Information theory and the spectrum of isotropic turbulence

1982 J. Phys. A: Math. Gen. 15 2285

(http://iopscience.iop.org/0305-4470/15/7/038)

J. Phys. A: Math. Gen. **15** (1982) 2285–2306. Printed in Great Britain

Information theory and the spectrum of isotropic turbulence

Timothy M Brown

High Altitude Observatory, National Center for Atmospheric Research†, High Altitude Observatory, PO Box 3000, Boulder, Colorado 80307, USA

Received 25 September 1981

Abstract. The problem of determining the spectrum of isotropic turbulence can be thought of as one of finding the most appropriate joint probability distribution for the flow taken as a whole. From the point of view of information theory, what one means by the most appropriate distribution is clearly defined and easily justified; it is the probability distribution that maximises the information theory entropy, subject to whatever constraints one can impose on the flow. In this work, the relevant constraints are taken to be the Reynolds number and energy dissipation rate of the flow, energy balance (on average) at every point in wavenumber space, and adherence to the Navier–Stokes equations. Using these constraints, it is shown that the maximum entropy formalism leads to a pair of coupled equations describing the distribution of energy in the turbulent spectrum, and the correlations between the amplitudes of velocity components with nearly identical wavenumbers. Although solutions to these equations are not presented, it develops that if a power-law solution exists, it can only be the Kolmogorov law $E(k) \propto k^{-5/3}$. In arriving at this result, a useful concept is that of the 'turbulent temperature', defined as the reciprocal of the derivative of the entropy with respect to the local energy dissipation rate. This quantity plays a role directly analogous to the thermodynamic temperature, governing the rate of energy exchange between different wavenumbers. It is found that, within the spectrum's inertial subrange, the turbulent temperature is virtually constant, with only a minute temperature gradient required to drive the energy cascade.

1. Introduction

It is well known that describing turbulent flows in terms of average correlation functions leads to an infinite set of coupled equations. Deriving any useful information about such flows therefore involves some (more or less *ad hoc*) means of closing the equation set so that only a finite amount of labour is required to get a solution. The merit of any such closure method then depends on its agreement with the experimental facts, and on the plausibility of the arguments justifying its adoption.

One closure method that has received little attention is to suppose that the flow behaves in such a way that a suitably defined entropy is maximised, subject to whatever constraints are imposed by boundary conditions and the equations of motion. Edwards and McComb (1969) discussed this method and showed that it leads to reasonable turbulence spectra, but their treatment has not been favoured because of its extreme complexity, and because its physical interpretation is obscure. In this paper, I will

† The National Center for Atmospheric Research is sponsored by the National Science Foundation.

0305-4470/82/072285+22$02.00

consider a similar approach, with the entropy defined in terms suggested by information theory. My aim is to clarify both the logical justification and the physical implications of this approach, and to show that it leads naturally to a Kolmogorov energy spectrum in the case of homogeneous, isotropic, turbulence in an incompressible fluid. Although this theory leads to equations that can in principle be solved for the exact turbulent spectrum, I will not do this here. Rather, the emphasis throughout will be on physically acceptable approximations, chosen to help illuminate the processes at work. It should be noted, however, that this approach differs from that taken by Edwards and McComb (1969), in that the current work ascribes central importance to the presence of small but non-zero viscous dissipation at all wavenumbers.

Since the arguments given here are drawn from two disparate areas of physics, a short preview is in order to help keep things straight. The next section contains a brief summary of the relevant equations for turbulent flow, and of the fundamental ideas about probability and information that will be required later on. Two related subjects will merit particular attention: the difference between traditional approaches to solving the turbulence equations and the current approach, and the logical justification for invoking a maximum entropy principle. Section 3 contains a naive application of the results of § 2 to the problem of finding a turbulent spectrum. Although fundamentally inconsistent, this example provides a useful reference point for later results. Section 4 is a short detour into probability and information theory. Its purpose is to establish a result that is crucial for later calculations of the correlation between velocity amplitudes at different wavenumbers in a turbulent spectrum. In § 5, a relation is derived giving the energy flow into a given wavenumber range in the turbulent spectrum from any other wavenumber range. This result leads to the aforementioned equations for the exact spectrum. These equations are, however, immediately dropped in favour of an approximation (the 'constant-temperature' approximation) that is both very accurate and much more tractable. Section 6 shows that by combining the constant-temperature approximation with the requirement that the turbulent energy should go smoothly to zero at large wavenumbers, one obtains the Kolmogorov spectrum $E(k) \propto k^{-5/3}$. Section 7 contains a validation of the constant-temperature approximation, estimates of global entropy relationships, and similar loose ends. Finally, § 8 is a discussion of the principal ideas already developed and their consequences, with analogies from other areas of physics.

2. The connection between turbulence and information theory

The Navier–Stokes equations provide the starting point for any discussion of turbulence. For an incompressible fluid, these may be written

$$\partial \boldsymbol{v}(\boldsymbol{x}, t)/\partial t + \boldsymbol{v}(\boldsymbol{x}, t) \cdot \nabla \boldsymbol{v}(\boldsymbol{x}, t) = -\nabla p(\boldsymbol{x}, t) + \nu \nabla^2 \boldsymbol{v}(\boldsymbol{x}, t), \tag{1}$$

$$\nabla \cdot \boldsymbol{v}(\boldsymbol{x}, t) = 0. \tag{2}$$

For the sake of definiteness, suppose that the fluid under consideration is contained within a cube of edge length L, and that the boundary conditions are periodic, i.e. that the flow pattern within the cube is repeated infinitely often outside it. Also suppose that the fluid within the cube is being stirred in a random fashion, with a characteristic stirring scale equal to the edge length L, and with a characteristic stirring speed u. In this case, the flow can be characterised by its Reynolds number (the ratio

of inertial to viscous forces) and by the rate of energy input per unit mass:

$$\mathrm{Re} = uL/\nu, \tag{3}$$

$$G = u^3/L. \tag{4}$$

In turbulent flows, the Reynolds numbers are commonly very large, so that the large-scale fluid motions are rapidly broken up by nonlinear interactions (the advective term in equation (1)). In this way, energy is transferred to smaller and smaller scales, until eventually a scale is reached at which viscous forces become important, and the energy is lost to dissipation. To describe this process more directly, it proves convenient to take the Fourier transform of equations (1)–(2) (see e.g. Orszag (1970), which this part of the discussion follows closely). First, write

$$\boldsymbol{v}(\boldsymbol{x}, t) = \sum \boldsymbol{a}(\boldsymbol{k}, t) \exp(\mathrm{i}\boldsymbol{k} \cdot \boldsymbol{x}), \tag{5}$$

where, by virtue of the assumed periodicity, the $\boldsymbol{k}$ are discrete and given by

$$\boldsymbol{k} = 2\pi\boldsymbol{n}/L, \qquad \boldsymbol{n} = (n_1, n_2, n_3), \qquad n_i = 0, \pm 1, \pm 2, \ldots . \tag{6}$$

Then the transformed Navier–Stokes equations become

$$\left(\frac{\partial}{\partial t} + \nu k^2\right) a_\alpha(\boldsymbol{k}, t) = -\frac{\mathrm{i}}{2} C_{\alpha\beta\gamma}(\boldsymbol{k}) \sum_{\boldsymbol{p}} a_\beta(\boldsymbol{p}, t) a_\gamma(\boldsymbol{q}, t), \tag{7}$$

$$k_\alpha a_\alpha(\boldsymbol{k}, t) = 0, \tag{8}$$

$$C_{\alpha\beta\gamma}(\boldsymbol{k}) = k_\beta C_{\alpha\gamma}(\boldsymbol{k}) + k_\gamma C_{\alpha\beta}(\boldsymbol{k}), \tag{9}$$

$$C_{\alpha\beta}(\boldsymbol{k}) = \delta_{\alpha\beta} - k_\alpha k_\beta / k^2. \tag{10}$$

Greek subscripts range from 1 to 3, summation over repeated indices is implied, and $\boldsymbol{q} = \boldsymbol{k} - \boldsymbol{p}$. Here and in what follows, $k^2 \equiv \boldsymbol{k} \cdot \boldsymbol{k}$, $a^2 \equiv \boldsymbol{a} \cdot \boldsymbol{a}^+$, and the explicit time dependence will usually be dropped. For convenience, the velocity basis functions $\exp \mathrm{i}\boldsymbol{k} \cdot \boldsymbol{x}$ will often be referred to as 'modes'. These equations contain no explicit forcing terms, although equation (4) requires that some such forcing be present. This omission is deliberate; the predictable characteristics of isotropic turbulent flow presumably do not depend on the details of the forcing, hence these details should not appear explicitly in the equations of motion. For the purposes of this paper, it will be sufficient to know that the mean energy input rate is given by equation (4), and that this energy is put into the system at the smallest wavenumbers.

