6

THE STATISTICAL DESCRIPTION OF TURBULENCE

Up to now, we have considered only average values of fluctuating quantities, such as U and $-\overline{uv}$. It is just as important to our understanding of turbulence to examine how fluctuations are distributed around an average value and how adjacent fluctuations (next to each other in time or space) are related. The study of distributions around a mean value requires the introduction of the probability density and its Fourier transform, the characteristic function. The study of the relation between neighboring fluctuations calls for the introduction of the autocorrelation and its Fourier transform, the energy spectrum. This chapter is devoted to the development of these mathematical tools; in the following two chapters, they are used in the study of turbulent transport ("diffusion") and of spectral dynamics. One other tool needed in the study of turbulent transport is the central limit theorem, which makes predictions about the shape of the probability density of certain quantities. The central limit theorem is introduced and discussed at the end of this chapter.

6.1
The probability density

We restrict the discussion to fluctuating quantities that are statistically steady, so that their mean values are not functions of time. Only under this condition does the idea of a time average make sense. A statistically steady function is called *stationary*; an example of a stationary function is given in Figure 6.1. The fluctuating $\tilde{u}(t)$ might be the streamwise velocity component measured in a wind tunnel behind a grid. We are interested in measuring the relative amount of time that $\tilde{u}(t)$ spends at various levels. We could get a crude idea of this by displaying $\tilde{u}(t)$ on the y axis of an oscilloscope, with a rapid sweep on the x axis. A time exposure would have a variable density, proportional to the time spent at each value of y. A more accurate measurement can be obtained by the use of a gating circuit, which turns on when the signal $\tilde{u}(t)$ is between two adjacent levels. In Figure 6.1 the levels are placed fairly close together in terms of the width of $\tilde{u}(t)$. The output of the gating circuit is shown below $\tilde{u}(t)$. If this is averaged, we obtain the percentage of time spent by $\tilde{u}(t)$ between the two levels. Adjusting the electronic "window" successively to different heights, we obtain a function similar to the one shown to the right of $\tilde{u}(t)$ in Figure 6.1.

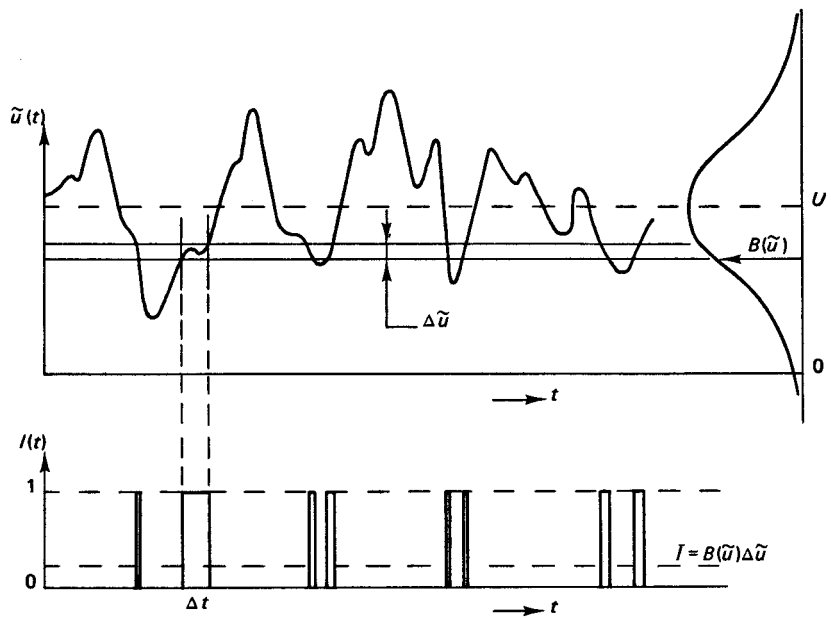

Figure 6.1. Measurement of the probability density of a stationary function. The function $I(t)$ is the discriminator output.

We expect that the averaged output of the gating circuit is proportional to the window width $\Delta\tilde{u}$, so that it is convenient to define a quantity $B(\tilde{u})$ by

$$B(\tilde{u})\,\Delta\tilde{u} \equiv \lim_{T\to\infty} \frac{1}{T}\Sigma(\Delta t). \tag{6.1.1}$$

The function $B(\tilde{u})$ is called a *probability density*; the probability of finding $\tilde{u}(t)$ between $\tilde{u}$ and $\tilde{u}+\Delta\tilde{u}$ is equal to the proportion of time spent there. Because $B(\tilde{u})$ represents a fraction of time, it is always positive, while the sum of the values of $B(\tilde{u})$ for all $\tilde{u}$ must be equal to one:

$$B(\tilde{u}) \geqslant 0, \quad \int_{-\infty}^{\infty} B(\tilde{u})\,d\tilde{u} = 1. \tag{6.1.2}$$

The shape of $B(\tilde{u})$ sketched in Figure 6.1 is typical of probability densities measured in turbulence. Many other shapes are possible; the probability density of a sine wave is sketched in Figure 6.2. This curve is zero beyond ± 1, because the sine wave has unit amplitude. Near ± 1, the slope goes to zero, so that the sine wave spends most time there, making the values of $B(\tilde{u})$ near ± 1 very large.

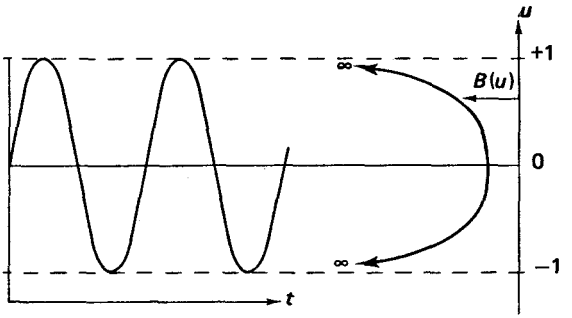

Figure 6.2. The probability density distribution of a sine wave.

We may express the averages we are familiar with in terms of $B(\tilde{u})$. Suppose we wish to average some function $f(\tilde{u})$. The time average

$$\bar{f} = \lim_{T\to\infty} \frac{1}{T}\int_{t_0}^{t_0+T} f(\tilde{u})\,dt \tag{6.1.3}$$

can be formed by adding all of the time intervals between t_0 and $t_0 + T$ during which $\tilde{u}(t)$ is between $\tilde{u}$ and $\tilde{u} + \Delta\tilde{u}$, multiplying this by $f(\tilde{u})$, and summing over all levels. The proportion of time spent between $\tilde{u}$ and $\tilde{u} + \Delta\tilde{u}$ is equal to $B(\tilde{u})\,\Delta\tilde{u}$, so that we can write

$$\bar{f} = \lim_{T\to\infty} \frac{1}{T}\int_{t_0}^{t_0+T} f(\tilde{u})\,dt = \int_{-\infty}^{\infty} f(\tilde{u})B(\tilde{u})\,d\tilde{u}. \tag{6.1.4}$$

The mean values of the various powers of $\tilde{u}$ are called *moments*. The first moment is the familiar mean value, which is defined by

$$U \equiv \int_{-\infty}^{\infty} \tilde{u}B(\tilde{u})\,d\tilde{u}. \tag{6.1.5}$$

In experimental work, the mean value is always subtracted from the fluctuating function $\tilde{u}(t)$. As in Chapter 2, we denote the fluctuations by u, so that $u = \tilde{u} - U$ and $\bar{u} = 0$. We then have $B(\tilde{u}) = B(U + u)$, so that it is convenient to use a probability density $B(u)$, which is obtained by shifting $B(\tilde{u})$ over a distance U along the $\tilde{u}$ axis. The moments formed with u^n and $B(u)$ are called *central moments*. The first central moment, of course, is zero.

