INTRODUCTORY
LECTURES on TURBULENCE
Physics, Mathematics and Modeling

J. M. McDonough

Departments of Mechanical Engineering and Mathematics
University of Kentucky

Contents

List of Figures

Chapter 1

Fundamental Considerations

In this chapter we first consider turbulence from a somewhat heuristic viewpoint, in particular discussing the importance of turbulence as a physical phenomenon and describing the main features of turbulent flow that are easily recognized. We follow this with an historical overview of the study of turbulence, beginning with its recognition as a distinct phenomenon by da Vinci and jumping to the works of Boussinesq and Reynolds in the 19^{th} Century, continuing through important 20^{th} Century work of Prandtl, Taylor, Kolmogorov and many others, and ending with discussion of an interesting paper by Chapman and Tobak [1] from the late 20^{th} Century. We then provide a final section in which we begin our formal study of turbulence by introducing a wide range of definitions and important tools and terminology needed for the remainder of our studies.

1.1 Why Study Turbulence?

The understanding of turbulent behavior in flowing fluids is one of the most intriguing, frustrating—and important—problems in all of classical physics. It is a fact that most fluid flows are turbulent, and at the same time fluids occur, and in many cases represent the dominant physics, on all macroscopic scales throughout the known universe—from the interior of biological cells, to circulatory and respiratory systems of living creatures, to countless technological devices and household appliances of modern society, to geophysical and astrophysical phenomena including planetary interiors, oceans and atmospheres and stellar physics, and finally to galactic and even supergalactic scales. (It has recently been proposed that turbulence during the very earliest times following the Big Bang is responsible for the present form of the Universe.) And, despite the widespread occurrence of fluid flow, and the ubiquity of turbulence, the "problem of turbulence" remains to this day the last unsolved problem of classical mathematical physics.

The problem of turbulence has been studied by many of the greatest physicists and engineers of the 19^{th} and 20^{th} Centuries, and yet we do not understand in complete detail how or why turbulence occurs, nor can we predict turbulent behavior with any degree of reliability, even in very simple (from an engineering perspective) flow situations. Thus, study of turbulence is motivated both by its inherent intellectual challenge and by the practical utility of a thorough understanding of its nature.

1.2 Some Descriptions of Turbulence

It appears that turbulence was already recognized as a distinct fluid behavior by at least 500 years ago (and there are even purported references to turbulence in the Old Testament). The following figure is a rendition of one found in a sketch book of da Vinci, along with a remarkably modern description:

> *"... the smallest eddies are almost numberless, and large things are rotated only by large eddies and not by small ones, and small things are turned by small eddies and large."*

Figure 1.1: da Vinci sketch of turbulent flow.

Such phenomena were termed "turbolenza" by da Vinci, and hence the origin of our modern word for this type of fluid flow.

The Navier–Stokes equations, which are now almost universally believed to embody the physics of all fluid flows (within the confines of the continuum hypothesis), including turbulent ones, were introduced in the early to mid 19^{th} Century by Navier and Stokes. Here we present these in the simple form appropriate for analysis of incompressible flow of a fluid whose transport properties may be assumed constant:

$$\nabla \cdot \boldsymbol{U} = 0\,, \tag{1.1a}$$

$$\boldsymbol{U}_t + \boldsymbol{U} \cdot \nabla \boldsymbol{U} = -\nabla P + \nu \Delta \boldsymbol{U} + \boldsymbol{F}_{_B}\,. \tag{1.1b}$$

In these equations $\boldsymbol{U} = (u, v, w)^T$ is the velocity vector which, in general, depends on all three spatial coordinates (x, y, z); P is the reduced, or kinematic (divided by constant density) pressure, and $\boldsymbol{F}_{_B}$ is a general body-force term (also scaled by constant density). The differential operators ∇ and Δ are the gradient and Laplace operators, respectively, in an appropriate coordinate system, with $\nabla\cdot$ denoting the divergence. The subscript t is shorthand notation for time differentiation, $\partial/\partial t$, and ν is kinematic viscosity.

These equations are nonlinear and difficult to solve. As is well known, there are few exact solutions, and all of these have been obtained at the expense of introducing simplifying, often physically unrealistic, assumptions. Thus, little progress in the understanding of turbulence can be obtained via analytical solutions to these equations, and as a consequence early descriptions of turbulence were based mainly on experimental observations.

O. Reynolds (*circa* 1880) was the first to systematically investigate the transition from laminar to turbulent flow by injecting a dye streak into flow through a pipe having smooth transparent walls. His observations led to identification of a single dimensionless parameter, now called the *Reynolds number*, and denoted Re,

$$Re = \frac{\rho U L}{\mu}\,, \tag{1.2}$$

that completely characterizes flow behavior in this situation. In this expression ρ and μ are, respectively, the fluid properties density and dynamic viscosity. U is a velocity scale (*i.e.*, a "typical" value of velocity,

say, the average), and L is a typical length scale, *e.g.*, the radius of a pipe through which fluid is flowing. We recall that Re expresses the relative importance of inertial and viscous forces. (The reader may wish to provide a first-principles demonstration of this as a review exercise.)

It is worth noting here that Eqs. (1.1b) can be rescaled and written in terms of Re as follows:

$$\boldsymbol{U}_t + \boldsymbol{U}\cdot\nabla\boldsymbol{U} = -\nabla P + \frac{1}{Re}\Delta\boldsymbol{U} + \widetilde{\boldsymbol{F}}_B\,, \tag{1.3}$$

where now pressure will have been scaled by twice the dynamic pressure, $\frac{1}{2}\rho U^2$, and $\widetilde{\boldsymbol{F}}_B$ is a dimensionless body force, often termed the Grashof number in mathematical treatments (see, *e.g.*, Constantin and Foias [2]), but which is more closely related to a Froude number under the present scaling. One can see from (1.3) that in the absence of body forces Re is the only free parameter in the N.–S. equations; hence, setting its value prescibes the solution.

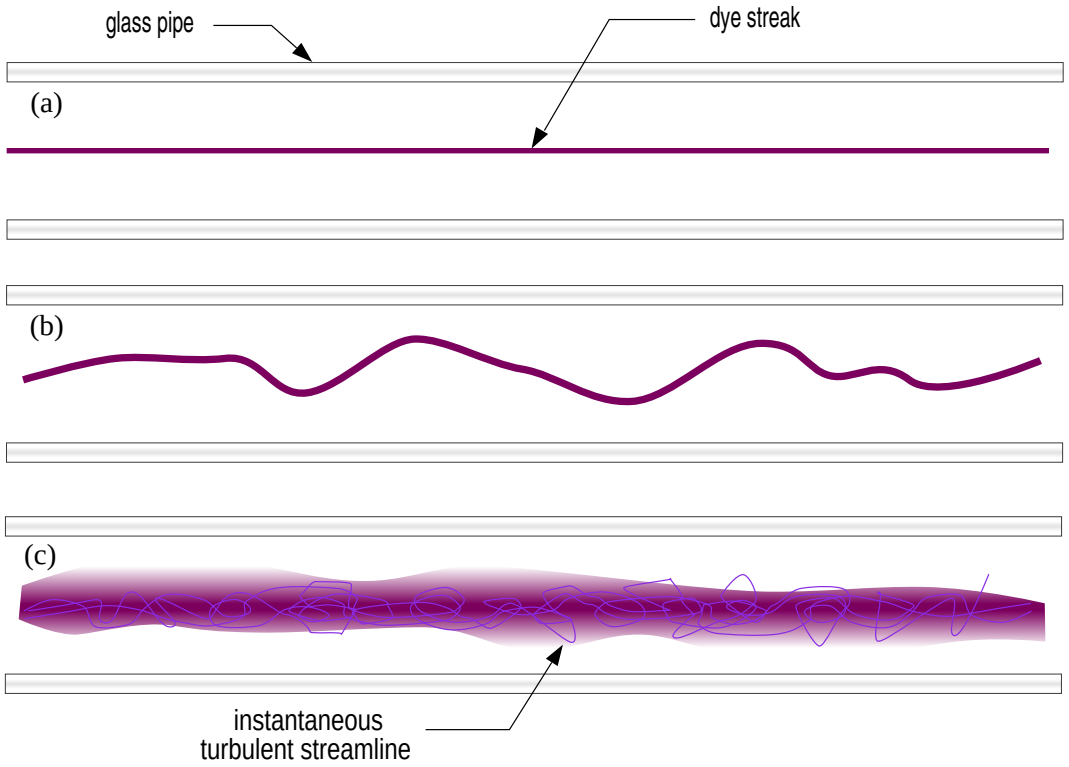

Figure 1.2: The Reynolds experiment; (a) laminar flow, (b) early transitional (but still laminar) flow, and (c) turbulence.

Figure 1.2 provides a sketch of three flow regimes identified in the Reynolds experiments as Re is varied. In Fig. 1.2(a) we depict laminar flow corresponding to $Re \lesssim 2000$ for which dye injected into the stream can mix with the main flow of water only via molecular diffusion. This process is generally very slow compared with flow speeds, so little mixing, and hence very little apparent spreading of the dye streak, takes place over the length of the tube containing the flowing water. Figure 1.2(b) shows an early transitional state of flow ($2000 \lesssim Re \lesssim 2300$) for which the dye streak becomes wavy; but this flow is still laminar as indicated by the fact that the streak is still clearly identifiable with little mixing of dye and water having taken place.

Turbulent flow is indicated in Fig. 1.2(c). Here we see that instantaneous streamlines (now different from the dye streak, itself) change direction erratically, and the dye has mixed significantly with the water. There are a couple things to note regarding this. First, the enhanced mixing is a very important feature of turbulence, and one that is often sought in engineering processes as in, for example, mixing of reactants in a combustion process, or simply mixing of fluids during stirring. The second is to observe that this

mixing ultimately leads to the same end result as molecular diffusion, but on a much faster time scale. Thus, turbulence is often said to "enhance diffusion," and this viewpoint leads to a particular approach to modeling as we will see later.

But we should recognize that although the final result of turbulent mixing is the same as that of diffusive mixing, the physical mechanisms are very different. In fact, turbulence arises when molecular diffusion effects are actually quite small compared with those of macroscopic transport. Here we should recall the form of Eqs. (1.1b) and note that the second term on the left-hand side corresponds to macroscopic transport (of momentum). These are the "convective" or "advective" terms of the N.–S. equations. The second term on the right-hand side represents molecular diffusion of momentum, as should be clear from the fact that its coefficient is a physical transport property (viscosity, in this case) of the fluid and does not depend on the flow situation (provided we neglect thermal effects), and the differential operator is second order. Clearly, if ν is small we should expect advective, nonlinear behavior to be dominant, and this is the case in a turbulent flow. In contrast to this, if ν is relatively large molecular diffusion will be dominant, and the flow will be laminar. If we recall that $\nu = \mu/\rho$, we see that the nonlinear, macroscopic transport case corresponding to turbulence occurs when the Reynolds number is large.

It is often claimed that there is no good definition of turbulence (see, *e.g.*, Tsinober [3]), and many researchers are inclined to forego a formal definition in favor of intuitive characterizations. One of the best known of these is due to Richardson [4], in 1922:

> *Big whorls have little whorls,*
> *which feed on their velocity;*
> *And little whorls have lesser whorls,*
> *And so on to viscosity.*

This reflects the physical notion that mechanical energy injected into a fluid is generally on fairly large length and time scales, but this energy undergoes a "cascade" whereby it is transferred to successively smaller scales until it is finally dissipated (converted to thermal energy) on molecular scales. This description underscores a second physical phenomenon associated with turbulence (recall that the first is diffusion): dissipation of kinetic energy. We comment that this process is also mediated by molecular viscosity, and historically utilization of this has been crucial in modeling efforts as we will outline in more detail in the sequel.

T. von Kármán [5] quotes G. I. Taylor with the following definition of turbulence:

> *"Turbulence is an irregular motion which in general makes its appearance in fluids, gaseous or liquid, when they flow past solid surfaces or even when neighboring streams of the same fluid flow past or over one another."*

Hinze, in one of the most widely-used texts on turbulence [6], offers yet another definition:

> *"Turbulent fluid motion is an irregular condition of the flow in which the various quantities show a random variation with time and space coordinates, so that statistically distinct average values can be discerned."*

It is readily seen that none of these definitions offers any precise characterization of turbulent flow in the sense of predicting, *a priori*, on the basis of specific flow conditions, when turbulence will or will not occur, or what would be its extent and intensity. It seems likely that this lack of precision has at least to some extent contributed to the inability to solve the turbulence problem: if one does not know what turbulence is, or under what circumstances it occurs, it is rather unlikely that one can say much of anything about it in a quantitative sense.

Chapman and Tobak [1] have described the evolution of our understanding of turbulence in terms of three overlapping eras: *i*) statistical, *ii*) structural and *iii*) deterministic. We shall further explore this viewpoint in the next section, but here we point out that a more precise definition of turbulence is now possible within the context of ideas from the deterministic era. Namely,

> *"Turbulence is any chaotic solution to the 3-D Navier–Stokes equations that is sensitive to initial data and which occurs as a result of successive instabilities of laminar flows as a bifurcation parameter is increased through a succession of values."*

While this definition is still somewhat vague, it contains specific elements that permit detailed examination of flow situations relating to turbulence. In the first place, it specifies equations—the Navier–Stokes equations—whose solutions are to be associated with turbulence. By now it is widely (essentially universally) accepted that the N.–S. equations may exhibit turbulent solutions, while previous definitions have failed to explicitly acknowledge this. Second, it requires that the fluid behavior be chaotic, *i.e.*, erratic, irregular, as required in earlier definitions, but deterministic and not random (because the N.–S. equations are deterministic). This is in strong contrast to, especially, Hinze's definition, but it is overwhelmingly supported by experimental data. Third, we require that turbulence be three dimensional. This is consistent with the common classical viewpoint (see, *e.g.*, Tennekes and Lumley [7]) where generation of turbulence is ascribed to vortex stretching which can only occur in 3D as will be considered in more detail below. But we comment that chaotic solutions have been obtained from 1-D and 2-D versions of the N.–S. equations, suggesting that chaos is not necessarily turbulence, even when associated with the N.–S. equations.

The modern definition also imposes a requirement of "sensitivity to initial data" which allows one to distinguish highly irregular laminar motion (such as arises in a quasiperiodic flow regime) from actual turbulence. This is lacking in older definitions despite the fact that experimental evidence has always suggested such a requirement. We also comment that this provides a direct link to modern mathematical theories of the N.–S. equations in the sense that sensitivity to initial conditions (SIC) is the hallmark characteristic of the "strange attractor" description of turbulence first put forward by Ruelle and Takens [8].

We remark that the notion of loss of stability of the laminar flow regime(s) has both classical and modern roots. Stability analyses in the context of, mainly, normal mode analysis has been a mainstay in studies of fluid motion for at least a century, and their connections to transition to turbulence were already made in boundary layer studies. The modern contribution is to embed such approaches within bifurcation theory, thus opening the way to use of many powerful mathematical tools of modern analysis of dynamical systems.

We close this section with a list of physical attributes of turbulence that for the most part summarizes the preceding discussions and which are essentially always mentioned in descriptions of turbulent flow. In particular, a turbulent flow can be expected to exhibit all of the following features:

1. disorganized, chaotic, seemingly random behavior;

2. nonrepeatability (*i.e.*, sensitivity to initial conditions);

3. extremely large range of length and time scales (but such that the smallest scales are still sufficiently large to satisfy the continuum hypothesis);

4. enhanced diffusion (mixing) and dissipation (both of which are mediated by viscosity at molecular scales);

5. three dimensionality, time dependence and rotationality (hence, potential flow cannot be turbulent because it is by definition irrotational);

6. intermittency in both space and time.

We note that there are several views of intermittency, as will be apparent later; but for now we simply take this to be related to the percentage of time a flow exhibits irregular temporal behavior at any selected spatial location.

1.3 A Brief History of Turbulence

In this section we will first briefly review some of the main highlights of the evolution of ideas associated with the problem of turbulence and its treatment, beginning with da Vinci and proceeding chronologically to the beginnings of the modern era of computational work by Orszag and coworkers (*e.g.*, Orszag and Patterson [9]) and Deardorff [10], and including mention of key theoretical and experimental works as well. We then consider somewhat more recent work such as that of Launder and Spalding [11], Rogallo and Moin [12], Kim *et al.* [13], Lesieur [14] and others. Finally, we summarize a paper by Chapman and Tobak [1] that casts this evolution in a particularly interesting light.

The goal of this section is to provide an indication of not only where we are in turbulence research in the present era, but just as important, also how we got to where we are.

1.3.1 General overview

As we have already noted, our earliest recognition of turbulence as a distinguished physical phenomenon had already taken place by the time of da Vinci (*circa* 1500). But there seems to have been no substantial progress in understanding until the late 19^{th} Century, beginning with Boussinesq [15] in the year 1877. His hypothesis that *turbulent stresses are linearly proportional to mean strain rates* is still the cornerstone of most turbulence models, and is likely to be invoked (sometimes subtley) at some point in the derivation even when it is not directly used. It is interesting to note that Boussinesq himself was quite wary of the hypothesis and presciently warned that determination of "eddy viscosities" (the "constant" of proportionality) linking turbulent stress to mean strain rate would be difficult, if not completely impossible; but this has not deterred the efforts of investigators for well over a century.

Osborne Reynolds' experiments, briefly described above, and his seminal paper [16] of 1894 are among the most influential results ever produced on the subject of turbulence. The experimental results led to identification of the Reynolds number as the only physical parameter involved in transition to turbulence in a simple incompressible flow over a smooth surface, and moreover they suggested that only a few transitions were required to reach a turbulent condition (a fact that was not fully recognized until late in the 20^{th} Century—and possibly still is not held universally). The views and analyses of the 1894 paper set the "way of seeing" turbulence for generations to come. In particular, Reynolds concluded that turbulence was far too complicated ever to permit a detailed understanding, and in response to this he introduced the decomposition of flow variables into mean and fluctuating parts that bears his name, and which has resulted in a century of study in an effort to arrive at usable predictive techniques based on this viewpoint. Beginning with this work the prevailing view has been that turbulence is a random phenomenon, and as a consequence there is little to be gained by studying its details, especially in the context of engineering analyses.

It is interesting to note that at approximately the same time as Reynolds was proposing a random description of turbulent flow, Poincaré [17] was finding that relatively simple nonlinear dynamical systems were capable of exhibiting chaotic random-in-appearance behavior that was, in fact, completely deterministic. Despite the fact that French, American and Russian mathematicians continued studies of such systems throughout the early to mid 20^{th} Century, it would be nearly 70 years before the American meteorologist Lorenz [18] would in 1963 first propose possible links between "deterministic chaos" and turbulence.

Following Reynolds' introduction of the random view of turbulence and proposed use of statistics to describe turbulent flows, essentially all analyses were along these lines. The first major result was obtained by Prandtl [19] in 1925 in the form of a prediction of the eddy viscosity (introduced by Boussinesq) that took the character of a "first-principles" physical result, and as such no doubt added significant credibility to the statistical approach. Prandtl's "mixing-length theory," to be analyzed in more detail later, was based on an analogy between turbulent eddies and molecules or atoms of a gas and purportedly utilized kinetic theory to determine the length and velocity (or time) scales needed to construct an eddy viscosity (analogous to the first-principles derivation of an analytical description of molecular viscosity obtained from the kinetic theory of gases). Despite the fact that this approach has essentially never been successful

at making true predictions of turbulent flow, it does a fairly good job at making "postdictions" of certain simple flows.

The next major steps in the analysis of turbulence were taken by G. I. Taylor during the 1930s. He was the first researcher to utilize a more advanced level of mathematical rigor, and he introduced formal statistical methods involving correlations, Fourier transforms and power spectra into the turbulence literature. In his 1935 paper [20] he very explicitly presents the assumption that turbulence is a random phenomenon and then proceeds to introduce statistical tools for the analysis of homogeneous, isotropic turbulence. It is clear that the impact of this has lasted even to the present. In addition, Taylor in this same paper analyzed experimental data generated by wind tunnel flow through a mesh to show that such flows could be viewed as homogeneous and isotropic. The success of this provided even further incentive for future application of the analytical techniques he had introduced. A further contribution, especially valuable for analysis of experimental data, was introduction of the "Taylor hypothesis" which provides a means of converting temporal data to spatial data. This will be presented in more detail later. Other widely-referenced works of this period include those of von Kármán [21], von Kármán and Howarth [22] and Weiner [23].

It is worthwhile to mention that just as Poincaré had provided a deterministic view of chaotic phenomena at the same time Reynolds was proposing a statistical approach, during the period of Taylor's most celebrated work the Frenchman Leray was undertaking the first truly mathematically-rigorous analyses of the Navier–Stokes equations [24], [25] that would provide the groundwork for developing analytical tools ultimately needed for the dynamical systems (deterministic) approach to study of the N.–S. equations and their turbulent solutions.

In 1941 the Russian statistician A. N. Kolmogorov published three papers (in Russian) [26] that provide some of the most important and most-often quoted results of turbulence theory. These results, which will be discussed in some detail later, comprise what is now referred to as the "K41 theory" (to help distinguish it from later work—the K62 theory [27]) and represent a distinct departure from the approach that had evolved from Reynolds' statistical approach (but are nevertheless still of a statistical nature). However, it was not until the late 20^{th} Century that a manner for directly employing the theory in computations was discovered, and until recently the K41 (and to a lesser extent, K62) results were used mainly as tests of other theories (or calculations).

During the 1940s the ideas of Landau and Hopf on the transition to turbulence became popular. They (separately) proposed that as Re is increased a typical flow undergoes an (at least countable) infinity of transitions during each of which an additional incommensurate frequency (and/or wavenumber) arises due to flow instabilities, leading ultimately to very complicated, apparently random, flow behavior (see Hopf [28] and Landau and Lifshitz [29]). This scenario was favored by many theoreticians even into the 1970s when it was shown to be untenable in essentially all situations. In fact, such transition sequences were never observed in experimental measurements, and they were not predicted by more standard approaches to stability analysis.

Throughout the 1940s there were numerous additional contributions to the study of turbulence; we mention only a few selected ones here, and refer the reader to the often extensive reference lists of various of these citations if more details are desired. For the most part, as noted by Leslie [30], this decade produced a consolidation of earlier statistical work (but with the exceptions already discussed above). Works of Batchelor [31], Burgers [32], Corrsin [33], Heisenberg [34], von Kármán [35], Obukhov [36], Townsend [37] and Yaglom [38] are among the most often cited, with those of Corrsin, Obukhov and Townsend involving experiments.

The first full-length books on turbulence theory began to appear in the 1950s. The best known of these are due to Batchelor [39], Townsend [40] and Hinze [6]. All of these treat only the statistical theory and heavily rely on earlier ideas of Prandtl, Taylor, von Kármán and Yaglom (as well as work of the authors themselves, especially in the first two cases), but often intermixed with the somewhat different views of Kolmogorov, Obukhov and Landau. Again, as was true in the preceding decade, most of this work represented consolidation of earlier ideas and served little purpose beyond codifying (and maybe

mystifying?) these notions so as to provide an aura of infallibility that to the present has been difficult to dispel. Moreover, the specific references [39], [40] and [6] presented the problem of turbulence as being so intractable that for several generations few researchers were willing to address it. But it is important to note that experimental work during this period, and even somewhat earlier, was beginning to cast some doubt on the consistency, and even the overall validity, of the random view of turbulence. In particular, already as early as the work of Emmons [41] it was clear that a completely random viewpoint was not really tenable, and by the late 1950s measurement techniques were becoming sufficiently sophisticated to consistently indicate existence of so-called "coherent structures" contradicting the random view of turbulence, as already foreseen in a review by Dryden [42] as early as 1948.

By the beginning of the 1960s experimental instrumentation was improving considerably, although the available techniques were rather crude by modern standards (laser doppler velocimetry and particle image velocimetry were yet to be invented). But the advance that would ultimately lead to sweeping changes in the treatment of turbulence was on the horizon—the digital computer. In 1963 the MIT meteorologist E. Lorenz published a paper [18], based mainly on machine computations, that would eventually lead to a different way to view turbulence. In particular, this work presented a deterministic solution to a simple model of the N.–S. equations which was so temporally erratic that it could not (at the time) be distinguished from random. Moreover, this solution exhibited the feature of sensitivity to initial conditions (later to be associated with a "strange attractor" in the mathematics literature), and thus essentially nonrepeatability. Furthermore, solutions to this model contained "structures" in a sense that much later would be exploited by McDonough *et al.* [43], Mukerji *et al.* [44] and Yang *et al.* [45] and which might, at least loosely, be associated with the coherent structures being detected by experimentalists—although this was not recognized in 1963. The important point to take from this is that a deterministic solution to a model of the N.–S. equations (albeit, a very simple one) had been obtained which possessed several notable features of physical turbulence.

It is useful to further recognize that throughout the 1960s progress was also being made on the mathematical understanding of the N.–S. equations, the long-term effects of which would be very significant. Such studies occurred both in the context of basic analysis (*i.e.*, existence, uniqueness and regularity of solutions) as exemplified in the landmark book of Ladyzhenskaya [46] and in the field of dynamical systems, where the works of Smale [47] in the U. S. and Arnol'd [48] in the Soviet Union are representative among many.

At the same time a new direction was being taken in the attack on the turbulence "closure problem"—*the existence of more unknowns than equations* in the statistical formulations of turbulence. A number of new techniques were introduced beginning in the late 1950s with the work of Kraichnan [49], [50] who utilized mathematical methods from quantum field theory in the analysis of turbulence. These involve use of Fourier representations (both series and transforms), Feynman diagrams and more fundamental (than N.-S.) equations such as the Liouville and Fokker–Planck equations, to approximate higher moments that occur each time an equation for any particular lower moment is constructed. We will not provide details of this work herein since for the most part it represents yet another "blind alley" that contributed more to mystification of the turbulence problem than to its solution. For the interested reader, a quite detailed treatment, written soon after much of the work was completed, can be found in the book by Leslie [30] published in 1973.

There was also significant progress in experimental studies of turbulence during the decade of the 60s. Efforts were beginning to address quite detailed technical aspects of turbulence such as decay rates of isotropic turbulence, return to isotropy of homogeneous anisotropic turbulence, details of boundary layer transitions, transition to turbulence in pipes and ducts, effects of turbulence on scalar transport, *etc.* These include the works of Comte-Bellot and Corrsin [51] on return to isotropy, Tucker and Reynolds [52] on effects of plain strain, Wygnanski and Fiedler [53] on boundary layer transition, Gibson [54] on turbulent transport of passive scalars, and Lumley and Newman [55], also on return to isotropy.

Publication of the seminal paper by Ruelle and Takens [8] in 1971 probably best delineates the beginning of what we will often term a "modern" view of turbulence. In this work it was shown that the N.–S.

equations, viewed as a dynamical system, are capable of producing chaotic solutions exhibiting sensitivity to initial conditions (SIC) and associated with an abstract mathematical construct called a strange attractor. Furthermore, this paper also presents the sequence of transitions (bifurcations) that a flow will undergo as Re is increased to arrive at this chaotic state, namely,

$$\text{steady} \longrightarrow \text{periodic} \longrightarrow \text{quasiperiodic} \longrightarrow \text{turbulent}$$

Availability of such a specific prediction motivated much experimentation during the 1970s and 80s to determine whether this actually occurred. It is to be emphasized that this short sequence of bifurcations directly contradicts the then widely-held Landau–Hopf scenario mentioned earlier.

Indeed, by the late 1970s and early 1980s many experimental results were showing this type of sequence. (In fact, as we mentioned earlier, such short sequences were always seen, but not initially recognized.) But in addition, other short sequences of transitions to turbulence were also confirmed in laboratory experiments. In particular, the period-doubling subharmonic sequence studied by Feigenbaum [56] as well as sequences involving at least some of the intermittencies proposed by Pomeau and Manneville [57] were observed repeatedly and consistently. It should be noted further that other transition sequences, usually involving a "phase-locking" phenomenon, were also observed, particularly in flows associated with natural convection heat transfer (see *e.g.*, Gollub and Benson [58]); but these were still short and in no way suggested validity of the Landau–Hopf view.