For a typical turbulent flow, the number of terms in the sum in equation (7) is exceedingly large. For this reason, tracing the time evolution of such a flow by means of equations (7)–(10) is a problem so complex as to defy computation, let alone comprehension. Fortunately, for practical purposes one is usually concerned only with certain systematic characteristics of a given flow (the average transport coefficients, for example). The detailed time evolution is then of no interest whatever. The flow is then most appropriately described in terms of a probability distribution, and the quantities of interest are obtained by taking moments of this distribution. To define the notion of the probability distribution more clearly, note that the instantaneous state of the flow can be described by a vector in a space Γ of very large dimension, with the components of the vector given by the $\boldsymbol{a}(\boldsymbol{k})$ from equation (5). The probability function P is then a probability density defined on Γ. An important feature of this

choice of coordinate space (i.e. the coordinates given by the amplitudes of the Fourier components of velocity) is that, in the absence of driving or dissipative terms in equations (7)–(10), points in Γ describing turbulent flows evolve in time in accordance with Liouville's theorem. That is, suppose that at some initial time one knows that a particular flow is described by a point within a volume element dv in Γ. Then as the flow evolves, the volume element within which the flow is to be found will move through Γ, and its shape will change, but its volume will remain dv. This fact allows one to ascribe equal *a priori* probabilities to equal volumes in Γ.

For conceptual clarity, it proves useful to discretise P. This is easily done by subdividing Γ into cells, with each cell as small as one likes. Then moments of the probability distribution becomes sums over all the cells in Γ. For example, if there are M distinct cells in Γ, then for a properly normalised distribution the zeroth moment is

$$\sum_{\boldsymbol{a}(\boldsymbol{k}_1)} \sum_{\boldsymbol{a}(\boldsymbol{k}_2)} \cdots \sum_{\boldsymbol{a}(\boldsymbol{k}_M)} P[\boldsymbol{a}(\boldsymbol{k}_1), \boldsymbol{a}(\boldsymbol{k}_2), \ldots, \boldsymbol{a}(\boldsymbol{k}_M)] = 1. \tag{11}$$

More compactly,

$$\sum_{\Gamma} P_{\Gamma} = 1. \tag{12}$$

When sums are done over Γ, they will be denoted as above with upper case Greek indices.

Within this framework, the traditional approach to the problem of a stationary turbulent flow is to imagine an ensemble of similarly prepared flows, with the vector representing each realisation independently evolving in time. The probability distribution, in this view, is proportional to the local density of points in Γ, and the turbulence problem amounts to finding a distribution of densities that, under the influence of equations (7)–(10), does not change with time. This is a very difficult problem, on which a great deal of work has been done (see Monin and Yaglom (1975) for a review). However, this is not the only interpretation that one can put on the probability distribution P, and a different interpretation leads to a significantly different strategy for attacking the problem.

According to another point of view, one considers the assertion that at the time in question, the flow may be found within cell Ξ in Γ. The probability P_{Ξ} is then the *rational degree of belief* to which this assertion is entitled, given all of the available information concerning the flow. Defined in this manner, the probability distribution is descriptive of one's knowledge of the flow, rather than dependent on the density of points in an ensemble. As applied to statistical mechanics, this view has been advanced by, among others, Katz (1967), Baierlein (1971), and particularly by Jaynes (1957a, b, 1979). In the context of turbulence, one imagines trying to estimate where in Γ a particular flow is to be found, given only the stationarity of the flow, the Reynolds number, the energy input rate, and the Navier–Stokes equations. The information at one's disposal is manifestly inadequate to isolate one cell in Γ in which the flow may be found, so a description in terms of probabilities is evidently in order. The question is what probability distribution one should pick.

The answer to this question comes in two parts. First, distributions that conflict with the available knowledge should be excluded. For example, many points in Γ correspond to flows that violate the continuity equation (8). These flows cannot occur, and must be assigned zero probability. A more interesting restriction arises from the stationary character of the flow. Since one's knowledge about the flow is not changing

with time, presumably the probabilities one assigns should be independent of time as well. Further, since the average properties of the flow are assumed to be independent of time, it must be true that

$$\left\langle \frac{\partial g}{\partial t} \right\rangle \equiv \sum_{\Gamma} P_{\Gamma} \left(\frac{\partial g}{\partial t} \right)_{\Gamma} = 0, \tag{13}$$

for any dynamical quantity g one cares to name. (Here, 'dynamical quantity' is used in the sense of any function of the instantaneous mode amplitudes, defined if need be by reference to the equations of motion.) In particular, the mean time derivative of the energy at wavenumber $\boldsymbol{k}$ should vanish, for any value of $\boldsymbol{k}$:

$$\langle \partial [\tfrac{1}{2} a^2(\boldsymbol{k})] / \partial t \rangle = 0. \tag{14}$$

Insisting on the time invariance of the energy alone is a choice that requires some discussion. However, in the interest of continuity, this discussion will be postponed until § 8.

Simply forcing agreement with known information is not generally sufficient to determine a unique probability distribution. For example, there are infinitely many distributions that are consistent with equation (14), each corresponding to a different distribution of energy among the various values of $\boldsymbol{k}$. An additional criterion is needed to isolate one probability distribution from all the rest. A reasonable choice is the maximum entropy principle. This states that when choosing among a number of probability distributions, all of which agree with the available information, one should prefer the distribution that maximises the entropy. The entropy used here is defined in a sense consistent with information theory:

$$S = -h \sum_{\Gamma} P_{\Gamma} \ln(P_{\Gamma}), \tag{15}$$

where h is a positive constant that may be chosen arbitrarily. With a suitable choice for h, S may be interpreted as the amount of information (in bits, say) required to specify the cell in phase space actually occupied by the flow, with prior information consisting only of the probability distribution. Further elaboration of the notion of entropy as missing information can be found in the original work of Shannon (Shannon and Weaver 1949), and in any of a number of treatises in physics and engineering (Brillouin 1962, Khinchin 1957, Goldman 1953). A final point concerning the entropy as just defined has to do with the coordinates $a(\boldsymbol{k})$ used to define the space Γ. Applying an arbitrary nonlinear transformation to these coordinates will in general change the value computed for the entropy, and indeed maximising the entropy in one such transformed coordinate system does not necessarily maximise it in others. Thus, one is obliged to ask whether one particular coordinate system is the appropriate one for performing entropy calculations, and if so, why? In the absence of dissipation, the coordinates used here are easily justified because of the Liouville property discussed above. However, when viscous forces are present, this property no longer applies, and the correct choice of coordinate is not clear. Lacking a proper resolution to this problem, the approach taken here will simply be to use the coordinates that are appropriate in the non-viscous case. This approach will lead to useful results, but one should remember that these results rest on an assumption that has not yet been justified.

To illustrate some of the properties of the information theory entropy, consider a probability distribution consisting of a product of independent gaussians. For a gaussian function of one variable, the entropy may be computed directly. Taking a_0

as the (one-dimensional) cell size in Γ, let

$$P_\Gamma = (2\pi\sigma^2 a_0^2)^{-1/2} \exp -\left(\frac{a^2}{2\sigma^2 a_0^2}\right), \qquad \ln P_\Gamma = -\tfrac{1}{2}\ln(2\pi\sigma^2 a_0^2) - \frac{a^2}{2\sigma^2 a_0^2}. \tag{16}$$

This probability distribution has been chosen so that it is properly normalised, and so that its second moment is $\sigma^2 a_0^2$:

$$\sum_\Gamma P_\Gamma = 1, \qquad \sum_\Gamma P_\Gamma a^2 \equiv \langle a^2 \rangle = \sigma^2 a_0^2. \tag{17}$$

Then, if σ is large enough for the sum over Γ to be approximated by an integral over a, one obtains an elementary integral

$$-h \sum_\Gamma P_\Gamma \ln P_\Gamma = \frac{-h}{\sigma a_0 \sqrt{2}} \int_{-\infty}^{\infty} \exp -\left(\frac{a^2}{2\sigma^2 a_0^2}\right)\left(-\tfrac{1}{2}\ln(2\pi) - \ln(a_0\sigma) - \frac{a^2}{2\sigma^2 a_0^2}\right) \mathrm{d}a$$

$$= h(\ln \sigma a_0 + \tfrac{1}{2}\ln 2\pi + \tfrac{1}{2}). \tag{18}$$

Thus, apart from constant additive terms, the entropy is proportional to the logarithm of the width of the gaussian. This is reasonable: σ is roughly the number of cells in which a is likely to be found. To isolate the one cell in which a actually resides, one needs to specify an integer between 1 and σ, requiring about $\ln \sigma$ bits of information.