The mean-square departure σ^2 from the mean value U is called the *variance*, or second (central) moment. It is defined by

$$\sigma^2 \equiv \overline{u^2} = \int_{-\infty}^{\infty} u^2 B(\tilde{u})\,d\tilde{u} = \int_{-\infty}^{\infty} u^2 B(u)\,du. \tag{6.1.6}$$

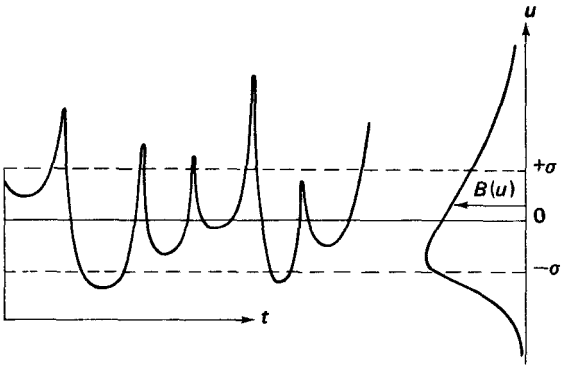

Figure 6.3. A function with positive skewness.

The square root of the variance, σ, is the familiar *standard deviation* (rms amplitude). The standard deviation is the most convenient measure of the width of $B(u)$.

The value of σ^2 is not affected by any lack of symmetry in $B(u)$ about the origin; if $B(u)$ is written as the sum of symmetric and antisymmetric parts, the latter does not contribute to σ^2. The *third moment,* however, defined by

$$\overline{u^3} \equiv \int_{-\infty}^{\infty} u^3 B(u)\, du, \tag{6.1.7}$$

depends only on the lack of symmetry in $B(u)$. If $B(u)$ is symmetric about the origin, $\overline{u^3} = 0$. It is customary to nondimensionalize $\overline{u^3}$ by σ^3, which gives a dimensionless measure of the asymmetry. This is called the *skewness* (S):

$$S \equiv \overline{u^3}/\sigma^3. \tag{6.1.8}$$

Figure 6.3 pictures a function with a positive value of S. The skewness is positive because large negative values of u^3 are not as frequent as large positive values of u^3.

The fourth moment, nondimensionalized by σ^4, is called *kurtosis* or *flatness factor*; it is represented by the symbol K:

$$K \equiv \frac{\overline{u^4}}{\sigma^4} = \frac{1}{\sigma^4} \int_{-\infty}^{\infty} u^4 B(u)\, du. \tag{6.1.9}$$

Two functions, one with a relatively small and the other with a relatively large kurtosis, are sketched in Figure 6.4. The value of the kurtosis is large if the values $B(u)$ in the tails of the probability density are relatively large. The

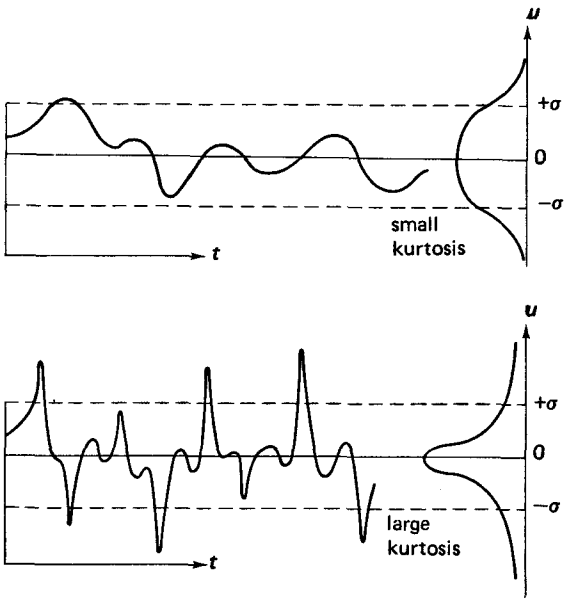

Figure 6.4. Functions with small and large kurtosis.

peaky function in Figure 6.4 frequently takes on values far away from the axis, so that its kurtosis is large. Because the fourth moment is nondimensionalized with σ^4, K contains no information on the width of the density $B(u)$.

6.2

Fourier transforms and characteristic functions

Although it is easy to see the physical significance of the probability density, it is often more convenient to work with another quantity, the *characteristic function*. This function is defined as the Fourier transform of $B(u)$. This means that we have to discuss Fourier transforms.

A Fourier-transform pair is defined by

$$\phi(k) \equiv \int_{-\infty}^{\infty} e^{iku} B(u)\, du, \quad B(u) \equiv \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-iku} \phi(k)\, dk. \tag{6.2.1}$$

We have used the probability density $B(u)$ and the corresponding characteristic function $\phi(k)$ as examples; we use other Fourier-transform pairs later. The conditions on the existence of $\phi(k)$ and on its ability to produce $B(u)$ upon integration are straightforward and need not concern us here.

In order to gain an appreciation for the usefulness of Fourier transforms, the behavior of $\phi(k)$ as reflected in $B(u)$ and conversely are explored. From the definition of the average of a function given in (6.1.4), and the definition of $\phi(k)$, it is evident that

$$\phi(k) = \overline{\exp\,[iku(t)]}. \tag{6.2.2}$$

As always, the overbar denotes a time average. This equation suggests that $\phi(k)$ can be measured by averaging the output of a function generator that converts $u(t)$ into $\sin u(t)$ and $\cos u(t)$. The experimental convergence of $B(u)$ is poor, because one must wait longer and longer to obtain a stable average as the window width Δu is decreased. The convergence of $\phi(k)$ is much better. Of course, there cannot be a net gain; to determine $\phi(k)$ accurately enough to obtain $B(u)$ from the Fourier transform is bound to take just as long as a direct measurement of $B(u)$.

If we have to deal with combinations of functions, say the sum of $u(t)$ and $v(t)$, the characteristic function of the sum (the *joint characteristic function*) is simply expressed by

$$\phi(k,\ell) = \overline{\exp\,[iku(t) + i\ell v(t)]}. \tag{6.2.3}$$

The corresponding probability density, which we encounter shortly, has no such simple form. This simplicity is one reason for the introduction of the characteristic function. We further discuss joint characteristic functions in Section 6.3.

The moments of $u(t)$ are related to $\phi(k)$ in a simple way. Differentiating the first of (6.2.1) with respect to k, we find that the moments are related to derivatives of $\phi(k)$ at the origin:

$$\left.\frac{d^n\phi(k)}{dk^n}\right|_{k=0} = i^n\,\overline{u^n}. \tag{6.2.4}$$

Because $\overline{u} = 0$, the slope of ϕ at the origin is zero. Because of (6.2.4), the characteristic function can be written as a Taylor series of the moments:

$$\phi(k) = \sum_{n=0}^{\infty} \frac{(ik)^n}{n!}\,\overline{u^n}. \tag{6.2.5}$$

Because no densities obtained in a laboratory have moments that are unbounded, the corresponding characteristic functions in principle have all derivatives. We say "in principle," because the larger the order of a moment is, the longer it takes to obtain a stable value. High-order moments are very

strongly affected by large excursions from the mean, which seldom occur. Therefore, moments higher than the fourth are seldom measured, so that we never have more than the first few derivatives of $\phi(k)$.

If $B(u)$ is symmetric, $\phi(k)$ is real. This can be seen if the first of (6.2.1) is written in terms of $\sin ku$ and $\cos ku$. This yields

$$\phi(k) = \int_{-\infty}^{\infty} \cos ku\, B(u)\, du + i \int_{-\infty}^{\infty} \sin ku\, B(u)\, du. \tag{6.2.6}$$

Only the antisymmetric part of $B(u)$ can contribute to the second integral. From (6.2.6) we also conclude that the real part of $\phi(k)$ is even in k, while the imaginary part is odd.