Two other aspects of turbulence experimentation in the 70s and 80s are significant. The first of these was detailed testing of the Kolmogorov ideas, the outcome of which was general confirmation, but not in complete detail. This general correspondence between theory and experiment, but lack of complete agreement, motivated numerous studies to explain the discrepancies, and similar work continues even to the present. The second aspect of experimentation during this period involved increasingly more studies of flows exhibiting complex behaviors beyond the isotropic turbulence so heavily emphasized in early work. By the beginning of the 1970s (and even somewhat earlier), attention began to focus on more practical flows such as wall-bounded shear flows (especially boundary-layer transition), flow over and behind cylinders and spheres, jets, plumes, *etc.* During this period results such as those of Blackwelder and Kovasznay [59], Antonia *et al.* [60], Reynolds and Hussain [61] and the work of Bradshaw and coworkers (*e.g.*, Wood and Bradshaw [62]) are well known.

From the standpoint of present-day turbulence investigations probably the most important advances of the 1970s and 80s were the computational techniques (and the hardware on which to run them). The first of these was large-eddy simulation (LES) as proposed by Deardorff [10] in 1970. This was rapidly followed by the first direct numerical simulation (DNS) by Orszag and Patterson [9] in 1972, and introduction of a wide range of Reynolds-averaged Navier–Stokes (RANS) approaches also beginning around 1972 (see *e.g.*, Launder and Spalding [11] and Launder *et al.* [63]). In turn, the last of these initiated an enormous modeling effort that continues to this day (in large part because it has yet to be successful, but at the same time most other approaches are not yet computationally feasible). We will outline much of this in a later chapter.

It was immediately clear that DNS was not feasible for practical engineering problems (and probably will not be for at least another 10 to 20 years beyond the present), and in the 70s and 80s this was true as well for LES. The reviews by Ferziger [64] and Reynolds [65] emphasize this. Thus, great emphasis was placed on the RANS approaches despite their many obvious shortcomings that we will note in the sequel. But by the beginning of the 1990s computing power was reaching a level to allow consideration of using LES for some practical problems if they involved sufficiently simple geometry, and since then a tremendous amount of research has been devoted to this technique. Recent extensive reviews have been provided, for example, by Lesieur and Métais [66] and Meneveau and Katz [67]. It is fairly clear that for the near future this is the method holding the most promise. Indeed, many new approaches are being explored, especially for construction of the required subgrid-scale models. These include the dynamic models of Germano *et al.* [68] and Piomelli [69], and various forms of "synthetic-velocity" models such as those of Domaradzki and coworkers (*e.g.*, Domaradzki and Saiki [70]), Kerstein and coworkers (*e.g.*, Echekki *et al.* [71]) and

McDonough and coworkers (*e.g.*, McDonough and Yang [72]). In these lectures we will later outline the basics of some of these approaches, but will not be able to give complete details. The interested reader can find much of this treated in detail by Sagaut [73].

1.3.2 Three eras of turbulence studies

In 1985 a little-known, but quite interesting, paper was published by Chapman and Tobak [1] in which a rather different view of the evolution of our views on turbulence was presented. The authors divide the century between Reynolds' experiments in 1883 to the then present time 1984 into three overlapping "movements" that they term *statistical*, *structural* and *deterministic*. Figure 1.3 provides a sketch similar to the one presented in [1] (as Fig. 1.3), but we have included additional entries consistent with discussions of the preceding section.

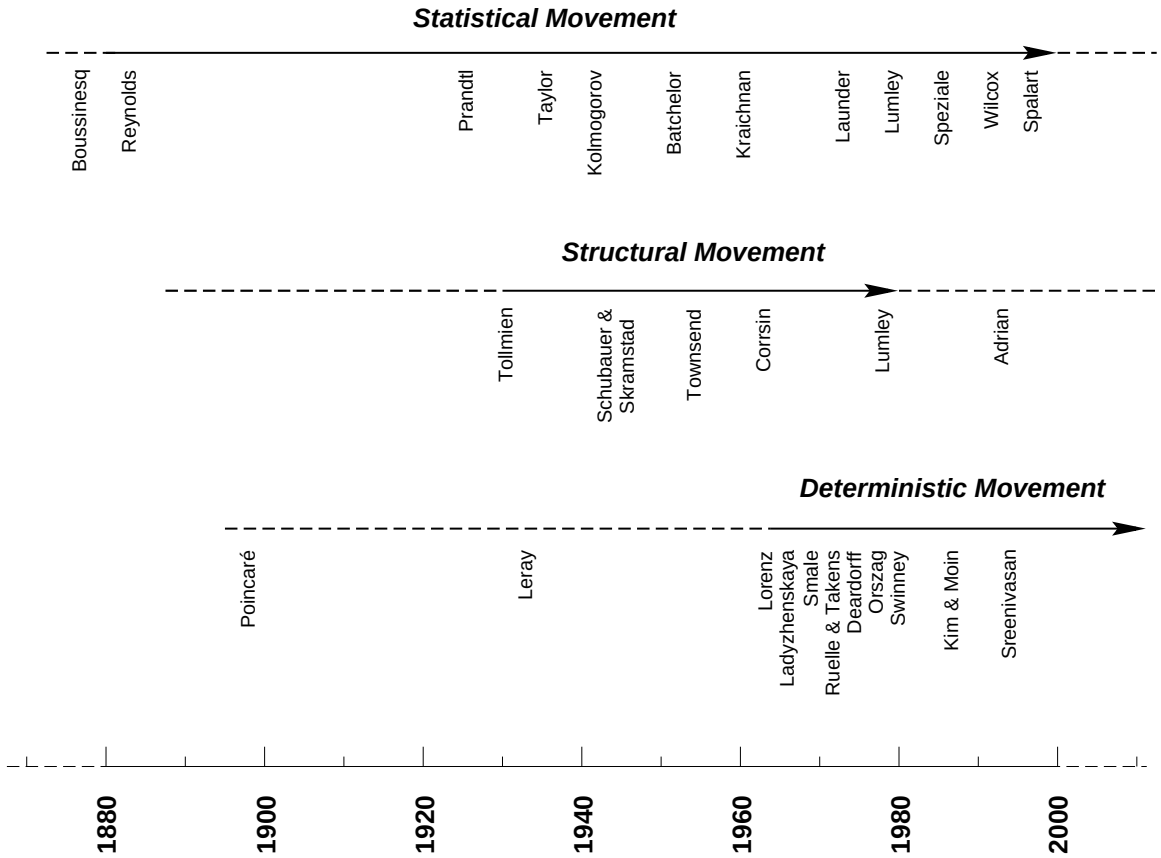

Figure 1.3: Movements in the study of turbulence, as described by Chapman and Tobak [1].

As we noted earlier, the statistical approach was motivated by the view that turbulence must surely be random, and despite repeated experimental contradictions of this interpretation we see the statistical movement extending all the way to the present with recent work now attempting to combine RANS and LES approaches. One of the more interesting contradictions of this era arises from the fact that very early many researchers were already accepting the N.–S. equations as being the correct formulation of turbulent flow. But these equations are deterministic, so a question that should have been asked, but evidently was not, is "How can a deterministic equation exhibit a random solution?" We comment that there are superficial answers, and the reader is encouraged to propose some; but in the end, solutions to the N.–S. equations are deterministic, leaving the following choice: either accept the N.–S. equations as the correct description of turbulent flow and admit turbulence is not random, or seek a completely different description, possibly

based on stochastic differential equations. Moreover, if we insist that turbulence is a random phenomenon, then averaging the N.–S. equations as is done in RANS approaches makes little sense—we would be starting with the wrong equations and yet ending with equations that are not stochastic.

The structural movement is viewed by Chapman and Tobak as having begun possibly with the Schubauer and Skramstad [74] observations of Tollmien–Schlicting waves in 1948; but as we have already noted, even the early experiments of Reynolds indicated existence of coherent structures and short bifurcation sequences. In any case, this movement too persists even to the present, and much research on detecting and analyzing coherent structures in turbulent flows continues.

In [1] the result of the statistical movement is summarized as "a structureless theory having little power of conceptualization," and we add, also little power of prediction at least in part as a consequence of the lack of structure. By way of contrast, the same authors characterize the structural movement as having produced "structure without theory." Because of the massive amounts of data that have arisen during experimentation it has been difficult to construct a theory, but in some respects it is not clear that there actually is much structure either.

Chapman and Tobak show the deterministic movement beginning with the work of Lorenz [18] mentioned earlier but also note that one could easily include studies as far back as those of Poincaré [17]. After describing a considerable body of research up to 1984 they conclude that while results of the deterministic movement are encouraging, as of that date they had not yet provided a successful approach to simulating turbulent flows. (Indeed, even to the present, deterministically-based techniques are often criticized for this same reason. DNS is too expensive for practical simulations, and essentially none of the efficiently-computed modeling techniques that might be directly linked to the deterministic approach have proven themselves.) The authors of [1] then conclude the paper by expressing the belief that future directions in the study of turbulence will reflect developments of the deterministic movement, but that they will undoubtedly incorporate some aspects of both the statistical and structural movements. We comment, that in a sense this is proving to be the case. Certainly, LES can be viewed as a product of the deterministic movement in that the large energy-containing scales are directly computed (as in DNS). On the other hand, LES might also be seen in the light of the statistical movement because the subgrid-scale (SGS) models are usually based on a statistical approach. At the same time, there are beginning to be other approaches to SGS model construction that do, in at least an indirect way, incorporate aspects of the structural and deterministic movements.

1.4 Definitions, Mathematical Tools, Basic Concepts

In this final section of the chapter we begin our formal studies of turbulence. We will do this by first introducing a quite broad range of definitions of various terms that are widely used (and too often not defined) in turbulence studies, many of which are by now completely taken for granted. Without these definitions a beginning student can find reading even fairly elementary literature rather difficult. We continue in a second subsection with a number of widely-used mathematical constructs including various forms of averaging, decompositions of flow variables, Fourier series and transforms, *etc.* Then, in a final subsection we introduce some further basic concepts that often arise in the turbulence literature; to some extent this will provide further discussions and applications of terminology appearing in the first two subsections.

1.4.1 Definitions

In this section we introduce many definitions and terminology to be used throughout these lectures. We recognize that there is a disadvantage to doing this at the beginning in that nearly all of these will of necessity be given out of context. On the other hand, they will all appear in one place, and they will be numbered for easy later reference; thus, this subsection provides a sort of glossary of terms. The reader will recognize that some of these will have already been used with little or no explanation in earlier sections;

so hopefully contents of this section will help to clarify some of the earlier discussions. In general, we will attempt to present ideas, rather than formulas, in the present section, and then provide further elaboration, as needed, in the remaining two subsections of the current section.

We will somewhat arbitrarily classify the terminology presented here in one of three categories, although it will be clear that some, if not all, of the terms given could (and sometimes will) appear in one or both of the remaining categories. These will be identified as: *i*) purely statistical, *ii*) dynamical systems oriented, and *iii*) physical and computational turbulence.

Purely Statistical

In this section we provide a few definitions of statistical terms often encountered in the study of turbulence (especially in the classical approach) and elsewhere.

Definition 1.1 (Autocorrelation) *Autocorrelation is a function that provides a measure of how well a signal "remembers" where it has been; it is an integral over time (or space) of the value of the signal at a given time multiplied times a copy with shifted time (or space) argument.*

Values of autocorrelation are usually scaled to range between -1 and $+1$.

Definition 1.2 (Cross correlation) *Cross correlation provides a measure of how closely two signals (or functions) are related; it is constructed as a scaled inner product of the two functions integrated over a domain (temporal or spatial) of interest.*

If two signals are identical, they have a cross correlation equal to unity, and if they everywhere have equal magnitudes but opposite signs, their cross correlation will be -1. All other possible values lie between these limits.

Definition 1.3 (Ergodicity) *Ergodicity implies that time averages and ensemble averages are equivalent.*

Note that this is a consequence of the definition, and not actually the definition, itself. (The formal definition is rather technical.) The averages mentioned here will be treated in detail in the next section.

Definition 1.4 (Flatness) *Flatness requires an equation for formal definition. Here we simply note that it represents the deviation from Gaussian in the sense that functions having large flatness values are more sharply peaked than are Gaussian distributions, and conversely.*

Flatness (sometimes called "kurtosis") is always greater than zero, and flatness of a Gaussian is exactly three (3).

Definition 1.5 (Probability density function (pdf)) *The probability density function of a random (or otherwise) variable expresses the probability of finding a particular value of the variable over the range of definition of the variable.*

As with most of the definitions provided here, we will treat this in more detail, and with more formal mathematics, later. The simplest "picture" to remember for a pdf is a histogram scaled so that its area is unity.

Definition 1.6 (Random) *A random variable, function, or number is one whose behavior at any later time (or place) cannot be predicted by knowledge of its behavior at the present time (or place).*

It is worth mentioning that the autocorrelation of a random signal decays to zero very rapidly.

Definition 1.7 (Skewness) *Skewness is a measure of asymmetry of a function with respect to the origin (or elsewhere).*

Skewness can take on both positive and negative values, and that observed in turbulence experiments is usually (but not always) negative. Just as will be the case for flatness, we will later provide a specific formula by means of which to calculate skewness.

Definition 1.8 (Stochastic) *A stochastic variable is one whose autocorrelation decays to zero exponentially fast.*

The notions of randomness and stochasticity are often used interchangeably, but this is not formally correct. Indeed, a deterministic behavior can exhibit stochasticity, but random behavior is always stochastic. Hence, random $\Rightarrow$ stochastic, but not conversely. A consequence of this is that it makes perfect sense to apply statistical tools in the analysis of deterministic dynamical systems, and this is often done.

Dynamical Systems Oriented

Here we introduce numerous definitions associated with dynamical systems, *per se*, and also with applied mathematics of the Navier–Stokes equations.

Definition 1.9 (Attractor) *An attractor is a region in phase space to which all trajectories starting from within the basin of attraction are drawn after sufficiently long time, and remain there.*

Various of the terms in this definition are, themselves, undefined. This will be taken care of below. We should also point out that there are formal technical mathematical definitions for attractor. We will not need that level of rigor in these lectures.

Definition 1.10 (Basin of attraction) *The set of initial data whose trajectories reach the associated attractor.*

It should be mentioned that the basin of attraction is not a trivial notion. Basins can be fractal in nature, implying that changing the value of the initial point by an infinitesimal amount might result in a drastic change in long-time behavior of the dynamical system. In particular, this is associated with the questions of uniqueness and stability of solutions to initial value problems.

Definition 1.11 (Bifurcation) *A bifurcation (usually termed, simply, a transition in the fluid dynamics literature) is a discontinuous qualitative change in system behavior as a (bifurcation) parameter (say, the Reynolds number) moves continuously through a critical value.*

There are formal mathematical definitions associated with bifurcation, and we will introduce these later, as needed.

Definition 1.12 (Cantor set) *The Cantor set is a set formed by starting with the unit interval, and then discarding the middle third. This leaves a one-third subinterval on each end. Now we remove the middle third from each of these, and continue this process ad infinitum.*

What remains at the end of this process is a set still containing an uncountable infinity of points, but which has zero measure (think length). Moreover, its fractal dimension can be shown to be $\sim 0.62\ldots$.

Definition 1.13 (Chaotic) *Chaotic is the terminology now used to connote erratic, irregular, disorganized, random-in-appearance behavior that is, in fact, deterministic.*

Definition 1.14 (Codimension) *Codimension refers to the dimension of the space formed by the bifurcation parameters of a dynamical system.*

As with most of our definitions in this section, this is not precise, and other definitions are sometimes used. In any case, for our purposes the codimension of a dynamical system is simply its number of bifurcation parameters. It is worthwhile to note that systems with codimension greater that two are presently essentially impossible to treat analytically.

Definition 1.15 (Critical value) *A value of a bifurcation parameter at which the qualitative behavior of the system changes, i.e., a "bifurcation point."*

This corresponds to the situation in which a particular solution to the equations representing a dynamical system is stable for values of the bifurcation parameter that are less than the critical value, and unstable for those that are greater. As the parameter value passes through the critical one, the original solution (which is still a solution) loses stability and is no longer observed (either physically or computationally) and is replaced with a different stable solution.

Definition 1.16 (Delay map) *A delay map is a phase space construction (generally representing an attractor) obtained by plotting the value of a variable against a second, shifted-in-time (delayed), value of the same variable as time evolves.*

We observe that delay maps are particularly valuable when only incomplete data associated with an attractor are available. In particular, there is a theorem due to Takens [75] that proves that successive time shifts of data, to each of which is associated an embedding dimension, allows recovery of the topology of an attractor when only a single variable (out of possibly many) is known completely. The use of this Takens "embedding theorem" can be especially valuable in treating experimental data where only limited measurements may have been taken.

Definition 1.17 (Deterministic) *Deterministic implies predictable, at least for short times. Any behavior described by differential and/or algebraic systems possessing no random coefficients or forcings can be expected to be deterministic.*

It is important to note here that predictability may, in fact, be for only a short time. Indeed, deterministic chaos is of precisely this nature—predictable, but not for very long.

Definition 1.18 (Dynamical system) *A dynamical system is a mathematical representation, via (usually) differential and/or algebraic equations, of a physical (or otherwise) system evolution starting from prescribed initial conditions.*

Associated with this dynamical system will be a formal solution operator (in the form of a semigroup) that maps initial data to results at later times. See Frisch [80] for a precise definition.

Definition 1.19 (Embedding dimension) *The embedding dimension is the dimension of a phase space reconstructed by delays of (usually) a single time series.*

An important question, associated with having only a single recording for the behavior of a multicomponent dynamical system, is "How large should the embedding dimension be (that is, how many time-delayed variables must be constructed) to guarantee that the topology of the phase space representation is equivalent to that of the original dynamical system?" This is discussed in considerable detail by Takens [75].

Definition 1.20 (Feigenbaum Sequence) *A Feigenbaum bifurcation sequence is one associated with period-doubling (subharmonic) bifurcations.*

We remark that such sequences had been know prior to Feigenbaum's studies, but previous work did not match the detail or depth of his investigations. Primarily through numerical experiments he was able to show that quadratic maps exibit universal behaviors that can be characterized by two constant determined by him (see [56]).

Definition 1.21 (Flow) *The set of trajectories of a dynamical system generated from all possible initial conditions associated with a particular attractor.*

Definition 1.22 (Fractal) *An object whose measured linear scales increase with increase in precision of the measurements is termed fractal.*

We remark that the length of the coastline of Great Britain is often cited as an example of this. Furthermore, for usual fractals as the scale on which measurements (or simply observations) is decreased, the object still "looks the same on every scale." This is due to what is termed "self similarity."

Definition 1.23 (Fractal dimension) *A fractal dimension is a generally non-integer dimension that in some way characterizes the structure of a fractal, and which collapses to the usual integer values of dimension when applied to ordinary, nonfractal objects.*

It is worth commenting here that there are numerous ways to compute fractal dimension, and none is particularly easy to carry out. Because a non-integer dimension is considered to be one of the characterizations of a strange attractor, much effort was expended on arriving at good computational procedures throughout the 1980s. However, none of these were completely successful, and other approaches to characterizing a strange attractor are now more often used.

Definition 1.24 (Hilbert space) *A Hilbert space is a complete, normed, linear space equipped with an inner product that induces the norm of the space.*

The notion of a Hilbert space is incredibly important in the mathematics of the N.–S. equations. In the present lectures, however, we will touch on this only lightly.

Definition 1.25 (Inertial manifold) *An inertial manifold is a finite-dimensional manifold associated with the solution operator (semigroup) of a dynamical system that is (positively) invariant under the semigroup and exponentially attracts all orbits of the dynamical system.*

The fact that the inertial manifold of a dynamical system is finite-dimensional implies the possibility of being able to actually compute trajectories on it. If no such object exists, it would be impossible to employ DNS to solve the N.–S. equations at high Re.

Definition 1.26 (Intermittency) *Intermittency, in the sense of dynamical systems, is one of three phenomena associated with switching between nearly steady and chaotic behavior, or between periodic and chaotic behavior.*

The detailed mathematical definition of intermittency requires more background than is expected of readers of the present lectures. The interested reader can consult Bergé *et al.* [77] and Frisch [80] and references therein.

Definition 1.27 (Invariant sets, manifolds, tori) *The term invariant can be applied to any of these mathematical objects, and others, and refers to the property that the dynamical system (viewed as a semigroup) can map the entire, say manifold of trajectories, back to itself. Then the manifold is said to be invariant.*

It is important to note that invariance is not as strong a property as might be thought. In particular, under invariance alone there is no guarantee that points of a manifold will be mapped back to themselves—only that all points of the (invariant) manifold will be mapped back to some point(s) of the manifold by the dynamical system.

Definition 1.28 (Limit cycle) *A limit cycle of a dynamical system is a closed trajectory that is followed repeatedly (cyclicly) for all time.*

Definition 1.29 (Lorenz attractor) *The Lorenz attractor is the attractor associated with solutions to the Lorenz [18] equations.*

The Lorenz attractor was the first strange attractor to be computed numerically, and when this was carried out in 1963 the terminology "strange attractor" had not yet appeared in the literature. The importance of this attractor is that it was the first to mathematically (albeit, numerically) suggest a connection between turbulence and deterministic behavior simply due to the fact that the Lorenz equations comprise a very low order Galerkin approximation to the N.–S. equations.

Definition 1.30 (Lyapunov exponent) *The Lyapunov exponent of a dynamical system provides a time-averaged measure of the rate of divergence of two trajectories that were initially nearby.*

We note that positive Lyapunov exponents are generally taken to be a sure sign of a strange attractor (in the case that there is an attractor), and negative ones imply decay of solution behavior to a steady state. Because of their significance in identifying strange attractors, much effort has been devoted to accurate, efficient computational techniques for their determination. Some theoretical details can be found in [77], as well as in many other references, and Wolf *et al.* [78] provide one of the most widely-used computational procedures. We comment that like fractal dimension calculations, Lyapunov exponent results are always subject to some uncertainty, and obtaining them tends to be very computationally intensive. As a consequence, they are not nearly as widely used today as was the case in the 1980s.

Definition 1.31 (Manifold) *A manifold is a set whose points can be (at least locally) put into correspondence with the points of a Euclidean space of the same dimension.*

There are very precise and complicated definitions of manifolds, and many types of manifolds (the interested reader can consult Wiggins [79] for some details), but here we note that identifying manifolds with smooth surfaces is not a bad first approximation—although it will fail in the case of a strange attractor. Also, n-dimensional Euclidean space is a manifold for any finite value of n. Thus, in the case $n = 3$ we have been living in a manifold all our lives.

Definition 1.32 (Multifractal) *Multifractal refers to a fractal that has multiple fractal character in the sense that scaling to smaller and smaller scales is not self similar as it is in the case of an ordinary fractal.*

Definition 1.33 (Orbit) *The term orbit refers to a trajectory of a dynamical system that has reached an attractor and thenceforth continues to, in some sense circle around (orbit) the attractor.*

The terms orbit and trajectory are obviously closely related, and while they are not completely equivalent they are often used interchangeably.

Definition 1.34 (Phase lock) *Phase lock is a type of periodicity that can arise after a system has reached a quasiperiodic state (two incommensurate frequencies), and further increases in a bifurcation parameter leads to returning to a two-frequency commensurate state in which the two frequencies are integrally related, and said to be "locked."*

In the physics literature, this condition is often termed *resonance*, and in some dynamical systems contexts (particularly discrete dynamical systems) the state is called *n-periodic* because the power spectrum will show n strong equally-spaced spikes starting with the lower of the two frequencies and ending with the higher. The $n-1$ highest frequencies might be viewed as harmonics of the lowest one, but this is not an accurate interpretation because of the existence of two frequencies to begin with.

Definition 1.35 (Phase portrait) *A phase portrait is a plot of two (or possibly three) components of a dynamical system against one another as time evolves.*

A simple case of this is the limit cycle often pictured as part of the analysis of stability of periodic solutions. It is also worth noting that the delay maps described above provide an approximation of the phase portrait.

Definition 1.36 (Phase space) *Phase space is the region in which a phase portrait (or delay map) is plotted.*

Definition 1.37 (PMNS equation) *The PMNS equation (poor man's Navier–Stokes equation) was the name given to a very simple discrete (algebraic) dynamical system by Frisch [80] when arguing that despite its simplicity, it had many of the aspects of the N.–S. equations.*

McDonough and Huang [81], [82] later showed that a somewhat more complicated discrete system could be derived directly from the N.–S. equations and, moreover, that this still relatively simple algebraic system was capable of exhibiting essentially any temporal behavior found in the N.–S. partial differential system.

Definition 1.38 (Poincaré map) *A Poincaré map gives the location of the next passing of a trajectory through a Poincaré section based on the current location of the trajectory.*

We note that in general Poincaré maps cannot be derived analytically; they must usually be constructed numerically. They provide mainly a qualitative description of overall behavior of a dynamical system; but because they are relatively simple (and easily computed), it has been suggested that they might hold potential for use in turbulence models.

Definition 1.39 (Poincaré section) *A Poincaré section is a slice (a plane) through the multi-dimensional torus containing the trajectories of a dynamical system.*

It is not hard to deduce that if the trajectories behave in a chaotic fashion, their points of intersection with any Poincaré section will show signs of irregularity as well.

Definition 1.40 (Pomeau–Manneville Sequence) *A Pomeau–Manneville bifurcation sequence is one in which chaos arises through a route of behaviors including so-called "intermittencies." There are at least three distinct types of these (see definition, above). More detailed information can be obtained from the earlier cited references, and from [57].*

Definition 1.41 (Power spectral density, PSD) *The power spectrum of a signal is a function that gives the power (amplitude) of the signal (usually in decibels, dB) as a function of frequency or wavenumber.*

Definition 1.42 (Quasiperiodic) *A quasiperiodic signal is one containing two incommensurate (not rationally related) frequencies (or wavenumbers).*

As we have indicated earlier, at least in the case of observations and theory associated with N.–S. flows, the quasiperiodic state often occurs as the second bifurcation beyond steady state.

Definition 1.43 (Recurrent) *A recurrent behavior in a dynamical system characterizes the fact the the system will repeatedly (recurrently) return to any, and all, states of a stationary configuration with an infinity of such occurrences.*

We observe that limit cycles are the simplest examples of a recurrent regime, but stationary turbulence is typically also of this nature.

Definition 1.44 (Ruelle–Takens Sequence) *The Ruelle and Takens bifurcation sequence consists of no more than four bifurcations leading to a strange attractor, but the sequence can be as short as three bifurcations. It typically consists of two Hopf bifurcations from steady behavior (the first to simple periodicity, and the second to quasiperiodicity), followed by a bifurcation to a strange attractor.*

We remark that laboratory experiments very early identified modifications of this sequence. In particular, the studies of Gollub and Benson [58] should many different forms of phase-locked behavior; these authors note that the specific forms of phase lock and their embeddings within the Ruelle–Takens sequence appear to not be repeatable in laboratory experiments, *i.e.*, they are sensitive to initial conditions.

Definition 1.45 (Self similarity) *Self similarity is a property often possessed by fractals leading to a structure that looks the same on all scales.*

This property leads to a particular mathematical representation, and one that must be altered in the case of multifractality.

Definition 1.46 (Sensitivity to initial conditions) *Sensitivity to initial conditions (SIC) implies that the behavior of a dynamical system can change drastically due to small changes in the initial data.*

SIC is one of the main attributes of a strange attractor, and in the context of actual physics is the reason weather cannot be predicted for more than a few days in advance.

Definition 1.47 (Shell model) *Shell models are dynamical systems constructed so as to minimize the number of equations (termed "modes" because they typically are obtained via a Galerkin procedure and Fourier analysis) needed to reproduce specific aspects of a physical pheneomenon.*

Shell models have often been constructed to efficiently mimic behavior of the N.–S. equations, at least locally in Fourier space. In this case they are usually designed to reproduce certain symmetries and conservation properties of the N.–S. equations. See Bohr *et al.* [83] for extensive analyses.

Definition 1.48 (Stable, and unstable, manifold) *A stable manifold consists of the set of points in the trajectories of a dynamical system that approach a specific behavior (say, a fixed point) as $t \longrightarrow \infty$. Corresponding to this is the notion of an unstable manifold which exhibits the same behavior as $t \longrightarrow -\infty$.*

We note that this definition is lacking precision, but we will not have much specific need for it in the sequel.