To extend this result to the product of many independent one-dimensional probability distributions, write

$$P_\Gamma = \prod_{\boldsymbol{k}} \eta[\boldsymbol{a}(\boldsymbol{k})], \qquad \ln P_\Gamma = \sum_{\boldsymbol{k}} \ln \eta[\boldsymbol{a}(\boldsymbol{k})]. \tag{19}$$

Here, there is no restriction on the form of the functions η. In particular, they need not be gaussian, and in fact they need not all be of the same form. The normalisation condition (11) must hold, however. From this joint probability distribution, one can compute the probability that a particular mode j has the amplitude $\boldsymbol{a}(j)$, regardless of the amplitudes of all the other modes. This is done by summing over all of the amplitudes except j:

$$P_j[\boldsymbol{a}(j)] = \eta[\boldsymbol{a}(j)] \sum_{\boldsymbol{a}(\boldsymbol{k}_1)} \eta[\boldsymbol{a}(\boldsymbol{k}_1)] \ldots \qquad (\text{all } \boldsymbol{k} \neq \boldsymbol{j}) \ldots \qquad \sum_{\boldsymbol{a}(\boldsymbol{k}_M)} \eta[\boldsymbol{a}(\boldsymbol{k}_M)]. \tag{20}$$

By invoking the normalisation condition, one immediately obtains

$$P_j[\boldsymbol{a}(j)] = \eta[\boldsymbol{a}(j)] / \sum_{\boldsymbol{a}(j)} \eta[\boldsymbol{a}(j)] \tag{21}$$

so that

$$\sum_{\boldsymbol{a}(j)} P_j[\boldsymbol{a}(j)] = 1. \tag{22}$$

i.e. the one-dimensional probability distributions are normalised individually. Since this is true, it is consistent and convenient to choose the normalisation for the η so that

$$\sum_{\boldsymbol{a}(\boldsymbol{k})} \eta[\boldsymbol{a}(\boldsymbol{k})] = 1 \tag{23}$$

valid for all values of $\boldsymbol{k}$. The entropy may now be written as

$$S = -h \sum_{\Gamma} P_{\Gamma} \ln P_{\Gamma} = -h \sum_{a(k_1)} \cdots \sum_{a(k_M)} \Big(\prod_{k} \eta[a(k)] \Big) \Big(\sum_{k} \ln \eta[a(k)] \Big). \tag{24}$$

This in turn may be rewritten in the form

$$S = -h \Big[\prod_{k} \Big(\sum_{a(k)} \eta[a(k)] \Big) \Big] \Big(\sum_{k} \ln \eta[a(k)] \Big). \tag{25}$$

Now for a given value of $\boldsymbol{k}$, the terms in $\Sigma \ln \eta[a(k)]$ combine with only one of the sums in the product over $\boldsymbol{k}$. By virtue of (23), all the other sums are equal to unity, and equation (25) may thus be written as

$$S = -h \sum_{k} \sum_{a(k)} \eta[a(k)] \ln \eta[a(k)]. \tag{26}$$

Thus, for a joint probability distribution in which the variations in different modes are independent of one another (a product distribution), the entropy is equal to the sum of the entropy contributions from the individual modes. Again, this is in agreement with one's intuitive notion of what constitutes information. Since the variations in different modes are uncorrelated, knowledge of the amplitude of one mode imparts no knowledge about the amplitude of any other mode. The information needed to specify the actual state of the flow is then the sum of the information required for the individual modes.

Finally, equations (18) and (26) may be combined to give the entropy associated with a probability distribution that is the product of many gaussians. As a rule, only entropy differences are of interest, so constant terms in the entropy may be dropped. This leads to

$$S = -h \sum_{k} \ln \frac{\langle a^2(k) \rangle^{1/2}}{a_0}, \tag{27}$$

a result that will prove useful shortly.

As this example indicates, the entropy is a quantitative measure of one's ignorance concerning the system in question. This viewpoint provides a simple justification for the application of the maximum entropy principle to fluid dynamics. Probability distributions with less than the maximum entropy imply more knowledge about the flow than does the maximum entropy distribution. However, all the distributions are based on the same prior knowledge. Therefore, to avoid making assertions about the state of the flow that cannot be supported by the available information, one should choose the probability distribution with maximum entropy. In this sense, the maximum entropy principle contains no physics at all. Rather, it is a logical principle, prescribing a way to reach consistent conclusions when faced with incomplete information.

3. A constant-dissipation model

To clarify the above notions, consider the example of a stationary flow in which all the energy is arbitrarily restricted to wavenumbers with magnitudes less than some maximum K. Let the energy input rate be G, and suppose that viscous dissipation is

significant, but that nothing is known about the mechanism by which energy is distributed among the various modes. According to the principles outlined above, the best estimate for the probability distribution for this flow is that which maximises the entropy, subject to the condition that the total dissipation rate equal the energy input rate. This constraint may be written

$$\sum_{\Gamma}\left(G-2\nu\sum_{\boldsymbol{k}}a^2(\boldsymbol{k})k^2\right)P_\Gamma=0. \tag{28}$$

The only other constraint on the probability distribution is the normalisation condition of equation (11). The entropy can be maximised by the standard technique of Lagrange multipliers. The function to be maximised is then

$$-h\sum_{\Gamma}P_\Gamma\ln P_\Gamma+\lambda_0\left(\sum_{\Gamma}P_\Gamma-1\right)+\lambda_1\left[\sum_{\Gamma}P_\Gamma\left(G-2\nu\sum_{\boldsymbol{k}}a^2(\boldsymbol{k})k^2\right)\right], \tag{29}$$

where the λ are constants to be determined. Differentiating with respect to the probabilities gives

$$\ln P_\Gamma=\left(\frac{\lambda_0}{h}+1\right)+\frac{\lambda_1}{h}\left(G-2\nu\sum_{\boldsymbol{k}}a^2(\boldsymbol{k})k^2\right), \tag{30}$$

while differentiating with respect to the λ recovers the constraint equations. Equation (30) shows that, except for constant factors, the probability distribution is a product of functions that apply to the individual modes. The mode amplitudes are thus independent of one another; this is a consequence of the express lack of knowledge about any interactions between modes. The probability distribution that applies to a single mode is

$$P_{\boldsymbol{k}}(\boldsymbol{a})\propto\exp-(2\lambda_1\nu k^2a^2/h), \tag{31}$$

i.e., the probability distribution is gaussian in the mode amplitude. It is significant that this distribution is the same as that derived by Thompson (1981) for randomly forced two-dimensional viscous flows, when the forcing is such as to produce no net energy transfer between modes.

The most interesting feature of this model is its energy spectrum. From equation (30), it is easy to show that the expectation value of νk^2a^2 is the same for all modes. This result is analogous to the equipartition theorem of statistical mechanics. However, in this case the quantity being partitioned equally among the various modes is not the energy, but rather the dissipation rate. Behaviour of this sort will recur in later sections, so it is useful to show in an explicit way why it happens. First note that from equation (27), the entropy can be written as

$$S=\sum_{\boldsymbol{k}}S_{\boldsymbol{k}}=-h\sum_{\boldsymbol{k}}4\ln\frac{\langle a^2(\boldsymbol{k})\rangle}{a_0}. \tag{32}$$

The expression takes this simple form because there are no correlations between modes. The factor of 4 within the sum occurs because, in an incompressible flow, $\boldsymbol{a}(\boldsymbol{k})$ may have no component along $\boldsymbol{k}$. Two complex vector components are therefore needed to describe $\boldsymbol{a}$, or four independent real numbers. The dissipation rate for one mode is

$$D(\boldsymbol{k})=2\nu k^2\langle a^2\rangle, \tag{33}$$

so

$$\frac{\partial S}{\partial D(\boldsymbol{k})} = \frac{2}{\nu k^2 \langle a^2 \rangle} = \frac{4}{D(\boldsymbol{k})}. \tag{34}$$

If the entropy is to be maximum, the derivative in equation (34) must be the same for all modes; otherwise, it would be possible to increase the entropy by transferring dissipation from one mode to another. Thus $D(\boldsymbol{k})$ must be the same for each of the allowed modes. Its numerical value depends on the value of G, and on how many distinct modes are included in the allowed wavenumber range.

The energy spectrum, as usually defined, is the total energy to be found per unit shell thickness in a shell of radius k:

$$E(k) \equiv 4\pi(\tfrac{1}{2}k^2 \langle a^2(\boldsymbol{k}) \rangle). \tag{35}$$

$E(k)$ is evidently proportional to $D(\boldsymbol{k})$, so for this model, the predicted energy spectrum is constant out to the arbitrarily chosen cut-off, and is zero thereafter. The reason for this odd behaviour is precisely the neglect of processes that transfer energy from one mode to another. These energy transfer processes arise from the nonlinear interactions described in equation (7), and consequently depend on the correlation between the amplitudes of different modes. To understand how these correlations affect the energy spectrum, one needs to have accurate expressions for the efficiency of energy transfer produced in this way, and for the effect of such correlations on the entropy.

4. The entropy decrease due to correlations between modes

In the simple model of § 3, the velocity amplitudes of different modes are uncorrelated, and there can be no mean energy transfer between modes. This occurs because no explicit energy balance constraint for individual modes is imposed on the probability distribution. It now becomes necessary to determine how the entropy changes when such constraints are applied.