The modulus of $\phi(k)$ is given by

$$|\phi(k)| = \left| \int_{-\infty}^{\infty} e^{iku} B(u)\, du \right| \leqslant \int_{-\infty}^{\infty} |e^{iku}|\, B(u)\, du = \int_{-\infty}^{\infty} B(u)\, du = 1, \tag{6.2.7}$$

because $B(u) \geqslant 0$ and because the modulus of the exponential is unity. The last integral in (6.2.7) is equal to $\phi(0)$, so that we can write

$$|\phi(k)| \leqslant 1 = \phi(0). \tag{6.2.8}$$

The widths of $\phi(k)$ and of $B(u)$ are inversely related. Let us nondimensionalize the fluctuations u by σ, so that we have $u/\sigma = \eta$. Let us define a new probability density B' by

$$B'(\eta) \equiv \sigma B(u) = \sigma B(\sigma\eta). \tag{6.2.9}$$

Defined this way, the integral of B', according to (6.1.2), is equal to one. The characteristic function then becomes

$$\phi(k) = \int_{-\infty}^{\infty} e^{ik\sigma\eta} B'(\eta)\, d\eta. \tag{6.2.10}$$

A measure for the width of $\phi(k)$ can be defined as the value of k where the right-hand side of (6.2.10) is equal to $\frac{1}{2}$. This value is clearly proportional to $1/\sigma$, because $B'(\eta)$ has unit width. The effective width of $\phi(k)$ thus increases if σ decreases. If $\phi(k)$ is narrow, $B(u)$ is broad, and vice versa.

The effects of spikes and discontinuities Suppose $B(u)$ has a very high, narrow spike at some value of u, which we denote by s. This is pictured in Figure 6.5. The flat spots in the function $u(t)$ might be caused by a "dwell"

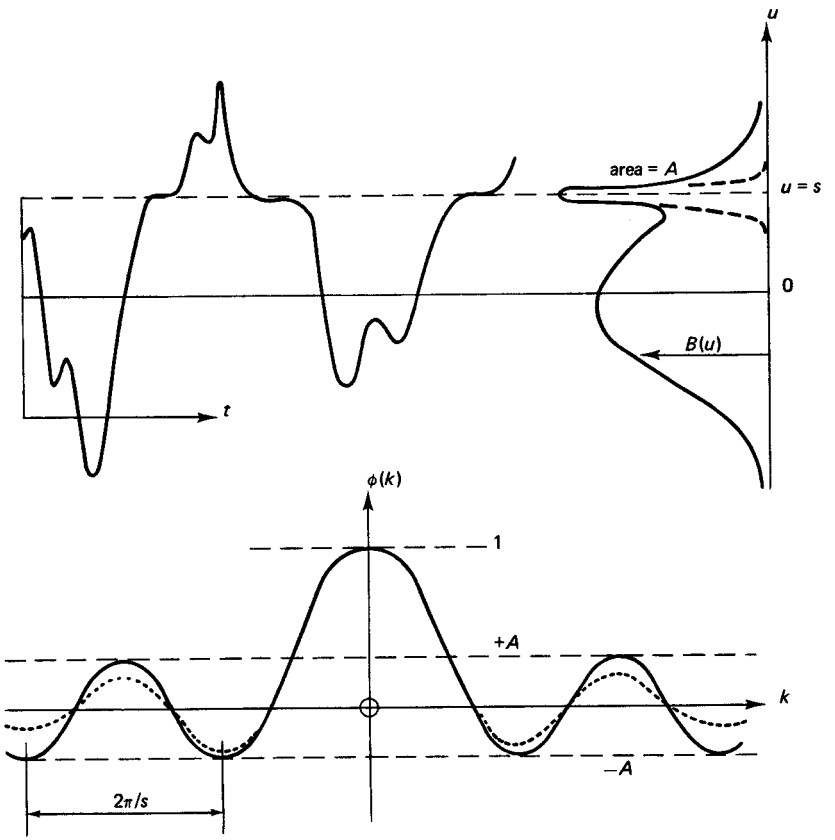

Figure 6.5. The characteristic corresponding to a probability density with a spike. The dotted line indicates the attenuation of $\phi(k)$ due to the finite width of the spike.

circuit of some kind. We assume that the area under the spike in $B(u)$ is A; the rest of the area enclosed by $B(u)$ is then $1 - A$. The spike in $B(u)$ produces a component of the characteristic function which behaves as $A \exp iks$. This component does not decay at infinity (Figure 6.5). In reality, of course, the spike is never infinitely high and narrow. If B_s is the spike component of B and ϕ_s is the spike component of ϕ, the latter can be written as

$$\phi_s(k) = \int_{-\infty}^{\infty} e^{iku} B_s(u)\, du = e^{iks} \int_{-\infty}^{\infty} e^{ik(u-s)} B_s(u)\, du$$

$$= e^{iks} \int_{-\infty}^{\infty} e^{ikx} B_s(s+x)\, dx = e^{iks} \phi_s'(k). \tag{6.2.11}$$

Here, $\phi_s'(k)$ is the transform of B_s, but shifted to the origin, so that it does not oscillate. Therefore, $\phi_s'(k)$ is a characteristic function with a width inversely proportional to the width of the spike. If the spike is infinitely narrow, $\phi_s'(k)$ is constant. If the spike has a finite width σ_s, $\phi_s'(k)$ decreases as k/σ_s, thus reducing the amplitude of $\exp iks$ (Figure 6.5).

If $B(u)$ has a discontinuity, so that its derivative has a spike, similar oscillations of $\phi(k)$ are generated. Integrating the first of (6.2.1) by parts once, we obtain

$$\phi(k) = -\int_{-\infty}^{\infty} \frac{e^{iku}}{k} \frac{dB(u)}{du} du. \tag{6.2.12}$$

If the spike in dB/du is infinitely narrow, we conclude that $\phi(k)$ behaves as $(1/k)\exp iks$ at large values of k. If the spike has finite width, $\phi(k)$ decreases somewhat faster. In general, if $B(u)$ and its first n derivatives are continuous, with a discontinuity in the $(n+1)$st, $\phi(k)$ is proportional to $k^{-(n+2)}\exp iks$ asymptotically.

Three pairs of Fourier transforms are sketched in Figure 6.6. In the first example, $B(u)$ itself has a discontinuity, so that ϕ decays as k^{-1}. In the second example, B has a discontinuity of slope, so that ϕ decays as k^{-2}. The third example is the probability density of a sine wave; here B has a spike, but it is not infinitely narrow, so that ϕ does decay, though rather slowly.

Parseval's relation Consider two functions, f and g, with Fourier transforms F and G:

$$F(k) = \int_{-\infty}^{\infty} e^{ikx} f(x)\, dx, \quad G(k) = \int_{-\infty}^{\infty} e^{ikx} g(x)\, dx. \tag{6.2.13}$$

With a little algebra it can be shown that

$$\int_{-\infty}^{\infty} F(k)G^*(k)\, dk = 2\pi \int_{-\infty}^{\infty} f(x)g^*(x)\, dx, \tag{6.2.14}$$

where asterisks denote the complex conjugates. This is known as Parseval's relation; it can be used to see how an operation carried out on a function affects its Fourier transform. For example suppose that $f(x)$ is being averaged over an interval $-X \leqslant x \leqslant X$. This amounts to evaluating the integral on the right-hand side of (6.2.14) with the use of a function $g^*(x)$ that looks like the "top-hat" function at the top left of Figure 6.6:

$$\begin{aligned} g^*(x) &= \tfrac{1}{2}X^{-1} \quad \text{for } -X \leqslant x \leqslant X, \\ g^*(x) &= 0 \qquad \text{otherwise.} \end{aligned} \tag{6.2.15}$$