Definition 1.49 (Stationary) *A dynamical system can be said to exhibit stationary behavior, or stationarity, if it is recurrent and can be assigned a well-defined average with respect to time.*

This, like most of our definitions, is rather imprecise and heuristic; but it contains the basic idea: stationary behavior is, in general, time dependent—it is usually not steady—but it can be viewed as fluctuations about some mean. Frisch [80] provides a precise and quite mathematically-complicated definition, while Tennekes and Lumley [7] use one even simpler than that given here. One should note that steady (time-independent) behavior implies stationarity, but not conversely.

Definition 1.50 (Strange attractor) *A strange attractor is an attractor of a dynamical system (see Def. 1.9 of this section) that can be viewed as having been constructed as the Cartesian product of a smooth manifold and a Cantor set.*

In particular, one might envision a simple strange attractor as being the interior of a two-torus that has had nearly all of its points removed in a process analogous to the Cantor set construction described in Def. 1.12. A more accurate view comes from recognizing that three main geometric processes occur in the generation of a manifold corresponding to a strange attractor: *i*) shrinking in one or more directions and stretching in one or more directions, *ii*) rotation, and *iii*) folding. Indeed, it can be shown that when these operations are applied recursively to the flow of a dynamical system, the outcome is a multi-dimensional version of a Cantor set. See Alligood *et al.* [85] and Lanford [86] for more details.

Definition 1.51 (Strong solution) *A strong solution (mainly in the context of partial differential equations) is one that satisfies the equations expressed in classical form almost everywhere (a.e.) on the problem domain of definition.*

We remark that the notion of strong solution takes on many different forms, but the important aspect, as the definition implies, is that the solution is sufficiently differentiable to make sense of the differential equation on all but a set of measure zero in the domain of the problem being considered.

Definition 1.52 (Subharmonic) *Subharmonic refers to the appearance of the power spectrum corresponding to a signal arising from a period-doubling (or, subharmonic) bifurcation.*

The power spectrum will exhibit frequencies at the half frequencies of each of the harmonics of the original signal (including, and especially) the fundamental. One can think of this in phase space as arising for an attractor consisting of a simple limit cycle (periodic behavior) after an increase in the bifurcation parameter changes the dynamics so that the trajectory does not quite return to its starting point after one orbit, and requires a second orbit to return. This doubles the period and halves the frequency.

Definition 1.53 (Trajectory) *A trajectory is the point set in phase space of the path followed by a dynamical system starting from a single point of initial data.*

Definition 1.54 (Weak solution) *A weak solution to a differential equation is one that does not have sufficient differentiability to permit substitution of the solution into the differential equation.*

This notion is an extremely important one in the analysis of the N.–S. equations, and it is associated with integral forms of these (and other) equations that permit moving derivatives off of formally non-differentiable solutions and onto infinitely-differentiable test functions via integration by parts within a linear functional. In fact, the formal definition of such solutions does not alone admit pointwise evaluation, but practical constructions such as use of Galerkin procedures do yield solutions that can be evaluated in the usual pointwise sense—but which, nevertheless could have no derivatives. Stakgold [87] provides a good treatment of these ideas.

Physical and Computational Turbulence

In this section we present a wide assortment of terms associated with both the physical and computational aspects of turbulence.

Definition 1.55 (Backscatter) *Backscatter is the name often given to the process of transferring turbulence kinetic energy from small scales to large ones, or equivalently, from high wavenumbers to low ones.*

It is interesting to note that the early theories of turbulence did not include backscatter (despite the fact that Fourier analysis of the N.–S. equations suggests it must occur), and instead viewed transfer of energy as going only from large scales to small.

Definition 1.56 (Boussinesq hypothesis) *The Boussinesq hypothesis states that small-scale turbulent stress should be linearly proportional to the mean (large-scale) strain rates.*

While this idea is an appealing one, since it is somewhat analogous to Newton's law of viscosity, there is no physical reason why it should hold, and, in fact, in almost all circumstances it does not.

Definition 1.57 (Buffer layer) *The buffer layer is the region in wall-bounded flows between the viscous sublayer and the log layer.*

Definitions of viscous sublayer and log layer are given below, and as might be expected, they are physical locations in wall-bounded flows. Intuitively, one should expect that these separate regions (flow behaviors) must be "matched" both physically and mathematically to obtain a well-defined global treatment of such a flow. The buffer layer is this matching, overlap region.

Definition 1.58 (Closure) *Closure refers to finding expressions for terms that arise in averaged and filtered versions of the N.–S. equations and for which there are no fundamental equations.*

The "closure problem" is the outstanding difficulty with any RANS or LES method: there are more unknowns than there are equations, and this leads to the need for modeling in order to produce a "closed" system of equations.

Definition 1.59 (Coherent structures) *Coherent structures are not-very-well-defined behaviors in a turbulent flow, but which are identified as being easy to "see," may or may not be of fairly large scale, and are somewhat persistent.*

As can be seen, there is little about these characterizations that is precise, and this among other things makes looking for coherent structures in experimental data a subjective activity. It is clear that such structures exist and are deterministic, but endowing them with a rigorous quantification is difficult.

Definition 1.60 (Cross stress) *Cross stresses arise in the formal analysis of LES and consist of products of resolved- and unresolved-scale quantities.*

These arise in LES specifically because a filtered fluctuating quantity is not generally zero. The analogous quantities do not occur (*i.e.*, they are identically zero) in RANS formulations.

Definition 1.61 (Deconvolution method) *Deconvolution methods are techniques employed to construct subgrid-scale models in large-eddy simulation by extracting information from the highest resolved wavenumber parts of a solution and using this to infer behavior of the lowest wavenumber unresolved parts.*

Definition 1.62 (Defect layer) *The defect layer, sometimes simply called the "outer layer," is the region of a turbulent boundary layer beyond the log layer (see below), and beginning approximately one-tenth the boundary layer thickness from the wall.*

Definition 1.63 (DES) *DES is an acronym for detached-eddy simulation. This is a fairly recent development in turbulence modeling that attempts to improve the efficiency of LES by using RANS approaches in near-wall regions where usual LES would require high resolution to achieve even moderate accuracy.*

It should be noted that there are at least two fundamental difficulties with such an approach. The first is that, as we will see in the next section, the RANS and LES decompositions of dependent variables correspond to very different mathematics, so trying to match their results at any particular set of points in a flow field is problematic. Second, and associated with this, is the need to determine where this match should occur. On the other hand, proponents of this approach argue that formally the equations of LES and RANS are almost identical (if we ignore the point made earlier that RANS must correspond to steady state), and, even more, the outcome of modeling in either case is increased (numerical) dissipation.

Definition 1.64 (Dissipation range) *The dissipation range refers to the length (or, equivalently, wavenumber) scales on which the dominant physical phenomenon is viscous dissipation.*

The dissipation range corresponds to small length scales (correspondingly, large wavenumbers), and at such scales kinetic energy of fluid motion is converted to thermal energy.

Definition 1.65 (DNS) *DNS stands for for the numerical procedure direct numerical simulation. This corresponds to solving the Navier–Stokes equations on a digital computer using sufficient resolution to capture all physically important scales from the largest to the dissipation scales.*

It is to be emphasized that no modeling is employed for DNS; there is no closure problem. Thus, the only problem with this approach is that current computers are not sufficiently large and fast to permit the necessary resolution if Re is high and/or the problem possesses other physical/geometric complications.

Definition 1.66 (Eddy turnover time) *Eddy turnover time is the time taken for a hypothesized turbulent eddy to perform one complete* 360° *rotation.*

Definition 1.67 (Eddy viscosity) *Eddy viscosity is, formally, the constant of proportionality between turbulent (Reynolds) stresses and mean (large-scale) strain rate, analogous to physical viscosity in Newton's law of viscosity.*

Definition 1.68 (Energy cascade) *Energy cascade refers to the transfer of kinetic energy from large, macroscopic scales of motion, where it is presumed to be input to the flow, through successively smaller scales, ending with viscous dissipation and conversion to heat (thermal energy).*

The specific physical mechanism proposed for this is interaction of vortices: large vortices break into smaller ones that, in turn, break into still smaller ones *à la* Richardson's limerick. This view has played a major role in development of turbulence models from the beginning, despite the fact that it is not clear such successions of vortices even exist in many flows. The reader is referred to Tsinober [3] and Kraichnan [88] for a contrary view. We also note here that the energy cascade is often analyzed in terms of dependence of turbulence kinetic energy on Fourier wavenumbers, and in this context each wavenumber is associated with a turbulent eddy size.

Definition 1.69 (Enstrophy) *Enstrophy is the* L^2 *norm (squared) of vorticity.*

Definition 1.70 (Equilibrium) *An equilibrium turbulent flow is one in which production of turbulence kinetic energy is precisely balanced by dissipation of this energy.*

It should be noted that this has nothing, *per se*, to do with thermodynamic equilibrium, or with equilibrium solutions in the sense of stability analyses.

Definition 1.71 (Friction velocity) *Friction velocity is a somewhat artificial construct having dimensions of velocity, but defined in terms of the wall shear stress in a wall-bounded flow.*

Definition 1.72 (Fully-developed turbulence) *Fully-developed turbulence occurs at high Re, and in simple terms can be taken to imply that statistical quantifications of the flow do not change in the flow direction.*

See Frisch [80] for a precise definition.

Definition 1.73 (Galilean invariance) *Galilean invariance implies that measured physics must be the same in any non-accelerating frame of reference. (The term frame invariance is also used.)*

The importance of this property cannot be overstated. If Galilean invariance did not hold for fluid flow (and, thus, for the N.–S. equations), experimentalists making measurements of identical flow situations at different locations on Earth would record different results.

Definition 1.74 (Helicity) *Helicity is the (vector) inner product of the velocity field with its vorticity field.*

We remark that this is not a widely-used concept. On the other hand, Tsinober [3] notes that it "provides a clear indication of direct coupling of large and small scales." Information of this sort may be valuable in construction of advanced turbulence modeling techniques.

Definition 1.75 (Homogeneous turbulence) *Homogeneous turbulence is such that statistical properties do not change with spatial translation, i.e., they do not change with position.*

It should be noted that it is only the turbulent, fluctuating quantities that are being considered. In particular, it is possible to have homogeneous turbulence in a flow field that is nonhomogeneous on large scales. Also, we observe that homogeneity is often viewed with respect to only specified (not necessarily all) directions.

Definition 1.76 (Implicit large-eddy simulation, ILES) *Implicit large-eddy simulation is a relatively new development in LES studies and consists of solving unfiltered equations; i.e., the N.–S. equations, themselves, are solved without recourse to any formal turbulence modeling—but on fairly coarse grids.*

It is easily recognized that dissipation must be introduced in some manner for this type of approach to succeed, and in the case of ILES this is done with highly-dissipative discretizations of nonlinear terms in the N.–S. equations. Thus, the dissipation that would have been produced by a turbulence model in more standard approaches is obtained directly from the numerical procedures in ILES.

Definition 1.77 (Inertial range) *The inertial range, or inertial subrange, is that range of length scales (or wavenumbers) in which viscous effects are essentially negligible.*

Definition 1.78 (Integral scale) *The integral scale is taken to be an* $\sim \mathcal{O}(1)$ *to* $\mathcal{O}(10^{-1})$*, but less than unity, multiple of the geometric scale of a flow field.*

In terms of the energy spectrum, the integral scale is often identified with the wavenumber corresponding to maximum energy. An analogous time scale can also be constructed.

Definition 1.79 (Intermittency) *Intermittency (or intermittency factor) is the fraction of total flow time during which turbulent fluctuations can be observed.*

Clearly, this must be viewed locally; the intermittency factor will not be constant throughout a flow field.

Definition 1.80 (Isotropic) *Isotropic turbulence exhibits statistics that are independent of rotations and reflections. When this is not the case, the turbulence is anisotropic.*

The term *local isotropy* refers to isotropy only on small scales, identified by their (high) Fourier wavenumbers. In addition, it should be noted that isotropy implies that statistics for each of the individual fluctuating velocity components and their normal direction strains are the same. In turn, because this need not be true in a homogeneous flow, we see that isotropy is a stronger condition than homogeneity; *i.e.*, a homogeneous flow may be anisotropic.

Definition 1.81 ($k-\varepsilon$ model) *The $k-\varepsilon$ turbulence models are a class of RANS models for which two additional (beyond those for mean flow quantities) partial differential equations must be solved (one for turbulence kinetic energy, k, and one for turbulence kinetic energy dissipation rate, ε) to obtain length and time scale information needed to construct local (in space) eddy viscosities.*

We observe that presently $k-\varepsilon$ models are among the most popular for practical engineering problems. They produce results that, basically, are no worse than those obtained from more sophisticated approaches, and they are much less expensive in terms of computer run times.

Definition 1.82 (K41 theory) *The K41 theory refers to results published by Kolmogorov in a series of papers [26] in 1941.*

The K41 theory provides two specific, testable results: the 2/3 law which leads directly to the prediction of a $k^{-5/3}$ decay rate in the inertial range of the energy spectrum, and the 4/5 law that is the only exact (contains no adjustable constants) result for N.–S. turbulence at high Re. See Frisch [80], and below, for more details.

Definition 1.83 (Kolmogorov scale) *Kolmogorov scale is another name for dissipation scales.*

These scales were predicted on the basis of dimensional analysis as part of the K41 theory.

Definition 1.84 (Kurtosis) *Kurtosis is another name for the statistical quantity flatness; it is actually flatness−3.*

Definition 1.85 (Landau–Hopf Theory) *The Landau–Hopf theory was one of the earlier mathematical theories for turbulence where it was hypothesized that turbulence resulted from an infinite sequence of bifurcations of solutions to the N.–S. equations in each of which a new incommensurate frequency appeared in the solution representation.*

It should be noted that this view was at one time very attractive to both mathematicians and turbulence theorists. It employed easily-understood mathematics to produce a behavior that seemed at least similar to laboratory observations. Experimentalists, however, quickly recognized that there did not appear to be infinite sequences—but, instead, rather short bifurcation sequences—leading to onset of turbulence. This scenario consisting of short sequences was first proven by Ruelle and Takens [8] and was quickly confirmed by numerous experimentalist specifically investigating this issue. We comment further that the work in [8], associating turbulence with a strange attractor, almost completely invalidates the Landau–Hopf theory because even though the fluid motion can be extremely complicated in this view, it cannot be sensitive to initial conditions—a well-known property of both turbulence and strange attractors.

Definition 1.86 (Large scale) *Large scale refers to that part of a turbulent flow that carries most of the energy and corresponds to length scales not extremely smaller than $\mathcal{O}(1)$, or correspondingly, wavenumbers that are $\sim \mathcal{O}(1)$. The small-scale part of the flow is then that part that does not have these properties; in particular, small scales correspond to high wavenumbers.*

Definition 1.87 (Law of the wall) *The law of the wall is a semi-empirical expression relating velocity to distance from the wall in a turbulent wall-bounded flow.*

There is also a "law of the wake." Further detailed information on these terms can be found in most general texts on turbulent flow (*e.g.*, Tennekes and Lumley [7]), and more details will be provided in later lectures.

Definition 1.88 (Leonard stress) *Leonard stresses arises in the analysis of subgrid-scale stresses in LES models as a product of resolved-scale velocity components times resolved-scale velocity components.*

Observe that these contributions arise due to the fact that a twice-filtered quantity is different from the once-filtered one for most filters. But it should be noted that this contribution to the SGS stresses can be directly computed without modeling.

Definition 1.89 (LES) *LES is the acronym for large-eddy simulation.*

This form of turbulence calculation lies between DNS and RANS methods, both with respect to form and cost to compute. In particular, in LES the large-scale, energy-carrying motion is directly simulated while the small (mainly dissipative) scales are modeled.

Definition 1.90 (Log layer) *The log layer is that part of the boundary layer in a wall-bounded flow in which a logarithmic semi-empirical result accurately describes the variation of velocity with distance from the wall. It corresponds to scaled distances that are relatively far from the wall.*

Definition 1.91 (MILES) *MILES stands for monotonically integrated large-eddy simulation.*

This is a form of LES for which no formal filtering is performed. Instead, a monotone numerical scheme, that introduces artificial viscosity, is applied to discretize the N.–S. equations. This added viscosity is claimed to be roughly equivalent to that supplied by SGS models in usual LES.

Definition 1.92 (Mixing length) *Mixing length is the distance over which a hypothesized turbulent eddy retains its identity.*

There are more precise definitions of this quantity (see, *e.g.*, Wilcox [90]) as will be discussed in the sequel.

Definition 1.93 (Moment methods) *Moment methods is a generic name sometimes used in reference to any turbulence computational approach based on RANS equations.*

This terminology is widely used in the turbulent combustion literature (see, *e.g.*, Libby [91]).

Definition 1.94 (RANS) *RANS is the acronym for Reynolds-Averaged Navier–Stokes approaches to turbulence calculation.*

As will be evident in later lectures, RANS approaches are essentially the opposite of DNS, namely, nearly all scales of the solution must be modeled under a RANS formalism; only (time) mean quantities are directly computed.

Definition 1.95 (Realizability) *Realizability is the term used to summarize physical constraints that must be placed on any turbulence model to avoid direct physical contradictions.*

Although it is not obvious, it is possible for RANS models that include a kinetic energy prediction, for example, to produce negative values of kinetic energy. This is, of course, nonphysical, and any model that can yield such a result does not satisfy realizability.

Definition 1.96 (Receptivity) *Receptivity is a measure of the tendency of a boundary-layer flow to be destabilized (starting a transition to turbulence) by freestream disturbances.*

Definition 1.97 (Return to isotropy) *The phrase return to isotropy is associated with the tendency, of at least the smaller scales of turbulence, to return to a state of approximately equal contributions to turbulence kinetic energy from each velocity component once any source of strain is removed.*

Definition 1.98 (RNG) *RNG is the acronym for renormalization group.*

Renormalization group approaches are adapted from their original use in analysis of phase transitions and have been intensely studied by McComb [92] and Yahot [93], among others.

Definition 1.99 (Resolved scale) *Resolved scale is a name used in LES for the part of the turbulent solution that is directly computed.*

This is often referred to as the large-scale part.

Definition 1.100 (Reynolds analogy) *Reynolds analogy provides a relationship between the Stanton number and skin friction coefficient for turbulent flows.*

Note that the Stanton number is a heat, or mass, transfer coefficient. Thus, the Reynolds analogy provides a direct way to handle turbulence effects on passive scalars once effects on the flow, itself, have been determined.

Definition 1.101 (Reynolds stress) *Reynolds stresses arise from the averaging procedure used in deriving the RANS equations. They are the additional unknowns that create the closure problem.*

Details of the derivation of the RANS equations will be given in Chap. 2 where modeling of the Reynolds stresses will also be discussed.

Definition 1.102 (Scale separation) *Scale separation implies that the largest scales in a turbulent flow do not strongly interact with the smallest scales.*

Note that the implication of this for modeling is extremely important, *viz.*, models can be constructed independent of the details of the large-scale flow behavior. See Tsinober [3] for refutation of such ideas.

Definition 1.103 (Scale similarity) *Scale similarity is an assumption widely employed in construction of LES subgrid-scale models, in which it is presumed that the behavior at the lowest wavenumbers of the unresolved part is "similar" to that of the highest wavenumbers of the resolved scale.*

This assumption permits at least coarse predictions of behavior on the unresolved scales which may be adequate at low Re, but probably not as Re becomes large.

Definition 1.104 (Second-order closure) *Second-order closure refers to a class of RANS models for which the second moments (the Reynolds stresses) are directly computed, usually as solutions to partial differential equations.*

Proponents of RANS models often argue that this approach is far superior because it contains "more physics" than do, for example, the $k-\varepsilon$ methods. But in order to solve the differential equations for the second moments, the third moments must be modeled; and the models for these are little more than guesses. In particular, they contain little or no verifiable physics.

Definition 1.105 (Self preservation) *Self preservation is a property occurring in flow situations for which the turbulence time scales are sufficiently short that as a turbulent disturbance moves downstream with the flow it has time to adjust to its new local environment, implying that the turbulence looks the same at all locations provided it is properly scaled with local variables.*

Definition 1.106 (Structure function) *Structure functions are averages of differences between turbulence quantities measured at different nearby locations in the flow.*

Structure functions can be defined for any integer order corresponding to the power to which differences are raised, and are a key part of the Kolmogorov K41 theory. In particular the theory predicts how structure functions should scale with distance between measurement points, *viz.*, in the form of a power law.

Definition 1.107 (Subgrid-scale (SGS) model) *Subgrid-scale models are used in LES to account for the unresolved part of the solution.*

By now there are many different approaches to constructing such models, some of which will be discussed later.

Definition 1.108 (Subgrid-scale stress) *Subgrid-scale stress is the turbulent stress in a LES model that is analogous to (i.e., arises in the same way as does) Reynolds stress in RANS models.*

As hinted in earlier definitions (Defs. 1.60 and 1.88), the SGS stress is more complicated than is the Reynolds stress and consists of three contributions: cross stress, Leonard stress, and stress that is analogous to Reynolds stress (and is often termed Reynolds stress) but not identical to it. Despite these complications, most modern SGS stress models treat the entire stress in a single model without decomposing it and treating different components separately as was once done.

Definition 1.109 (Synthetic-velocity model) *Synthetic-velocity models employ direct modeling of the SGS velocities rather than modeling the stresses.*

There are several variants of this approach, and it is believed by its proponents that more physics of the sub-grid scales is captured in this way. In particular, among other things this approach permits direct account of SGS interactions between turbulence and other physics, *e.g.*, chemical kinetics, which is not possible with usual LES SGS models, or with RANS models. See References [70], [71] and [72] for more information on this approach.

Definition 1.110 (Taylor's hypothesis) *The Taylor's (frozen-flow) hypothesis states that in a turbulent flow for which the magnitude of the fluctuations is not too great, it is possible to deduce spatial turbulence quantities from time series measured at a single point in the flow.*

Definition 1.111 (Taylor microscale) *The Taylor microscale corresponds to length scales between the integral scale and the dissipation scale.*

From this description we see that the Taylor microscale length should be of approximately the same order of magnitude as length scales corresponding to the inertial subrange of the energy spectrum.

Definition 1.112 (Transition) *Transition refers to a change from one qualitative state of flow to another, and in particular in the context of turbulence studies, the transitions from laminar to turbulent flow.*

Use of the term transition is essentially interchangeable with bifurcation from the dynamical systems view.

Definition 1.113 (Turbulence kinetic energy) *Turbulence kinetic energy is the kinetic energy calculated with turbulent fluctuating velocities.*

It will be clear when we discuss various forms of decomposition of flow variables that exactly what is meant by "turbulent fluctuation" will depend somewhat on the particular decomposition being considered; but in the absence of any specific mention of the decomposition, use of fluctuating velocities from a RANS decomposition will be assumed.

Definition 1.114 (Turbulence energy dissipation rate) *Turbulence energy dissipation rate is the rate at which turbulence energy is being converted to thermal energy by viscous effects on small scales.*

Definition 1.115 (URANS) *URANS stands for unsteady Reynolds-averaged Navier–Stokes.*

This is a currently-popular approach to obtaining time-dependent turbulent solutions rather inexpensively; but the approach is inconsistent with the mathematics of constructing the RANS equations, and as can be reasonably expected in such a case, computed results have not been reliable—even qualitatively.

Definition 1.116 (Viscous sublayer) *Viscous sublayer is the name given to a very thin layer of fluid immediately adjacent to the wall in a wall-bounded turbulent flow.*

The terminology arises from the fact that the viscous terms are dominant in the N.–S. equations for the conditions of such regions. This layer is sometimes mistakenly termed the "laminar" sublayer, but as noted in [7], this is not correct because there are turbulent fluctuations within this region.

Definition 1.117 (VLES) *VLES is the acronym for very large-eddy simulation.*

This corresponds to a coarse-grid LES and now is viewed as being equivalent to URANS. But it will be evident when decompositions are studied in the next section that this is not really the case. Moreover, VLES has a potential for good accuracy, at least in a qualitative sense, if accurate SGS models can be developed.

Definition 1.118 (Vortex stretching) *Vortex stretching refers to a physical phenomenon involving deformation of vortices, and which only occurs in three space dimensions.*

Vortex stretching is often, in classical turbulence analyses, shown to directly give rise to the Reynolds stresses (see [7]), and since it can occur only in 3-D it is often argued that there can be no 2-D turbulence. While the conclusion is correct (for other reasons), the entire notion of vortex stretching, *per se*, is difficult to justify simply because we do not know whether there are any vortices to be stretched at any particular location at a given time. On the other hand, vorticity (as a flow property) exists throughout a flow field, and analyses involving "vortex stretching" actually only make use of vorticity. See Ref. [3] for more details.

Definition 1.119 (zero-, one-, two-equation models) *Zero-, one-, two-equation models refer to the number of partial differential equations needed (beyond the equations for mean quantities) to complete a RANS model.*

Zero-equation models are strictly algebraic, one-equation models involve a single additional partial differential equation (PDE), usually providing a length scale. The $k-\varepsilon$ model described earlier in Def. 1.81 is the most widely used two-equation model.

1.4.2 Mathematical tools

In this section we present an assortment of mathematical tools that are often used in the study of turbulence; we will use most, if not all, of these in the sequel. In some cases the discussions provided here will include further elaboration on definitions presented in the preceding subsection, and in particular, equations will now usually be given. We begin with an introduction to Hilbert spaces because these provide a framework into which essentially everything else we do can be cast. This will include definition and discussion of Fourier series and transforms. We next consider forms of averaging and other statistical quantities, and continue on with a discussion of decomposition of flow variables. Finally, we introduce the Cartesian tensor notation that is so widely used in classical treatments of turbulence.

Introduction to Hilbert Spaces

Because Hilbert spaces play a crucial role in the modern analysis of the N.–S. equations, and such analyses have contributed significantly to the understanding of turbulence and its simulation within the deterministic framework, we will provide a brief, pedestrian introduction to this important area of applied

mathematics that we hope will be sufficient for the present lectures. We first recall Def. 1.24 of a Hilbert space from the preceding section: *a complete, normed, linear space equipped with an inner product that induces the norm of the space.* We assume familiarity with the notions of norm and linearity, and proceed to a discussion of inner product. Once this is available we will consider Fourier representations of functions. Such representations are important in both analytical and computational studies of the N.–S. equations, and thus for turbulence.

Inner product. The inner product of two functions is given in the following definition.

Definition 1.120 *Let u and v be elements of the Hilbert space $L^2(\Omega)$, $\Omega \subseteq \mathbb{R}^d$, $d < \infty$. Then the inner product of u and v is defined as*

$$\langle u, v \rangle \equiv \int_\Omega u(\boldsymbol{x}) v(\boldsymbol{x})\, d\boldsymbol{x}\,. \tag{1.4}$$

This provides a generalization to infinite-dimensional function spaces of the usual finite-dimensional "dot" product of two vectors, and at the same time is a special case of the *scalar product* defined on more general function spaces.

It is clear from the definition that

$$\langle u, u \rangle = \int_\Omega u(\boldsymbol{x}) u(\boldsymbol{x})\, d\boldsymbol{x} = \int_\Omega u^2(\boldsymbol{x})\, d\boldsymbol{x} \equiv \|u\|_{L^2}^2\,,$$

where we are taking u to be real. We see that, as required by the definition of a Hilbert space, the inner product induces a norm on the space. We note that any function u is in the space of functions $L^2(\Omega)$, the canonical Hilbert space, if $\|u\|_{L^2} < \infty$. Furthermore, we observe that the notation being used here for emphasis and clarity is not standard in the context of analysis of the N.–S. equations. In particular, the L^2 norm is often denoted in the same way as is the absolute value function and complex modulus, *viz.*, $|\cdot|$. We will sometimes use this notation in the sequel.

There is a well-known theorem that further relates the inner product to the L^2 norm:

Theorem 1.1 (Cauchy–Schwarz) *Let $u, v \in L^2(\Omega)$. Then*

$$\langle u, v \rangle \leq \|u\|_{L^2} \|v\|_{L^2}\,. \tag{1.5}$$

The proof of this theorem is elementary and can be found, for example, in Stakgold [87].

With this information in hand, we are now prepared to consider Fourier series representations of functions in L^2.

