L, 0)$ is non-degenerate for $i = 3$ and non-regular for $i = 1$, it is irregular for all i.

4.7. *Mixed classes*

Since $\alpha_1 < \alpha_2 < \alpha_3$, the classes to which a cascade ($i = 3$) and its cross-sections ($i = 1$ and $i = 2$) belong may be different. Neglecting the behaviour for $i = 2$, 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 $C^{-1} < p < C^{-\frac{1}{3}}$. Here $\phi_3(h) < 0$ but $\phi_1(h) > 0$ for all $h > 1$. That is, the cascade is regular for $i = 3$ but degenerate for $i = 1$.

Comments. I doubt whether this mixed class ever is encountered in practice, because it suggests that the spatial distribution of dissipation is sparser than I think likely.

An example from the mixed regular–irregular class. $C = 27$, so that $C^{\frac{1}{3}} = \Gamma = 3$, and the random variable W is either equal to 3·7, with probability 0·1, or equal to 0·7, with probability 0·9. Since $W < C$, a three-dimensional cascade with this W is regular and may correspond to a canonical approximation to a micro-canonical cascade. On the other hand, it is not true that $W < \Gamma$, while it is true that $\phi_1'(1) < 0$, and as a result a one-dimensional cascade with 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 lognormal, even approximately. *One task for the experimental study of turbulence should be this: to check whether or not such a mixture ever occurs. Also, if under different circumstances one observes either this mixture or the uniformly regular class, it is a task to classify such circumstances according to the class 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: Δ_i as a fractional intrinsic dimension*

Select two arbitrary small thresholds ω and γ. It can be shown (see § A 3) that when $\Delta_i > 0$ the eddies of side r can be divided into two groups. The eddies of the first group contain a proportion above $1-p$ of the whole dissipation, but their number lies between $(L/r)^{\Delta_i-\gamma}$ and $(L/r)^{\Delta_i+\gamma}$, which makes the group relatively very small. Thus, almost all eddies belong to the second group; but, taken together, they include a proportion of the total dissipation at most equal to ω, which makes the dissipation negligible.

It is convenient to call Δ_i an intrinsic dimension, alternatively (because it need not be an integer), a fractional dimension.

The notion that a geometric figure can have a fractional dimension was conceived in 1919 by a pure mathematician, Hausdorff. It is closely related to the Cantor set, and both have the reputation of lacking any conceivable application, in fact of 'turning off' any natural scientist. I believe that this reputation is no longer deserved and hope to show (elsewhere) that in fact fractional dimension is something very concrete and that different aspects of it are useful measurable physical characteristics. Examples are the degree of the wiggliness of coastlines (Mandelbrot 1967), the degree of clustering of galaxies and the intensity of the intermittency of turbulence. See also Mandelbrot (1974).

In these applications, it is best to use a semi-formal variant called the 'self-similarity dimension', which is of more limited validity than Hausdorff's concept, but incomparably simpler. It derives from elementary features of the usual concept of dimension for segments of a straight line for rectangles and for parallelepipeds. A line has dimension $\Delta = 1$, and for every positive integer N, the segment where $0 \leqslant x < X$ can be exactly decomposed into N non-overlapping segments of the form $(n-1)\,X/N \leqslant x < nX/N$, where n runs from 1 to N. Each of these parts is deducible from the whole by a similarity of ratio $\rho(N) = N^{-1}$. Similarly a plane has dimension $\Delta = 2$ and for every integer $\sqrt{N}$, the rectangle where $0 \leqslant x < X$ and $0 \leqslant y < Y$ can be decomposed exactly into N non-overlapping rectangles of the form

$$(k-1)\,N/\sqrt{N} \leqslant x < kX/\sqrt{N} \quad \text{and} \quad (h-1)\,Y/\sqrt{N} \leqslant y < hY/\sqrt{N},$$

where k and h run from 1 to N. Each of these parts is deducible from the whole by a similarity of ratio $\rho(N) = 1/N^{\frac{1}{2}}$. More generally, a Δ-dimensional rectangular parallelepiped can for every integer $N^{1/\Delta}$ be decomposed into N parallelepipeds deducible from the whole by a similarity of ratio $\rho(N) = 1/N^{1/\Delta}$.