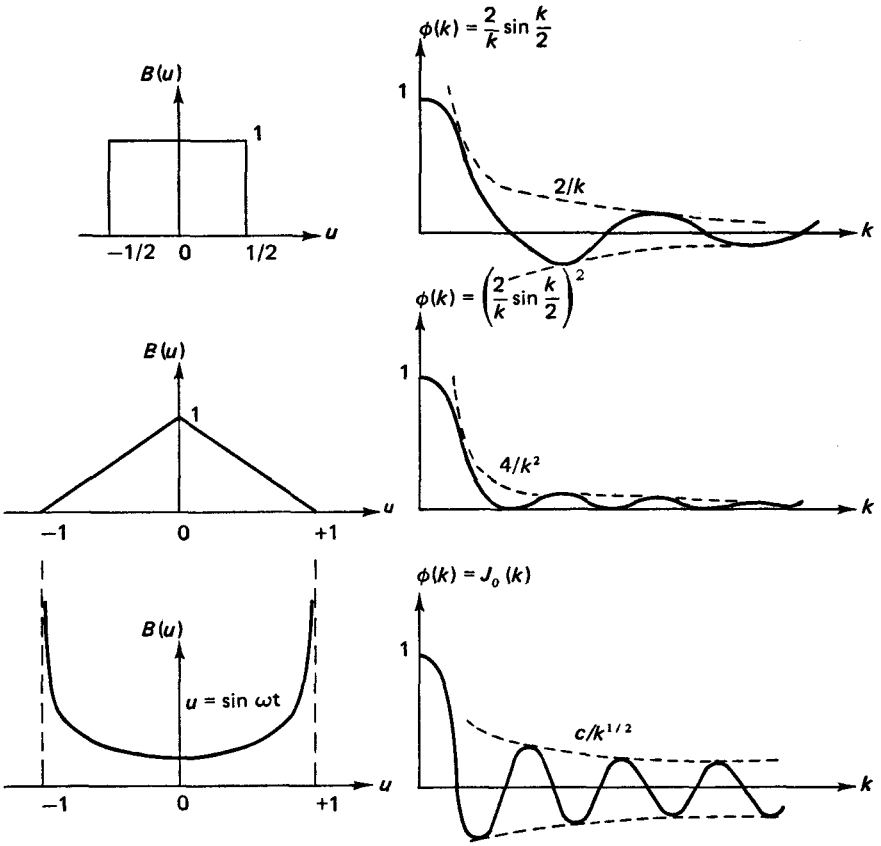

Figure 6.6. Some Fourier-transform pairs. Note that $\phi(k) = \phi(-k)$ because $B(u)$ is real.

The integrand on the left-hand side of (6.2.14) consists of the product of $F(k)$ and $G^*(k)$. The latter looks like the function on the top right of Figure 6.6. Now, as $g^*(x)$ becomes wider, $G^*(k)$ becomes narrower, as we saw earlier. If the averaging interval is quite long so that $G^*(k)$ is quite narrow, the integral on the left-hand side of (6.2.14) may be approximated by $F(0)$ times the integral of $G^*(k)$. Apparently, averaging a function is equivalent to selecting the value of its Fourier transform at the origin. If the physical variable is time, the transform variable is frequency; the origin in transform space corresponds to zero frequency. If we average something, the only thing left is the component at zero frequency; all other components become zero.

Similar problems arise when random variables are measured with sensors of finite dimensions. For example, a hot wire of finite length spatially averages

the velocity fluctuations that are measured. The effects of this averaging on the output of the hot-wire instrument can be described in terms of the first two Fourier-transform pairs in Figure 6.6 (Uberoi and Kovasznay, 1953).

6.3
Joint statistics and statistical independence

Let us consider the probability density for two variables simultaneously. A simple way to visualize this is to imagine that one variable $u(t)$ is displayed on the x axis of an oscilloscope, while the other variable $v(t)$ is displayed on the y axis (Figure 6.7). We assume that u and v are variables with zero mean, for simplicity. The *joint probability density* $B(u,v)$ is proportional to the fraction of time that the moving spot in Figure 6.7 spends in a small window between u and $u+\Delta u$, v and $v+\Delta v$. If we took a time exposure of the screen, the intensity at a point would be proportional to the joint probability density. As before, the sum of all the amounts of time spent at all locations must be equal to the total time, and the time fractions cannot be negative. Thus,

$$B(u, v) \geqslant 0, \quad \int\int_{-\infty}^{\infty} B(u, v)\, du\, dv = 1. \tag{6.3.1}$$

Also, if all of the values of v at a given value of u are combined, we should get the density of $u(t)$, which we call $B_u(u)$. On the oscilloscope, this amounts to turning the gain to zero on the y axis, so that the figure collapses to a horizontal line. A similar statement can be made about $B_v(v)$, so that we can write

$$\int_{-\infty}^{\infty} B(u, v)\, dv = B_u(u), \quad \int_{-\infty}^{\infty} B(u, v)\, du = B_v(v). \tag{6.3.2}$$

The moments of $u(t)$ and $v(t)$ can be obtained separately, or with (6.3.2). The most important *joint moment* is $\overline{uv}$, which is defined as

$$\overline{uv} \equiv \int\int_{-\infty}^{\infty} uvB(u, v)\, du\, dv. \tag{6.3.3}$$

This is called the *covariance* or *correlation* between u and v. Students of mechanics will recognize that the covariance is equivalent to the product of inertia of a distribution of mass. The correlation is thus a measure of the asymmetry of $B(u, v)$. If the value of $B(-u, v)$ is the same as that of $B(u, v)$, then $\overline{uv} = 0$. A few examples are given in Figure 6.8.

As we discussed in Section 2.1, if $\overline{uv} = 0$, $u(t)$ and $v(t)$ are said to be

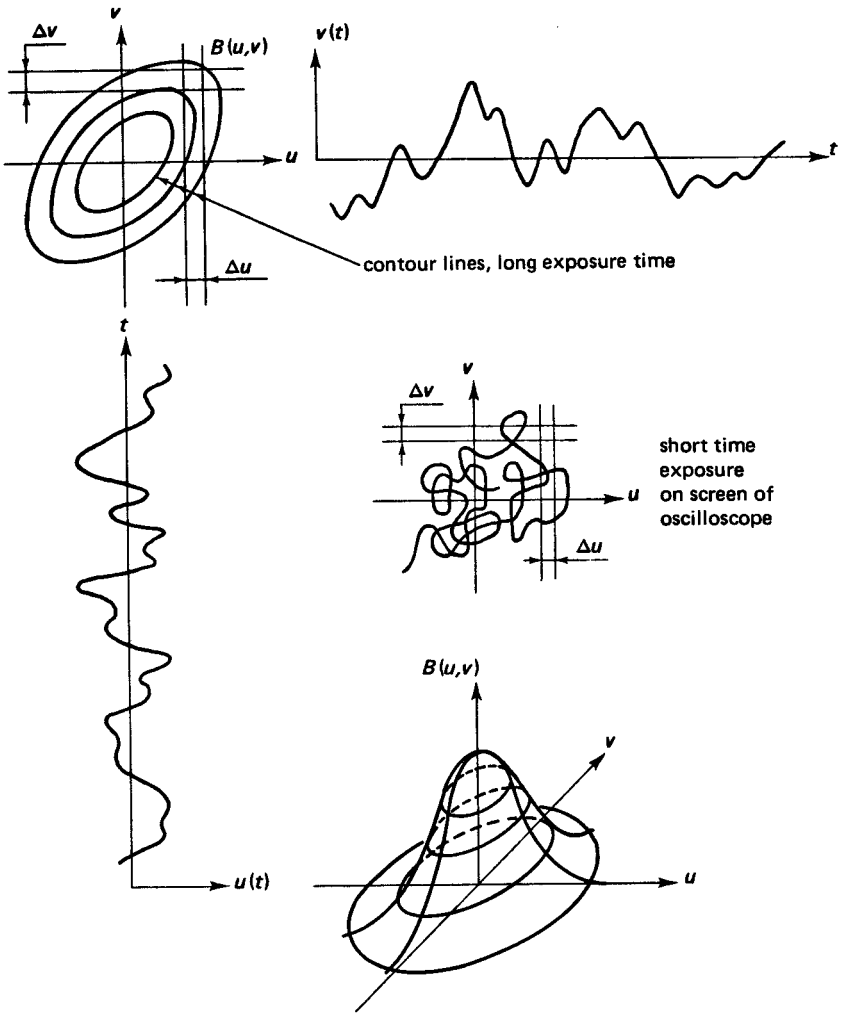

Figure 6.7. The joint probability density.