Fourier series. We begin with a somewhat informal statement of the result that justifies use of Fourier series, the Riesz representation theorem (or, in the form we present, sometimes referred to as the Riesz–Fischer theorem), and then discuss some properties of Fourier series.

Theorem 1.2 (Riesz representation) *Any function $f \in L^2(\Omega)$, $\Omega \subset \mathbb{R}^d$, $d < \infty$, can be expressed in the form of a Fourier series,*

$$f(\boldsymbol{x}) = \sum_{\boldsymbol{k}=-\infty}^{\infty} a_{\boldsymbol{k}} \varphi_{\boldsymbol{k}}(\boldsymbol{x})\,, \tag{1.6}$$

with wavevectors $\boldsymbol{k} \equiv (k_1, \ldots, k_d)^T$, and $\{\varphi_{\boldsymbol{k}}\}$ being a complete (in L^2), orthonormal basis. The $a_{\boldsymbol{k}}$s are the Fourier coefficients.

In particular, for any $f \in L^2$, the $a_{\boldsymbol{k}}$s exist, are uniquely defined, and decay to zero sufficiently fast with increasing $|\boldsymbol{k}|$ to guarantee convergence of the series in a particular sense.

The $a_{\boldsymbol{k}}$s are calculated using the inner product defined above:

$$a_{\boldsymbol{k}} \equiv \int_\Omega f(\boldsymbol{x}) \varphi_{\boldsymbol{k}}(\boldsymbol{x})\, d\boldsymbol{x}\,, \tag{1.7}$$

which by the Cauchy–Schwarz inequality is guaranteed to exist for $f \in L^2$ and $\varphi_{\boldsymbol{k}}$ no less smooth than required for inclusion in L^2. Usually, the $\varphi_{\boldsymbol{k}}$s are much smoother, typically being eigenfunctions of a Sturm–Liouville problem (see, *e.g.*, [87]) such as sines and cosines, and complex exponentials. Moreover, we note that completeness of the $\varphi_{\boldsymbol{k}}$s in $L^2(\Omega)$ implies that they will be sufficient to construct any function in $L^2(\Omega)$.

We comment that what we have called the Riesz representation theorem is actually a special case of that theorem, *i.e.*, a corollary; the complete theorem is far more general. Furthermore, we note that convergence of the series in Eq. (1.6) can be deduced directly from the Parseval identity which we now state.

Theorem 1.3 (Parseval identity) *Let $f \in L^2(\Omega)$, and let $\{\varphi_{\boldsymbol{k}}(\boldsymbol{x})\}$ be a complete orthonormal basis with respect to $L^2(\Omega)$. Then*

$$\int_\Omega |f(\boldsymbol{x})|^2 \, d\boldsymbol{x} = \sum_{\boldsymbol{k}} |a_{\boldsymbol{k}}|^2 \,. \tag{1.8}$$

Clearly, since $f \in L^2(\Omega)$, the integral on the left-hand side exists, so the summation on the right-hand side converges, implying that the Fourier representation (1.6) converges to f in the L^2 norm. That is,

$$\int_\Omega \left| f - \sum_{\boldsymbol{k}}^{N} a_{\boldsymbol{k}} \varphi_{\boldsymbol{k}} \right|^2 d\boldsymbol{x} \longrightarrow 0$$

as $N \to \infty$. We leave proof of this as an exercise for the reader.

There are several key points to be made regarding the Fourier representation introduced here. First, we have at no time assumed the function f to be periodic. Although it may well be that the $\varphi_{\boldsymbol{k}}$s are periodic, this does not imply that f must also be because we are requiring convergence only in L^2. Obviously, a finite sum of $a_{\boldsymbol{k}}\varphi_{\boldsymbol{k}}$ would be periodic if this were true of the $\varphi_{\boldsymbol{k}}$s, but this does not generally hold for an infinite sum. (It is a myth, and a badly misleading and incorrect one, that only periodic functions have Fourier series expansions.) It should also be mentioned that due to details of the Lebesgue integral, it is possible that the series does not converge at all at (at least) a countable number of points $\boldsymbol{x} \in \Omega$ (but this does not matter in the sense of L^2 convergence).

Finally, if we consider a function that also depends on time, say $u(\boldsymbol{x}, t)$, then the corresponding Fourier representation is

$$u(\boldsymbol{x}, t) = \sum_{\boldsymbol{k}=-\infty}^{\infty} a_{\boldsymbol{k}}(t) \varphi_{\boldsymbol{k}}(\boldsymbol{x}) \,, \tag{1.9}$$

with

$$a_{\boldsymbol{k}}(t) \equiv \int_\Omega u(\boldsymbol{x}, t) \varphi_{\boldsymbol{k}}(\boldsymbol{x}) \, d\boldsymbol{x} \,. \tag{1.10}$$

In particular, we note that the Fourier coefficients are now time dependent, functions and the representation Eq. (1.9) is reminiscent of "separation of variables" from elementary PDE analysis (see, *e.g.*, Berg and McGregor [95]). Indeed, Eq. (1.9) in some sense provides both a generalization and a justification of this simple solution technique.

Fourier transforms. Because such a large body of analysis of the N.–S. equations has been performed using Fourier transforms, for the sake of completeness, we briefly describe these here. But we note that these apply to what is often a different type of function space than present knowledge of the N.–S. equations would suggest is appropriate (see Foias *et al.* [89] for more details). Thus it is doubtful that much is to be gained from an exhaustive study, and we will simply provide the definition and a few comments.

Definition 1.121 (Fourier transform) *Let $f \in L^2(\mathbb{R}^d)$. Then the Fourier transform of f is*

$$\mathcal{F}(f)(\boldsymbol{\omega}) = \int_{\mathbb{R}^d} f(\boldsymbol{x}) e^{-2\pi i \boldsymbol{\omega} \cdot \boldsymbol{x}} \, d\boldsymbol{x} \,. \tag{1.11}$$

There is, of course, an inverse Fourier transform defined (with appropriate normalization) and denoted by $\mathcal{F}^{-1}$ so that

$$\mathcal{F}^{-1}\big(\mathcal{F}f(\boldsymbol{\omega})\big)(\boldsymbol{x}) = f(\boldsymbol{x})\,.$$

We remark that Fourier transforms have been widely used in the renormalization group (RNG) analyses of N.–S. turbulence by *e.g.*, McComb [92] and Yahot and Orszag [93]. The latter of these has served as the starting point for the once highly-touted RNG turbulence models now available in many commercial CFD codes. But there are fundamental flaws in the analysis presented in [93], as is well known, and results presented by Freitas [94] show that RNG turbulence models are "unstable to implementation details." That is, two different codes of the same model can produce drastically different results when applied to the same problem. We thus do not recommend use of such an approach to turbulence modeling.

Forms of Averaging and Filtering, and Other Statistical Quantities

Both classical (statistical) and modern (deterministic) approaches to the analysis of turbulence generally must employ averaging or filtering at some level. Here, we present the various possible methods: temporal, spatial and ensemble averaging, and spatial filtering. We then discuss various other mathematical tools of a, mainly, statistical nature.

Time average. Let $u(\boldsymbol{x},t)$ be an integrable function with respect to t for $t\to\infty$ and defined for any desired $\boldsymbol{x}\in\Omega\subset\mathbb{R}^d\,,\ d=1,2,3$. Then the *time average* of u at the point $\boldsymbol{x}$ is defined as

$$\overline{u}(\boldsymbol{x}) \equiv \lim_{T\to\infty}\frac{1}{T}\int_0^T u(\boldsymbol{x},t)\,dt\,. \tag{1.12}$$

We note that the lower limit of integration is arbitrary and may be shifted as needed. It is clear that the average can no longer be a function of time, and it follows that all time derivatives of $\overline{u}$ are identically zero. We comment that this is often (usually!) ignored in modern treatments of RANS modeling (see below, and Def. 1.115), but without the formal limit taken in Eq. (1.12), the average is not well defined, as emphasized, *e.g.*, by Tennekes and Lumley [7].

Spatial average. Let $\Omega\subset\mathbb{R}^d\,,\ d=1,2,3$ with $u(\boldsymbol{x},t)$ integrable on Ω at any desired time t. Then the *spatial average* of u at time t is

$$\widetilde{u}(t) \equiv \frac{1}{V_\Omega}\int_\Omega u(\boldsymbol{x},t)\,d\boldsymbol{x}\,, \tag{1.13}$$

where V_Ω denotes the "volume" of Ω. Clearly this is merely the length of an interval when $d=1$, and an area when $d=2$. We remark that the " $\widetilde{}$ " notation will also be used for formal filtering (see below) which when applied spatially, as is usually done in the context of turbulence analyses associated with LES, represents a straightforward generalization of (1.13).

Ensemble average. Let $\big\{u^{(i)}(\boldsymbol{x},t)\big\}_{i=1}^N$ be a sequence of realizations of a function $u(\boldsymbol{x},t)$ defined for $\boldsymbol{x}\in\Omega\subset\mathbb{R}^d$ and $t\in[0,t_f]$. Then the *ensemble average* of u is defined as

$$\langle u(\boldsymbol{x},t)\rangle \equiv \frac{1}{N}\sum_{i=1}^N u^{(i)}(\boldsymbol{x},t)\,. \tag{1.14}$$

Formally one can consider taking the limit $N\to\infty$, but this clearly imposes additional mathematical difficulties and requirements on the sequence $\big\{u^{(i)}(\boldsymbol{x},t)\big\}_{i=1}^N$; moreover, in practice—in the context of either experiments or numerical simulations—N would necessarily be finite. Nevertheless,

there still is a convergence question, namely, "How large must N be before $\langle u \rangle$ no longer changes significantly with increasing N?" Practical experience has shown that such convergence occurs very slowly, making use of ensemble averages difficult in either experimental or computational contexts. On the other hand, observe that an ensemble-averaged variable is still a function of both $\boldsymbol{x}$ and t, despite the averaging—an often desirable property.

We also mention that the angular-bracket notation, $\langle \cdot \rangle$, will sometimes be used later to denote a generic averaging process that might be any one of the above forms.

Spatial filtering. For any function $u(\boldsymbol{x}, t)$ in a Hilbert space we define the *spatially-filtered* function $\widetilde{u}(\boldsymbol{x}, t)$ over a domain Ω as

$$\widetilde{u}(\boldsymbol{x}, t) \equiv \frac{1}{V_\Omega} \int_\Omega u(\boldsymbol{\xi}, t) G(\boldsymbol{x}|\boldsymbol{\xi}) \, d\boldsymbol{\xi} \,, \tag{1.15}$$

where $G(\boldsymbol{x}|\boldsymbol{\xi})$ is the *kernel* of the filter, presumed to be at least in $L^2(\Omega)$. Ω and V_Ω are defined as in the case of spatial averaging (corresponding to $G \equiv 1$), but we observe that in normal usage the functions employed for G decay rapidly near the boundaries of subdomains of Ω. As a consequence, in contrast to the spatially-averaged case which is usually global (at least in specific directions), the filtered function is now local and still dependents on both $\boldsymbol{x}$ and t, just as is the case for ensemble averaging. We note that an analogous definition holds for temporal filtering. This is seldom used in turbulence modeling but is widely used in the context of experimental data analysis.

Autocorrelation. Here we provide more details associated with Def. 1.1. Let $u(\boldsymbol{x}, t) \in L^2(0, t_f)$ be stationary with respect to t and bounded for each $\boldsymbol{x} \subset \mathbb{R}^d$. Then we define the *autocorrelation coefficient* of u at the point $\boldsymbol{x}$ to be

$$c_a(\boldsymbol{x}, \tau) \equiv \frac{\langle u(\boldsymbol{x}, t), u(\boldsymbol{x}, t + \tau) \rangle}{\|u\|^2_{L^2}} \tag{1.16}$$

for each increment $\tau < t_f$. Clearly, by the Schwarz inequality, $c_a(\boldsymbol{x}, \tau)$ takes on values on the interval $[-1, 1]$.

We remark that it is possible to also define autocorrelation in terms of spatial variables in an analogous fashion. We leave this as an exercise for the reader. Also observe, as discussed in considerable detail in [7], that the autocorrelation coefficient leads to an *integral time scale* $\mathcal{T}$ defined as

$$\mathcal{T} \equiv \int_0^\infty c_a(\tau) \, d\tau \,. \tag{1.17}$$

It should be noted that the infinite upper limit in the integral is formal. The integral time scale is a measure of how long turbulent fluctuations remain correlated, and in practice is taken to be some fraction (~ 0.5, or less) of the time to the first zero of c_a. Hence, $\mathcal{T} > 0$ always holds.

Cross correlation. Analogous to the above, for two stationary functions $u(\boldsymbol{x}, t), v(\boldsymbol{x}, t) \in L^2(0, t_f)$ and bounded for each $\boldsymbol{x} \subset \Omega$, we can define the *cross-correlation coefficient* for any point $\boldsymbol{x}$ as

$$c_c(\boldsymbol{x}) \equiv \frac{\langle u(\boldsymbol{x}, t), v(\boldsymbol{x}, t) \rangle}{\|u\|_{L^2} \|v\|_{L^2}} \,. \tag{1.18}$$

We remark that although c_a and c_c can be computed for any functions in the appropriate spaces, within the context of applications to turbulence analysis, they are always used for functions having zero mean. Thus, in the preceding definitions we would replace u with $u - \overline{u}$, for example. If we then calculate the cross correlation without the above normalization, but instead as an average over time, we have

$$\overline{u'v'} = \frac{1}{t_f} \int_0^{t_f} u'(t) v'(t) \, dt \,,$$

as $t_f \to \infty$, and where, *e.g.*, $u' = u - \overline{u}$ as will be introduced in the next section. We will later see that this correlation is (up to scaling for dimensional consistency) one of the components of the Reynolds stress tensor arising from averaging the N.–S. equations in the process of obtaining the RANS equations.

Probability density function. The probability density function (pdf), also termed "probability distribution function," describes the frequency of occurrence of values of a given function over the range of the function, as noted in Def. 1.5. We will here denote the pdf of a function u as $P(u)$. More precisely, $P(u)$ is the probability that a specified value will be found between u and $u + du$. Since $P(u) \in [0, 1]$ (because it is a probability), and the sum of all $P(u)$ is taken to be unity, it follows that

$$\int_{-\infty}^{\infty} P(u)\, du = 1\,. \tag{1.19}$$

One of the important uses of the pdf is in construction of various "moments" of turbulence quantities. In particular, if the pdf is somehow known for a variable u, then the *first moment* (*i.e.*, the average) is given by

$$\langle u \rangle = \int_{-\infty}^{\infty} uP(u)\, du\,.$$

It is clear that this is just a typical "weighted" average with $P(u)$ being the weighting function. Similarly, the second moment, termed the *variance*, is calculated as

$$\sigma^2 = \int_{-\infty}^{\infty} (u')^2 P(u')\, du' \quad \left(= \overline{\left(u'\right)^2} \right)$$

with u' as given above. Observe that u' is used here because it is usually associated with deviation from the mean.

It is not hard to imagine that if all moments of a statistical variable could be constructed such results would constitute complete information about the variable, and we note that there are *pdf methods* for modeling turbulence based on this idea (see, for example, Pope [96]) which we will not treat in these lectures. But it should be clear from the preceding formulas that this would require knowledge of the pdf, and hence of the variable itself.

Flatness. *Flatness*, also called *kurtosis* (but the precise definition of kurtosis is $F - 3$) is described heuristically in Def. 1.4. The formula for calculating flatness of any particular quantity, say $u \in L^4$, is

$$F(u) = \frac{\langle u^4 \rangle}{\langle u^2 \rangle^2}\,, \tag{1.20}$$

where $\langle \cdot \rangle$ denotes any desired average, but in most cases is temporal. It is noted by Tennekes and Lumley [7] that values of flatness are large (compared with the value three (3) for a Gaussian distribution) if the pdf has relatively large values in its tails, *i.e.*, it does not go to zero as fast as Gaussian as its argument approaches $\pm\infty$. This occurs when time series of the function contain significant numbers of sharp peaks, and is related to intermittency, as discussed by Frisch [80]; in particular, as noted in that work, a high-pass filtered function is said to be *intermittent on small scales* if its flatness grows without bound with increasing filter frequency. We also note that flatness is the fourth moment of a function divided by the square of the second moment. Hence, it can be calculated in terms of the pdf, as is done in [7].

Skewness. *Skewness* of a function $u \in L^3$ is computed from the formula

$$S(u) = \frac{\langle u^3 \rangle}{\langle u^2 \rangle^{3/2}}\,. \tag{1.21}$$

As noted in Def. 1.7, this is a measure of the asymmetry of a function (with respect to its mean—usually zero in the context of turbulence studies) as indicated either by direct observation of its time series or from the shape of its pdf. If the pdf of a function is symmetric (as is the case for a Gaussian), the skewness is zero. Furthermore, if a time series exhibits more negative values than positive ones, skewness will be negative, and conversely.

We remark that flatness and skewness of velocity derivatives are widely studied in turbulence analyses.

Power spectral density. As indicated in Def. 1.41, the power spectrum provides a representation of the dependence of energy on frequency (or wavenumber), where energy should be viewed in the generalized sense of the L^2 norm of the function being considered. In particular, if we recall the definition of L^2 norm following Eq. (1.4) and consider the relationship between this norm and the Fourier coefficients as given by Parseval's identity, we see that the PSD can be viewed as a distribution of squared Fourier coefficients as a function of wavenumber (or frequency). This is the interpretation used in calculating PSDs via fast Fourier transform (FFT).

Structure functions. *Structure functions of order p* are defined, somewhat heuristically, by

$$S_p(r) = \langle (u(\boldsymbol{x}+\boldsymbol{r}) - u(\boldsymbol{x}))^p \rangle \ , \tag{1.22}$$

where $\boldsymbol{r}$ is a vector pointing between two nearby locations of "measurement" of the quantity u; r is the magnitude of $\boldsymbol{r}$, and $\langle \cdot \rangle$ denotes any convenient average, but in any case performed over all samples having the same value of r. This is the form of structure function typically applied to scalar quantities. We have suppressed temporal notation, but we observe that, in general, S_p will depend on time unless $\langle \cdot \rangle$ includes temporal averaging.

For vector quantities (*e.g.*, velocity or vorticity), both transversal and longitudinal structure functions can be defined. For the former of these, the distances r are taken perpendicular to the direction of separation of measurement points, and the notation $S_{p_\perp}$ is used. These *transversal structure functions* are calculated as

$$S_{p_\perp}(r) = \langle [(\boldsymbol{U}(\boldsymbol{x}+\boldsymbol{r}) - \boldsymbol{U}(\boldsymbol{x})) \cdot \boldsymbol{e}_\perp]^p \rangle \ , \tag{1.23}$$

with $\boldsymbol{e}_\perp$ a unit vector in the transversal (perpendicular) direction with respect to the direction of $\boldsymbol{U}(\boldsymbol{x})$. The *longitudinal structure function* is defined in an analogous way:

$$S_{p_\parallel}(r) = \left\langle \left[(\boldsymbol{U}(\boldsymbol{x}+\boldsymbol{r}) - \boldsymbol{U}(\boldsymbol{x})) \cdot \boldsymbol{e}_\parallel \right]^p \right\rangle \ . \tag{1.24}$$

We remark that it is the latter of these that is more often measured in laboratory experiments. Moreover, we again emphasize that structure functions can vary in both space and time. On the other hand, most experimental data associated with these have been at least time averaged, and sometimes averaged in space as well.

Structure functions are extremely important elements in Kolmogorov's theory of turbulence, as we will later see. We can gain a hint of this already by noting that if u represents a velocity component, and r is not too large, we would expect via Taylor expansion that

$$u(\boldsymbol{x}+\boldsymbol{r}) - u(\boldsymbol{x}) \simeq \frac{\partial u}{\partial x} r \, ,$$

and further, we might view this as a fluctuation representing small-scale behavior since it corresponds to a change in the velocity over a short distance. This, in turn, suggests that we might view the second-order structure function,

$$S_2(r) = \left\langle (u(\boldsymbol{x}+\boldsymbol{r}) - u(\boldsymbol{x}))^2 \right\rangle \ ,$$

as representing energy of fluctuations. If we now recognize that wavenumbers in Fourier space are related to distances in physical space by $k \sim 1/r$, we see that if a formula for S_2 in terms of r can be found, we can readily predict the distribution of energy as a function of wavenumber. We will later see that this is precisely one of the things accomplished by the Kolmogorov K41 theory.

Forms of Function Decomposition

It will be apparent as we proceed that essentially all forms of turbulence modeling in current use involve some form of decomposition of dependent variables representing the physical situation. The two main ones are Reynolds decomposition and the LES decomposition. There are others, but we will not make use of these in the present lectures with but one exception, a Hilbert space decomposition.

Reynolds decomposition. Let $u(\boldsymbol{x}, t)$ be well defined in a domain $\Omega \subset \mathbb{R}^d,\ d = 1, 2, 3$ and for $t \in [0, t_f]$, and suppose further that $\overline{u}(\boldsymbol{x})$ exists in the sense of the time average, Eq. (1.12). Then the *Reynolds decomposition* of $u(\boldsymbol{x}, t)$ is

$$u(\boldsymbol{x}, t) = \overline{u}(\boldsymbol{x}) + u'(\boldsymbol{x}, t)\,, \tag{1.25}$$

where $u'(\boldsymbol{x}, t)$ is termed the "fluctuating part." A key observation made earlier is that $\overline{u}(x)$ is independent of time, implying that any equations derived for computing this quantity must be steady state. Figure 1.4 depicts the temporal behaviors of $\overline{u}$ and u'. It is important to note

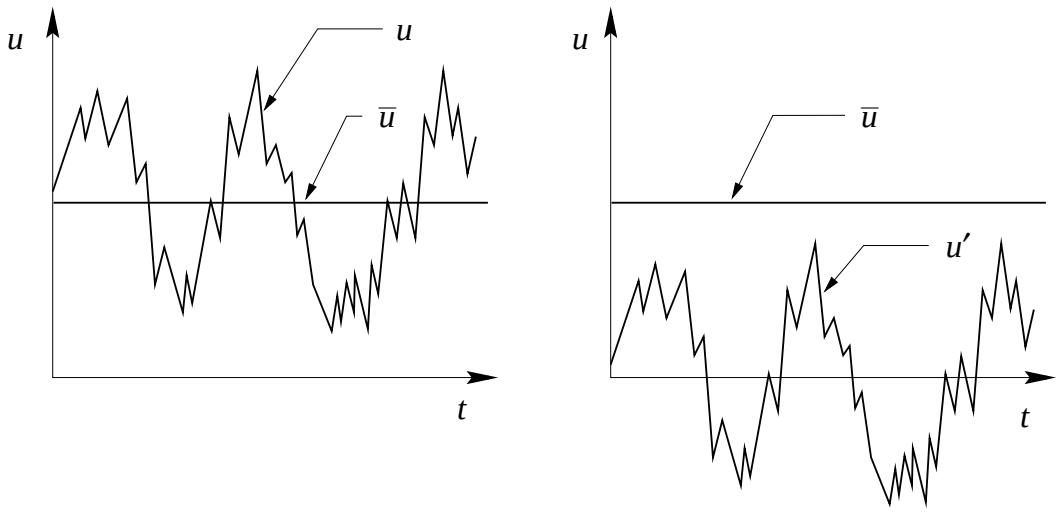

Figure 1.4: Plots of parts of Reynolds decomposition.

that for the type of temporal behavior shown here (probably noisy quasiperiodic, from a dynamical systems point of view) the Reynolds decomposition is likely to provide an inadequate representation. In particular, the mean is seldom attained by the actual signal, and the fluctuating part carries essentially all of the dynamics; in fact, the magnitude of the fluctuations can easily be essentially high as the mean. We will later see that this is one of the major shortcomings of RANS approaches to turbulence modeling.

It is clear from the definition of time average that the Reynolds decomposition possesses the following two properties:

$$\overline{\overline{u}} = \overline{u}\,, \qquad \text{and} \qquad \overline{u'} = 0\,, \tag{1.26}$$

with the first of these being obvious, and the second following from the first. These will be crucial in our later derivation of the RANS equations, and the reader is encouraged to show that these equalities hold.

LES decomposition. The LES decomposition was introduced by Deardorff [10] and was first analyzed in detail for the incompressible N.–S. equations by Leonard [76]. It is constructed by applying a local spatial filter (or in the simplest case, spatial average) to all appropriate variables. It is interesting to note that the averaging originally proposed by Reynolds [16] was spatial rather than temporal; hence, it can be seen that Deardorff's approach was rather close to Reynolds' original formulation. Despite this, current applications of Reynolds averaging employ time averaging, and LES uses a spatial filter, as we will now describe.

We will see below that the LES decomposition can be interpreted as a Hilbert space decomposition, so we begin by assuming $u(\boldsymbol{x},t) \in L^2(\Omega) \times C^1(0,t_f)$. Then we write the *LES decomposition* as

$$u(\boldsymbol{x},t) = \widetilde{u}(\boldsymbol{x},t) + u'(\boldsymbol{x},t)\,. \tag{1.27}$$

In this decomposition $\widetilde{u}$ is usually termed the *large-* or *resolved-scale* part of the solution, and u' is called the *small-scale*, or *subgrid-scale*, or *unresolved* part. It is important to note that both resolved and unresolved scales depend on both space and time, and this is a major distinction and advantage compared with the Reynolds decomposition. In Eq. (1.27) $\widetilde{u}$ is formally the filtered solution corresponding to Eq. (1.15). From this it is easily shown that, in general,

$$\widetilde{\widetilde{u}} \neq \widetilde{u}\,, \qquad \text{and} \qquad \widetilde{u'} \neq 0\,, \tag{1.28}$$

(although there are special cases for which equalities do hold). The reader may wish to investigate this.

This lack of equality in Eqs. (1.28) significantly complicates derivation of the usual LES formulations in comparison with RANS approaches and, in particular, leads to Leonard and cross stresses in addition to the usual Reynolds stress as will be seen in Chap. 3. On the other hand, once the formulation has been developed there is little additional complication, and the LES equations are quite similar to the RANS equations with one major exception: the RANS equations are formally independent of time, while the LES equations are valid for time-dependent calculations.

Hilbert space decomposition. The final decomposition we consider here is a mathematical one that is closely related to the LES decomposition. It is the ability to construct such decompositions that is crucial to obtaining numerical approximations to solutions of the N.–S. equations. In particular, because of the properties of the Hilbert space L^2 we can express the Fourier representation, Eq. (1.6), as

$$f(\boldsymbol{x}) = \sum_{|\boldsymbol{k}|\geq 0}^{N} a_{\boldsymbol{k}}\varphi_{\boldsymbol{k}}(\boldsymbol{x}) + \sum_{|\boldsymbol{k}|=N+1}^{\infty} a_{\boldsymbol{k}}\varphi_{\boldsymbol{k}}(\boldsymbol{x})\,. \tag{1.29}$$

Each series on the right-hand side is an element of a subspace of L^2, and due to orthogonality of the $\varphi_{\boldsymbol{k}}$s these subspaces do not intersect. Furthermore, as $N \to \infty$, the first term on the right-hand side, alone, more closely approximates f, and we view the second term as the remainder of the series representation.

If we draw an analogy between the terms of the LES decomposition (1.27) with those in the above expression, we expect that the large-scale part of LES will converge to a solution of the N.–S. equations as the discrete resolution of the solution procedure is increased provided the SGS model is constructed so that the small-scale part (corresponding to the series remainder) approaches zero in this same limit. We will later see that, indeed, this is the case, thus providing assurance that LES is mathematically well founded. We remark that no such convergence result can be deduced

for RANS methods. In particular, it should be clear from Fig. 1.4 that u' cannot be viewed as the remainder of a Fourier series expansion and moreover, in light of its physical interpretation, it cannot be considered to be small, in general.

Cartesian Tensor Notation

Cartesian tensor notation provides a short-hand that is widely used in fluid dynamics, especially in the context of classical turbulence analyses. We will provide a brief treatment of this in the current section since it will be necessary to occasionally use it later.