To do this, consider a maximum-entropy probability distribution P_0, that is appropriate to one's knowledge of a system in the presence of certain constraints that are linear in the P_Γ, but are otherwise unspecified. P_0 is assumed to be properly normalised, and has an associated entropy:

$$\sum_\Gamma P_{0\Gamma} = 1, \tag{36}$$

$$S_0 = -h \sum_\Gamma P_{0\Gamma} \ln P_{0\Gamma}. \tag{37}$$

Suppose that f is some dynamical quantity satisfying

$$\langle f \rangle = \sum_\Gamma P_{0\Gamma} f = 0. \tag{38}$$

We now wish to know how the entropy changes if an additional constraint is added to the probability distribution, namely that $\langle f \rangle = F$, where F is some definite value.

If the new probability distribution is denoted by P, then equation (30) shows that P must be of the form

$$P = \kappa P_0 \exp -\lambda (F - f)/h, \tag{39}$$

where κ is a normalisation constant and λ is an undetermined constant multiplier. This formal solution leads to a formal difficulty, namely that many possible forms for f, e.g. f a trilinear function of the $a(\boldsymbol{k})$, lead to distributions P that are not normalisable. However, in many interesting cases (including all the applications considered here), $\lambda F/h$ is very small compared to unity, and so is $\lambda F/h$ for regions of Γ where P_0 is significantly different from zero. In this case, P is significantly different from zero in two regions: one near the origin where P_0 is large, and one far from the origin where the small high-order terms succeed in dominating the (usually gaussian) decrease of P_0. If one considers only the amplitude of a single mode, the separation between these regions in amplitude space is indeed very large; for the trilinear mode interactions that will appear in the next section, it is typically Re times the width of P_0. This suggests that, in order to preserve continuity of P as λF goes to zero, one should simply ignore that part of P that is far removed from the origin. On the other hand, one may expect significant correlations between the amplitudes and phases of the very large number of modes contributing to the instantaneous energy transfer. Since the complete probability distribution consists of the product of many exponentials like that in equation (39), it is possible to encounter normalisation difficulties without any need of the velocity amplitudes growing unduly large. These and similar problems were discussed by Orszag (1966), but he did not succeed in resolving them. For the current purposes, it will simply be assumed that one is justified in truncating the probability distribution at amplitudes that are large enough to include most of the guassian core of the probability distribution, but small enough to avoid problems with the normalisation. Then equation (39) may be approximated by

$$P = \kappa(1 - \lambda F/h)P_0(1 + \lambda f/h). \tag{40}$$

Since P itself must be normalised,

$$\kappa\left(1 - \frac{\lambda}{h}F\right)\sum_{\Gamma} P_{0\Gamma}\left(1 + \frac{\lambda}{h}f\right) = 1. \tag{41}$$

But by virtue of equation (38), the term containing f vanishes, so that

$$P = P_0(1 + \lambda f/h). \tag{42}$$

The value of λ is determined by the constraint on f

$$\sum_{\Gamma} P_{\Gamma} f = \sum_{\Gamma} P_{0\Gamma}\left(1 + \frac{\lambda}{h}f\right)f = \frac{\lambda}{h}\langle f^2\rangle = F, \tag{43}$$

where the angle brackets indicate an average over the initial probability distribution P_0. The new probability distribution is then given by

$$P = P_0(1 + fF/\langle f^2\rangle). \tag{44}$$

The entropy of this distribution is approximately

$$S = -h \sum_{\Gamma} P_{0\Gamma}\left(1 - \frac{fF}{\langle f^2\rangle}\right)\left(\ln P_{0\Gamma} + \frac{fF}{\langle f^2\rangle}\right) \tag{45}$$

$$= S_0 - h\frac{F^2}{\langle f^2\rangle} - h\sum_{\Gamma} \ln P_{0\Gamma}\frac{fF}{\langle f^2\rangle}. \tag{46}$$

Ordinarily, $\langle f\rangle$ vanishes because f is strictly antisymmetric with respect to P_0. If this is true, the last term in equation (46) vanishes, and the new entropy is given by

$$S = S_0 - hF^2/\langle f^2\rangle. \tag{47}$$

This is a plausible result. An additional constraint must cause the entropy to decrease, since the extra constraint implies additional information about the flow. The decrease must be at least quadratic in F, since the original probability distribution maximised the entropy with $F = 0$. Finally, $\langle f^2\rangle$ is a measure of how serious a dislocation in P_0 one requires to produce a given value of $\langle f\rangle$. If $\langle f^2\rangle$ is large, then P_0 need not be modified very much to conform to the new constraint, and the entropy change is rather small.

From the point of view of determining the turbulent spectrum, the qualitative importance of equation (47) is that the correlations required to transfer energy between modes entail a loss of entropy. It will develop that the correlations of interest for energy transfer are those between sets of three modes, at least two of which are widely separated in $\boldsymbol{k}$ space. Although there is no inherent difference, it proves convenient to distinguish between correlations of this sort and correlations involving two or more modes that are close neighbours in $\boldsymbol{k}$. In the latter case, the entropy loss due to correlations can best be described in terms of N, the effective number of independent velocity values per mode. If there are no local correlations between modes, then $N = 4$, as in equation (32). On the other hand, if M adjacent modes are perfectly correlated, four values suffice to describe the entire group, and $N = 4/M$. Thus, in the absence of explicit long-range constraints, the entropy per mode is

$$S_{\boldsymbol{k}} = Nh \ln(\langle a^2(\boldsymbol{k})\rangle^{1/2}/a_0). \tag{48}$$

It is important to note that, unlike the three-mode correlations, correlations between pairs of modes described by the $N(\boldsymbol{k})$ may not persist indefinitely. This is so because such correlations necessarily feed energy into (or extract it from) modes with particular wavenumbers. For such a process to continue in perpetuity implies a preferred direction in the flow, and hence a breakdown of isotropy. It is consistent to consider the groups of correlated modes described by the $N(\boldsymbol{k})$ as wavepackets, with each group of modes associated with a localised disturbance somewhere in the fluid volume. Then, if one knows the position of the disturbance and the complex amplitude of a single mode in its associated wavepacket, the amplitudes of all the other modes in the packet are determined as well. Thus, the $N(\boldsymbol{k})$ describe a sort of conditional correlation, with a temporal persistence that is only as long as the lifetime of the associated disturbances.

5. The energy flow between modes

It is now possible to approach the problem of energy transport between modes. The approach will be to choose maximum-entropy probability distributions for each

wavenumber $\boldsymbol{k}$, using the $N(\boldsymbol{k})$ and the mean energy transfer rates between modes as unspecified free parameters. The values taken by these parameters will then be chosen so as to maximise the entropy globally. In the arguments that follow, the expressions are simplified if sums over all of the modes are approximated by integrals over $\boldsymbol{k}$ space. From equation (6), the number of modes per unit volume in k is

$$(L/2\pi)^3 \equiv V, \tag{49}$$

so that

$$\sum_{\boldsymbol{k}} \to V \int \mathrm{d}^3 k. \tag{50}$$

In a similar manner, quantities like the dissipation rate will no longer be referred to as dissipation per mode, but rather as dissipation per unit volume in k space (or 'specific dissipation'). Because of the isotropy of the flow quantities averaged over the probability distribution can be functions of the magnitude of $\boldsymbol{k}$, but not of its direction. This fact will be used explicilty from time to time.

With these notational conventions, the instantaneous rate of energy transfer into a single mode due to nonlinear interactions can be determined from equation (7):

$$\frac{\partial}{\partial t}(\tfrac{1}{2}a^2) = \tfrac{1}{2}\,\mathrm{Im}\Big(C_{\alpha\beta\gamma}(\boldsymbol{k})\,V \int \mathrm{d}^3 p\, a_\alpha^+(\boldsymbol{k}, t) a_\beta(\boldsymbol{p}, t) a_\gamma(\boldsymbol{q}, t)\Big). \tag{51}$$

Dropping the explicit time dependence, the specific rate of energy transfer from the modes near $\boldsymbol{p}$ to those near $\boldsymbol{k}$ is thus

$$f_{\boldsymbol{pk}} \equiv \tfrac{1}{2} V^2\, \mathrm{Im}[C_{\alpha\beta\gamma} a_\alpha^+(\boldsymbol{k}) a_\beta(\boldsymbol{p}) a_\gamma(\boldsymbol{q})]. \tag{52}$$

Two average quantities of interest can be formed from $f_{\boldsymbol{pk}}$: the mean and mean square specific rates of energy transfer. These are

$$\langle f_{\boldsymbol{pk}} \rangle = \tfrac{1}{2} V^2 C_{\alpha\beta\gamma} \langle \mathrm{Im}[a_\alpha^+(\boldsymbol{k}) a_\beta(\boldsymbol{p}) a_\gamma(\boldsymbol{q})] \rangle, \tag{53}$$

$$\langle f_{\boldsymbol{pk}}^2 \rangle = \tfrac{1}{4} V^4 C_{\alpha\beta\gamma}^2 \langle [\mathrm{Im}\, a_\alpha^+(\boldsymbol{k}) a_\beta(\boldsymbol{p}) a_\gamma(\boldsymbol{q})]^2 \rangle. \tag{54}$$