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

$$\Delta = -\log N/\log \rho(N),$$

and this is the property that suggests a generalization of the concept of dimension to the set on which the bulk of intermittent turbulence is concentrated. Here $1/\rho(N) = L/r$ and for every $\gamma > 0$, $(L/r)^{\Delta_i-\gamma} < N < (L/r)^{\Delta_i+\gamma}$, i.e.

$$\Delta_i - \gamma < \log N/\log \rho(N) < \Delta_i + \gamma.$$

In other words the dimension is Δ_i. Using the inequality $\Delta_i < i$, the intuitive notion that turbulence concentrates on an extremely sparse set is expressed numerically. By choosing W appropriately, the dimension Δ_i can take any value between 0 and i. Note also that, when $\Delta < 3$, $\frac{1}{3}\Delta_3 = \frac{1}{3}\Delta$ is greater than

$$\tfrac{1}{2}\Delta_2 = \tfrac{1}{2}(\Delta - 1),$$

which in turn is greater than $\Delta_i = (\Delta - 2)$. The inequality $\Delta < 3$ expresses that a figure does not fill the space dimensionally, and $\frac{1}{3}\Delta > \Delta - 2$ 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 f(h)*

Suppose that W is non-lognormal and bounded with $\sigma^2 \log W < \infty$, and let W^* be its lognormal approximation, the corresponding determining functions being $f(h)$ and $f^*(h)$, with the obvious definitions for $\phi_i^*(h)$. We have already noted that, since $f(h)$ has (for $h \to \infty$) a finite asymptotic slope $\log_C \max W$, while $f^*(h)$ is parabolic, their asymptotic behaviours differ qualitatively. On the other hand, the behaviour of $f(h)$ and the $\phi_i(h)$ for small h only depends on $\langle \log W \rangle$ and $\sigma^2 \log W$, and therefore remains unchanged when $\log W$ is replaced by its normal approximation. Hence the following consequences.

The moment of order $h = \frac{2}{3}$ is likely to be covered by this approximation. Therefore, the conclusions of Kolmogorov and Yaglom, obtained by applying the '$\frac{2}{3}$-law', may well be essentially unchanged.

For $h = 1, f^*(1)$ need not equal 1. Also, $\phi_i'(1)$ need not equal $\phi_i^{*\prime}(1)$. In extreme instances, like the one reported at the end of § 4.2, they may have different signs. The 'real' $\phi_i'(1)$ may be negative, meaning that the cascade is non-degenerate, while $\phi_i^{*\prime}(1) > 0$ suggests the cascade is degenerate. The approximations of the values of Q and the α_i may be even poorer.

A different lognormal approximation to W, to be called W^{**}, is achieved by approximating $f(h)$ by a parabola $f^{**}(h)$ passing through $f(0) = f(1) = 0$ and having the correct slope $f'(1)$. The mean and variance of $\log W^{**}$ are determined by the properties of this $f^{**}(h)$ near $h = 0$. Since, from the practical viewpoint, all that is of direct interest is the portion of $\phi_i(h)$ that lies between $h = 1$ and $h = \alpha_i$, we see that, when $f(h)$ is smooth and the original α_i is small, the error introduced by the lognormal approximation W^{**} may well be acceptable. Whenever such is the case, the various properties of $\bar{\epsilon}(D, L, 0)$ linked to Δ_i, Q and α_i turn out after all to be related. That is, given the inaccuracy inherent in experimental work, one may be brought back to the situation prevailing when the single characteristic parameter Q was thought sufficient. When on the contrary the original α_i is large, and especially when it is infinite, the error in using W^{**} is very big, meaning that the process of approximation changes significantly the class to which such a cascade belongs.†

The author has greatly benefited 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 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 him 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 being published in the following references: *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), *Proceedings of the Polytechnic Institute of Brooklyn Symposium on Turbulence of Fluids and Plasmas*, New York, 1968 (John Wiley, 1969).

Appendix

A 1. *Differences between approximate and strict lognormality are deep, and in particular the use of approximate lognormality to calculate moments is unsafe*

Let us consider a normal 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 $B_k = \log R_k$ bounded by $\beta < \infty$. Since G, P and B will be assumed to be nearly identical, and since the mean and the variance are equal in the case of P, they must be assumed equal for G and B also, and near identity also requires their common value δ to be large. It follows that

$$\langle (e^G)^h \rangle = \exp(h\delta + \tfrac{1}{2}\delta h^2) = \exp[\delta(h + \tfrac{1}{2}h^2)],$$
$$\langle (e^P)^h \rangle = \exp(-\delta + \delta e^h) = \exp[\delta(e^h - 1)],$$
$$\langle (e^B)^h \rangle \leqslant \exp(hK\beta).$$

Thus, $\langle (e^B)^h \rangle$ increases at most exponentially with h, $\langle (e^G)^h \rangle$ increases more rapidly than any exponential, and $\langle (e^P)^h \rangle$ more rapidly still. The expectations, equal respectively to $\langle (e^G) \rangle = \exp(1{\cdot}5\delta)$ and $\langle (e^P) \rangle = \exp(1{\cdot}7\delta)$, are already very different. The coefficients of variation

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

differ even more, and higher order moments differ strikingly. In short, although B and P are nearly normal from the usual viewpoint (which is that of the 'weak

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

topology'), from the present viewpoint they are poor approximations to the normal distribution. However, when $h < 1$, i.e. when $h = \frac{2}{3}$ as in the calculation of spectra, the discrepancy is less great.