We begin by noting that the mathematical description of a tensor (over a vector space) is a *multi-linear form* (see, *e.g.*, Bishop and Goldberg [97]), meaning that a tensor is a multi-variable construct that satisfies the definition of linearity in each of its variables, separately. Within this framework we might represent a tensor as $F(V_1, V_2, \ldots, V_N)$ with F being linear with respect to each of the N vector spaces, $V_i\,,\ i = 1, \ldots, N$, appearing in its argument. But Cartesian tensors are much less general (as the name suggests), and unlike the preceding expression which is coordinate free, Cartesian tensors are presented in terms of Cartesian coordinates, the specific orientation and origin of which must be specified (in terms of a basis set) to completely define such a tensor. Indeed, from a purely mathematical perspective, what are often called tensors are actually a specific representation of a tensor in terms of coordinates. We note that in our present context, scalars are tensors of rank zero; vectors have rank one, and matrices have rank two. The term "tensor" will thus imply rank two, *i.e.*, a matrix, unless noted otherwise.

Cartesian tensor notation is concise because of its use of indexing and operations specifically associated with this indexing. For example, the velocity vector is represented as

$$\boldsymbol{U} = (u_1, u_2, u_3)^T \quad \left(= (u, v, w)^T\right)\,,$$

and spatial coordinates as $\boldsymbol{x} = (x_1, x_2, x_3)^T$. Operations with Cartesian tensors involve implied summation over repeated indices (*Einstein summation*); so, for example, the divergence of the velocity field can be expressed as

$$\nabla \cdot \boldsymbol{U} = \frac{\partial u_i}{\partial x_i} \equiv \sum_{i=1}^{3} \frac{\partial u_i}{\partial x_i} = \frac{\partial u_1}{\partial x_1} + \frac{\partial u_2}{\partial x_2} + \frac{\partial u_3}{\partial x_3} \quad \left(= \frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} + \frac{\partial w}{\partial z}\right)\,.$$

We observe that the index i appears in both numerator and denominator of the tensor representation $\partial u_i/\partial x_i$, and so is "repeated." But this is not the only form of repeated indices. For example, if $\boldsymbol{S}$ denotes the *strain rate tensor* with elements

$$S_{ij} \equiv \frac{1}{2}\left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i}\right)\,,$$

then $S_{ik}S_{kj}$ denotes the matrix multiplication

$$\boldsymbol{S}\,\boldsymbol{S} = \sum_{k=1}^{3} S_{ik}S_{kj} \qquad \forall\ i, j = 1, 2, 3.$$

Using these constructs permits expressing the incompressible N.–S. equations given in Eqs. (1.1) as

$$\frac{\partial u_i}{\partial x_i} = 0\,, \tag{1.30a}$$

$$\frac{\partial u_i}{\partial t} + u_j\frac{\partial u_i}{\partial x_j} = -\frac{\partial P}{\partial x_i} + \nu\frac{\partial^2 u_i}{\partial x_j \partial x_j}\,, \qquad i = 1, 2, 3\,. \tag{1.30b}$$

Although this notation seems not to provide any significant advantages in expressing the N.–S. equations, themselves (and, in fact, it doesn't), it is very effective in concisely representing the RANS equations and various quantities associated with turbulence models constructed for use with these equations. This will be apparent in the sequel.

We end this description of Cartesian tensors by introducing two specific items that are widely used to further collapse notation. The first is the *Kronecker* δ which, in the present context, is used to represent the identity matrix:

$$\delta_{ij} \equiv \begin{cases} 1 & \text{if } i = j\,, \quad i,j = 1,2,3\,, \\ 0 & \text{otherwise}\,. \end{cases} \tag{1.31}$$

The second is what is called the *permutation tensor*, defined as

$$\epsilon_{ijk} \equiv \begin{cases} 1 & \text{for any permutation of } i,j,k = 1,2,3\,, \\ -1 & \text{for non-repeating combinations of } i,j,k \text{ that are not permutations}\,, \\ 0 & \text{if any two indices are repeated}\,. \end{cases} \tag{1.32}$$

We leave as an exercise to the reader demonstration that the curl of the velocity field can be very concisely expressed in terms of the permutation tensor as

$$(\nabla \times \boldsymbol{U})_i = \epsilon_{ijk} \frac{\partial u_k}{\partial x_j}\,.$$

Finally, following Wilcox [98], we note that there is an identity relating δ_{ij} and ϵ_{ijk}:

$$\epsilon_{ijk}\epsilon_{ist} = \delta_{js}\delta_{kt} - \delta_{jt}\delta_{ks}\,. \tag{1.33}$$

Proof of this is left as an exercise.

1.4.3 Further basic concepts

In this section we consider some further, fairly general, basic ideas associated either indirectly, or directly, with turbulence. In the first of two subsections we treat some basic properties of the N.–S. equations since, as we will emphasize throughout these lectures, these equations are now almost universally accepted as providing the correct description of turbulent flows. Much of what is presented in this subsection is extracted from Frisch [80] or from Foias *et al.* [89].

In a second subsection we present a number of somewhat more physically-oriented concepts that, while they are not necessarily related to any particular turbulence model, have played significant roles in setting the "way of seeing" turbulence that has contributed to the form of a number of models. This information comes from a variety of sources, but mainly from [80] and [7].

Mathematical Properties of the N.–S. Equations

In this section we consider some mathematical aspects of the N.–S. equations that have had major impacts on recent ideas concerning turbulence, and also some concepts that have influenced the nature of models as well. We begin with a discussion of symmetries of the N.–S. equations as described in [80], the most important of which is Galilean invariance. We follow this with a brief discussion of current knowledge on existence, uniqueness and regularity of N.–S. solutions, and we conclude the section by constructing the Galerkin approximation to the N.–S. equations.

Symmetries, symmetry groups. Frisch [80] notes that *symmetry* is a term employed by theoretical physicists when referring to *invariance groups* of a dynamical theory. (For each symmetry there

is an associated "conservation law.") We observe that *group* has a precise, and somewhat abstract, mathematical meaning that we do not need to explore for the present lectures. In any case, suppose $\boldsymbol{G}$ is such a group, and suppose $\boldsymbol{U}$ is a solution to the (incompressible) N.–S. equations. If $g \in \boldsymbol{G}$ and $g\boldsymbol{U}$ is also a solution to the N.–S. equations, then $\boldsymbol{G}$ is a *symmetry group* for these equations. We remark that the "similarity transformations" arising in boundary-layer theory provide a familiar example of such groups.

A total of six symmetries of the N.–S. equations are discussed in [80]; but, as noted there, some do not hold for high Re, and others do not hold for low Re. There are three that are valid symmetry groups independent of Re (but with other restrictions) and that are of particular importance. These are the following:

i) space translations $\quad g^{space}_{\boldsymbol{\rho}} : t, \boldsymbol{r}, \boldsymbol{U} \longmapsto t, \boldsymbol{r} + \boldsymbol{\rho}, \boldsymbol{U}\,, \ \boldsymbol{\rho} \in \mathbb{R}^d \ \ d = 1, 2, 3;$

ii) time translations $\quad g^{time}_{\tau} : t, \boldsymbol{r}, \boldsymbol{U} \longmapsto t + \tau, \boldsymbol{r}, \boldsymbol{U}\,, \ \tau \in \mathbb{R}\ ;$

iii) Galilean transformations $\quad g^{Gal}_{\boldsymbol{V}} : t, \boldsymbol{r}, \boldsymbol{U} \longmapsto t, \boldsymbol{r} + \boldsymbol{V}t, \boldsymbol{U} + \boldsymbol{V}\,, \ \boldsymbol{V} \in \mathbb{R}^d.$

In these mappings $\boldsymbol{\rho}, \tau$ and $\boldsymbol{V}$ represent any bounded constant with the specified dimension. Thus, proof that space and time translations are symmetry groups of the N.–S. equations is direct, and we leave this as an exercise for the reader.

Galilean invariance, described briefly in Def. 1.73, is one of the most important properties of the N.–S. equations because, as observed earlier, without it we would find it impossible to compare fluid experiments performed in different parts of the world. Indeed, if the N.–S. equations were not Galilean invariant, this would be a sure sign that they could not properly describe fluid motion. It turns out that this universal symmetry is often violated by RANS turbulence models (see, *e.g.*, Speziale [99]), and much effort has gone into modifying such models to avoid this physical discrepancy—at the expense of relative simplicity. Frisch notes that the key to proving this symmetry is to show cancellation of terms arising from $\boldsymbol{U}_t$ and $\boldsymbol{U} \cdot \nabla \boldsymbol{U}$ under the transformation. We will demonstrate this here for the 1-D Burgers' equation and leave the proof for the complete N.–S. system as an exercise.

To begin, recall that Burgers' equation is a 1-D model equation possessing all of the same types of terms as do the momentum equations of the N.–S. system. Thus, it can be expressed as

$$u_t + uu_x = \nu u_{xx}\,, \tag{1.34}$$

where ν can be viewed as a viscosity, or the reciprocal of the Reynolds number. Now for a general transformation, say T, we have

$$T : \begin{pmatrix} x \\ t \end{pmatrix} \longrightarrow \begin{pmatrix} \xi(x,t) \\ \tau(x,t) \end{pmatrix} \qquad \text{with} \qquad T^{-1} : \begin{pmatrix} \xi \\ \tau \end{pmatrix} \longrightarrow \begin{pmatrix} x(\xi,\tau) \\ t(\xi,\tau) \end{pmatrix}.$$

Equation (1.34) is invariant under the transformation T if we can show that

$$u_\tau + uu_\xi = \nu u_{\xi\xi}\,. \tag{1.35}$$

That is, the <u>form</u> of the equation remains the same after transformation. We leave it to the reader to demonstrate that this does not happen, in general, that is, for arbitrary general transformations.

In the specific case of the Galilean transformation we have

$$g^{Gal}_v : \xi = x + vt\,, \quad \tau = t\,,$$

and

$$\left(g^{Gal}_v\right)^{-1} : x = \xi - v\tau\,, \quad t = \tau\,.$$

But since the dependent variable is also transformed in this case, for invariance to hold we must have

$$w_\tau + ww_\xi = \nu w_{\xi\xi}\,,$$

with $w = u + v$. Thus, we must show that

$$(u+v)_\tau + (u+v)(u+v)_\xi = \nu(u+v)_{\xi\xi} \tag{1.36}$$

is the same as Eq. (1.34).

We first observe that since v is a constant, the above collapses to

$$u_\tau + (u+v)u_\xi = \nu u_{\xi\xi}\,. \tag{1.37}$$

Next, from properties of transformations rather generally, and the inverse function theorem in particular, we know that

$$u(\xi,\tau) = u(\xi(x,t),\tau(x,t)) = u(x(\xi,\tau),t(\xi,\tau))\,,$$

under reasonable smoothness assumptions, which clearly hold for the Galilean transformation, itself, although possibly not for dependent variables of the differential equation(s). Hence, we have

$$\frac{\partial u}{\partial \tau} = \frac{\partial u}{\partial x}\frac{\partial x}{\partial \tau} + \frac{\partial u}{\partial t}\frac{\partial t}{\partial \tau} = -v\frac{\partial u}{\partial x} + \frac{\partial u}{\partial t}\,,$$

and similarly,

$$\frac{\partial u}{\partial \xi} = \frac{\partial u}{\partial x}\frac{\partial x}{\partial \xi} + \frac{\partial u}{\partial t}\frac{\partial t}{\partial \xi} = \frac{\partial u}{\partial x}\,.$$

From this it follows that $u_{\xi\xi} = u_{xx}$. Then Eq. (1.37) becomes

$$-v\frac{\partial u}{\partial x} + \frac{\partial u}{\partial t} + (u+v)\frac{\partial u}{\partial x} = \nu\frac{\partial^2 u}{\partial x^2}\,,$$

and we see that the cancellation mentioned earlier indeed occurs, completing proof of Galilean invariance of Burgers' equation. A quite similar proof can be applied for the N.–S. equations.

As is noted in [80], Chap. 1, with the exception of Galilean invariance the above symmetries are broken as *Re* is increased. But once *Re* is sufficiently high to admit fully-developed turbulent behavior, the symmetries are restored in a statistical sense provided there is no directional forcing. However, we remark that such forcing is nearly always present in actual flows, so it is only in the case of high-*Re* homogeneous turbulence that symmetries, even in the statistical sense, are truly restored.

Existence, uniqueness, regularity of N.–S. solutions. Here we briefly describe current knowledge of the nature of solutions to the Navier–Stokes equations. This is particularly important because these equations must be solved by numerical methods implemented on digital computers, and the nature of solutions directly influences what types of methods will be effective, and what types will not. We encourage the reader to consult Foias *et al.* [89] for a more thorough treatment corresponding to the present discussions. We begin by stating the main results in an informal way, without defining terms. We then proceed to, again rather informally, define terms and explain consequences of the results.

Results for the N.–S. equations can be classified as pertaining to 2-D or to 3-D flows, and as to whether they relate to weak or strong solutions. With regard to these classifications, it is usually said that essentially everything worth proving for 2-D problems has been proven. In particular, existence and uniqueness of solutions, both weak and strong, have been proven for all time beyond

any specified initial time for quite reasonable (physically) problems. In 3D, long-time existence can be demonstrated for weak solutions, but uniqueness has not been proven for this case. On the other hand, only short-time existence has been proven for 3-D strong solutions, but it is known that these are unique. Especially with regard to existence of strong solutions, the constraints that must be imposed on the shape of the domain Ω (in particular, smoothness of $\partial\Omega$), the boundary and initial conditions, and especially on body-force terms and Reynolds number ($\sim$viscosity) can be quite stringent if solutions are to be proven to exist for any but very short times.

We recall that a *weak solution* is one that is not sufficiently differentiable to be substituted into the differential form of the equations, and instead only satisfies an integral (weak) form of the equations. The Galerkin approximation to be presented next is one such form. A *strong solution* is one that is sufficiently smooth to satisfy the original differential equation(s) in the sense of L^2. It is important to note, however, that this is not the same as a *classical solution* which is sufficiently differentiable to permit substitution into the differential equation at all points of the problem domain. In the case of a strong solution there can exist a nonempty set of points within the domain at which this cannot be done.

In summary, in 2D it can be proven that both weak and strong solutions to the N.–S. equations exist for all time and are unique; hence, they are equivalent. In 3D, weak solutions exist for all time, but may not be unique; and strong solutions are unique, but existence can be proven for only finite times.

There are several important consequences of these results. First, since it is widely accepted that turbulence is, indeed, three dimensional, inability to prove long-time existence of strong solutions to the 3-D N.–S. equations might cast some doubt on acceptance of these equations as the correct representation of turbulent flow. In particular, without long-time existence it is difficult to argue that stationary solutions even exist, and analysis of these is just as important in modern theories as in classical ones. But it is usually felt that lack of the desired long-time existence proof reflects more on inadequacy of current mathematical techniques than suggests an inappropriateness of the N.–S. equations. In support of this is the general success of DNS, not only in the sense of being able to quite accurately reproduce known turbulence results but also in terms of making predictions later confirmed in laboratory experiments—although only for relatively low Re. (DNS has yet to be performed for high-Re flows for reasons we have mentioned previously and which we will consider in detail below.)

But even if it should happen that one day it is proven that long-time strong solutions do <u>not</u> exist, this removes only uniqueness and regularity; we already know 3-D weak solutions exist for all time. They simply are not very smooth, and they might not be unique.

Second, even in the context of strong solutions, which satisfy the N.–S. equations only in the sense of L^2, *i.e.*,

$$\int_\Omega (\boldsymbol{U}_t + \boldsymbol{U}\cdot\nabla\boldsymbol{U} + \nabla P - \nu\Delta\boldsymbol{U})^2 \, d\boldsymbol{x} \to 0$$

as the number of modes in a Fourier representation (or, equivalently, the number of grid points in a finite-difference, or finite-volume, or finite-element approximation) $\to \infty$, is the fact that higher derivatives beyond second that appear in representations of the truncation error of typical discrete methods may not exist, and thus cannot be bounded. This implies that attempting to employ high-order methods for discretization of the N.–S. equations in strong form is likely doomed to fail, at least at high Re, unless fairly elaborate filtering is also used. On the other hand, spectral methods will not suffer from this difficulty and are of "infinite-order" accuracy for smooth solutions. Furthermore, finite-element methods (FEMs) are oftened formulated for a weak form of the equations being solved, so higher-order FEMs should work satisfactorily if properly formulated.

Finally, we observe that if it turns out that only weak solutions exist for long times, and in addition they cannot be proven to be unique, it will then become very important for experimentalists

to attempt to find multiple solutions arising from the same set of conditions (and there will be many significant difficulties with such attempts). We note that this situation is distinct from sensitivity to initial conditions which has always been observed in laboratory experiments, but, on the other hand, could still be influenced by this phenomenon. Instead, it corresponds to a situation for which, in the dynamical systems sense, there exists at least two attractors for the "same" basin of attraction. This can occur for a basin of attraction whose topology is so fractally intertwined that extremely small changes in initial data lead to long-time behavior associated with entirely different attractors—as opposed to simply different trajectories on the same attractor as occurs in the usual sense of SIC. Clearly, there are stability issues to be addressed here—probably via mathematics instead of laboratory experiments—and this might offer an approach to proving uniqueness of weak solutions; in particular, possibly there is only a single stable one. At the same time, there seem to be a number of published results (usually not very well supported) suggesting such a nonuniqueness property. Both experimental and computational studies have occasionally indicated this.

Galerkin approximation to the N.–S. solutions. For simplicity we treat the 2-D dimensionless form of Eqs. (1.1) in the absence of body forces. We express these here as

$$u_x + v_y = 0\,. \tag{1.38a}$$

$$u_t + (u^2)_x + (uv)_y = -p_x + \frac{1}{Re}\Delta u\,, \tag{1.38b}$$

$$v_t + (uv)_x + (v^2)_y = -p_y + \frac{1}{Re}\Delta v\,, \tag{1.38c}$$

and we represent the dependent variables in Fourier series:

$$u(x,y,t) = \sum_{\boldsymbol{k}}^{\infty} a_{\boldsymbol{k}}(t)\varphi_{\boldsymbol{k}}(x,y)\,, \tag{1.39a}$$

$$v(x,y,t) = \sum_{\boldsymbol{k}}^{\infty} b_{\boldsymbol{k}}(t)\varphi_{\boldsymbol{k}}(x,y)\,, \tag{1.39b}$$

$$p(x,y,t) = \sum_{\boldsymbol{k}}^{\infty} c_{\boldsymbol{k}}(t)\varphi_{\boldsymbol{k}}(x,y)\,, \tag{1.39c}$$

with $\boldsymbol{k} \equiv (k_1, k_2)^T$. The lower bound for components of this wavevector is typically one of $-\infty$, 0 or 1. Also, an analogous expansion would be needed for dependent variables contained in a body force if one were present.

For convenience we will assume the domain Ω is a rectangle and that the boundary conditions used with Eqs. (1.38) are periodic. Although this is a quite restrictive situation, our main goal here is to introduce some mathematical notions associated with the N.–S. equations when viewed in Fourier space, so the overall simplicity of this arrangement is an advantage. It is important to note, however, that it is difficult to lift these conditions and retain a numerically efficient procedure; as a consequence, most computer implementations of this method are quite similar to the version we will discuss here.

In light of the periodicity conditions it is natural to employ complex exponentials as basis functions; that is, we set

$$\varphi_{\boldsymbol{k}}(x,y) = e^{i\boldsymbol{k}\cdot\boldsymbol{x}} = e^{i(k_1x+k_2y)} = e^{ik_1x}e^{ik_2y}\,. \tag{1.40}$$

The last form on the right is often termed a "tensor product" basis because its two factors are defined in an uncoupled way on subsets of two separate copies of $\mathbb{R}^1$. (Clearly, an analogous form exists for 3D.)

To begin construction of the Galerkin form of Eqs. (1.38) we start with the simplest, Eq. (1.38a), the divergence-free condition, or continuity equation. We substitute Eqs. (1.39a) and (1.39b) into this equation and commute summation and differentiation to obtain

$$i\sum_{\boldsymbol{k}}^{\infty}\left(k_1 a_{\boldsymbol{k}} + k_2 b_{\boldsymbol{k}}\right)\varphi_{\boldsymbol{k}} = 0\,.$$

Since this must hold at (almost) all points of Ω, and $\varphi_{\boldsymbol{k}} \not\equiv 0$, we must have

$$k_1 a_{\boldsymbol{k}} + k_2 b_{\boldsymbol{k}} = 0 \qquad \forall\ \boldsymbol{k}\,. \tag{1.41}$$

Similarly, if we substitute the expansions (1.39) into the x-momentum equation (1.38b), we obtain

$$\frac{\partial}{\partial t}\sum_{\boldsymbol{\ell}} a_{\boldsymbol{\ell}}\varphi_{\boldsymbol{\ell}} + \frac{\partial}{\partial x}\sum_{\boldsymbol{\ell},\boldsymbol{m}} a_{\boldsymbol{\ell}} a_{\boldsymbol{m}}\varphi_{\boldsymbol{\ell}}\varphi_{\boldsymbol{m}} + \frac{\partial}{\partial y}\sum_{\boldsymbol{\ell},\boldsymbol{m}} a_{\boldsymbol{\ell}} b_{\boldsymbol{m}}\varphi_{\boldsymbol{\ell}}\varphi_{\boldsymbol{m}} =$$

$$-\frac{\partial}{\partial x}\sum_{\boldsymbol{\ell}} c_{\boldsymbol{\ell}}\varphi_{\boldsymbol{\ell}} + \frac{1}{Re}\left[\frac{\partial^2}{\partial x^2}\sum_{\boldsymbol{\ell}} a_{\boldsymbol{\ell}}\varphi_{\boldsymbol{\ell}} + \frac{\partial^2}{\partial y^2}\sum_{\boldsymbol{\ell}} a_{\boldsymbol{\ell}}\varphi_{\boldsymbol{\ell}}\right].$$

Again commuting summation and differentiation yields

$$\sum_{\boldsymbol{\ell}} \dot{a}_{\boldsymbol{\ell}}\varphi_{\boldsymbol{\ell}} + i\sum_{\boldsymbol{\ell},\boldsymbol{m}}(\ell_1 + m_1) a_{\boldsymbol{\ell}} a_{\boldsymbol{m}}\varphi_{\boldsymbol{\ell}}\varphi_{\boldsymbol{m}} + i\sum_{\boldsymbol{\ell},\boldsymbol{m}}(\ell_2 + m_2) a_{\boldsymbol{\ell}} b_{\boldsymbol{m}}\varphi_{\boldsymbol{\ell}}\varphi_{\boldsymbol{m}} =$$

$$-i\sum_{\boldsymbol{\ell}}\ell_1 c_{\boldsymbol{\ell}}\varphi_{\boldsymbol{\ell}} - \frac{1}{Re}\sum_{\boldsymbol{\ell}}\left(\ell_1^2 + \ell_2^2\right) a_{\boldsymbol{\ell}}\varphi_{\boldsymbol{\ell}}\,.$$

It can be seen that the solutions (the $a_{\boldsymbol{\ell}}$s, $b_{\boldsymbol{\ell}}$s, $c_{\boldsymbol{\ell}}$s) to this equation can be complex, but we are only interested in those that are real if we are considering solutions corresponding to actual physics of fluid flow. Indeed, if we had used sine and cosine as basis functions rather than the otherwise more convenient complex exponentials, this would not have been a concern. Furthermore, the reader will recall that sines and cosines may be expressed in terms of complex exponentials in any case. This suggests that a means of avoiding the potentially complex solutions should be available. In fact, all that is required is setting $a_{-\boldsymbol{k}} = \overline{a}_{\boldsymbol{k}}$, where the overbar here denotes complex conjugate, and the imaginary parts become zero (see, *e.g.*, [89]).

This permits us to consider the preceding equation without the imaginary factors i and proceed formally. (This description is lacking some details, but it is sufficient for our purposes.) We now use orthonormality of $\{\varphi_{\boldsymbol{k}}\}$ and form inner products of each of these with the above equation to obtain

$$\dot{a}_{\boldsymbol{k}} + \sum_{\boldsymbol{\ell},\boldsymbol{m}} A^{(1)}_{\boldsymbol{k}\boldsymbol{\ell}\boldsymbol{m}} a_{\boldsymbol{\ell}} a_{\boldsymbol{m}} + \sum_{\boldsymbol{\ell},\boldsymbol{m}} B^{(1)}_{\boldsymbol{k}\boldsymbol{\ell}\boldsymbol{m}} a_{\boldsymbol{\ell}} b_{\boldsymbol{m}} = -k_1 c_{\boldsymbol{k}} - \frac{|\boldsymbol{k}|^2}{Re} a_{\boldsymbol{k}}\,, \quad \forall\ -\infty < \boldsymbol{k} < \infty\,. \tag{1.42}$$

In these equations the $A^{(1)}_{\boldsymbol{k}\boldsymbol{\ell}\boldsymbol{m}}$s and $B^{(1)}_{\boldsymbol{k}\boldsymbol{\ell}\boldsymbol{m}}$s, sometimes termed *Galerkin triple products*, are defined as, for example,

$$A^{(1)}_{\boldsymbol{k}\boldsymbol{\ell}\boldsymbol{m}} \equiv (\ell_1 + m_1)\int_{\Omega}\varphi_{\boldsymbol{k}}\varphi_{\boldsymbol{\ell}}\varphi_{\boldsymbol{m}}\,d\boldsymbol{x}\,, \tag{1.43}$$

where the (1) superscript denotes the x-momentum equation. Clearly, an analogous result holds for the y-momentum equation:

$$\dot{b}_{\boldsymbol{k}} + \sum_{\boldsymbol{\ell},\boldsymbol{m}} A^{(2)}_{\boldsymbol{k}\boldsymbol{\ell}\boldsymbol{m}} b_{\boldsymbol{\ell}} a_{\boldsymbol{m}} + \sum_{\boldsymbol{\ell},\boldsymbol{m}} B^{(2)}_{\boldsymbol{k}\boldsymbol{\ell}\boldsymbol{m}} b_{\boldsymbol{\ell}} b_{\boldsymbol{m}} = -k_2 c_{\boldsymbol{k}} - \frac{|\boldsymbol{k}|^2}{Re} b_{\boldsymbol{k}}\,. \tag{1.44}$$

We remark that this is the step (constructing the inner products) that casts the N.–S. equations in a weak, integral form, but we should also note that the Fourier representations only guarantee convergence in L^2 in any case; so we would not necessarily expect strong solutions, in general.

At this point we should recall that it is generally possible to eliminate pressure from the momentum equations. If we view Eqs. (1.42) and (1.44) as a vector equation for the vector $(a_{\boldsymbol{k}}, b_{\boldsymbol{k}})$ of Fourier coefficients, we can take the dot product of this with $\boldsymbol{k} = (k_1, k_2)^T$ to obtain

$$k_1\dot{a}_{\boldsymbol{k}} + k_2\dot{b}_{\boldsymbol{k}} + \sum_{\boldsymbol{\ell},\boldsymbol{m}} \left[k_1\left(A^{(1)}_{\boldsymbol{k\ell m}}a_{\boldsymbol{\ell}}a_{\boldsymbol{m}} + B^{(1)}_{\boldsymbol{k\ell m}}a_{\boldsymbol{\ell}}b_{\boldsymbol{m}}\right) + k_2\left(A^{(2)}_{\boldsymbol{k\ell m}}b_{\boldsymbol{\ell}}a_{\boldsymbol{m}} + B^{(2)}_{\boldsymbol{k\ell m}}b_{\boldsymbol{\ell}}b_{\boldsymbol{m}}\right)\right]$$

$$= \quad -\left(k_1^2 + k_2^2\right)c_{\boldsymbol{k}} - \frac{|\boldsymbol{k}|^2}{Re}\left(k_1 a_{\boldsymbol{k}} + k_2 b_{\boldsymbol{k}}\right) \qquad \forall \ -\infty < \boldsymbol{k} < \infty\,.$$

We observe that this merely corresponds to constructing the divergence of the N.–S. equations in Fourier space.