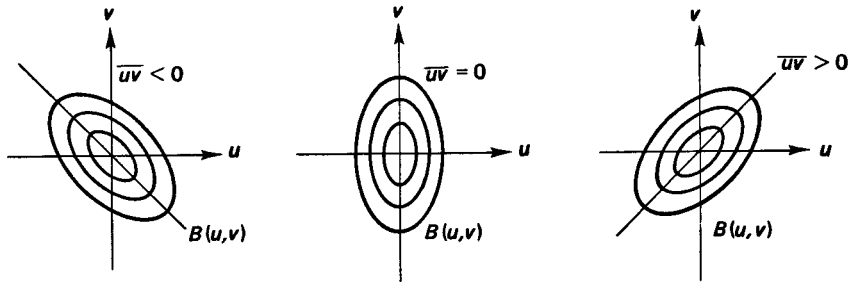

Figure 6.8. Examples of joint densities with various correlations.

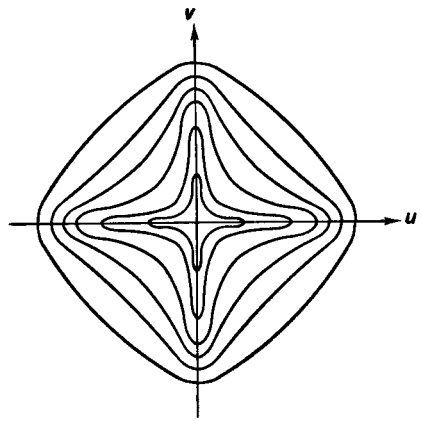

Figure 6.9. Probability density for two uncorrelated variables that tend to inhibit each other.

uncorrelated. Uncorrelated variables, however, are not necessarily independent of each other. The correlations in Figure 6.8 can be made zero by rotating the figures on the left and right until they are symmetric about one of the axes. In other words, it is possible to select two linear combinations of $u(t)$ and $v(t)$ and to create two new variables $u'(t)$ and $v'(t)$ which are uncorrelated. Clearly, the absence of correlation is no clue for the presence or absence of a dependence between the variables.

Two variables are *statistically independent* if

$$B(u, v) = B_u(u)\, B_v(v). \tag{6.3.4}$$

The probability density of one variable is then not affected by the other variable, and vice versa. For variables that depend on each other, the joint density cannot be written as a product. An example of the joint density of

uncorrelated, but dependent, variables is shown in Figure 6.9. Here, one variable tends to inhibit the other, so that they are seldom large simultaneously.

The joint characteristic function, defined by (6.2.3), is clearly the two-dimensional Fourier transform of the joint density, $B(u, v)$. In the case of statistically independent variables, the joint characteristic function is a simple product:

$$\phi(k,\ell) = \overline{\exp\,[iku + i\ell v]} = \overline{\exp\,[iku]}\ \overline{\exp\,[i\ell v]} = \phi_u(k)\,\phi_v(\ell). \tag{6.3.5}$$

6.4
Correlation functions and spectra

If we want to describe the evolution of a fluctuating function $u(t)$, we need to know how the values of u at different times are related. This question could be answered by forming a joint density for $u(t)$ and $u(t')$. However, as we have seen, the correlation provides much of the required information. The correlation $\overline{u(t)u(t')}$ between the values of u at two different times is called the *autocorrelation*. Because we are working with stationary variables, the autocorrelation gives no information on the origin of time, so that it can depend only on the time difference $\tau = t' - t$. Also, because $\overline{u(t)u(t')} = \overline{u(t')u(t)}$, the autocorrelation must be a symmetric function of τ.

Schwartz's inequality states that

$$|\overline{u(t)u(t')}| \leqslant [\overline{u^2(t)} \cdot \overline{u^2(t')}]^{1/2}. \tag{6.4.1}$$

For stationary variables, $\overline{u^2(t)} = \overline{u^2(t')} = \text{const}$, so that it is convenient to define an *autocorrelation coefficient* $\rho(\tau)$ by

$$\frac{\overline{u(t)u(t')}}{\overline{u^2}} \equiv \rho(\tau) = \rho(-\tau). \tag{6.4.2}$$

With (6.4.1) and (6.4.2), we obtain

$$|\rho| \leqslant 1 = \rho(0). \tag{6.4.3}$$

An autocorrelation coefficient similar to $\rho(\tau)$ was used in Section 2.3. The *integral scale* $\mathscr{T}$ is defined by

$$\mathscr{T} \equiv \int_0^\infty \rho(\tau)\,d\tau. \tag{6.4.4}$$

In turbulence, it is always assumed that the integral scale is finite. The value

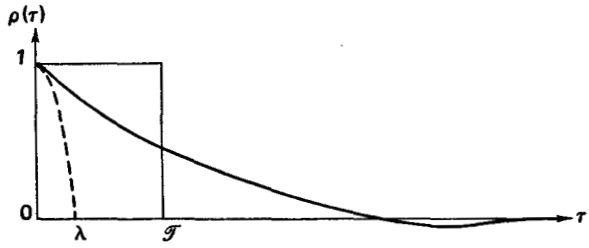

Figure 6.10. Sketch of an autocorrelation coefficient.

of $\mathscr{T}$ is a rough measure of the interval over which $u(t)$ is correlated with itself. A sketch of $\rho(\tau)$ is given in Figure 6.10.

Also shown in Figure 6.10 is the *microscale* λ, which is defined by the curvature of the autocorrelation coefficient at the origin:

$$d^2\rho/d\tau^2\big|_{\tau=0} \equiv -2/\lambda^2. \tag{6.4.5}$$

Expanding ρ in a Taylor series about the origin, we can write, for small τ,

$$\rho(\tau) \cong 1 - \tau^2/\lambda^2. \tag{6.4.6}$$

The microscale is thus the intercept of the parabola that matches $\rho(\tau)$ at the origin (Figure 6.10). Because $u(t)$ is stationary, we can write

$$0 = \frac{d^2}{dt^2}(\overline{u^2}) = 2\,\overline{u\frac{d^2u}{dt^2}} + 2\,\overline{\left(\frac{du}{dt}\right)^2}. \tag{6.4.7}$$

From (6.4.5) and (6.4.7) we obtain

$$\overline{\left(\frac{du}{dt}\right)^2} = \frac{\overline{2u^2}}{\lambda^2}. \tag{6.4.8}$$

In Chapter 3, the Taylor microscale, defined in a similar way from the spatial velocity autocorrelation, was extensively used.

The convergence of averages Suppose we want to obtain the average value of a function $\tilde{u}(t)$ in the laboratory. Of course, we cannot integrate over an infinitely long time interval, so that we have to consider the error due to finite integration time. The average is

$$U_T = \frac{1}{T}\int_{t_0}^{t_0+T} \tilde{u}(t)\,dt. \tag{6.4.9}$$

The difference between U_T and the true mean value U (recall that $\tilde{u} = U + u$) is given by

$$U_T - U = \frac{1}{T}\int_0^T [\tilde{u}(t) - U]\, dt = \frac{1}{T}\int_0^T u(t)\, dt. \tag{6.4.10}$$

Here we took $t_0 = 0$ for convenience. The mean-square value of (6.4.10) is

$$\overline{(U_T - U)^2} = \frac{\overline{u^2}}{T^2}\iint_0^T \rho(t' - t)\, dt\, dt' = \frac{\overline{2u^2}}{T}\int_0^T \left(1 - \frac{\tau}{T}\right)\rho(\tau)\, dt. \tag{6.4.11}$$

If the integrating time T is much longer than the integral scale $\mathscr{T}$, $\tau/T \sim 0$ in the range of values of τ where $\rho(\tau) \neq 0$, so that, by virtue of (6.4.4), the mean-square error may be approximated by

$$\overline{(U_T - U)^2} \cong 2\,\overline{u^2}\;\mathscr{T}/T. \tag{6.4.12}$$

It is clear that the average value can be determined to any accuracy desired if the integral scale is finite.