Now write $a_\alpha = a(\boldsymbol{k}) \exp \mathrm{i}\phi_{\boldsymbol{k}}$, with similar expressions for the other amplitudes. If the three-mode correlations are not very strong, then one may expect these (real) mode amplitudes and phases to be unrelated, and

$$\langle f_{\boldsymbol{pk}}^2 \rangle = \tfrac{1}{4} V^4 C_{\alpha\beta\gamma}^2 \langle a^2(\boldsymbol{k}) \rangle \langle a^2(\boldsymbol{p}) \rangle \langle a^2(\boldsymbol{q}) \rangle \langle \sin^2(-\phi_{\boldsymbol{k}} + \phi_{\boldsymbol{p}} + \phi_{\boldsymbol{q}}) \rangle. \tag{55}$$

Then, under the same assumption, the average value of the $\sin^2$ function may be taken to be $\frac{1}{2}$, and

$$\langle f_{\boldsymbol{pk}}^2 \rangle = \tfrac{1}{8} V^4 C_{\alpha\beta\gamma}^2 a_{\boldsymbol{k}}^2 a_{\boldsymbol{p}}^2 a_{\boldsymbol{q}}^2. \tag{56}$$

The specific entropy $S(\boldsymbol{k})$ depends on $\langle a^2(\boldsymbol{k}) \rangle$, and on the character of the local and long-range correlations described by $N(\boldsymbol{k})$, $\langle f_{\boldsymbol{pk}} \rangle$, and $\langle f_{\boldsymbol{pk}}^2 \rangle$. The effect of the local correlations is readily determined from equation (48), but that of the long-range correlations responsible for energy transfer is more complicated. Suppose that $\langle f_{\boldsymbol{pk}} \rangle$ is required to take the particular value $\varepsilon_{\boldsymbol{pk}}$. If there were no local correlations, this requirement would imply V^2 constraints per unit volume of $\boldsymbol{k}$ and $\boldsymbol{p}$, one for each pair of modes in the volumes under consideration. However, the presence of local correlations means that not all of the modes are independent, and fewer constraints need be applied. The number of independent modes per unit volume near $\boldsymbol{k}$, $\boldsymbol{p}$ are $VN(\boldsymbol{k})$, $VN(\boldsymbol{p})$, respectively. The number of distinct constraints put on the probability

distribution per unit volume of $\boldsymbol{k}$ and $\boldsymbol{p}$ is then $V^2N(\boldsymbol{k})N(\boldsymbol{p})$, and each constraint is of the form $\langle f_{pk}\rangle = \varepsilon_{pk}$. The specific entropy is then

$$S(\boldsymbol{k}) = hVN(\boldsymbol{k}) \ln \frac{\langle a^2(\boldsymbol{k})\rangle^{1/2}}{a_0} - \frac{h}{2}\int \mathrm{d}^3p\, V^2N(\boldsymbol{k})N(\boldsymbol{p})\frac{\varepsilon_{pk}^2}{\langle f_{pk}^2\rangle}. \tag{57}$$

The factor of $\frac{1}{2}$ in the integral term corrects for the double counting of constraints that takes place if $\boldsymbol{k}$ and $\boldsymbol{p}$ are interchanged.

Now consider a small region of wavenumber space near wavenumber $\boldsymbol{k}$. On average, this region gains energy from some parts of wavenumber space and loses it to other parts; the difference between the gain and loss rates is the local dissipation rate. Since the dissipation rate for a particular mode depends only on $\langle a^2\rangle$, the mean-square mode amplitude can be determined if all of the ε_{pk} are known. Thus in equation (57), S may be considered a function of the $N(\boldsymbol{k})$ and the ε_{pk}.

One can imagine changing the probability distribution so as to vary the energy transfer between $\boldsymbol{k}$ and some other wavenumber region ($\boldsymbol{p}$, say), without changing any of the other energy transfer rates. In general, this variation will modify the entropy not only through its direct effects in the integral in equation (57), but also through its influence on the $\langle a^2(\boldsymbol{k})\rangle$ and $\langle f_{pk}^2\rangle$. Recalling that ε_{pk} is defined in the sense of an energy flow from $\boldsymbol{p}$ to $\boldsymbol{k}$, the variations in dissipation rates for a given variation in ε_{pk} are

$$\mathrm{d}D(\boldsymbol{k}) = \mathrm{d}\varepsilon_{pk}, \qquad \mathrm{d}D(\boldsymbol{p}) = -\mathrm{d}\varepsilon_{pk}. \tag{58}$$

These in turn imply variations in the amplitudes, e.g.

$$\partial\langle a^2(\boldsymbol{k})\rangle^{1/2}/\partial\varepsilon_{pk} = 1/4\, V\nu k^2\langle a^2(\boldsymbol{k})\rangle^{1/2}. \tag{59}$$

The variations in mode amplitudes resulting from (58), (59) in turn cause variations in $\langle f_{pk}^2\rangle$, but one can demonstrate that for large Reynolds numbers, these have a negligible effect on the entropy. The sum of these variations must result in no net change in the entropy if the probability distribution is to be in accord with the maximum entropy principle. From equation (57), the contributions to the specific entropy from $\boldsymbol{k}$, $\boldsymbol{p}$, and their mutual interaction is given by

$$S = hV\left(N(\boldsymbol{k}) \ln \frac{\langle a^2(\boldsymbol{k})\rangle^{1/2}}{a_0} + N(\boldsymbol{p}) \ln \frac{\langle a^2(\boldsymbol{p})\rangle^{1/2}}{a_0}\right) - hV^2N(\boldsymbol{k})N(\boldsymbol{p})\frac{\varepsilon_{pk}^2}{\langle f_{pk}^2\rangle}. \tag{60}$$

Thus, to maximise the entropy,

$$\frac{\partial S}{\partial \varepsilon_{pk}} = \frac{h}{4}\left(\frac{N(\boldsymbol{k})}{\nu k^2\langle a^2(\boldsymbol{k})\rangle} - \frac{N(\boldsymbol{p})}{\nu p^2\langle a^2(\boldsymbol{p})\rangle}\right) - 2hV^2N(\boldsymbol{k})N(\boldsymbol{p})\frac{\varepsilon_{pk}}{\langle f_{pk}^2\rangle} = 0. \tag{61}$$

The quantity $\nu k^2\langle a^2(\boldsymbol{k})\rangle/N(\boldsymbol{k})$ plays a very important role here. It is closely analogous to the usual temperature seen in statistical mechanics, since its reciprocal is a measure of the entropy gain per unit dissipation. Since this analogy is so suggestive, and since no confusion with the thermodynamic temperature will arise here, this grouping of terms will henceforth be called the 'turbulent temperature' (or just 'temperature'), and will be denoted by $T(\boldsymbol{k})$. With this definition, the energy flow is

$$\varepsilon_{pk} = \frac{1}{8V^2}\frac{\langle f_{pk}^2\rangle}{N(\boldsymbol{k})N(\boldsymbol{p})}\left(\frac{1}{T(\boldsymbol{k})} - \frac{1}{T(\boldsymbol{p})}\right). \tag{62}$$

It is evident from equation (62) that energy flows from regions of high T to regions

of low T, as the terminology entitles one to expect. As with the example of § 3, this happens because transferring dissipation from a higher- to a lower-temperature region increases the net entropy. However, in the current case, the very process of transferring energy decreases the entropy. The balance of these two effects determines the energy transfer rate between modes. The net specific energy input rate is now determined by integrating equation (62) over p:

$$\varepsilon(\boldsymbol{k}) = \frac{1}{8V^2}\int \mathrm{d}^3p \frac{\langle f_{pk}^2\rangle}{N(\boldsymbol{k})N(\boldsymbol{p})}\left(\frac{1}{T(\boldsymbol{k})} - \frac{1}{T(\boldsymbol{p})}\right). \tag{63}$$

Before proceeding, it is useful to give an expression for the temperature in terms of the known parameters of the flow. The energy driving the flow is, by assumption, inserted at very small wavenumbers. Further, since the fluid is being stirred randomly, it is consistent to suppose that the mode amplitudes at these wavenumbers are uncorrelated. Then for $k = 2\pi/L$, $N = 4$, and $\langle a^2\rangle = u^2$. The temperature at very low wavenumbers is

$$T = \nu(4\pi)^2u^2/L^2 \propto G/\mathrm{Re}. \tag{64}$$

One may obtain an additional equation relating $N(\boldsymbol{k})$ and ε_{pk} by maximising the entropy with respect to variations in $N(\boldsymbol{k})$. Once more allowing for double counting of interactions between $\boldsymbol{p}$ and $\boldsymbol{k}$ as $\boldsymbol{k}$ runs over its range,

$$\frac{\partial S}{\partial N(\boldsymbol{k})} = hV \ln\frac{\langle a^2(\boldsymbol{k})\rangle^{1/2}}{a_0} - h\int \mathrm{d}^3p\, V^2N(\boldsymbol{p})\frac{\varepsilon_{pk}^2}{\langle f_{pk}^2\rangle} = 0. \tag{65}$$