From Eq. (1.41) we see that the first term on the left-hand side and the second on the right-hand side are both identically zero, and moreover, we can solve what remains for $c_{\boldsymbol{k}}$ in terms of only the $a_{\boldsymbol{k}}$s and $b_{\boldsymbol{k}}$s:

$$c_{\boldsymbol{k}} = -\frac{1}{|\boldsymbol{k}|^2}\sum_{\boldsymbol{\ell},\boldsymbol{m}}\left[k_1\left(A^{(1)}_{\boldsymbol{k\ell m}}a_{\boldsymbol{\ell}}a_{\boldsymbol{m}} + B^{(1)}_{\boldsymbol{k\ell m}}a_{\boldsymbol{\ell}}b_{\boldsymbol{m}}\right) + k_2\left(A^{(2)}_{\boldsymbol{k\ell m}}b_{\boldsymbol{\ell}}a_{\boldsymbol{m}} + B^{(2)}_{\boldsymbol{k\ell m}}b_{\boldsymbol{\ell}}b_{\boldsymbol{m}}\right)\right]\,. \tag{1.45}$$

This implies, as we expected, that Eqs. (1.42) and (1.44) can be expressed in a form that is independent of the pressure. Beyond this it explicitly demonstrates the quantitative dependence of pressure on the velocity field—in Fourier space. It is interesting to note the tendency to view pressure as the driving force that sets velocity components of the N.–S. equations, and this is to some extent true for externally-applied (boundary) pressures. But we see from Eq. (1.45) that internal static pressures are set by the velocity field within the confines of the divergence-free condition—just the opposite of the usual (physical) interpretation.

It is worthwhile to examine some of the details of the Galerkin form of the N.–S. equations. We first make the obvious observation that this form comprises a system of equations for the time evolution of the Fourier coefficients of the velocity components, and as such is a dynamical system recall (Def. 1.18). Furthermore, at each instant in time for which $a_{\boldsymbol{k}}(t)$ and $b_{\boldsymbol{k}}(t)$ are known, the Fourier coefficients for pressure can be calculated directly from Eq. (1.45); and these coefficients can be inserted into Eqs. (1.39) to obtain values of u, v and p at any point $\boldsymbol{x} \in \Omega$. Of course, in practice these Fourier representations can contain only a finite number N of terms, so the results are only approximate. But at least for solutions possessing a high degree of regularity, Fourier series converge very rapidly (in fact, exponentially), so not many terms are needed to obtain accurate approximations. For the reader interested in this feature, as well as numerous other details associated with these types of approximations, the monograph by Canuto *et al.* [100] is recommended.

Our main purpose in presenting the Galerkin form of the N.–S. equations is to demonstrate how this Fourier-space representation can be used to deduce qualitative mathematical (and physical) features of N.–S. flows. We consider only the x-momentum equation (1.42); but the same treatment applies to y-momentum, and extension to 3-D flows occurs in a natural way. We begin by temporarily neglecting all nonlinear terms in Eq. (1.42), with $c_{\boldsymbol{k}}$ also eliminated in light of (1.45). Then what remains is

$$\dot{a}_{\boldsymbol{k}} = -\frac{|\boldsymbol{k}|^2}{Re}a_{\boldsymbol{k}}\,, \tag{1.46}$$

the solution of which is

$$a_{\boldsymbol{k}}(t) = a_{\boldsymbol{k}}(0)e^{-\frac{|\boldsymbol{k}|^2}{Re}t}\,, \tag{1.47}$$

where

$$a_{\boldsymbol{k}}(0) \equiv \int_\Omega u_0(\boldsymbol{x}) e^{-i\boldsymbol{k}\cdot\boldsymbol{x}}\, d\boldsymbol{x}\,, \tag{1.48}$$

with $u_0(\boldsymbol{x})$ denoting the x component of initial velocity.

Clearly, this solution decays in time and approaches zero as $t \to \infty$. Moreover, the rate at which this occurs is $|\boldsymbol{k}|^2/Re$. In particular, for fixed Re higher wavenumber Fourier modes decay faster than do lower ones. From a mathematical perspective this suggests (but does not prove) convergence of the Fourier series representation and, hence, existence of solutions of the form Eqs. (1.39). On the other hand, if $|\boldsymbol{k}|$ is fixed, then the rate of decay of $a_{\boldsymbol{k}}$ decreases with increasing Re. Now if we recall that the right-hand side of Eq. (1.46) is precisely the Fourier-space representation of the viscous terms of Eq. (1.38b), we see that we can associate the rate of decay of $a_{\boldsymbol{k}}$ with physical viscous dissipation; in particular, increasing Re implies decreasing viscous dissipation, and conversely. While we will later obtain a specific formula for viscous dissipation rate, *per se*, in physical space (which is not the same as the viscous terms of the N.–S. equations), it is clear that these terms—and only these terms—contribute to decay of solutions directly.

We now consider effects of the nonlinear terms. To do this we first drop the linear viscous dissipation term from Eq. (1.42), and we also drop the term containing $c_{\boldsymbol{k}}$ arising from the pressure gradient in Eq. (1.38b) since, as can be seen from Eq. (1.45), this is directly related to the nonlinear terms; so nothing is lost at the qualitative level by ignoring it. Then we are left with

$$\dot{a}_{\boldsymbol{k}} = -\sum_{\boldsymbol{\ell},\boldsymbol{m}} \left[A^{(1)}_{\boldsymbol{k}\boldsymbol{\ell}\boldsymbol{m}} a_{\boldsymbol{\ell}} a_{\boldsymbol{m}} + B^{(1)}_{\boldsymbol{k}\boldsymbol{\ell}\boldsymbol{m}} a_{\boldsymbol{\ell}} b_{\boldsymbol{m}} \right]\,. \tag{1.49}$$

We first observe that when $\boldsymbol{\ell} = \boldsymbol{m}$ a quadratic term appears in the equation. It is worthwhile to consider the effects of this alone since an analytical solution can be obtained, and this will provide at least some insight into the qualitative behavior of the nonlinear terms in general. Thus, we solve the initial-value problem

$$\dot{a}_{\boldsymbol{k}} = -A^{(1)} a^2_{\boldsymbol{k}}\,, \tag{1.50}$$

with $a_{\boldsymbol{k}}(0)$ again given by Eq. (1.48) and subscript notation for $A^{(1)}_{\boldsymbol{k}\boldsymbol{\ell}\boldsymbol{m}}$ suppressed. We leave as a simple exercise to the reader demonstration that

$$a_{\boldsymbol{k}}(t) = \frac{1}{A^{(1)} t + 1/a_{\boldsymbol{k}}(0)}\,.$$

It is clear from this that if $A^{(1)} a_{\boldsymbol{k}}(0) > 0$, then $|a_{\boldsymbol{k}}| \to 0$ as $t \to \infty$, although only algebraically. But if $A^{(1)}$ and $a_{\boldsymbol{k}}(0)$ are of opposite signs, $a_{\boldsymbol{k}}(t) \to \infty$ can occur in finite time. This implies a potential for very ill behavior (including nonexistence after only a finite time) of N.–S. solutions. In 2D it is known that this does not actually occur (which implies this simplification is not an extremely accurate model); but as we have already indicated, this possibility has not been eliminated in 3D—it is possible that 3-D solutions to the N.–S. equations fail to exist after a finite time.

There is a further aspect of the behavior of the terms on the right-hand side of Eq. (1.49) that deserves mention. It is that such nonlinearities can generate new Fourier modes not present in the original representation (or in initial data for the problem). Here we consider only the first advective term of the x-momentum equation, $(u^2)_x$, and recall its Fourier representation:

$$(u^2)_x = \frac{\partial}{\partial x} \sum_{\boldsymbol{\ell},\boldsymbol{m}}^{N} a_{\boldsymbol{\ell}} a_{\boldsymbol{m}} \varphi_{\boldsymbol{\ell}} \varphi_{\boldsymbol{m}}\,,$$

written now for only a finite number N of Fourier modes, as would be required for computer implementation.

Now if the basis set $\{\varphi_{\boldsymbol{k}}\}$ is similar to complex exponentials or trigonometric functions we see that, *e.g.*,

$$\varphi_{\boldsymbol{\ell}}(\boldsymbol{x})\varphi_{\boldsymbol{m}}(\boldsymbol{x}) = e^{i\boldsymbol{\ell}\cdot\boldsymbol{x}}e^{i\boldsymbol{m}\cdot\boldsymbol{x}} = e^{i(\boldsymbol{\ell}+\boldsymbol{m})\cdot\boldsymbol{x}}\,.$$

Since each of $\boldsymbol{\ell}$ and $\boldsymbol{m}$ can be as large as N, their sum is often greater than N, and the corresponding nonlinear term $a_{\boldsymbol{\ell}}a_{\boldsymbol{m}}$ generates solution behaviors that cannot be resolved by the given representation. We note here that this does not occur for the Galerkin procedure due to the values taken on by the Galerkin triple products, and due to the general global nature of the Fourier coefficients; but it does occur for the various Fourier collocation methods that are widely used as a more efficient alternative to the Galerkin procedure (see [100]) in the context of both DNS and LES. Moreover, it is easily argued that this must occur for finite-difference and finite-volume methods as well.

Finally, we emphasize that the actual time evolution of each of the $a_{\boldsymbol{k}}$s is effected by combinations of all the abovementioned behaviors, and as a consequence solutions can be very complicated—and very difficult to simulate. In particular, it should be clear that the $\boldsymbol{k}^{th}$ Fourier coefficient, $a_{\boldsymbol{k}}$, can be affected by coefficients corresponding to wavenumbers both less than and greater than $\boldsymbol{k}$ through the nonlinear terms of the Galerkin representation (including those arising from the pressure gradient). But the specific outcomes of such interactions are strongly influenced by the dissipation induced by the linear (viscous) terms. We recall that the degree of this is set by the combination of Reynolds number and the specific wavenumber under consideration.

General (mostly) Physical Concepts from Turbulence

In this section we begin by presenting some mostly heuristic ideas concerning the nature and causes of turbulence, relating these to the mathematical constructs of the preceding section. We then describe the various well-known length and time scales that occur in turbulent flows, again showing how these are associated with the mathematics—in particular, where they reside in the energy spectrum. In connection with the relationship between length and time scales we will introduce the often-used (especially in analysis of experimental data) Taylor hypothesis. Then we provide more details on homogeneous and isotropic turbulence, and we conclude the discussions with some basic analysis of the law of the wall.

What is turbulence?—and where does it come from? In some respects, it is not clear that the answer to either of these questions is completely known. Certainly, from a physical standpoint there are many unresolved issues. Earlier in this chapter we gave an extensive list of characterizations of turbulence and noted that turbulent flows always exhibit some (and maybe all) of these features; but this does not really define physical turbulence. On the other hand, we were able to at least propose a mathematical definition, namely, 3-D, chaotic solutions to the Navier–Stokes equations at high Reynolds number. But this "definition" is certainly not universally accepted for reasons we will discuss later in these lectures. So, the answer to the first question seems to yet be lacking, at least on purely physical grounds. We remark, that part of the difficulty seems to be that there are so many possible characterizations that do not all occur in every flow situation intuitively viewed as turbulent that evidently turbulence may be more than a single phenomenon. If so, this would suggest that the mathematical definition might be the best we can obtain.

Before addressing the second question we here introduce a list of misconceptions associated with turbulence. These have been extracted from a somewhat longer list provided by Tsinober [3] in Appendix D of his book.

'Statistical' and 'structural' contrapose each other.

Turbulence possesses a random (quasi)-Gaussian background.

Kolmogorov picture is structureless and quasi-Gaussian.

Large scales and small scales are decoupled.

'Eddy viscosity' and 'eddy diffusivity' *explain* the enhanced transfer rates of momentum, energy and passive objects.

Spatial fluxes represent 'cascade' in physical space.

Vorticity amplification is a result of the kinematics of turbulence.

Strain rate in turbulent flows is irrotational.

Enhanced dissipation in turbulent flows is due to vortex stretching.

Each of these has been touted as an important characterization of turbulence by numerous authors, but in [3] they are all shown to imply serious logical (and physical—based on experimental observations) flaws.

The second question can be answered somewhat more satisfactorily from a physical standpoint; and, fortunately, the answer is in accord with the mathematics of the N.–S. equations. In particular, it is well understood that turbulence only occurs in flows exhibiting fairly large velocity gradients. The simplest familiar example is pipe flow, but the descriptions we will provide work equally well for essentially any flow; only the mechanism for producing the velocity gradients would change.

We first observe that pipe flow is laminar until a Reynolds number of approximately 2000 is reached, whereupon transition begins. Recall that, in a pipe, the velocity satisfies the no-slip condition at the walls and adjusts to the centerline velocity across the radius of the pipe. Hence, there is a velocity gradient that is largest at the pipe wall. Furthermore, it is easy to imagine that if the wall of the pipe is rough the velocity gradients might be locally (but not necessarily uniformly) higher than in the smooth-wall case; and, indeed, for pipes having extremely smooth walls, it is possible to delay transition to turbulence until the Reynolds number is well above 50,000. Furthermore, there are no known uniform flows—either physically or mathematically—that exhibit transition.

From a purely mathematical standpoint, it is well known (and not difficult to show) that Fourier series converge far more slowly for functions having large derivatives than for very smooth, well-behaved functions. This, in turn, implies that larger wavenumbers will still provide significant contributions, and in the context of the Galerkin approximation to the N.–S. equations given above, this means that even higher wavenumber information will be generated spontaneously. Then, if Re is large, such wavenumber data will not be readily damped by the viscous terms, and chaos generated in the nonlinear terms will dominate the behavior—hence, turbulence. Conversely, if a flow has only small gradients, the linear viscous terms of the Fourier representation very effectively damp the solutions, thus maintaining, or even further reducing, the small gradients.

It is important to also mention here the classical view of this interpretation as given, for example, in Tennekes and Lumley [7]. In that reference, as well as many others, the role of vorticity is stressed, in conjunction with the energy cascade described earlier in Def. 1.68. In particular, it is argued that energy is usually supplied to a turbulent flow through mechanisms that create large vortices. It is hypothesized (*á la* Richardson [4]) that these vortices, usually referred to as "eddies," are somehow broken into smaller ones, that themselves split into still smaller ones, and so on, until they are sufficiently small as to be dissipated by viscosity. We remark, however, that vorticity, the quantification of the strength of such vortices, is not actually physics—vorticity is a purely mathematical definition. Indeed, vorticity is constructed from the velocity gradients described above—which are physics: the amount velocity changes over a given distance. So, it is not clear that a viewpoint expressed in terms of vorticity and eddies contributes to physical understanding even if it happens to be correct. But in addition, careful examination of most physical flows

does not clearly show sequences of sizes of vortices simply breaking or splitting as suggested by this simplistic scenario. Moreover, this view is not completely consistent with mathematics of the Galerkin approximation to the N.–S. equations. We noted in earlier discussions that Fourier coefficients (think "energy"—*i.e.*, the L^2 norm) are influenced by behaviors at both higher and lower wavenumbers; hence, energy might be transferred in either direction, at least locally, in wavenumber space. But this is not included (permitted?) in the energy cascade description involving breakup of eddies.

Length and time scales of turbulence. One of the items in our earlier list of characterizations of a turbulent flow at the end of Sec. 1.2 was its wide range of length and time scales. In fact, if such scales did not cover wide ranges, the "turbulence problem" would have been solved long ago. In such a case, the N.–S. equations could be easily and efficiently solved on a digital computer for essentially any physical flow. But as will be evident as we proceed, it is precisely this feature of turbulent flow that precludes use of DNS in most flow situations. Thus, it is especially important to understand some of the details.

We begin by noting that there are, in general, four main sets of scales in a turbulent flow (there may be more if other physical phenomena, *e.g.*, heat transfer and/or combustion are important); these are:

i) the *large scale*, based on the problem domain geometry,

ii) the *integral scale*, which is an $\mathcal{O}(1)$ fraction (often taken to be ~ 0.2) of the large scale (and termed the "outer scale," especially in Russian literature),

iii) the *Taylor microscale* which is an intermediate scale, basically corresponding to (actually, within) Kolmogorov's inertial subrange, and

iv) the *Kolmogorov* (or "dissipation") *scale* which is the smallest of turbulence scales (called the "inner scale" in Russian works).

Before considering details, it is worthwhile to first qualitatively compare these scales in terms of (spatial) wavenumbers (and we will later indicate that essentially the same description holds for temporal frequencies). Figure 1.5 displays the essential details as one would deduce from analysis of experimental measurements. We note, in passing, that this figure would correspond to a quite high Re because the range of wavenumbers encompassed by the inertial subrange is relatively large. In particular, at low Re there is essentially no inertial range, and as Re increases the length of this range increases. Observe that the range of wavenumbers appearing in the large scale and integral scale is only mildly effected, if at all, by increasing Re. This should be expected because the first of these is set by the size and shape of the flow domain, and the second is influenced by aspects of turbulence already present at relatively low wavenumbers, and by energy input. On the other hand, the wavenumber corresponding to beginning of the dissipation scales is strongly influenced by Re, as we have already seen in our Fourier analysis of the N.–S. equations (recall Eq. (1.47)). Thus, the wavenumber range covered by the inertial scales must increase with increasing Re. (Observe that we also have already indicated in the figure the $-5/3$ slope in the spectrum of this range in anticipation of later results.) If we now recall that $|\boldsymbol{k}| \sim 1/|\boldsymbol{r}|$ and $E(k) \sim a_k^2 \sim u^2$, we can also easily deduce the physical-space correspondences—in particular, the highest energy is associated with larger scales of motion.

We now treat each of these scales in more detail, beginning with the largest. On this scale we take the characteristic length to be L, basically the size of the physical problem domain (*e.g.*, the radius or diameter of a pipe, or the chord length of an airfoil), and we use U (the mean, or possibly centerline, velocity in a pipe, or the freestream velocity over an airfoil) for the characteristic velocity.

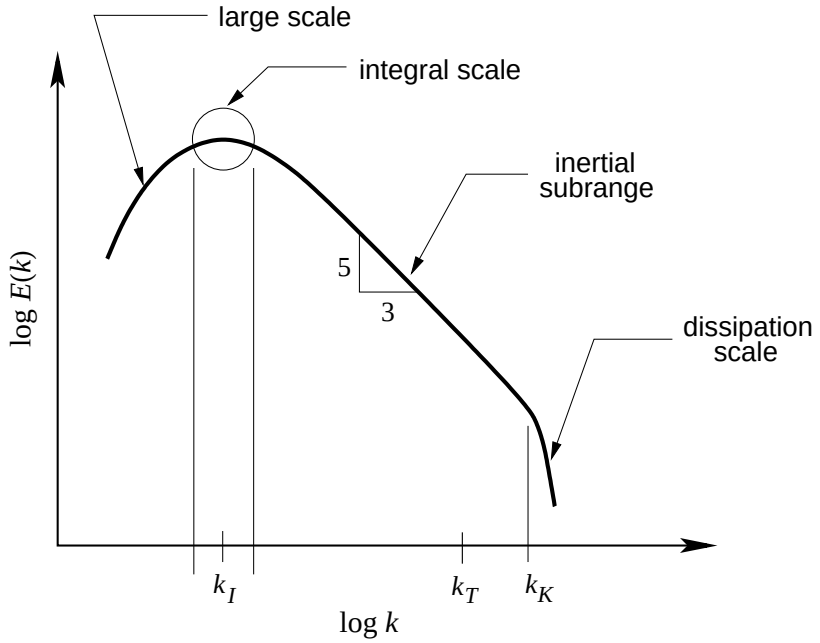

Figure 1.5: Turbulence energy wavenumber spectrum.

Then, if ν is the kinematic viscosity, the Reynolds number will be

$$Re_L = \frac{UL}{\nu}\,.$$

Furthermore, from the length and velocity scales we can construct a "convective" time scale given by $t_c = L/U$. There is a second time scale that can also be obtained from the given physical quantities; it is the "diffusive" time scale. Recall that kinematic viscosity has generalized units L^2/T, so it follows that $t_d = L^2/\nu$ is another possible time scale. It is of physical interest to form the ratio of t_d and t_c, as done in [7], to compare the rate at which flow properties are transferred by (molecular) diffusion, as compared to (macroscopic) convection. We find

$$\frac{t_d}{t_c} = \frac{L^2/\nu}{L/U} = Re_L\,,$$

indicating that Reynolds number can be viewed as a ratio of these time scales in addition to the usual interpretation as the ratio of inertial to viscous forces. Thus, if Re is large, the diffusive time is very long compared with time for convection, and hence diffusive effects are essentially negligible in the high-Re limit.

We next consider the integral scales. As hinted in Fig. 1.5, these do not cover a wide range; indeed, they are sometimes associated with a single wavenumber, k_I, the one corresponding to the maximum in turbulence energy. We have already indicated a statistical approach to estimating the integral time scale $\mathcal{T}$ in Eq. (1.17), and we can similarly obtain an integral length scale as

$$\ell = \frac{1}{\|u'\|^2_{L^2}} \int_{-\infty}^{\infty} u'(\boldsymbol{x},t)u'(\boldsymbol{x}+\boldsymbol{r},t)\,d\boldsymbol{r}\,, \tag{1.51}$$

where, as in Eq. (1.25), u' denotes a turbulent fluctuating component of velocity (with zero mean), and $\|\cdot\|^2_{L^2}$ is taken with respect to the spatial domain Ω. We note here that the integrations required to estimate ℓ and $\mathcal{T}$ give rise to the terminology. Clearly, once ℓ (or $\mathcal{T}$) is known, the other can

be estimated using some norm of $\boldsymbol{u}'$ as a velocity scale. We note here that $\boldsymbol{u}'$ in this case is not generally the same as that in the Reynolds decomposition, but rather is more closely related to the subgrid-scale part of the LES decomposition. It is probably best to view $\boldsymbol{u}'$ as a high-pass filtered quantity still retaining a significant portion of the large-scale motion. In this view, employing $\boldsymbol{u}'$ from a Reynolds decomposition is not extremely inaccurate. Figure 1.6 provides a "cartoon" of what is usually intended for this quantity. In particular, notice that much of the large-scale "structure" of

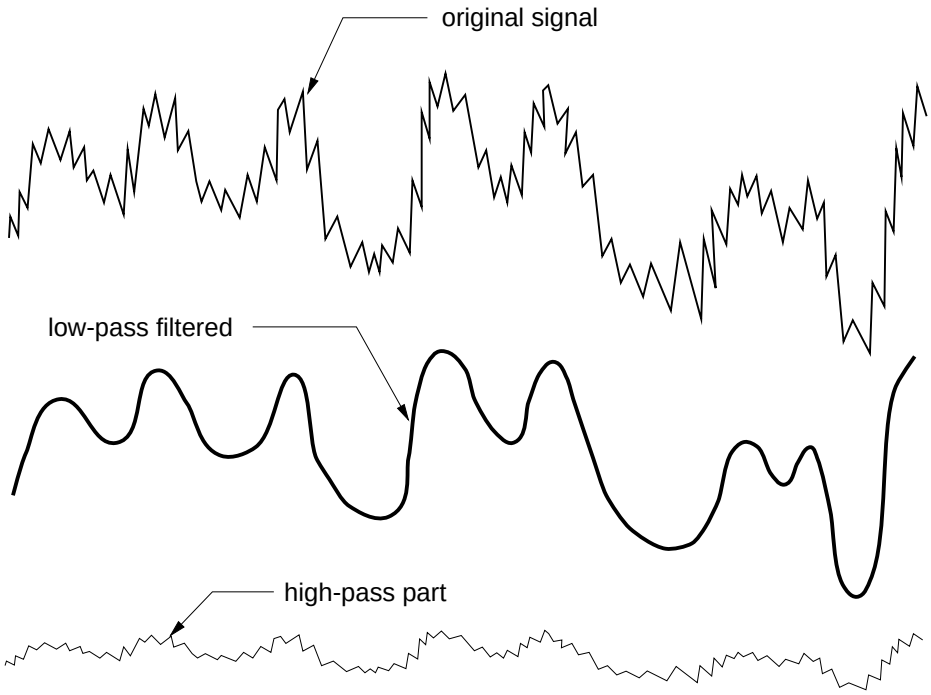

Figure 1.6: Low-pass and high-pass filtered parts of a signal.

the low-pass filtered signal still appears in the high-pass part, indicating presence of low-frequency (or wavenumber) behavior in addition to the obvious high-frequency content. Thus, u' is somewhere between the RANS fluctuations which carry all temporal information except the mean, and LES which, ideally, corresponds mainly to dissipation scales.

We should note that without further averaging the integral scales are not constant (and so do not correspond well with what is indicated in Fig. 1.5). In homogeneous turbulence spatial averaging is justified, and temporal averaging may be appropriate for stationary flows; but generally ℓ must be a function of $\boldsymbol{x}$ and t in complex turbulent flows. In any case we can now define the integral scale Reynolds number (often called the "turbulence" Reynolds number) as

$$Re_\ell \equiv \frac{|u'|\ell}{\nu}\,. \tag{1.52}$$

Here, $|u'|$ is usually taken to be the square root of the turbulence kinetic energy (per unit mass) consistent with the interpretation provided by Fig. 1.6; *i.e.*, $|u'| = k^{1/2}$ with $k = \frac{1}{2}(u'^2 + v'^2 + w'^2)$, or $|u'|$ may be the (square root of) *turbulence intensity* q^2, usually given as $q^2 = 2k$, but in some cases (mainly associated with isotropic turbulence) this value divided by three, corresponding to averaging the three fluctuating components.

It is of interest to also consider a somewhat different approach to obtaining the integral scale length, as provided by Lesieur [14]. In this treatment a velocity scale analogous to that discussed

above is used, and in conjunction with this, turbulence energy dissipation rate ε, usually given as

$$\varepsilon = 2\nu\|\boldsymbol{S}\|^2\,, \tag{1.53}$$

is employed to construct a length scale with elements of the strain rate tensor $\boldsymbol{S}$ given earlier in the discussion of Cartesian tensors. In [14] a slightly different definition of ε is used, namely, $\varepsilon = \nu\langle\boldsymbol{\omega}\cdot\boldsymbol{\omega}\rangle$. In Eq. (1.53) $\|\cdot\|$ is the usual matrix two norm, and in the preceding formula $\boldsymbol{\omega}$ is the vorticity vector. We note that these formulations are equivalent in the sense of norm equivalence since both $\boldsymbol{S}$ and $\boldsymbol{\omega}$ are obtained from rearrangement of $\nabla\boldsymbol{U}$ into a sum of symmetric ($\sim\boldsymbol{S}$) and skew-symmetric ($\sim\boldsymbol{\omega}$) parts as follows for any arbitrary portion of the velocity field.

$$\begin{aligned}
\nabla\boldsymbol{U} &= \begin{pmatrix} u_x & u_y & u_z \\ v_x & v_y & v_z \\ w_x & w_y & w_z \end{pmatrix} \\
&= \frac{1}{2}\left[\begin{pmatrix} 2u_x & u_y+v_x & u_z+w_x \\ v_x+u_y & 2v_y & v_z+w_y \\ w_x+u_z & w_y+v_y & 2w_z \end{pmatrix} + \begin{pmatrix} 0 & u_y-v_x & u_z-w_x \\ v_x-u_y & 0 & v_z-w_y \\ w_x-u_z & w_y-v_z & 0 \end{pmatrix}\right].
\end{aligned}$$

Also, we observe that the generalized units of ε must be $\mathrm{L}^2/\mathrm{T}^3$. Then it follows that a length scale can be constructed as

$$\ell = \frac{|u'|^3}{\varepsilon}\,. \tag{1.54}$$

Clearly, this is based entirely on fluctuating velocity still containing some relatively low wavenumber behavior, and hence, consistent with the physical interpretation of integral scale as inferred from Fig. 1.5. But we remark that if $\boldsymbol{S}$ is decomposed into mean and fluctuating components, $\|\boldsymbol{S}\|$ will be dominated by the latter, implying that ε is mainly a small-scale quantity. We would expect this based on our earlier Fourier analysis of the N.–S. equations.

We now derive the length and time scales associated with the Taylor microscale. We begin with a definition for the *Taylor microscale length* provided in [14]:

$$\lambda^2 = \frac{\langle|u'|^2\rangle}{\langle\|\boldsymbol{S}\|^2\rangle}\,, \tag{1.55}$$

where we are using the velocity scale discussed previously and are now formally indicating an averaging process with $\langle\cdot\rangle$, and again using norm of the strain rate tensor in place of norm of vorticity since this is more often done (see [7]). Then from Eq. (1.53) (with the factor of two suppressed, in accord with [14], we obtain the Taylor microscale length expressed as

$$\lambda = \left[\frac{\nu\langle|u'|^2\rangle}{\varepsilon}\right]^{1/2}. \tag{1.56}$$

A time scale can be obtained using this length scale and the velocity $|u'|$, and the Taylor microscale Reynolds number is calculated as

$$Re_\lambda = \frac{|u'|\lambda}{\nu}\,. \tag{1.57}$$

We again note that the Taylor microscale length is roughly consistent with the Kolmogorov inertial subrange scales.