Ergodicity The requirement that a time average should converge to a mean value, that is, that the error should become smaller as the integration time increases and that the mean value found this way should always be the same, is called *ergodicity*. A variable is called *ergodic* if averages of all possible quantities formed from it converge. An ergodic variable not only becomes uncorrelated with itself at large time differences ($\tau \to \infty$), but it also becomes statistically independent of itself. A variable is ergodic if all integral scales that can be formed from it exist. Actually, this condition is not quite necessary; more general statements could be made. Let us consider a laboratory average of $\exp iku(t)$, which should differ little from the characteristic function $\phi(k) = \overline{\exp iku(t)}$ defined by (6.2.2). If the integral scale of $\exp iku(t)$ exists, the autocorrelation between $\exp iku(t)$ and $\exp iku(t')$ should vanish for large $t' - t$. Thus, for large $t' - t$,

$$\overline{[\exp iku(t) - \phi(k)]\,[\exp iku(t') - \phi(k)]} \to 0, \tag{6.4.13}$$

so that

$$\overline{\exp\,[iku(t) + iku(t')]} \to \phi(k)\phi(k). \tag{6.4.14}$$

From the definition (6.2.3) of a joint characteristic function, and the form (6.3.5) which it takes for statistically independent variables, it is clear that

the left-hand side of (6.4.14) would not approach a simple product unless the joint density itself were a simple product. Thus, by virtue of (6.3.4), $u(t)$ and $u(t')$ are statistically independent at large time differences.

It is reasonable to expect that all the integral scales associated with $u(t)$ are about the same, because they are determined by the scale of the physical process that produces $u(t)$. The integral scale $\mathscr{T}$ of $u(t)$ itself is thus not only a measure of the time over which $u(t)$ is correlated with itself but also a measure of the time over which $u(t)$ is dependent on itself. For time intervals large compared to $\mathscr{T}$, $u(t)$ becomes statistically independent of itself, so that $\mathscr{T}$ is a measure for the time interval over which $u(t)$ "remembers" its past history.

Another look at this concept is obtained if the output of the discriminator circuit in Figure 6.1 is considered. Let us call this function $I(t)$; it is equal to one if $u(t)$ appears in the window between u and $u + du$, and zero otherwise. The mean value of $I(t)$ is the value of $B(u)\,\Delta u$ we wish to determine:

$$\overline{I(t)} = B(u)\,\Delta u. \tag{6.4.15}$$

The mean-square error in the measurement of $B(u)\,\Delta u$ is obtained as follows. The variance σ^2 of $I(t)$ is given by

$$\begin{aligned}\sigma^2 &= \overline{[I(t) - B\,\Delta u]^2} = \overline{I^2(t)} - 2\overline{I(t)}\,B\,\Delta u + (B\,\Delta u)^2 \\ &= \overline{I^2(t)} - (B\,\Delta u)^2 = B\,\Delta u - (B\,\Delta u)^2.\end{aligned} \tag{6.4.16}$$

The last step in (6.4.16) could be taken because $I(t)$ and $I^2(t)$ always have the same value (either one or zero). Applying the error estimate (6.4.12) to the laboratory average I_T (obtained by integrating $I(t)$ over a time T), we find, if T is large and $B\,\Delta u$ is small,

$$\overline{(I_T - B\,\Delta u)^2} = 2\,\mathscr{T} B\,\Delta u/T. \tag{6.4.17}$$

The mean-square relative error is then given by

$$\overline{(I_T/B\,\Delta u - 1)^2} = 2\mathscr{T}/(TB\,\Delta u). \tag{6.4.18}$$

Now $TB\,\Delta u$ is the amount of time spent by $u(t)$ between u and $u + \Delta u$ if the averaging time is T. Hence, (6.4.18) shows that the error is small if the averaging time is so long that the amount of time spent in the window Δu is large compared to the integral scale $\mathscr{T}$.

Another way to obtain $B\Delta u$ is to sample $I(t)$ at time intervals large enough

to make the samples statistically independent of each other. With this procedure, the mean-square relative error is

$$\frac{1}{(B\,\Delta u)^2}\frac{1}{N}\overline{\sum_{n=1}^{N}[I(t_n)-B\,\Delta u]^2} \cong \frac{1}{NB\,\Delta u}. \tag{6.4.19}$$

Here, N is the total number of independent samples taken. If we compare (6.4.19) with (6.4.18), we see that $T/2\mathscr{T}$ may be regarded as the number of independent samples in a record of length T. Therefore, sampling once every two integral scales is adequate. We conclude that averages converge and integral scales exist if $u(t)$ may be regarded as consisting of a series of records of length $2\mathscr{T}$ (say, pieces of an analog tape), which are approximately statistically independent of each other.

The Fourier transform of $\rho(\tau)$ The autocorrelation coefficient $\rho(\tau)$ is a function that is equal to unity at the origin and is majorized by that value, that is real and symmetric, and that goes to zero faster than $1/\tau$, so that its integral scale exists. Referring back to Section 6.2, we conclude that $\rho(\tau)$ must be the Fourier transform of a continuous, symmetric, positive, real function $S(\omega)$ whose integral is unity. The transform of $\rho(\tau)$ must be continuous because ρ goes to zero faster than $1/\tau$; it must be symmetric because ρ is real; it must be real because ρ is symmetric; it must have a unit integral because $\rho = 1$ at the origin; it must be positive because ρ is majorized by its value at the origin.

The Fourier transform $S(\omega)$ of $\rho(\tau)$ is known as the *power spectral density*, or simply *spectrum;* it is defined by

$$\rho(\tau) = \int_{-\infty}^{\infty} e^{i\tau\omega}s(\omega)\,d\omega, \quad S(\omega) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-i\tau\omega}\rho(\tau)\,d\tau. \tag{6.4.20}$$

An appreciation for the relevance of $S(\omega)$ can be obtained by attempting to formulate a Fourier transform of $u(t)$ itself. Let us define

$$a_T(\omega, t) \equiv \frac{1}{T}\int_{t}^{t+T} e^{i\omega t'}u(t')\,dt'. \tag{6.4.21}$$

Let us recall the discussion on Parseval's relation at the end of Section 6.2. In this case, the function multiplying $u(t')$ is $g^*(t')$, which is given by

$$\begin{aligned} g^*(t') &= (1/T)\exp i\omega t' \quad \text{for } t \leqslant t' \leqslant t+T, \\ g^*(t') &= 0 \quad \text{otherwise.} \end{aligned} \tag{6.4.22}$$

The transform of $g^*(t')$ is

$$G^*(\omega') = \frac{\sin(\omega' - \omega)T/2}{(\omega' - \omega)T/2} \exp\left[-i(\omega' - \omega)(t + T/2)\right]. \tag{6.4.23}$$

The exponential has an absolute value of unity; it is present only because the midpoint of the integration interval T is a running variable. The first factor on the right-hand side of (6.4.23) is exactly the same as the function on the top right of Figure 6.6 but displaced to the center frequency ω. The average in (6.4.21) thus selects the value of the Fourier transform of $u(t)$ at the frequency ω rather than at the origin, if the time interval T is large enough.

Apparently, (6.4.21) is obtained by passing $u(t)$ through a filter that admits only frequences near ω. The width of the filter is about $1/T$. If we think of $u(t)$ as being synthesized from contributions at many frequencies, only the contributions close to ω form a square with $\exp i\omega t'$, so that only for those contributions does the integrand in (6.4.21) not oscillate. The contributions to $u(t)$ from all other frequencies cause the integrand to oscillate, so that they do not contribute to $a_T(\omega,t)$ if the integration time T is large (that is, if the bandwidth $1/T$ is small).