Substituting for ε_{pk}, this becomes

$$\ln\frac{\langle a^2(\boldsymbol{k})\rangle^{1/2}}{a_0} = \frac{1}{16V^3}\frac{1}{N^2(\boldsymbol{k})}\int \mathrm{d}^3p \frac{\langle f_{pk}^2\rangle}{N(\boldsymbol{p})}\left(\frac{1}{T(\boldsymbol{k})} - \frac{1}{T(\boldsymbol{p})}\right)^2. \tag{66}$$

Determining the spectrum of isotropic turbulence can now be formulated as a transport problem. One wishes to find the run of $\langle a^2(k)\rangle$, $N(k)$ satisfying the constraint of local energy balance in $\boldsymbol{k}$. This is simply a statement that, at every wavenumber, the net energy input rate from nonlinear interactions with other modes is balanced by the viscous dissipation rate:

$$\varepsilon(k) = D(k), \tag{67}$$

with the additional constraint that the total dissipation balances the total energy input:

$$V\int_0^\infty 4\pi k^2D(k)\,\mathrm{d}k = G. \tag{68}$$

The boundary conditions are that $\langle a^2\rangle$ and N at small k must agree with the values imposed by the stirring process.

A clearer picture of the problem now results if $\langle f_{pk}^2\rangle$ is written out explicitly. Equations (33), (56), (63) and (67) can be combined to give

$$2V\nu k^2a^2(k) = \frac{V^2}{64N(k)}\int \mathrm{d}^3p \frac{C_{\alpha\beta\gamma}^2(\boldsymbol{k})\langle a^2(\boldsymbol{k})\rangle\langle a^2(\boldsymbol{p})\rangle\langle a^2(\boldsymbol{q})\rangle}{N(p)}\left(\frac{1}{T(\boldsymbol{k})} - \frac{1}{T(\boldsymbol{p})}\right). \tag{69}$$

Now the form of $C_{\alpha\beta\gamma}(\boldsymbol{k})$ requires attention. Without loss of generality, $\boldsymbol{k}$ can be taken to lie along the x axis. Then $\boldsymbol{k}$, $\boldsymbol{p}$, and $\boldsymbol{q}$ are as shown in figure 1. Consideration of equations (9)–(10) shows that with this choice of axes, the only non-zero components

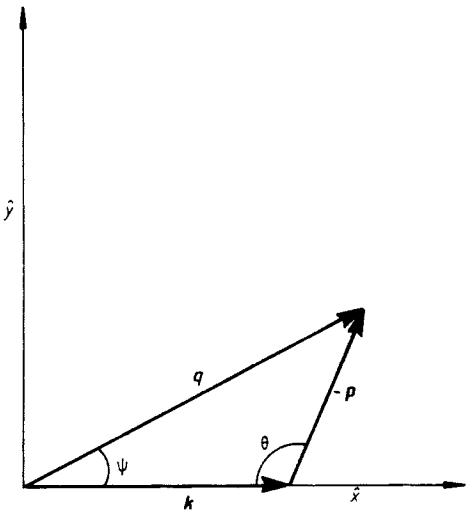

Figure 1. The geometry of wavenumbers occurring in the triple product in equation (69).

of $C_{\alpha\beta\gamma}$ are C_{212}, C_{313}, C_{221}, and C_{331}; these are all equal to k. The velocity component of a mode parallel to its wavevector must be zero, because of the continuity equation (8). On the other hand, symmetry requires that the two components perpendicular to the wavevector have identical average properties. Thus, $C^2_{\alpha\beta\gamma}(\boldsymbol{k})$ may be written

$$C^2_{\alpha\beta\gamma}(\boldsymbol{k}) = k^2[\sin\theta(1+\cos\psi)+\sin\psi(1+\cos\theta)]^2. \tag{70}$$

Using the law of sines, this is

$$C^2_{\alpha\beta\gamma}(\boldsymbol{k}) = k^2\sin^2\theta[1+\cos\psi+(p/q)(1+\cos\theta)] \equiv k^2\Theta^2. \tag{71}$$

By substituting this result into equations (69) and (66), using the definition of $T(p)$, and writing the integration in terms of p and θ, one obtains

$$\frac{64}{V}\nu^2 N(k) = \int \mathrm{d}p \int \sin\theta\,\mathrm{d}\theta\,\Theta^2\langle a^2(q)\rangle\left(\frac{T(p)}{T(k)}-1\right), \tag{72}$$

$$\frac{64}{V}\nu^2 \ln\frac{\langle a^2(k)\rangle^{1/2}}{a0} N(k) = \int \mathrm{d}p \int \sin\theta\,\mathrm{d}\theta\,\Theta^2\langle a^2(q)\rangle\left(\frac{T(p)}{T(k)}+\frac{T(k)}{T(p)}-2\right). \tag{73}$$

These are a coupled pair of integral equations describing the variation of the mode amplitudes $\langle a^2(k)\rangle$, and, through $N(k)$, the strength of local correlations between modes. It appears that a unique solution of equations (72)–(73) can be found, but this will not be done here. Some of the properties of such a solution merit attention, however, and are accessible without going into unnecessary details.

Equation (72) indicates two important things. The first of these is most readily seen by dividing both sides by u^2. Then using the definition of Re, one gets the approximate expression

$$N(k) \sim \mathrm{Re}^2 L \int \mathrm{d}k \frac{\langle a^2(\boldsymbol{q})\rangle}{u^2}\left(\frac{T(p)}{T(k)}-1\right). \tag{74}$$

$N(k)$ can grow no larger than 4, by definition, but the right-hand side contains a term of order Re^2. Also, for most wavenumbers of interest, one expects the magnitude of

$a(q)/u$ to be substantially larger than Re^{-1}. If this relation is to be satisfied, one must therefore have $T(k)$ very nearly equal to $T(p)$; i.e. the turbulent temperature of the flow must be very nearly constant with wavenumber. This holds except where the velocity amplitudes are very small compared with u. This may easily be understood by noting that the nonlinear interactions can transfer energy between modes with great efficiency, but for moderate wavenumbers the energy gain needed to offset dissipative losses is quite small. Only a small temperature difference is then required to drive the necessary energy flux.

The second point to notice about equation (72) is that it suggests that the energy transport between modes is semi-local. This is meant in the sense that two modes are able to exchange energy efficiently only if the magnitudes of their wavenumbers differ by a factor of less than about 2. Unless $\langle a^2(q)\rangle$ or the temperature decrease very rapidly with increasing q, the $\sin^2\theta$ dependence of Θ^2 assures that significant contributions to the integral in equation (72) occur only for $\boldsymbol{p}$ near a plane perpendicular to $\boldsymbol{k}$. This prevents significant energy transfer from modes for which p is much less than k. On the other hand, unless $\langle a^2(q)\rangle$ decreases very slowly with increasing q, then contributions to the integral must also drop sharply as p becomes greater than k. Thus, if the spectrum of $\langle a^2(k)\rangle$ lies within a reasonable range of steepness, the semi-local character of the energy transport holds good.

6. The constant-temperature model

The above considerations give enough information to construct a plausible model of isotropic turbulence giving a Kolmogorov spectrum. To do this, one requires three assumptions.

(1) The turbulent temperature is constant out to some maximum value of k called $k_{\max}$, beyond which it drops rapidly and smoothly to zero. $k_{\max}$ will ultimately be identified with the turbulent dissipation length.

(2) For $k < k_{\max}$, the mode amplitudes are described by a power law in k, explicitly $a(k) = Ak^m$, where $A = u(L/2\pi)^m$.

(3) The total dissipation at wavenumbers less than $k_{\max}$ is equal to the energy input rate G.

Of these assumptions, (1) has been partially justified above, and will be considered further in the next section. (3) is simply necessary for any self-consistent model, although there may be some question about the importance of dissipation at wavenumbers slightly larger than $k_{\max}$. Finally, (2) is plausible, and assuming a power-law dependence for the amplitudes allows a direct comparison with previous results. However, it will not be demonstrated that a power law is a unique (or even correct) solution to equations (72)–(73).