We now find expressions for the smallest scales of turbulence. These were derived by Kolmogorov under the assumption that at these scales mainly dissipation would be important, so the only two

physical parameters needed to describe behavior from a dimensional standpoint are viscosity ν and dissipation rate ε of turbulence kinetic energy. We will see that combinations of only these parameters lead to all three of length, time and velocity scales in the dissipation range. Again, recall that the generalized units of kinematic viscosity are L^2/T, and those for energy dissipation rate are $\mathrm{L}^2/\mathrm{T}^3$. It follows that a length scale can be obtained by eliminating time between these two sets of units. This can be done by forming the ratio of ν^3 with ε, and taking the fourth root of the result. Thus, we obtain the *Kolmogorov length scale*

$$\eta = \left(\frac{\nu^3}{\varepsilon}\right)^{1/4} . \tag{1.58}$$

Similarly, a *Kolmogorov time scale* can be constructed as the square root of the ratio of ν to ε:

$$\tau = \left(\frac{\nu}{\varepsilon}\right)^{1/2} . \tag{1.59}$$

Finally, the *Kolmogorov scale velocity* is just the ratio of the preceding quantities:

$$\upsilon = (\nu\varepsilon)^{1/4} . \tag{1.60}$$

One can readily see that the Reynolds number formed from this length and velocity scale equals unity, as might be expected on scales where viscous dissipation dominates all other phenomena.

It is of interest to compare some of these various scales. We observe that the length scales and Reynolds numbers can be related as follows. First, we can compare the Kolmogorov length scale η with the integral scale length ℓ using Eq. (1.58) with Eq. (1.54) solved for ε to write

$$\eta \sim \left(\frac{\nu^3}{|u'|^3/\ell}\right)^{1/4} ;$$

then

$$\frac{\eta}{\ell} \sim \left(\frac{\nu^3}{|u'|^3\ell^3}\right)^{1/4} \sim Re_\ell^{-3/4} ,$$

or

$$\frac{\ell}{\eta} \sim Re_\ell^{3/4} . \tag{1.61}$$

We have used the $\sim$ symbol in deriving this result to emphasize that both of Eqs. (1.54) and (1.58) are merely based on dimensional arguments. Nevertheless, Eq. (1.61) has very important consequences for computation because it implies that the dissipation scales, which must be resolved in a DNS of the N.–S. equations, scale like the integral scale Re to the 3/4 power—and this is for only one direction. Thus, in a 3-D problem the gridding requirements, and hence the computational work, must scale like $Re_\ell^{9/4}$ for a single time step. As noted by Frisch [80] a typical time step for numerical stability is of the order of the space step (Courant condition), implying $\sim \mathcal{O}(Re_\ell^{3/4})$ required time steps; so a very optimistic (under)estimate of total arithmetic for DNS is $\sim Re_\ell^3$; for Re (which, recall, is somewhat larger than Re_ℓ) of any reasonable size this is still a very formidable computation even on modern parallel supercomputers.

We next compare the Taylor and Kolmogorov scales, λ and η, respectively. Again starting with Eq. (1.58) we now use Eq. (1.56) with $\langle|u'|^2\rangle$ replaced with $|u'|^2$ for notational convenience to write

$$\eta \sim \left(\frac{\nu^3\lambda^2}{\nu|u'|^2}\right)^{1/4} \sim \left(\frac{\nu^2\lambda^2}{|u'|^2}\right)^{1/4} .$$

Then

$$\frac{\eta}{\lambda} \sim \left(\frac{\nu^2}{\lambda^2 |u'|^2}\right)^{1/4} \sim Re_\lambda^{-1/2}\,,$$

or

$$\frac{\lambda}{\eta} \sim Re_\lambda^{1/2}\,. \tag{1.62}$$

This shows that length scales in the inertial subrange are a factor of square root of the Taylor microscale Reynolds number larger than those in the dissipation range. For realistic flow situations this tends to be much greater than an order of magnitude.

We can now use results (1.61) and (1.62) to estimate Re_λ in terms of Re_ℓ. First, from the definitions of these two Reynolds numbers, Eqs. (1.57) and (1.52), respectively, we have

$$\frac{Re_\lambda}{Re_\ell} = \frac{\lambda}{\ell}\,. \tag{1.63}$$

But from (1.61) and (1.62) we have

$$\frac{\lambda}{\ell} \sim \frac{Re_\lambda^{1/2}}{Re_\ell^{3/4}} = \left(\frac{\lambda}{\ell}\right)^{1/2} Re_\ell^{1/2} Re_\ell^{3/4} = \left(\frac{\lambda}{\ell}\right)^{1/2} Re_\ell^{-1/4}\,,$$

or

$$\left(\frac{\lambda}{\ell}\right)^{1/2} \sim Re_\ell^{-1/4} \quad \Rightarrow \quad \frac{\lambda}{\ell} \sim Re_\ell^{-1/2}\,.$$

Then using Eq. (1.63) for λ/ℓ yields

$$Re_\lambda \sim Re_\ell^{1/2}\,. \tag{1.64}$$

Finally, again using Eq. (1.62) shows that

$$\frac{\lambda}{\eta} \sim Re_\ell^{1/4}\,,$$

or

$$\eta \sim \lambda Re_\ell^{-1/4}\,. \tag{1.65}$$

Thus, we conclude that

$$\lambda \sim \ell Re_\ell^{-1/2}\,, \qquad \text{and} \qquad \eta \sim \lambda Re_\ell^{-1/4}\,, \tag{1.66}$$

and substitution of the first of these into the second leads back to (1.61), as it must. Finally, note that for large Re_ℓ (or Re_L), we have $\eta \ll \lambda \ll \ell$.

Taylor's hypothesis. As noted in Def. 1.110, Taylor's hypothesis can be employed to deduce spatial information about turbulent fluctuations using time series of measurements at a single point, or at a sequence of points at which measurements have not been taken simultaneously. The preceding discussions of length and time scales suggest that these can typically be related through some velocity scale, and this is what is involved when invoking Taylor's hypothesis. In particular, as described in [7], measurements of fluctuating velocities are sometimes collected by traversing a probe through the flow field so rapidly that the nature of the turbulence does not change significantly during the measurement process. This permits construction of spatial derivatives of the fluctuating quantities at an ostensibly fixed time. If the speed of traversal U of the probe is sufficiently high, then a fluctuating velocity signal $u'(t)$ at a fixed location can be identified with fluctuations at a different location a distance x away by substituting $t = x/U$. This is often termed a "frozen

turbulence" approximation, and it is shown in Hinze [6], among other places, that $|u'|/U \ll 1$ must hold for results obtained from Taylor's hypothesis to be valid.

A somewhat more applicable description of Taylor's hypothesis is provided by Garde [101]. Assume the average flow velocity in the x direction is U_{avg}. Then turbulent fluctuations u' observed at a fixed point in space can be approximately interpreted as resulting from a frozen turbulence pattern being convected through the point from elsewhere by a mean flow velocity U_{avg}. This suggests we can relate temporal and spatial derivatives of flow quantities as

$$\frac{\partial}{\partial t} = -U_{avg}\frac{\partial}{\partial x}\,, \tag{1.67}$$

with the negative sign arising from the fact that a measurement at the current time and spatial location corresponds to the frozen pattern that earlier was upstream (negative x) of the current location. It is noted in [101] that for shear flows $\sqrt{\overline{u'^2}} \ll 0.45$ must hold, and for isotropic, homogeneous turbulence a value on the order of 10^{-2} is needed for use of Taylor's hypothesis to provide accurate results, an implicit assumption being that $U_{avg} \sim \mathcal{O}(1)$.

Homogeneous and/or isotropic turbulence. In Def. 1.75 we noted that homogeneous turbulence is such that statistics are invariant under spatial translations, and Def. 1.80 implies that isotropic turbulence is invariant under rotations and reflections. Because rotations and reflections can always be constructed as combinations of translations, one might suppose that homogeneity implies isotropy; but this is <u>not</u> the case. Instead, isotropy is far more restrictive than homogeneity, as we will now describe.

It is important to first note the consequences of homogeneity. The requirement that statistical properties remain invariant under arbitrary translations implies, for example, that $\overline{u'^2}(\boldsymbol{x}) = \overline{u'^2}(\boldsymbol{x}+\boldsymbol{r})$ for any possible vector $\boldsymbol{r} \in \mathbb{R}^3$ within the domain Ω of the flow field. On the other hand, homogeneity <u>does not</u> imply that, *e.g.*, $\overline{u'^2}(\boldsymbol{x}) = \overline{v'^2}(\boldsymbol{x})$, and this provides the distinction between homogeneity and isotropy.

Isotropy requires invariance of statistical quantities under rotations and reflections of the coordinate system. Suppose for definiteness that we consider rotating the coordinates through a counter-clockwise angle of 90°. Then in this new system, what had been the physical u component of velocity will now be the v component, and conversely (up to a sign). Thus, if statistics are to remain invariant under rotations it must be that

$$\overline{u'^2} = \overline{v'^2} = \overline{w'^2} \tag{1.68}$$

throughout the flow field. But to enforce statistical invariance under arbitrary rotations, beyond (1.68) must be the requirement that derivatives of these quantities in the normal coordinate direction relative to each also be invariant. This implies that

$$\frac{\partial\overline{u'^2}}{\partial x} = \frac{\partial\overline{v'^2}}{\partial y} = \frac{\partial\overline{w'^2}}{\partial z} \tag{1.69}$$

must hold.

It is thus fairly easy to see that a uniform strain in one direction on homogeneous isotropic turbulence will destroy the isotropy, but not the homogeneity. On the other hand, few if any physical shear flows are uniform, nor are they typically aligned with a single coordinate direction; so generally, introduction of shear (strain) will eliminate both isotropy and homogeneity. That is, real turbulent flows are inhomogeneous and anisotropic. Nevertheless, as is argued in Batchlor [39], the simplifications that can be achieved via assuming homogeneity and isotropy are sufficient to permit some pure analysis of turbulence, and this is considered worthwhile.

It is often argued that the smallest scales of turbulence must be isotropic; this is referred to as "local" isotropy, meaning local in wavenumber. This view is made reasonable by the Fourier analysis of an earlier section. In particular, we showed that for very high wavenumbers the viscous terms dominate the nonlinear terms (including pressure gradient terms), and it is also easily checked that for this situation there is essentially no coupling between momentum equations. Moreover, all such equations are of the same form so it might be expected that they would all have the same solution. Hence, small-scale behavior would be isotropic. (We will later see that this is an argument used to suggest that constructing SGS models for LES ought not be too difficult.)

This expectation of at least local isotropy leads to the concept of "return to isotropy" described in Def. 1.97. That is, removal of a strain field that is inducing anisotropy should allow turbulence to relax back to an isotropic state, at least on small scales. It is clear, again from Fourier representations, that this must occur at a finite rate, which for certain flows can be measured in laboratory experiments. This provides a quite detailed test of turbulence models (most RANS models fail this test), and some representative experimental results of Choi and Lumley [102] are compared with a well-known second-order closure model in a paper by Speziale [103].

It should be observed, however, that the notion of local isotropy is not precisely correct mathematically, nor is it strictly observed in laboratory experiments. Results coming from the rather heuristic Fourier analysis are too simplified. Indeed, Brasseur and Yeung [104] have theoretically shown via a detailed analysis of the Fourier representation of the N.–S. equations that, in fact, the smallest scales inherit much of the anisotropic structure of the large scales. While the work of these authors has at times been viewed as somewhat controversial, there have been numerous experimental results also indicating a lack of isotropy on small scales, for example, Shen and Warhaft [105]. Thus, strict local isotropy must in general be viewed with some skepticism and accepted only as an idealization.

Law of the wall. Here we provide a mainly heuristic (and pictorial) treatment of one of the best-known and widely-used concepts from the classical theory of turbulence. We will not present derivations at this time because these can best be done after the RANS equations have been obtained. But we will be able to deduce some fairly useful results having direct links to the physics of turbulent flows. At the same time, as is done in [7], we connect the various physical regions associated with the law of the wall to the scales introduced by Kolmogorov.

We begin by noting that there are two quite general and distinct types of turbulent flow beyond the homogeneous isotropic flow treated in the preceding subsection: namely, free shear flows and wall-bounded shear flows. We will consider the theory of both of these in Chap. 2 after we have derived the RANS equations. Here, we note that examples of the former include wakes, jets and shear layers, and that theoretical treatment of these is simpler than that of the latter because they possess only a single length scale.

As the name suggests, the law of the wall is asociated with wall-bounded shear flows, and depending upon just how one counts, these might be viewed as having two, three or even four different length scales represented in their physical behaviors. Such flows are found in boundary layers, and thus also in pipes, ducts and channels. Our treatment in these lectures will follow that of [7] to a great extent.

It is useful to begin by recalling the difference in the nature of velocity profiles between laminar and turbulent flow in a duct. This is depicted in Fig. 1.7. The parabolic profile of part (a) corresponds to a fully-developed Poiseuille flow in a duct for which it can be seen that the velocity gradient at the wall, and hence also the wall shear stress, τ_w, is not nearly so large as in the turbulent case of part (b) representing the (time) mean flow for fully-developed turbulence. The latter closely approximates uniform flow beginning quite close to the walls, so in light of the no-slip condition we

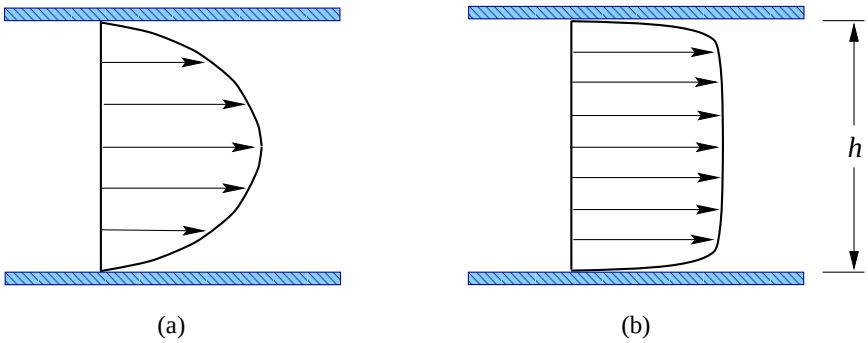

Figure 1.7: Comparison of laminar and turbulent velocity profiles in a duct; (a) laminar, and (b) turbulent.

see that

$$\tau_w = \mu \frac{\partial u}{\partial y}\bigg|_w$$

must be fairly large. Figure 1.7(b) also clearly demonstrates that there must be at least two length scales associated with this flow; one corresponds to the rather thin region adjacent to the walls in which the velocity profile is nonuniform, with large gradients, and the other can be related to the nearly uniform part of the flow farther from the walls.

The region very close to the wall exhibits a nearly linear velocity profile in the turbulent case, and in light of this we might recall Couette flow which is completely dominated by viscous effects. Indeed, this is the case for this inner layer, termed the *viscous sublayer*, in which velocity varies linearly with distance from the wall. Furthermore, because viscous effects are so dominant within this region, it is reasonable to associate the flow behavior with the Kolmogorov, or dissipation, scales defined earlier in Defs. 1.64 and 1.83.

The so-called "outer region" shows nearly constant velocity with distance from the wall, and we will not specifically analyze this here. But we note that in the context of perturbation analysis we recognize that this outer layer velocity cannot satisfy the no-slip condition at the walls, and at the same time the inner (linear) profile which does satisfy no slip will not correctly asymptote to the outer solution. This suggests that a third solution is needed to match these two results; formally, such solutions are derived via the method of matched asymptotic expansions. Here, we follow [7] and present an heuristic, physically-based treatment not requiring this formalism.

Let $\overline{u}(y)$ denote the time mean velocity, and let u_τ denote a velocity scale for the inner region. (This should generally correspond to turbulent velocity fluctuations and might, for example, be the square root of the turbulence kinetic energy. Here, as the notation suggests, and will be evident later, we use the friction velocity.) Now observe that the two length scales are a large advective scale associated with $\overline{u}$ (say, $h/2$, the half-height of the duct) and a viscous scale corresponding to u_τ, *viz.*, ν/u_τ. In order for an intermediate scale to make sense, it must be the case that the ratio of these two length scales be large; *i.e.*,

$$\frac{h/2}{\nu/u_\tau} = \frac{hu_\tau}{2\nu} \gg 1\,.$$

Then we are able to identify a range of distances y from the wall(s) such that

$$\frac{yu_\tau}{\nu} \gg 1\,, \qquad \text{and simultaneously} \qquad \frac{y}{h} \ll 1\,.$$

(We have dropped factors of two from these formulas because this argument is approximate, and an $\mathcal{O}(1)$ constant will be introduced subsequently in any case.)

Within this range one argues that ν/u_τ is too small to control flow dynamics, and h is too large to result in effective interactions. Hence, y itself is then the only length scale for this region; but there are two velocity scales, namely, $\overline{u}$ and u_τ. On purely dimensional grounds we argue that these must all be related according as

$$\frac{d\overline{u}}{dy} = C_1 u_\tau / y \,,$$

where C_1 is a constant that ultimately will be determined from experimental data. We now define the dimensionless quantities

$$y_+ \equiv \frac{y u_\tau}{\nu} \,, \qquad \text{and} \qquad u_+ \equiv \overline{u}/u_\tau \,, \tag{1.70}$$

and in terms of these express the above as

$$\frac{du_+}{dy_+} = C_1 / y_+ \,.$$

This can be directly integrated to yield

$$u_+ = C_1 \ln y_+ + C_2 \,, \tag{1.71}$$

where C_2 is an integration constant which also will need to be found from experimental data.

Modulo a few details which will be supplied later, Eq. (1.71) is the well-known "log law" that matches the inner to the outer layer. As noted in [7], the range of length scales over which the log law is valid corresponds to the inertial subrange of the Kolmogorov theory or, equivalently, to approximately the Taylor microscales.

It is worthwhile at this point to summarize these results in the usual way, as shown in Fig. 1.8. We remark that the inset to this figure is a more detailed representation of a turbulent velocity

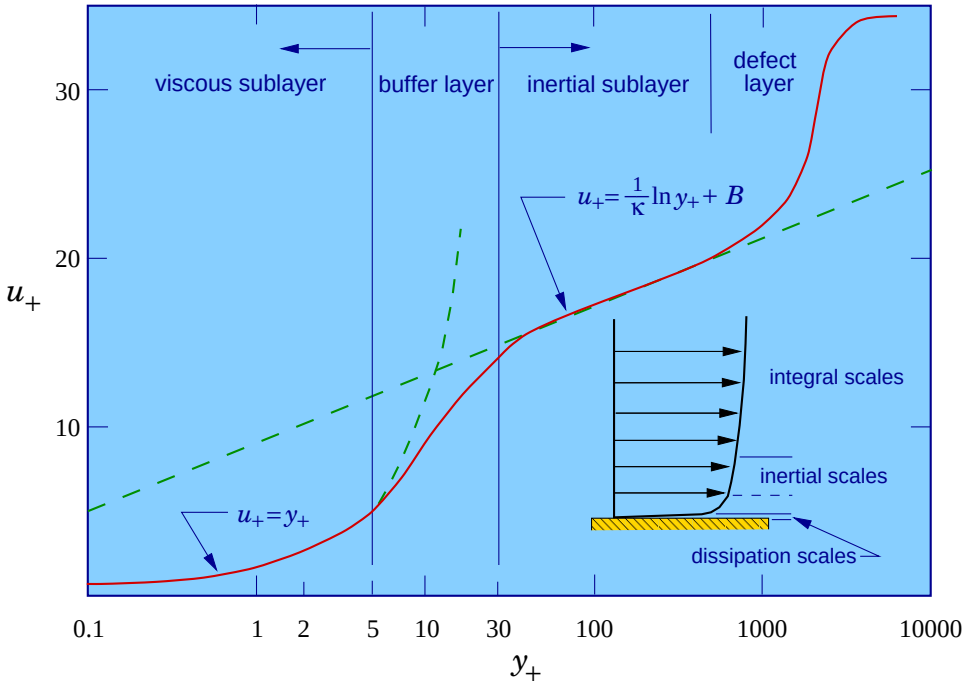

Figure 1.8: Law of the wall.

profile than that of Fig. 1.7(b). Moreover, we note that the somewhat arbitrary characteristic

turbulent velocity of the viscous sublayer is usually taken to be the *friction velocity* as hinted earlier and alluded to in Def. 1.71. This is defined as

$$u_\tau = \sqrt{\frac{\tau_w}{\rho}}\,. \tag{1.72}$$

Furthermore, we define in the usual way corresponding to the notation of the figure. The formula in the figure corresponding to the logarithmic part of the velocity profile contains the usual notation κ for the von Kármán constant and B for the integration constant given earlier as C_2. In Chap. 2 we will discuss these in more detail and, in particular, supply values for them.

We observe that this figure exhibits four different length scales as we hinted earlier might be the case. The one not previously discussed is usually termed the "buffer layer," and it (smoothly) connects the the viscous sublayer to the inertial sublayer. As indicated on Fig. 1.8 with dashed lines these two regions match mathematically, but the match is not smooth. Moreover, neither experimental nor computational data show the sharp change indicated by the dashed lines. The buffer layer might be viewed, physically, as a small range of scales over which inertial and dissipation effects are nearly balanced, corresponding to the range of wavenumbers in the energy spectrum just beyond that of the inertial subrange where the decay rate is beginning to increase (recall Fig. 1.5). In addition, the terminology "defect layer" is not universally employed for the outermost (furthest from the walls) region. It is suggested by the viewpoint that a turbulent boundary layer might be considered to be a turbulent wake with a solid wall on one side. In any case, it is clear from Fig. 1.8 that the log law does not hold in this part of a flow; typical length scales are of the order of the integral scale.

Finally, we note that the importance of the law of the wall in the present era is that it provides an explicit formula for the mean turbulent velocity profile near a solid boundary. With this in hand it seems not necessary to employ the fine gridding required to capture behavior on the scales of the buffer layer and smaller, as is formally necessary in LES; in particular, one can evaluate the log law formula and use the result as a velocity boundary condition at the outer edge of the buffer layer, or even farther from the wall if extremely coarse gridding is employed. This significantly reduces the required amount of total arithmetic for simulating wall-bounded shear flows and is widely used both in RANS methods and in LES. But it is important to recall the assumption that led to the log-law result, *viz.*, fully-developed flow. As a consequence, the simple law of the wall described here is not valid for nonself-similar boundary layers, and furthermore cannot be used accurately in the presence of flow separation. There have, however, been numerous attempts to extend the law of the wall to other flow situations beginning already with Tennekes and Lumley [7], and some of these have been at least moderately successful. But treatment of these is beyond the intended scope of these lectures, and the reader is referred to the extant literature.

1.5 Summary

In this introductory chapter we have begun by first attempting to provide some heuristic views of turbulence, an historical overview of its study and the progress to date, to some extent contrasting the classical, statistical view of turbulence as a random phenomenon with the modern, deterministic viewpoint. We next included an extensive glossary of terms associated with turbulence, many of which will be used repeatedly as we proceed through the following lectures. Then, as further preparation for these, we presented fairly detailed discussions of numerous mathematical and physical ideas that are of a fundamental nature and thus will be of use independent of whether we are considering the classical or modern view of turbulence. These included several alternative decompositions of flow variables, Fourier analysis of the N.–S. equations, length and time scales of turbulence, and the law of the wall.

We again emphasize that it is now almost universally accepted that the N.–S. equations are capable of producing turbulent solutions. Moreover, such solutions obtained via DNS agree exceptionally well with experimental observations within the range of Reynolds numbers currently accessible by this technique. This suggests that independent of what form of modeling might be attempted, tests of model results and/or analysis of model structure must include direct comparisons with the N.–S. equations and their solutions.

Chapter 2

Statistical Analysis and Modeling of Turbulence

Statistical analyses of turbulence have been employed from the beginning, certainly already in Reynolds' 1894 paper [16], but at least implicitly in the earlier work of Boussinesq [15] in 1877. The works of Reynolds, Prandtl, Taylor and others emphasized the perceived randomness of turbulent flows with the implication that statistical approaches were the only possibility for analysis, and this view was dominant during the early years of development of turbulence modeling procedures necessitated by averaging the nonlinear Navier–Stokes (N.–S.) equations.

In a sense, modeling began with the work of Boussinesq cited above, but in the absence of the Reynolds decomposition formalism. Prandtl's mixing length theory represents the first turbulence model in a more modern sense; in particular, it attempts to estimate values of eddy viscosity, introduced by Boussinesq, to close the Reynolds-averaged Navier–Stokes (RANS) equations. It is sometimes claimed that Kolmogorov [106] already by 1942 had introduced the ideas underlying k–ε models that we will treat in the sequel, and similar ideas were proposed later by Rotta [107] in 1951. But Jones and Launder [108] introduced the basic form of these models that has been studied to the present. In addition, Launder *et al.* [109] presented the form of so-called *second-order*, or *second-moment closures*, also called *Reynolds stress models*, that until rather recently received considerable attention.

In this chapter we begin by deriving the RANS equations, and then discuss a number of mathematical consequences that are independent of modeling. Following this we consider some details of a number of specific, widely-used models including Prandtl's mixing length, $k-\varepsilon$ models, eddy viscosity transport and the basics of second-order closures. We then provide a section in which the Kolmogorov theories are discussed. These also are statistical, but they do not rely on a Reynolds decomposition, and until only recently little effort had been expended in attempts to make direct use of them in modeling. We remark that this was probably at least in part due to the fact that this theory did not seem to provide anything of use for typical RANS approaches that, as we will see, usually require construction of an eddy viscosity.

2.1 The Reynolds-Averaged Navier–Stokes Equations

In this section we begin by employing the Reynolds decomposition, Eq. (1.25), to derive the RANS equations. We then obtain related results for vorticity transport and examine the relationship between vortex stretching and Reynolds stresses. We follow this with analyses of attempts to include time dependence in the RANS equations, and in particular consider a currently-popular procedure termed unsteady RANS (URANS). We then provide a fairly detailed analysis of the

inherent difficulties of RANS approaches in general, both with respect to formal mathematics, and with respect to physics. Finally, we present descriptions and comparisons of some of the more widely-implemented RANS procedures.

2.1.1 Derivation of the RANS equations

The starting point for this derivation is, of course, the N.–S. equations,

$$\nabla \cdot \boldsymbol{U} = 0\,, \tag{2.1a}$$

$$\boldsymbol{U}_t + \boldsymbol{U} \cdot \nabla \boldsymbol{U} = -\nabla P + \nu \Delta \boldsymbol{U}\,, \tag{2.1b}$$

given earlier in Eqs. (1.1), and written here without the body-force term. We next recall Eq. (1.25) for the Reynolds decomposition and substitute this into the first of the above equations:

$$\nabla \cdot \boldsymbol{U} = \nabla \cdot (\overline{\boldsymbol{u}} + \boldsymbol{u'}) = \nabla \cdot \overline{\boldsymbol{u}} + \nabla \cdot \boldsymbol{u'} = 0\,. \tag{2.2}$$

Then averaging this equation results in

$$\nabla \cdot \overline{\overline{\boldsymbol{u}}} + \nabla \cdot \overline{\boldsymbol{u'}} = 0\,, \tag{2.3}$$

and from Eq. (1.26) we deduce that

$$\nabla \cdot \overline{\boldsymbol{u}} = 0\,. \tag{2.4}$$

Then it follows from the far right-hand side of Eq. (2.2) that

$$\nabla \cdot \boldsymbol{u'} = 0 \tag{2.5}$$

also holds.

We remark that if the averaging performed in Eq. (2.3) had been omitted, we might conclude that only

$$\nabla \cdot \overline{\boldsymbol{u}} = -\nabla \cdot \overline{\boldsymbol{u'}}$$

is required to satisfy the divergence-free constraint of the overall velocity field. But we should recall that the Galerkin form of the N.–S. equations implies that the divergence-free condition must hold on a mode-by-mode basis (recall Eq. (1.41)), and this would be contradicted by the above.