With a little algebra, it can be shown that the mean-square value of $a_T(\omega,t)$ is related to the spectrum $S(\omega)$ by

$$\lim_{T\to\infty} T\,\overline{|a_T(\omega,t)|^2} = \overline{u^2}\,S(\omega). \tag{6.4.24}$$

For a similar calculation, see Hinze (1959), Section 1-12. The spectrum thus represents the mean-square amplitude of the filtered signal or the mean-square amplitude of the Fourier coefficient of $u(t)$ at ω; it may be thought of as the energy in $u(t)$ at that frequency.

From (6.4.20) we conclude that the value of $S(\omega)$ at the origin is given by $S(0) = \mathscr{T}/\pi$. Also, if $\rho(\tau) \geqslant 0$ everywhere, $S(\omega)$ is maximized by its value at the origin. Conversely, if $S(\omega)$ has a peak away from the origin, then $\rho(\tau)$ must have negative regions. However, this does not imply that $S(\omega)$ must have a peak away from the origin if $\rho(\tau)$ is negative somewhere, as the Fourier transform pairs in Figure 6.6 demonstrate.

The spectrum of the derivative of a function is related to the spectrum of the function in a simple way. The autocorrelation of du/dt is given by

$$\overline{\frac{du(t)}{dt}\frac{du(t')}{dt'}} = \overline{u^2}\,\frac{d^2}{dt\,dt'}\,\rho(t' - t) = -\overline{u^2}\,\frac{d^2\rho}{d\tau^2}. \tag{6.4.25}$$

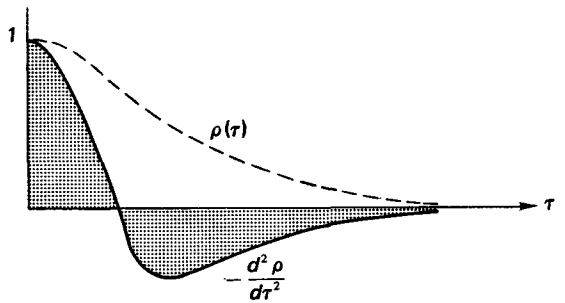

Figure 6.11. The autocorrelation coefficient of the first derivative of $u(t)$.

Differentiating the first of (6.4.20) twice, we obtain

$$-\frac{d^2\rho}{d\tau^2} = \int_{-\infty}^{\infty} e^{i\omega\tau}\omega^2 S(\omega)\, d\omega. \tag{6.4.26}$$

From (6.4.25, 6.4.26) we conclude that the spectrum of the first derivative is proportional to $\omega^2 S(\omega)$. This relation, of course, also can be applied to higher derivatives; for example, the spectrum of the second derivative is proportional to $\omega^4 S(\omega)$. Because $S(0)$ is finite (it is equal to $\mathscr{T}/\pi$, as we have seen), the spectra of derivatives vanish at the origin. This means that the integral scales of derivatives are zero. An example is given in Figure 6.11; the area under the curve is zero:

$$-\int_0^{\infty} \frac{d^2\rho}{d\tau^2}\, d\tau = \left.\frac{d\rho}{d\tau}\right|_0^{\infty} = 0. \tag{6.4.27}$$

6.5
The central limit theorem

In the analysis of turbulence, many quantities can be written as averages of stationary variables. In Chapter 7 we find that such quantities frequently arise in the discussion of turbulent transport (diffusion, mixing). The question arises, do averages of stationary variables have a probability density that is independent of the nature of the variable that is being averaged? In other words, we wonder if the very process of averaging introduces its own characteristic pattern, which masks the characteristics of the variable that is averaged. Subject to some simple conditions, the answer to this question is yes; the probability density of averages of stationary variables always tends to have the same shape.

Let us consider N statistically independent quantities $x_n(t)$. We assume that all $x_n(t)$ have identical probability densities and that their mean values are zero. It is convenient to work with the characteristic function $\phi(k)$, which is defined by

$$\phi(k) \equiv \overline{\exp ikx_n(t)}. \tag{6.5.1}$$

Because the densities of all $x_n(t)$ are the same, (6.5.1) holds for all n between 1 and N. Let us define the sum $z(t)$ of all $x_n(t)$ by

$$z(t) \equiv \sum_{n=1}^{N} x_n(t). \tag{6.5.2}$$

The variance of z is given by

$$z^2 = \sum_{n=1}^{N} \sum_{m=1}^{N} \overline{x_n x_m} = \sum_{n=1}^{N} \overline{x_n^2} = N\sigma^2. \tag{6.5.3}$$

Here, σ^2 is the variance of x_n, which is the same for each x_n because they have identical densities. The double sum becomes a single sum because x_n and x_m are statistically independent and have zero mean, so that they are uncorrelated. The variance of z increases as N increases, so that it is more convenient to define a new quantity $w(t)$ by

$$w(t) \equiv N^{-1/2} z(t). \tag{6.5.4}$$

The variance of $w(t)$ is equal to σ^2, no matter how large N becomes. Can we predict the probability density of $w(t)$? First it is convenient to compute the characteristic function $\phi_w(k)$ of $w(t)$. We obtain

$$\phi_w(k) \equiv \overline{\exp ikw(t)} = \overline{\exp\left(\frac{ik}{N^{1/2}} \sum_{n=1}^{N} x_n\right)} = [\phi(kN^{-1/2})]^N. \tag{6.5.5}$$

The last step in (6.5.5) could be taken because the x_n are statistically independent, so that the mean of the product of all $\exp(ikx_nN^{-1/2})$ is equal to the product of all ϕ. If the first few moments of the probability density of x_n exist, $\phi(kN^{-1/2})$ may be expanded in a Taylor series:

$$\phi(kN^{-1/2}) = 1 - k^2\sigma^2/2N + \mathcal{O}(k^3 N^{-3/2}). \tag{6.5.6}$$

This expansion is based on (6.2.5); the last term in (6.5.6) indicates that the remainder is of order $k^3 N^{-3/2}$, so that it can be made as small as desired by

selecting a sufficiently large value of N. Substituting (6.5.6) into (6.5.5), we obtain, for very large N,

$$\phi_w(k) = \lim_{N\to\infty} (1 - k^2\sigma^2/2N)^N = \exp(-k^2\sigma^2/2). \tag{6.5.7}$$

This is called the *central limit theorem;* the characteristic function $\phi_w(k)$ is called a *Gaussian characteristic function.* The probability density $B(w)$ corresponding to $\phi_w(k)$ can be computed from the definition (6.2.1) of the Fourier transform pair and the shape (6.5.7) of $\phi_w(k)$; the result is

$$B(w) = \frac{\exp(-w^2/2\sigma^2)}{(2\pi\sigma^2)^{1/2}}. \tag{6.5.8}$$

This is called a *Gaussian probability density*. The function $\exp -k^2$ is the only one that preserves its shape under a Fourier transformation. We conclude that asymptotically (as $N\to\infty$), the sum of a large number of identically distributed independent variables has a Gaussian probability density, regardless of the shape of the density of the variables themselves.

The statistics of integrals Let us now consider an integral of $u(t)$ over a time interval T. Because $u(t)$ is a stationary random variable, the value of the integral is also a stationary random variable which depends on the origin of the time interval. If the integration is performed in the laboratory, the probability distribution of the integral could be obtained by repeating the experiment many times.

An integral is like a sum, so that the central limit theorem may govern its probability distribution under suitable conditions. If the integration time T is large compared to the integral scale $\mathscr{T}$, the integral may be broken up into sections of length larger than $2\mathscr{T}$, so that the sections are approximately independent (recall the discussion of (6.4.18) and (6.4.19)):

$$\int_0^T u(t)\,dt = \int_0^{n\mathscr{T}} u(t)\,dt + \int_{n\mathscr{T}}^{2n\mathscr{T}} u(t)\,dt + \ldots. \tag{6.5.9}$$

As n increases, the sections of integral become more nearly independent, because adjacent sections depend on each other only near the ends. If the length of each section is $n\mathscr{T}$ and the total integration time is T, the number of sections is $T/n\mathscr{T}$. It is easy to arrange this in such a way that both $n\mathscr{T}$ and $T/n\mathscr{T}$ go to infinity as $T\to\infty$. We then have more and more sections, and they become less and less dependent, so that the probability distribution of the

integral on the left-hand side of (6.5.9) becomes Gaussian under favorable conditions.