A 2. *On Orszag's remark concerning the determination of turbulence by its moments*

Homogeneous turbulence is presumed to be determined by its moments, to the derivation of which the bulk of the theory based on the Navier–Stokes equations is devoted. Is intermittent turbulence also so determined? To answer, note that it is classical that a random variable can have the same moments as the lognormal distribution, without itself being lognormal ('classical' in probability theory means 'has been recorded by Feller (1971)'; this particular example is in vol. 2 (2nd edn.), p. 227, where it is credited to C. C. Heyde). The reason for this indeterminacy is that the moments of lognormal e^G increase so fast that $\Sigma[\langle \exp(2hG)\rangle]^{-\frac{1}{2}h} < \infty$, meaning that the lognormal distribution fails to satisfy a criterion 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, let Yaglom's $\bar{\epsilon}_r$ be examined not through its lognormal approximation, but directly, as a product of independent factors $R_k < \beta$. If those factors are bounded ($R_k < \beta$), then it follows that the moments of intermittent turbulence satisfy the Carleman criterion and therefore the indeterminacy noted by Orszag vanishes.†

A 3. *The exponent of dimension 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 $(L/r)^{\Delta_i}$ subeddies.

Preliminary example: binomial weights. Let

$$\Pr(W = 0) = 1-p \quad \text{and} \quad \Pr(W = p^{-1}) = p,$$

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

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

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

Second example: lognormal weights and cubic eddies D. Let us begin with the low frequency factor $\bar{\epsilon}_r(\mathbf{x}, L, r)$. With W lognormal as in §4.3, log $\bar{\epsilon}_r$ is Gaussian with variance $\mu \log(L/r)$ and expectation $-\frac{1}{2}\mu \log(L/r)$. To simplify the notation, we shall denote $\bar{\epsilon}_r$ by V. When $L/r \gg 1$, this lognormal factor has the feature that its expectation is overwhelmingly due to occasional large

† *Note added during revision.* Novikov (1971, p. 235) has made a remark to the same effect.

values, meaning that 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 V^* by

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

then EV^* is arbitrarily close to 1. It is claimed that such a result is achieved when N is a function (otherwise arbitrary) such that $\lim_{r\to 0} N(L/r)/[\log(L/r)]^{\frac{1}{2}} = 0$, and when 'threshold' is chosen to satisfy

$$\text{threshold } (L/r) = (L/r)^{\frac{1}{2}\mu} \exp\{-N(L/r)\,[\mu\log(L/r)]^{\frac{1}{2}}\}.$$

Indeed,
$$\langle V\rangle = \frac{1}{[2r\mu\log(L/r)]^{\frac{1}{2}}}\int \exp\left\{x - \frac{[x+\frac{1}{2}\mu\log(L/r)]^2}{4\mu\log(L/r)}\right\}dx,$$

with an integration range from $\log[\text{threshold}\,(r, L)]$ to infinity. The expression in braces transforms into

$$-[x-\tfrac{1}{2}\mu\log(L/r)]^2/4\mu\log(L/r)$$

and by changing the variable of integration to

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

we obtain
$$\langle V^*\rangle = (2\pi)^{-\frac{1}{2}}\int \exp(-\tfrac{1}{2}z^2)\,dz,$$

with an integration range from $-N(L/r)$ to infinity. As $L/r \to \infty$, $\langle V\rangle \to 1$, which shows that the contribution of other values of V to $\bar{\epsilon}_r$ is asymptotically negligible, and that the above choice of N has been appropriate to make V arbitrarily closely approximated by 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

$$[\log V + \tfrac{1}{2}\mu\log(L/r)]\,[\mu\log(L/r)]^{\frac{1}{2}} = G,$$

the above probability becomes

$$\Pr\{G > [\mu\log(L/r)]^{\frac{1}{2}} - N(L/r)\}.$$

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

$$(L/r)^3\frac{\exp[-\frac{1}{2}\mu\log(L/r)]}{[2\pi\mu\log(L/r)]^{\frac{1}{2}}} = \frac{(L/r)^{3-\frac{1}{2}\mu}}{[2\pi\mu\log(L/r)]^{\frac{1}{2}}} = \frac{(L/r)^{\Delta}}{[2\pi\mu\log(L/r)]^{\frac{1}{2}}}.$$

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

With cubic eddies replaced by straight segments, the only change is that in the above formulae the factor $(L/r)^3$ is replaced by L/r and hence $3-\frac{1}{2}\mu$ by $1-\frac{1}{2}\mu = \Delta_1$.