If the amplitudes follow a power law, it is simple to integrate the dissipation and obtain a relation for $k_{\max}$:

$$G = 2\nu VA^2 \int_0^{k_{\max}} k^{(2m+4)}\,\mathrm{d}k = \frac{2\nu VA^2}{(2m+5)} k_{\max}^{2m+5}. \tag{75}$$

Substituting explicit expressions for V and A, and invoking the definitions of Re and G,

$$k_{\max}^{(2m+5)} \propto \mathrm{Re}\, L^{-(2m+5)}. \tag{76}$$

Another expression for k_{max} can be obtained from equation (72). Because the temperature goes smoothly to zero, there must be wavenumbers immediately above k_{max} for which $[T(p)/T(k)-1]$ is of order unity. Because of the semi-local character of the energy transfer, the right-hand side of equation (72) will be dominated by contributions from modes with $q<k_{max}$. As $T(k)$ falls, the right-hand side must therefore rise to very large values, unless $a^2(q)$ is small. However, the magnitude of the left-hand side of equation (72) is limited, because $N(k)$ may not take values greater than 4. Thus, (72) can be satisfied only if the mode amplitudes near k are sufficiently small. This requirement may be written

$$(A/u)k_{max}^{m} \propto \mathrm{Re}^{-1}k_{max}^{-1/2}L^{-1/2}. \tag{77}$$

Once again using the explicit expression for A, this becomes

$$k_{max}^{(m+1/2)} \propto \mathrm{Re}^{-1}L^{-(m+1/2)}. \tag{78}$$

Equations (76) and (78) may be combined to eliminate Re, giving

$$k_{max}^{(11/2+3m)} \propto L^{-(11/2+3m)}. \tag{79}$$

This is a relation between k_{max} and L, which must hold independent of the Reynolds number. This can only be true if the exponent vanishes, i.e. if

$$m = -\tfrac{11}{6}. \tag{80}$$

This leads immediately to the familiar Kolmogorov energy spectrum

$$E(k) = 2\pi k^2 a^2 \propto k^{-5/3}. \tag{81}$$

The constant of proportionality cannot be determined without a detailed treatment of the behaviour of the temperature and the velocity amplitudes in the neighbourhood of k_{max}. It is important to know the value of this constant, since it may be compared directly with experimental results. However, this comparison will not be attempted here.

7. Validation of certain assumptions, and interpretation

The development in the last two sections was based on certain assumptions that have not been fully justified, and has implications that require further discussion. The two principal assumptions underlying the derivation of a Kolmogorov spectrum are the constant-temperature approximation and the claim that at some maximum value of k, the turbulent temperature drops rapidly to zero. These matters will be discussed below.

It is easy to verify the consistency of the constant-temperature approximation. This is done by assuming that $T(p)$ can be written

$$T(p) = T(k) + (\mathrm{d}T/\mathrm{d}k)(p-k), \tag{82}$$

which is then substituted into equation (72). By using power-law expressions for the amplitudes, and noting that in the constant-temperature approximation $N(k) = Bk^2a^2$, equation (77) can be put in the form

$$V\nu A^2 k^{(2m+2)} = \frac{\mathrm{d}T}{\mathrm{d}k}\frac{1}{T}\frac{\pi}{4\nu}\frac{A^4}{B}k^{(2m+2)}V^2 I(m). \tag{83}$$

Here $I(m)$ is a non-dimensional integral of order unity that depends only on the power-law exponent m. Equation (83) can be solved for $(1/T)\,\mathrm{d}T/\mathrm{d}k$, with the result

$$\frac{1}{T}\frac{\mathrm{d}T}{\mathrm{d}k}=\frac{4\nu^2}{\pi}\frac{B}{A^2V}\frac{1}{I(m)}. \tag{84}$$

The significance of this may be seen by considering $(k_{\max}/T)\,\mathrm{d}T/\mathrm{d}k$, a measure of the relative variation of T over the entire range of wavenumbers. Using either equation (76) or equation (78) for $k_{\max}$, one finds

$$(k_{\max}/T)\,\mathrm{d}T/\mathrm{d}k \propto \mathrm{Re}^{-5/4}. \tag{85}$$

Thus, for large Reynolds numbers and reasonable wavenumbers, the constant-temperature approximation should be extremely accurate.

The second major assumption made to derive the Kolmogorov spectrum was that the turbulent temperature drops rapidly to zero above $k_{\max}$. Examination of equation (72) shows that this cannot be strictly true. Once $N(k)$ has risen to its maximum value, further variations in $T(k)$ can occur only if there are comparable variations in the integral of $\langle a^2(q)\rangle$. Since this integral samples modes over a range in k space that is roughly comparable to k itself, the characteristic scale for changes in T must be about $k_{\max}$. Though scarcely a step function, this rate of decrease can be quite steep by comparison with that which applies in the constant-temperature region. The arguments of the last section do not appear to be affected in any fundamental way by this slow temperature drop-off, but an accurate solution to the equations clearly must take it into account.

A few correspondences between the results of this maximum-entropy approach and other treatments of turbulence are worth mentioning. First, it is possible to form length, velocity, and time scales from $k_{\max}$ and $a(k_{\max})$. The length scale is simply the reciprocal of $k_{\max}$:

$$l \propto k_{\max}^{-1} \propto L\,\mathrm{Re}^{-3/4} = (\nu^3/G)^{1/4}. \tag{86}$$

The relevant velocity scale is the RMS velocity at scales of about l. This is roughly

$$v_{\mathrm{RMS}} \sim [V^{-1}a^2(k_{\max})k_{\max}^3]^{1/2}. \tag{87}$$

By substituting for $a(k_{\max})$ and $k_{\max}$, one finds that for a given value of L, v_{RMS} scales as

$$v_{\mathrm{RMS}} \propto u\,\mathrm{Re}^{-1/4} = (\nu G)^{1/4}. \tag{88}$$

The time scale is the overturning time for the smallest length scales. This is simply

$$\tau \propto l/v_{\mathrm{RMS}} \propto (L/u)\,\mathrm{Re}^{-1/2} = (\nu/G)^{1/2}. \tag{89}$$

Unsurprisingly, these are the same as the Kolmogorov microscales for a flow with the same values of ν and G. To some extent this is just a dimensional necessity, but the correspondence provides a limited assurance that no serious errors have been committed in the analysis.

A more interesting issue centres around the behaviour of $N(k)$. In the constant-temperature region of k space, N must vary as k^2a^2, i.e. $N \propto k^{-5/3}$. Thus, within that region, the number of adjacent modes that are well correlated with one another grows rapidly with increasing k. The regions over which mode amplitudes are well correlated may be thought of as wavepackets, with each packet giving rise to a localised disturbance in the fluid flow. The characteristic width of these disturbances is inversely

proportional to the width of the corresponding wavepacket in wavenumber space. This suggests something of the intermittent behaviour of real turbulent flows: short wavelength disturbances tend to be concentrated in small, isolated regions occupying relatively little of the fluid volume. This is only a qualitative relation, however. A quantitative treatment of this effect would have to deal not only with the local correlations described by N, but also with the correlations that transport energy among widely separated modes.

Finally, it is interesting to see how the entropy associated with the spectrum just derived is distributed in wavenumber space. An expression for the specific entropy can be obtained by substituting equation (62) for the energy flow between modes in equation (57). Using the power-law form for $a(k)$, the constant temperature approximation, and equation (82) for $T(p)$, the result can be written in the form

$$\frac{1}{V}S = \int \mathrm{d}^3k\, N(k)\left(\ln\frac{\langle a^2(k)\rangle^{1/2}}{a_0} - \frac{Lk}{\mathrm{Re}^2}\frac{J}{I^2}\right) \tag{90}$$

where I is the same integral as in equation (83), and J is a related non-dimensional integral. The first term in brackets gives the entropy due to the uncertainty in the instantaneous mode amplitudes $a(k)$. The second term gives the correction resulting from the long-range correlations responsible for energy transport. Both terms are modified by the effects of the local correlations described by $N(k)$. Both I and J are roughly of order unity, so it is evident that the long-range correlations ordinarily have little effect on the total entropy. The entropy within the constant-temperature region may then be approximated by

$$S(k) = hV\int_0^{k_{\max}} \mathrm{d}k\, Bk^{1/3}\ln\frac{\langle a^2(k)\rangle^{1/2}}{a_0}. \tag{91}$$

This is easily evaluated; taking terms that vary no faster than ln(Re) to be effectively constant, the result is $S \propto \mathrm{Re}^{4/3}$. For the most part, this comparatively weak dependence on Re results from the rapid decline of $N(k)$. However, for k somewhat larger that $k_{\max}$, previous arguments show that N must increase rather rapidly to values near unity, while the mode amplitudes remain substantially constant. Thus, for k values greater than $k_{\max}$, the entropy contribution may be estimated as

$$S \propto hV\ln\frac{\langle a^2(k_{\max})\rangle^{1/2}}{a_0}(k_{\max}^3). \tag{92}$$

Again taking terms in ln(Re) to be constant, this gives $S \propto \mathrm{Re}^{9/4}$. This shows that while almost all of the energy and energy dissipation are found at wavenumbers less than $k_{\max}$, almost all of the entropy resides at larger wavenumbers.