The primary question is whether the sections of the integral become independent fast enough. It is possible to show (although we cannot do it here) that, as long as all integral scales exist and are nonzero, the sections become independent fast enough for the central limit theorem to apply. For a full discussion, see Lumley, 1972a.

The condition on the behavior of the correlation at large separations may be translated into a condition on the behavior of the spectrum near the origin, as we recall from the discussion in Section 6.2. From the definition (6.4.20) of $S(\omega)$ we conclude, in analogy with (6.2.4), that the derivatives of the spectrum near the origin are the moments of the correlation coefficient $\rho(\tau)$; if the moments exist, the derivatives do too, and vice versa. The condition that the correlation should be integrable to a value $\neq 0$ then becomes the condition that the spectrum near the origin be finite and nonzero.

A secondary question, which is not apparent in terms of correlations, becomes clear when stated in terms of the spectrum. We know from the discussion following Parseval's relation (6.2.14) that the average of $u(t)$ is equivalent to an operation on the Fourier transform of $u(t)$. In fact, the top-hat function at the top left of Figure 6.6 corresponds to an average. Evidently, averaging $u(t)$ is equivalent to multiplying the Fourier transform of $u(t)$ by the "filter" function at the top right of Figure 6.6. As the top-hat function representing the average becomes wider, the filter function on the right becomes narrower. The requirement that the spectrum be nonzero at the origin guarantees that the product of the Fourier transform and the filter function gets narrower as the integration time increases.

It is easy to find a violation of this condition. Consider du/dt; near the origin, its spectrum is proportional to ω^2, because $S(\omega)$ is approximately constant at small ω. The Fourier transform of du/dt must then be proportional to ω near the origin. However, the filter function on the top right of Figure 6.6 behaves as ω^{-1}. Hence, the product remains of constant width; it does not become narrower as the integration time increases. Therefore, we do not expect that the integral of du/dt will become Gaussian. This is obvious, because the integral of du/dt is $u(t)$ itself, which certainly does not need to be Gaussian.

A generalization of the theorem On the basis of the preceding discussion the central limit theorem can be simplified and generalized. Any variable having finite integral scales takes on a Gaussian distribution if it is filtered with a filter that is narrow enough; it becomes more Gaussian as the filter becomes progressively narrower. Clearly, we are not limited to simple averages. A variable $u(t)$ may be multiplied by any function before it is integrated; the only condition is that the Fourier transform of that function be a filter that, multiplied by the Fourier transform of $u(t)$, makes the product progressively narrower.

For example, a second integral may be written as

$$\int_{-T}^{T} dt \int_{-t}^{t} u(t')\,dt' = 2T \int_{-T}^{T} \left(1 - \frac{|t|}{T}\right) u(t)\,dt. \tag{6.5.10}$$

The factor $2T$ in front of the integral on the right-hand side need not concern us here. It is merely a normalizing factor that affects the variance of the double integral but not the applicability of the central limit theorem. The multiplying function in (6.5.10) has the same shape as the triangular function at the center left of Figure 6.6. Hence, the corresponding filter function decreases as ω^{-2}. If the Fourier transform of $u(t)$ rises more slowly than ω^2, the integral (6.5.10) becomes asymptotically Gaussian. This implies that a double integral of the first derivative of a stationary function $u(t)$ becomes Gaussian, even though a single integral of du/dt does not.

More statistics of integrals In the derivation (6.5.1–6.5.8) of the central limit theorem, the sum of the variables was normalized, so that the variance of $\omega(t)$ remained finite. That was a matter of convenience only; if the sum were not normalized, it would still have a Gaussian distribution, but with a variance that would increase with N.

Let us define an integral $X(T)$ of a stationary variable $u(t)$ by

$$X(T) \equiv \int_0^T u(t)\,dt. \tag{6.5.11}$$

The variance of $X(T)$ becomes (see (6.4.11))

$$\overline{X^2} = \overline{u^2} \iint_0^T \rho(t' - t)\,dt\,dt' = 2T\,\overline{u^2} \int_0^T \left(1 - \frac{\tau}{T}\right) \rho(\tau)\,d\tau \cong 2T\,\overline{u^2}\,\mathscr{T}. \tag{6.5.12}$$

The characteristic function $\phi_X(k)$ of $X(T)$ is Gaussian:

$$\phi_X(k) = \overline{\exp ikX(T)} \cong \exp(-k^2 \overline{u^2}\, T\mathscr{T}). \tag{6.5.13}$$

The probability density $B(X)$ corresponding to (6.5.13) is

$$B(X) = \frac{\exp(-X^2/4\overline{u^2}T\mathscr{T})}{(4\pi\, \overline{u^2}T\mathscr{T})^{1/2}}. \tag{6.5.14}$$

If a double integral $W(T)$ is defined by

$$W(T) \equiv \int_0^T dt \int_0^t u(t')\, dt', \tag{6.5.15}$$

it can be shown that the variance of $W(T)$ is given by

$$\overline{W^2} = \frac{2}{3}\, \overline{u^2}\, T^3 \int_0^T \left(1 - \frac{3\tau}{2T} + \frac{\tau^3}{T^3}\right) \rho(\tau)\, d\tau \cong \frac{2}{3}\, \overline{u^2}\, T^3 \mathscr{T}. \tag{6.5.16}$$

The characteristic function of $W(T)$ is Gaussian:

$$\overline{\exp ik\, W(T)} = \exp\,(-k^2 \overline{u^2}\, T^3\, \mathscr{T}/3). \tag{6.5.17}$$

We use these relations in Chapter 7.

Problems

6.1 Fluctuating velocity derivatives are associated with vorticity and strain-rate fluctuations. Will the skewness of a velocity-derivative signal ever be zero? Experiments have shown that the kurtosis of velocity derivatives is large if the Reynolds number is large. Use the simple model of Problem 3.2 to make estimates of the skewness and kurtosis.

6.2 Consider a stationary random variable with zero mean and a Gaussian probability density. Derive an approximate expression for the probability of exceeding amplitudes much larger than the standard deviation σ. What is the probability of exceeding 3σ? What is the probability of exceeding 10σ?

6.3 Compute the autocorrelation curve of a sine wave. What is the corresponding Fourier transform? What is the value of the integral scale?

6.4 In turbulent flow at large Reynolds numbers, the Taylor microscale λ is very small compared to the integral scale $\mathscr{T}$, and some investigators find it convenient to approximate the autocorrelation coefficient by $\rho(\tau) =$

$\exp(-|\tau|/\mathscr{T})$. What is the shape of the spectrum corresponding to this approximation? Also, is the spectrum of the derivative well behaved? Compare your results with the spectra given in Chapter 8.

6.5 Estimate the form of the spectrum of ocean waves in the range of frequencies where the Fourier coefficients of the wave amplitudes are determined by the frequency and the acceleration of gravity only.

6.6 Consider a sum of two statistically independent Gaussian variables, one of much lower frequency content than the other, both having zero mean. What do the autocorrelation and the spectrum look like? Suppose there is a gap between the spectra of the two, and the averaging time is long enough to average the fast one but not the slow; what do the correlation and spectrum look like in this case? What is the integral scale?

6.7 Consider one Gaussian variable modulated by another. The variables are statistically independent of each other; the second has a lower frequency content than the first. Both variables have zero mean. The product of the two variables appears to be "intermittent," that is, the low-frequency modulation appears to turn the high-frequency signal on and off. What is the kurtosis? What is the spectrum? Also consider a product of three independent variables, or of four. What is the kurtosis? If there are gaps between the spectra of the individual spectra, how does the measured kurtosis depend on the averaging time? Try to construct a continuous model, in which the logarithm of the signal is represented as the integral of a stationary process. Use the central limit theorem.