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 $\bar{\epsilon}_r(\mathbf{x}, r, \eta)$ 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 $(L/r)^{\Delta}$. 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 (1965).

A 4. *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 (Mandelbrot 1972) involved a departure from the assumption of § 2.1: the grid was itself made random, being generated by the same model as the distribution of dissipation. The purpose of this section of the appendix is to serve to readers of this paper as an introduction to my earlier work.

As a preliminary, let us consider a prescribed grid of eddies, and a canonical cascade with lognormal weight W: $\log \epsilon(\mathbf{x}, L, \eta)$ is Gaussian with variance $\mu \log (L/\eta)$ and expectation $-\frac{1}{2}\mu \log (L/\eta)$. Also, it was shown by Yaglom that the correlation of $\bar{\epsilon}$ is approximately proportional to d^{-Q}, with $Q = \mu$. Because the eddies were prescribed, the random function $\epsilon(\mathbf{x}, L, \eta)$ is non-stationary and discontinuous: it varies between an eddy and its neighbours, by jumps that may be very large. Both non-stationarity and discontinuity are of course quite unrealistic. One may instead demand that $\log \epsilon(\mathbf{x}, L, \eta)$ 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 $\bar{\epsilon}(\mathbf{x}, L, \eta)$ is nearly identical to $\epsilon(\mathbf{x}, L, \eta)$. It remains to ensure that $\epsilon(\mathbf{x}, L, \eta)$ has the d^{-Q} correlation. The simplest way is to require $\epsilon(\mathbf{x}, L, \eta)$ to have a truncated self-similar spectral density, namely a spectral density equal to $\eta/2\omega$ when the frequency ω satisfies $1/L < \omega < 1/\eta$ and equal to zero elsewhere. The resulting model may be viewed, alternatively, as combining self-similarity with a portion of the Kolmogorov third hypothesis, seemingly the maximum retrievable portion.

The properties of $\bar{\epsilon}(D, L, \eta)$ relative to this model can be summarized as follows. The dimensions continue to be $\Delta_i = i - \frac{1}{2}\mu$ and the cascades are never regular: for $\mu < 2/i$, they are irregular with $\alpha_i = 2/\mu$, while for $\mu > 2/i$, they are degenerate. Compared with a canonical cascade with a lognormal W, the main differences involve the values of certain numerical constants.

A 5. *Remarks on Kolmogorov's third hypothesis of lognormality*

This hypothesis can be viewed as splitting into the family of assumptions that, for every cube of centre $\mathbf{x}$ and side $r > \eta$, $\bar{\epsilon}_r(\mathbf{x}, L, \eta)$ follows the lognormal distribution, $\log \bar{\epsilon}_r$ having a variance equal to $\mu \log (L/r)$. First, it will be shown that within Yaglom's context of prescribed eddies, either canonical or microcanonical, Kolmogorov's third hypothesis cannot hold. Then, in a wider context, it will be shown to be tenable only under unlikely additional conditions.

In the context of prescribed microcanonical eddies, the difficulty is the following. To say that $\log \bar{\epsilon}_r$ 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 (Levy–Cramer) 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 to any fixed value of r, not even for $r = \eta$.

In the context of prescribed canonical eddies, the difficulty is different. One can show that, in a canonical context, the correlation of every pair of ϵ_r is positive.† On the other hand, we shall see that Kolmogorov's set of hypotheses implies that at least some of those correlations are negative. Thus, the portion of Kolmogorov's set of hypotheses concerning $r = \eta$ might conceivably hold, but the different portions corresponding to several values of r are incompatible, meaning the hypotheses are internally inconsistent.