8. Discussion

At this point, a brief recapitulation is in order. The problem of determining the spectrum of isotropic turbulence has been formulated as one of finding the probability distribution that agrees with all known facts about a given flow, but that is maximally non-committal about details of the flow concerning which no facts are available. In this case, the known facts are taken to be the governing equations of the flow (equations

(7)–(10)), the externally determined flow parameters (energy input rate G and Reynolds number Re), and the time invariance of the mean turbulent energy at every point in wavenumber space. 'Maximally non-committal' is interpreted in the sense of maximising the information theory entropy associated with the probability distribution. This is equivalent to maximising the number of bits of information needed to specify (with some stated precision) the instantaneous state of the flow, starting with only the facts given above. The entropy defined in this way depends on the amplitudes of the various modes, on how many modes are excited, and on the correlations that exist between different modes. The correlations between modes are particularly important as regards the dynamics of the flow, because these correlations determine the rate of energy transport between modes.

The energy transfer between two regions of wavenumber space depends on a subtle interplay between local correlations, long-range correlations, and the mode amplitudes. Other things being equal, the total entropy can be increased by transferring energy from modes with large amplitudes to those with small amplitudes. However, such energy transfer can only occur by means of correlations between the amplitudes of the modes in question. These correlations necessarily decrease the entropy, since knowing the amplitude of one mode imparts some additional knowledge about all of the modes with which the first is correlated. These considerations lead to a unique rate of energy exchange between widely separated regions of wavenumber space, at which rate the total entropy of the two regions is a maximum. The rate at which energy is actually transferred between two regions depends on the number of independent modes in each, i.e. on the local correlations between modes within each region. This is because the more independent modes there are, the more independent constraints must be put on the probability distribution in order to transfer energy at a given rate. Thus, energy transfer is very efficient between regions of wavenumber space where local correlations are strong and extend over many individual modes. These local correlations do, however, substantially decrease the local contribution to the entropy.

In discussing the energy transfer, a useful quantity turns out to be the derivative of the specific entropy with respect to the specific dissipation rate. The reciprocal of this quantity, denoted here by T, is called the 'turbulent temperature' or just 'temperature', and is directly analogous to the temperature used in thermodynamics and statistical mechanics. In wavenumber space, energy flows from high-temperature regions to low-temperature regions, and if the temperatures are nearly the same, the rate of energy flow is proportional to the temperature difference. Seen in this light, the turbulence problem becomes a transfer problem, with the net energy transferred into each wavenumber region dependent on the temperature of the surrounding regions, and on some geometrical factors. In order to maintain the stationary nature of the flow, the net energy input at each point must equal the net dissipation, and the total dissipation must equal the total energy input G. It develops that for large Reynolds numbers, the temperature gradient needed to maintain energy balance is very small. This leads to the constant-temperature approximation, in which, except at the very largest wavenumbers, the temperature gradient is ignored.

The resulting model of the turbulent spectrum consists of three regions. In the region below a critical wavenumber k_{max}, the temperature is constant and the amplitudes are described by a power law in k. In the region well above k_{max}, the temperature (and therefore the mean mode amplitude) is near zero. Finally, there is a matching region, in which the temperature goes smoothly to zero, over a distance comparable

to k_{max}. To determine what power law applies in the constant-temperature region, one first notes that virtually all the dissipation must take place at wavenumbers less than k_{max}, and this total dissipation must equal G. In addition, the rules governing energy transfer between modes allow the temperature to go to zero in the matching region only if the mode amplitudes at k_{max} take a particular value. Both these conditions give relations for k_{max} in terms of the power-law exponent; the relations can be simultaneously satisfied only if the exponent is $-\frac{11}{6}$, corresponding to the Kolmogorov energy spectrum.

In this model, the correlations between modes (particularly the local correlations) grow more and more important with increasing k, until k_{max} is reached. Above k_{max}, the local correlations rapidly diminish, until at wavenumbers of a few times k_{max}, individual modes are essentially uncorrelated. One result of this behaviour is that while virtually all the dissipation is contained in the constant-temperature region, virtually all the entropy is in the matching region. This entropy distribution provides a clue as to why local correlations occur at all. By virtue of the definition of T, one can increase the total entropy by causing dissipation to take place in low-temeprature regions rather than in high-temperature ones. It thus proves advantageous to make energy transfer to large wavenumbers as efficient as possible, even though this entails a loss of entropy in the constant-temperature region. An interesting side-effect of the local correlations that lead to efficient energy transfer is that they tend to concentrate the high-wavenumber energy in small, isolated regions of the fluid. This is qualitatively similar to the known intermittent behaviour of turbulent flows, but a quantitative treatment has not been worked out.

The arguments just summarised provide an encouraging indication that a maximum entropy approach to turbulence is a useful one. However, it must be emphasised that they do not by any means constitute a proof that this formulation is a correct description of turbulent flows. Both the advantages and the limitations of the current treatment (and the maximum entropy formalism itself) should be clearly understood.

First, the chief aim of this paper was to clarify physical concepts, rather than to give rigorous demonstrations of the implications of the maximum entropy formalism. The characteristics of the model used here have been justified in ways that seem plausible, but accurate solutions of the equations are still lacking. However, it is evident that the equations themselves are far simpler and more simply arrived at than in most theories of turbulence. Indeed, the solutions appear to be quite attainable, and they are being pursued. Second, and probably more significant, the maximum entropy approach itself is intrinsically better suited to some problems than to others. In problems concerning the average properties of stationary flows, maximum entropy ideas can likely be used to great advantage. Such problems include a large and important fraction of those found in physics, astrophysics, and engineering. However, for problems concerning the time evolution of individual flows, the morphology of short-lived flow patterns, or the mathematical nature of the turbulence problem, maximum entropy techniques will probably be of little help.

Finally, in cases where it is appropriate, the maximum entropy approach offers a significant advantage over other methods in that even its failures provide information about the system under consideration. This point has been discussed by Jaynes (1957b, 1979), and may be illustrated by reference to an issue raised in § 2. The question there was how to justify the choice of time invariance of the energy as the sole constraint to be satisfied by the probability distribution. Why not insist on the time invariance of some other quantity (or even a large number of such quantities), since stationarity requires that all of them must remain constant in time? To answer this

question, one must draw a sharp distinction between the logical content of the maximum entropy principle, and the physical content of the current application of that principle to turbulence. The maximum entropy principle simply provides a consistent and tractable means of incorporating *some* known facts into the description of systems for which one's information is incomplete. The physics of any such description lies in determining *which* of the available facts are important in determining the behaviour of the system. Thus, in the current example, the logical (or methodological) content is that the maximum entropy method, with reasonable constraints, leads to a well posed description of turbulent flows. The physical content is less general, but more useful: the correct energy spectrum can be found by invoking only the constraint of local energy balance in wavenumber space.

One can imagine being concerned with more complicated aspects of isotropic turbulence than the energy spectrum. In such cases, it is entirely possible that probability distributions based on local energy balance alone would fail to reproduce the observed effects. Then, as pointed out by Jaynes (1957b, 1979), the failure itself is conclusive evidence that some other constraint is important in the problem, while the details of the failure are likely to provide direct clues as to the nature of the missing constraint. The efficiency and simplicity of the maximum entropy technique arise because one need deal only with the constraints that are relevant to the processes one wants to predict. From this viewpoint, the chief conclusion reached here may be that one can say a great deal about turbulent flows on the basis of a very modest number of dynamical constraints.

Acknowledgments

It is a pleasure to acknowledge my debt to Drs James Crawford, Ralph Baierlein and Jack Herring, for their careful reading of early versions of this paper, to Drs Robert Kraichnan and Philip Thompson for valuable discussions and especially to Dr Bernard Durney, for his many perceptive comments and early encouragement.

References

Baierlein R 1971 *Atoms and Information Theory* (San Francisco: Freeman)
Brillouin L 1962 *Science and Information Theory* (New York: Academic)
Edwards S F and McComb W D 1969 *J. Phys. A: Gen. Phys.* **2** 157
Goldman S 1953 *Information Theory* (New York: Dover)
Jaynes E T 1957a *Phys. Rev.* **106** 620
—— 1957b *Phys. Rev.* **108** 171
—— 1979 in *The Maximum Entropy Formalism* ed R D Levine and M Tribus (Cambridge, Mass.: MIT) p 15
Katz A 1967 *Principles of Statistical Mechanics; the Information Theory Approach* (San Francisco: Freeman)
Khinchin A I 1957 *Mathematical Foundations of Information Theory* (New York: Dover)
Monin A S and Yaglom A M 1975 *Statistical Fluid Mechanics: Mechanics of Turbulence* ed J L Lumley (Cambridge, Mass.: MIT) ch 7
Orszag S A 1966 *Dynamics of Fluid Turbulence, Princeton University Plasma Physics Laboratory / Air Force Office of Scientific Research, Scientific Report PPLAF-13* p 45 ff
—— 1970 *J. Fluid Mech.* **41** 363
Shannon C E and Weaver W 1949 *The Mathematical Theory of Communication* (Urbana: University of Illinois)
Thomson P D 1972 *J. Fluid Mech.* **55** 711
—— 1981 *Non-equilibrium Probability Distributions for Randomly Forced Two-dimensional Flows* NCAR MS 0501/81-10