More generally, the joint assumptions that the random variables $\log \bar{\epsilon}_r(\mathbf{x}, L, \eta)$ are normal for every r, with $\sigma^2 \log \bar{\epsilon}_r = \mu \log (L/r)$ and $\langle \log \bar{\epsilon}_r \rangle = -\frac{1}{2}\mu \log (L/r)$, are incompatible with any model that leads to a positive reduced covariance for $\bar{\epsilon}_\eta$. Indeed it would follow from the assumptions that $r^3\bar{\epsilon}_r$, the mass of turbulence in a cube, satisfies

$$\langle [r^3\bar{\epsilon}_r(\mathbf{x}, L, \eta)]^h \rangle = r^{3h-\frac{1}{2}h(h-1)\mu} L^{\frac{1}{2}h(h-1)\mu}.$$

When r reaches its maximum value, which is $r = L$, then for all h the above moment will reduce to r^{3h}, as it should. But the nature of convergence to this limit must be examined more closely, by subdividing our cube into (say) 2^3 portions. When $\frac{1}{3}h > 2\mu$, the exponent of r in the above expression is negative, and as a result the value of the ratio $\langle (r^3\bar{\epsilon}_r)^h \rangle / \langle [(\frac{1}{2}r)^3 \bar{\epsilon}_r]^h \rangle$ 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 $\bar{\epsilon}_\eta(\mathbf{x}, L, \eta)$.‡

This is the moment to point out that in my limiting lognormal variant of Kolmogorov's model (see § A 4) the above inconsistency is avoided because every formerly misbehaving moment turns out to be infinite.

REFERENCES

Billingsley, P. 1965 *Ergodic Theory and Information*. Wiley.

de Wijs, H. J. 1951 Statistics of ore distribution. *Geol. en Mijnbouw* (Netherlands), **13**, 365–375.

de Wijs, H. J. 1953 Statistics of ore distribution. *Geol. en Mijnbouw* (Netherlands), **15**, 12–24.

Doob, J. L. 1953 *Stochastic Processes*. Wiley.

Feller, W. 1971 *An Introduction to the Theory of Probability and its Applications*. Wiley.

† Yaglom's rough derivation of the correlation suggests that it 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 §3.6.

‡ *Note added during revision.* This inconsistency enters into the discussion of Novikov (1971, p. 236) of the contradictory behaviour of the quantities he designates as μ_ρ. However, having noted the contradiction, Novikov did not resolve it.

HARRIS, T. 1963 *Branching Processes*. Springer.

KAHANE, J. P. 1973 Sur le modèle de turbulence de Benoit Mandelbrot, *Comptes Rendus*, in press.

KOLMOGOROV, A. N. 1962 A refinement of previous hypotheses concerning the local structure of turbulence in a viscous incompressible fluid at high Reynolds number. *J. Fluid Mech.* **13**, 82–85. (See also 1962 *Mécanique de la Turbulence*, pp. 447–458 (in French and Russian).)

MANDELBROT, B. 1967 How long is the coast of Britain? Statistical self-similarity and fractional dimension. *Science*, **156**, 636–638.

MANDELBROT, B. 1972 Possible refinement of the lognormal hypothesis concerning the distribution of energy dissipation in intermittent turbulence. In *Statistical Models and Turbulence* (ed. Rosenblatt & Van Atta), pp. 333–351. Springer.

MANDELBROT, B. 1973 Multiplications aléatoires iterées, et distributions invariantes par moyenne pondérée. *Comptes Rendus*, in press.

MANDELBROT, B. 1974 Physical objects with fractional dimension: seacoasts, galaxy clusters, turbulence and soap. *IBM Res. Ext.* Report. (Submitted to *Bull. Inst. Math. Applics.*)

NOVIKOV, E. A. 1969 Scale similarity for random fields. *Dokl. Akad. Nauk SSSR*, **184**, 1072–1075. (English trans. *Sov. Phys. Dokl.* **14**, 104–107.)

NOVIKOV, E. A. 1971 Intermittency and scale similarity in the structure of a turbulent flow. *Prikl. Mat. Mekh.* **35**, 266–277. (See also *P.M.M. Appl. Math. Mech.* **35**, 231–241.)

NOVIKOV, E. A. & STEWART, R. W. 1964 Intermittency of turbulence and the spectrum of fluctuations of energy dissipation. *Isv. Akad. Nauk SSSR, Seria Geofiz.* **3**, 408.

OBOUKHOV, A. M. 1962 Some specific features of atmospheric turbulence. *J. Fluid Mech.* **13**, 77–81.

ORSZAG, S. A. 1970 Indeterminacy of the moment problem for intermittent turbulence. *Phys. Fluids*, **13**, 2211–2212.

YAGLOM, A. M. 1966 The influence of the fluctuation in energy dissipation on the shape of turbulent characteristics in the inertial interval. *Dokl. Akad. Nauk SSSR*, **166**, 49–52. (English trans. *Sov. Phys. Dokl.* **2**, 26–29.)