We next consider Eqs. (2.1b). We begin by substituting Eq. (1.25) for each entry of $\boldsymbol{U}$, with an analogous expression for P, to obtain

$$(\overline{\boldsymbol{u}} + \boldsymbol{u'})_t + (\overline{\boldsymbol{u}} + \boldsymbol{u'}) \cdot \nabla (\overline{\boldsymbol{u}} + \boldsymbol{u'}) = -\nabla (\overline{p} + p') + \nu \Delta (\overline{\boldsymbol{u}} + \boldsymbol{u'})\,.$$

Now recall that $\overline{\boldsymbol{u}}$ is independent of t, by definition, so the first term in this expression is identically zero. Then upon expansion of the dot product on the left-hand side we have

$$\boldsymbol{u'}_t + \overline{\boldsymbol{u}} \cdot \nabla \overline{\boldsymbol{u}} + \overline{\boldsymbol{u}} \cdot \nabla \boldsymbol{u'} + \boldsymbol{u'} \cdot \nabla \overline{\boldsymbol{u}} + \boldsymbol{u'} \cdot \nabla \boldsymbol{u'} = -\nabla (\overline{p} + p') + \nu \Delta (\overline{\boldsymbol{u}} + \boldsymbol{u'})\,.$$

Next, just as we did with the continuity equation, we time average the entire equation. This results in

$$\overline{\boldsymbol{u'}_t} + \overline{\overline{\boldsymbol{u}} \cdot \nabla \overline{\boldsymbol{u}}} + \overline{\overline{\boldsymbol{u}} \cdot \nabla \boldsymbol{u'}} + \overline{\boldsymbol{u'} \cdot \nabla \overline{\boldsymbol{u}}} + \overline{\boldsymbol{u'} \cdot \nabla \boldsymbol{u'}} = -\nabla \overline{(\overline{p} + p')} + \nu \Delta \overline{(\overline{\boldsymbol{u}} + \boldsymbol{u'})}\,. \tag{2.6}$$

Here, we have already commuted time averaging and spatial differentiation in the linear terms on the right-hand side. Proof of validity of this is trivial. We next commute temporal averaging and differentiation in the first term on the left-hand side, and use the second of Eqs. (1.26) to conclude that

$$\overline{\boldsymbol{u'}_t} = \left(\overline{\boldsymbol{u'}}\right)_t \equiv 0\,. \tag{2.7}$$

We leave formal proof of validity of this, under the assumption that $\boldsymbol{u'} \in C^1(\mathbb{R}_+)$ with respect to time as an exercise for the reader.

Using Eqs. (1.26) in the right-hand side of (2.6) and (2.7) on the left-hand side leaves

$$\overline{\overline{\boldsymbol{u}} \cdot \nabla \overline{\boldsymbol{u}}} + \overline{\overline{\boldsymbol{u}} \cdot \nabla \boldsymbol{u'}} + \overline{\boldsymbol{u'} \cdot \nabla \overline{\boldsymbol{u}}} + \overline{\boldsymbol{u'} \cdot \nabla \boldsymbol{u'}} = -\nabla \overline{p} + \nu \Delta \overline{\boldsymbol{u}} \,. \tag{2.8}$$

We now consider the details of $\overline{\overline{\boldsymbol{u}} \cdot \nabla \boldsymbol{u'}}$. We have, by definition of the time average, that

$$\begin{aligned}
\overline{\overline{\boldsymbol{u}} \cdot \nabla \boldsymbol{u'}} &= \lim_{T\to\infty} \frac{1}{T} \int_0^T \overline{\boldsymbol{u}} \cdot \nabla \boldsymbol{u'} \, dt \\
&= \overline{\boldsymbol{u}} \cdot \left[\lim_{T\to\infty} \frac{1}{T} \int_0^T \nabla \boldsymbol{u'} \, dt \right] \\
&= \overline{\boldsymbol{u}} \cdot \nabla \left[\lim_{T\to\infty} \frac{1}{T} \int_0^T \boldsymbol{u'} \, dt \right] \\
&= \overline{\boldsymbol{u}} \cdot \nabla \overline{\boldsymbol{u'}} \\
&= 0 \,.
\end{aligned}$$

The second equality follows because $\overline{\boldsymbol{u}}$ is independent of t. Similarly, the third occurs because ∇, as an operator, is also independent of time, and the final equality arises from application of the second equation in (1.26). A similar analysis can be carried out for $\overline{\boldsymbol{u'} \cdot \nabla \overline{\boldsymbol{u}}}$, with the same result holding. Thus, Eq. (2.8) becomes

$$\overline{\overline{\boldsymbol{u}} \cdot \nabla \overline{\boldsymbol{u}}} + \overline{\boldsymbol{u'} \cdot \nabla \boldsymbol{u'}} = -\nabla \overline{p} + \nu \Delta \overline{\boldsymbol{u}} \,.$$

We next observe that by the divergence-free conditions Eqs. (2.4) and (2.5), the terms on the left-hand side of the above can be expressed as $\overline{\nabla \cdot \overline{\boldsymbol{u}}^2}$ and $\overline{\nabla \cdot \boldsymbol{u'}^2}$, respectively, where $\boldsymbol{u}^2$ is shorthand notation for $(u, v, w)^T(u, v, w)$. We leave demonstration of this as an exercise. Then it easily follows that

$$\overline{\nabla \cdot \overline{\boldsymbol{u}}^2} = \nabla \cdot \overline{\overline{\boldsymbol{u}}^2} = \nabla \cdot \overline{\boldsymbol{u}}^2 \,,$$

and

$$\overline{\nabla \cdot \boldsymbol{u'}^2} = \nabla \cdot \overline{\boldsymbol{u'}^2} \,.$$

(Again, supplying detailed arguments is left to the reader.) As a consequence of these identities we can now write Eq. (2.8) as

$$\nabla \cdot \overline{\boldsymbol{u}}^2 + \nabla \cdot \overline{\boldsymbol{u'}^2} = -\nabla \overline{p} + \nu \Delta \overline{\boldsymbol{u}} \,, \tag{2.9}$$

the Reynolds-averaged Navier–Stokes equations. It is clear that this vector equation contains far more unknowns than equations; that is, the second term on the left-hand side does not possess an equation for any of its six independent components to be described in more detail below. This constitutes our first direct encounter with the so-called "turbulence closure problem" which has led to extensive modeling efforts from the early 20^{th} Century to the present day.

We note that there are numerous formulations for the RANS equations, with no particular notation being completely standard. We present a few of the more common ones here. We begin by noting that $\overline{\boldsymbol{u'}^2}$ is usually called the *Reynolds stress tensor* although a check of dimensions will show that it is not actually a stress; it must be multiplied by density ρ, as done consistently in [7], in order to have dimensions corresponding to the terminology. (On the other hand, since $\rho \equiv$ const. we might set it to unity, thus obtaining implicit dimensional correctness. Moreover, because we

typically employ kinematic viscosity, there is an implied division by ρ in any case.) The matrix representation of this tensor is the following:

$$\overline{\boldsymbol{u'}^2} \;=\; \begin{pmatrix} \overline{u'^2} & \overline{u'v'} & \overline{u'w'} \\ \overline{u'v'} & \overline{v'^2} & \overline{v'w'} \\ \overline{u'w'} & \overline{v'w'} & \overline{w'^2} \end{pmatrix} . \tag{2.10}$$

In addition, in Cartesian tensor notation this is often denoted simply as $\boldsymbol{R}$ with components sometimes denoted both by R_{ij} and r_{ij} and expressed as

$$\boldsymbol{R} \;=\; \begin{pmatrix} \overline{u'_1u'_1} & \overline{u'_1u'_2} & \overline{u'_1u'_3} \\ \overline{u'_1u'_2} & \overline{u'_2u'_2} & \overline{u'_2u'_3} \\ \overline{u'_1u'_3} & \overline{u'_2u'_3} & \overline{u'_3u'_3} \end{pmatrix} . \tag{2.11}$$

In this context the individual components are often also denoted τ_{ij}, especially after multiplication by ρ as described above. We will not use this notation herein as we prefer to employ this for components of the complete (undecomposed) stress tensor. Finally, the functional tensor notation $\boldsymbol{R}(\boldsymbol{u'},\boldsymbol{u'})$ will sometimes be found. Note that this usually contains the divergence operator as well as the Reynolds stress tensor to simplify the notation; that is

$$\boldsymbol{R}(\boldsymbol{u'},\boldsymbol{u'}) \;=\; \nabla\cdot \begin{pmatrix} \overline{u'^2} & \overline{u'v'} & \overline{u'w'} \\ \overline{u'v'} & \overline{v'^2} & \overline{v'w'} \\ \overline{u'w'} & \overline{v'w'} & \overline{w'^2} \end{pmatrix} .$$

Thus, Eq. (2.9) might be written in any one of the following alternative (and equivalent) ways:

$$\nabla\cdot\overline{\boldsymbol{u}}^2 = -\nabla\overline{p} + \nu\Delta\overline{\boldsymbol{u}} - \boldsymbol{R}(\boldsymbol{u'},\boldsymbol{u'})\,, \tag{2.12a}$$

$$\frac{\partial}{\partial x_j}\overline{u}_i\overline{u}_j = -\frac{\partial\overline{p}}{\partial x_i} + \nu\frac{\partial^2\overline{u}_i}{\partial x_j\partial x_j} - \frac{\partial}{\partial x_j}\overline{u'_iu'_j}\,, \quad i=1,2,3\,, \tag{2.12b}$$

$$\frac{\partial}{\partial x_j}\overline{u}_i\overline{u}_j = -\frac{\partial\overline{p}}{\partial x_i} + \nu\frac{\partial^2\overline{u}_i}{\partial x_j\partial x_j} - \frac{\partial R_{ij}}{\partial x_j}\,, \quad i=1,2,3\,, \tag{2.12c}$$

or with R_{ij} in Eq. (2.12c) replaced with either r_{ij} or τ_{ij}. We will mainly employ the first two of these notations, or Eq. (2.9), itself, in the sequel.

Finally, as we have already noted, it is common to ignor the fact that $\overline{\boldsymbol{u}}$ is independent of time and write, *e.g.*,

$$\overline{\boldsymbol{u}}_t + \nabla\cdot\overline{\boldsymbol{u}}^2 = -\nabla\overline{p} + \nu\Delta\overline{\boldsymbol{u}} - \boldsymbol{R}(\boldsymbol{u'},\boldsymbol{u'})\,, \tag{2.13}$$

or in Cartesian tensor notation,

$$\overline{u}_{i,t} + \frac{\partial}{\partial x_j}\overline{u}_i\overline{u}_j = -\frac{\partial\overline{p}}{\partial x_i} + \nu\frac{\partial^2\overline{u}_i}{\partial x_j\partial x_j} - \frac{\partial}{\partial x_j}\overline{u'_iu'_j}\,, \quad i=1,2,3\,. \tag{2.14}$$

We will examine this in more detail in the following section.

2.1.2 Time-dependent RANS equations

In essentially all modern practical formulations of the RANS equations the time derivative term $\partial\overline{\boldsymbol{u}}/\partial t$ is included, despite the fact that $\overline{\boldsymbol{u}} = \overline{\boldsymbol{u}}(\boldsymbol{x})$ only is independent of time. In this section we briefly discuss two of the widely-used arguments justifying retention of this term. The first of these comes from an appeal to ergodicity of turbulent flows while the second is associated with time scales actually needed to obtain reasonably accurate averages. We consider some details of each of these and present consequences that follow from these arguments.

Appeal to Ergodicity

Recall from Def. 1.3 that the ergodic hypothesis implies "time averaging is equal to ensemble averaging." There are several precise statements of this, and the reader is referred to Frisch [80] for more details. Here, we first indicate how this might be employed, and then return to some of the details.

We first recall from Chap. 1 that the ensemble average of N realizations of a variable $u(\boldsymbol{x},t)$, say

$$\langle u(\boldsymbol{x},t)\rangle = \frac{1}{N}\sum_{i=1}^{N} u^{(i)}(\boldsymbol{x},t)\,,$$

is formally time dependent. Consequently, we might write the ensemble-averaged N.–S. equations as

$$\langle \boldsymbol{U}\rangle_t + \langle \boldsymbol{U}\rangle\cdot\nabla\langle \boldsymbol{U}\rangle = -\nabla\langle P\rangle + \nu\Delta\langle \boldsymbol{U}\rangle - \boldsymbol{R}(\boldsymbol{u'},\boldsymbol{u'})\,, \tag{2.15}$$

with the components of $\boldsymbol{R}$ now computed with ensemble averaging.

This form was often employed in early analyses of turbulence (see, *e.g.*, Batchelor [39]), but more recently it has sometimes been used in conjunction with the ergodic hypothesis to justify the form of the RANS equations given in Eq. (2.13). That is, if retaining the time dependence in ensemble-averaged equations is valid, and ensemble averaging is equivalent to time averaging, then it follows that

$$\overline{\boldsymbol{u}}_t + \overline{\boldsymbol{u}}\cdot\nabla\overline{\boldsymbol{u}} = -\nabla\overline{p} + \nu\Delta\overline{\boldsymbol{u}} - \boldsymbol{R}(\boldsymbol{u'},\boldsymbol{u'})$$

is a valid equation to consider.

There is, however, a basic flaw in this argument, and this arises from not recalling the hypotheses required in the ergodic theorems. There are two basic ones. The first is stationarity of the flow. Bradshaw [110] comments on the importance of this in turbulent flow data. Indeed, most (but not all) turbulent flows considered in practice are stationary. The second hypothesis is that the formal limit $T\to\infty$ be observed in the time-averaging process just as implied by the definition. The consequence of this is that formal time dependence of the ensemble-averaged equations cannot be used to deduce time dependence of the usual time-averaged equations, but rather precisely the opposite must hold—solutions to the ensemble-averaged equations, themselves, are <u>not</u> time dependent. In fact, this can be inferred also from the discussions in [80] although it is not explicitly stated there. (But note that this does <u>not</u> invalidate use of ensemble averages to obtain time-dependent data. It simply negates use of the ergodic hypothesis.)

Use of Multiple Time Scales

Probably the most commonly-used justification for retaining the time-derivative operator in the RANS equations comes from arguments associated with multiple time scales. These arguments are usually rather heuristic, but Wilcox [98] provides a particularly good treatment of this, and we will follow this, to some extent, in the present discussions.

The basic hypothesis used in such arguments is that the flow under consideration exhibits at least two widely-separated time scales, say T_1 and T_2 such that $T_2 \gg T_1$. Then one expects that there will be a range if times T such that $T_1 \ll T \ll T_2$. Furthermore, one might then associate turbulent fluctuations with the time scale T_1 and large-scale fluid motions with T_2. In the context of modern computing environments, the latter motions should be easily resolvable, so there is no need to indirectly account for them via averaging. Thus, we now define Reynolds averaging in terms of the time scales T as, *e.g.*,

$$\overline{\boldsymbol{u}}(\boldsymbol{x},t) = \frac{1}{T}\int_0^T \boldsymbol{U}(\boldsymbol{x},t)\,dt\,,$$

where $T \ll T_2$, and no limit is taken. In particular, $\boldsymbol{u}$ now depends on time with respect to the scale $T_2 \gg T$; and since T is assumed to satisfy $T \gg T_1$, we have evidently accurately averaged the turbulent fluctuations while retaining the slow time scale of the large-scale motions. This is shown pictorially in Fig. 2.1 which also suggests a correspondence with low-pass temporal filtering.

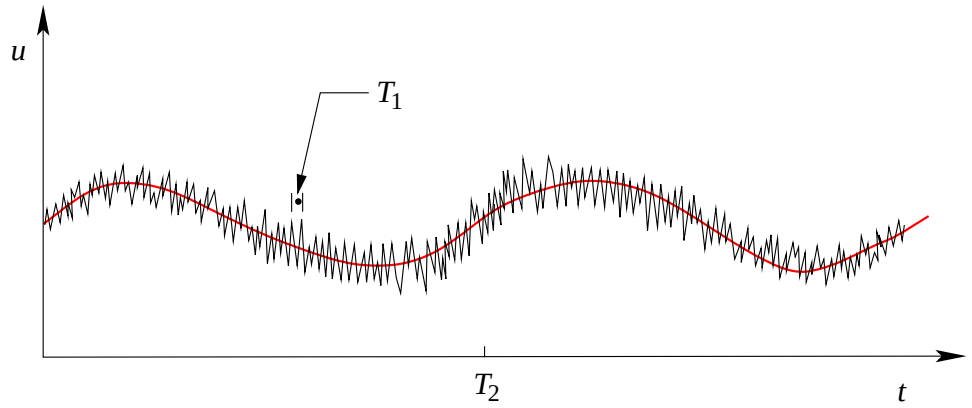

Figure 2.1: Multiple time scales for construction of time-dependent RANS equations.

On the surface the preceding is a quite compelling argument, and it is now widely accepted. But like all other proposed justifications for retaining the $\overline{\boldsymbol{u}}_t$ term in the RANS equations, it is seriously flawed. We first note, as does Wilcox [98], that most turbulent flows contain far more than two time scales. One would expect this from the nature of the Galerkin approximation to the N.–S. equations given in Chap. 1. In this case the above formalism cannot lead to accurate results (essentially any choice of T will be inadequate), and Wilcox recommends use of LES rather than attempting to employ RANS modeling. Second, we should observe that even in the case of two distinct time scales, the scale T with respect to which averaging is actually performed, is not well defined—we never say what T is, or how one might determine it. From a purely-mathematical perspective this is an extremely serious (essentially fatal) flaw, but even in practice this poses a significant problem. In particular, we will later see that determination of various constants present in any turbulence model requires experimental data. But if the time scale of the averaging is not well defined, the model constants, themselves, are not well defined, leading to very significant difficulties in employing experimental data (or results from DNS).

We conclude from this that, in general, little is likely to be gained by attempting to compute time-dependent solutions using RANS equations. Doing so is in conflict with the formal mathematics used to derive the equations, and at best it results in an undefined time scale that ultimately works its way into model constants.

2.1.3 Importance of vorticity and vortex stretching to turbulence

We have earlier observed that turbulence is inherently three dimensional; indeed, even in flows that are at least nearly 2D on large scales, it is observed in laboratory experiments that small-scale fluctuations are generally three dimensional. Furthermore, it is widely argued that turbulence, itself, is intimately connected with vorticity. References as otherwise widely disparate as Tennekes and Lumley [7] and Tsinober [3] make this same claim—albeit for different reasons and backed by different supporting arguments. Tennekes and Lumley argue that turbulence must be 3D on the basis of demonstrating a relationship between vortex stretching and production of Reynolds stresses, and the fact that vortex stretching can only occur in 3D; hence, Reynolds stresses can only be generated in 3D flows. Moreover, they argue that vortex stretching is the physical mechanism leading to the hypothesized energy cascade from large to small scales. On the other hand, Tsinober [3] notes first that this cascade is not exactly what happens and second that it is in fact vortex compression—not stretching—that would result in energy transfer to smaller scales. But in addition, he emphasizes that strain rate is equally as important as vorticity in the context of creation and maintenance of turbulence.

In this section we first present a somewhat different view of vorticity and its importance in fluid flow, generally. We then outline arguments provided in [7], but balanced by those in [3] and interpreted in light of the viewpoint given here.

What Is Vorticity, and How Important Is It—Really?

Vorticity and its study have been extremely important topics in almost any fluid dynamics course for the past century, and as already noted vorticity is considered to be a key ingredient of a turbulent flow. But what, exactly, is vorticity? It certainly is not a property of the fluid; but moreover, it is not a directly-measurable physical property of the flow. In fact, vorticity is nothing but a mathematical definition—the curl of the velocity field:

$$\boldsymbol{\omega} \equiv \nabla \times \boldsymbol{U} \,. \tag{2.16}$$

Thus, mathematical vorticity is constructed from gradients of the physical flow property velocity, and these gradients are calculated from measured velocity fields, leading to a rather indirect determination of vorticity. Furthermore, the same is true of strain rates, and this should be of no surprise because both rotation (essentially one half of the vorticity) and strain rate arise in a straightforward decomposition of the velocity gradient tensor (the Jacobian matrix of the velocity field) as is well known and is shown in Chap. 1.

What is clear from this is that the fundamental physical quantity is the set of velocity gradients; vorticity and strain rates are derived quantities. This is not, however, intended to imply that they are unimportant, or that they are not useful, but rather that they have possibly been overemphasized. Moreover, vorticity does not occur naturally in the N.–S. equations (or even in their derivation), and while strain rate is an important element of the derivation it does not, *per se*, appear in the final form of the equations unless variable transport properties are considered. Thus, especially in the case of vorticity, an overemphasis of its use in deducing flow behaviors can, potentially, result in misleading notions.

Relationship between Vorticity and Reynolds Stress

Here we will closely follow the treatment found in [7] to demonstrate the connection of vorticity to Reynolds stress, and thus to turbulence in the context of a Reynolds-averaged Navier–Stokes

interpretation. As in that reference, we begin by writing the N.–S. equations in terms of Cartesian tensor notation given in Eqs. (1.30):

$$\frac{\partial u_i}{\partial t} + u_j \frac{\partial u_i}{\partial x_j} = -\frac{\partial p}{\partial x_i} + \nu \frac{\partial^2 u_i}{\partial x_j \partial x_j}\,, \qquad i = 1, 2, 3\,. \tag{2.17}$$

We now express the advective term as

$$\begin{aligned} u_j \frac{\partial u_i}{\partial x_j} &= u_j \left(\frac{\partial u_i}{\partial x_j} - \frac{\partial u_j}{\partial x_i} \right) + u_j \frac{\partial u_j}{\partial x_i} \\ &= -\epsilon_{ijk} u_j \omega_k + \frac{\partial}{\partial x_i} \left(\frac{1}{2} u_j u_j \right)\,, \end{aligned}$$

and substitute this into Eq. (2.17) to obtain

$$\frac{\partial u_i}{\partial t} + \frac{1}{2} \frac{\partial}{\partial x_i}(u_j u_j) = -\frac{\partial p}{\partial x_i} + \epsilon_{ijk} u_j \omega_k + \nu \frac{\partial^2 u_i}{\partial x_j \partial x_j}\,, \tag{2.18}$$

We next introduce the Reynolds decompositions

$$u_i = \overline{u}_i + u_i'\,, \quad \omega_i = \overline{\omega}_i + \omega_i'\,, \quad p = \overline{p} + p'\,, \qquad i = 1, 2, 3\,,$$

and use these in Eq. (2.18) followed by averaging of the result to arrive at

$$\frac{1}{2} \frac{\partial}{\partial x_i}(\overline{u}_j \overline{u}_j) = -\frac{\partial \overline{p}}{\partial x_i} - \frac{1}{2} \frac{\partial}{\partial x_i} \left(\overline{u_j' u_j'} \right) + \epsilon_{ijk} \left(\overline{u}_j \overline{\omega}_k + \overline{u_j' \omega_k'} \right) + \nu \frac{\partial^2 \overline{u}_i}{\partial x_j \partial x_j}\,,$$

or

$$\frac{1}{2} \frac{\partial}{\partial x_i} \left(\overline{u_j' u_j'} \right) - \epsilon_{ijk} \overline{u_j' \omega_k'} = -\frac{\partial}{\partial x_i} \left(\overline{p} + \frac{1}{2} \overline{u}_j \overline{u}_j \right) + \epsilon_{ijk} \overline{u}_j \overline{\omega}_k + \nu \frac{\partial^2 \overline{u}_i}{\partial x_j \partial x_j}\,. \tag{2.19}$$

Here, we have expressed the fluctuating quantities related to Reynolds stresses in terms of mean quantities, including mean vorticity, thus emphasizing the importance of vorticity in generation of the Reynolds stresses.

There are, however, two serious flaws in this argument. The first is that introduction of vorticity into the N.–S. equations as done here is contrived and artificial; it in no way reflects any actual physics. (It was obtained by adding zero to the momentum equations.) But the second is more damaging, for even if introduction of vorticity could be argued on physical grounds, there is no justification for the implicit assumption that the Reynolds stresses are turbulence. In the first place, they are merely artifacts of having averaged the N.–S. equations—again, simply a mathematical operation; but beyond this is the fact that, formally, Reynolds stresses could be constructed for any stationary time-dependent flow, whether or not it happens to be turbulent. Finally, the Reynolds stresses, themselves, are time-averaged quantities and hence independent of time, and beyond this is the fact that only the normal stresses actually appear in the formulation. It is thus difficult to rationalize a direct equivalence between turbulence, which is characterized as always being time dependent, and a simple velocity correlation that by definition does not depend on time.

Vortex Stretching and Turbulence

As mentioned earlier, one of the proposed mechanisms for transferring energy from large to small scales is via vortex stretching. The basic physical idea is that a vortex tube subjected to strain from local velocity gradients of the surrounding flow field will be stretched, thus shrinking

its diameter. Then the energy associated with this vortex is acting at smaller length scales (and higher wavenumbers). For the moment we will assume validity of this argument and examine the mathematical representation of the N.–S. equations to identify terms able to generate this behavior. In particular, we will be seeking terms involving interactions of strain rate and vorticity in the context of the vorticity equation.

The 2-D vorticity equation. We first consider the two-dimensional case. If one begins with the 2-D N.–S. equations and computes the curl of these, the result is the 2-D vorticity transport equation,

$$\omega_t + \boldsymbol{U}\cdot\nabla\omega = \nu\Delta\omega\,. \tag{2.20}$$

Observe that in 2D only the third component of the vorticity vector is nonzero, and we denote this simply as the scalar ω.

If we now recall that components of the strain rate are of the form, *e.g.*,

$$S_{12} = \frac{1}{2}\left(\frac{\partial u_2}{\partial x_1} + \frac{\partial u_1}{\partial x_2}\right) = \frac{1}{2}(v_x + u_y)\,,$$

we readily see that Eq. (2.20) contains no such factors. (Recall that they do appear in the formal derivation of the N.–S. equations but disappear via the divergence-free constraint in the case of constant-viscosity flows even for these more fundamental equations.) From this it is concluded that turbulence cannot be 2D. That is, there is no mechanism in 2D to cause vortex stretching in the vorticity transport equation, and hence no physics to facilitate the energy cascade.

The 3-D vorticity equation. We now provide a more detailed analysis for the 3-D case. Just as in 2D we take the curl of the momentum equations (2.1b), but in contrast to the 2-D case the result will now be a vector. Formally, we write

$$\nabla\times[\boldsymbol{U}_t + \boldsymbol{U}\cdot\nabla\boldsymbol{U}] = -\nabla\times\nabla P + \nabla\times\Delta\boldsymbol{U}\,.$$

Since the curl of a gradient is identically zero (under mild smoothness conditions) as is easily checked, the first term on the right-hand side is zero, as in 2D. Furthermore, assuming some smoothness, we commute $\nabla\times$ and $\partial/\partial t$, and $\nabla\times$ and Δ. Then the above becomes

$$(\nabla\times\boldsymbol{U})_t + \nabla\times(\boldsymbol{U}\cdot\nabla\boldsymbol{U}) = \nu\Delta(\nabla\times\boldsymbol{U})\,,$$

or

$$\boldsymbol{\omega}_t + \nabla\times(\boldsymbol{U}\cdot\nabla\boldsymbol{U}) = \nu\Delta\boldsymbol{\omega}\,,$$

where $\boldsymbol{\omega}\equiv(\omega_1,\omega_2,\omega_3)^T$ is the 3-D *vorticity vector*.

We now need to consider the curl of the advective terms in more detail. We have

$$\begin{aligned}\nabla\times(\boldsymbol{U}\cdot\nabla\boldsymbol{U}) &= (\nabla\times\boldsymbol{U})\cdot\nabla\boldsymbol{U} + \boldsymbol{U}\cdot(\nabla\times\nabla\boldsymbol{U})\\ &= (\nabla\times\boldsymbol{U})\cdot\nabla\boldsymbol{U} + \boldsymbol{U}\cdot\nabla(\nabla\times\boldsymbol{U})\\ &= \boldsymbol{\omega}\cdot\nabla\boldsymbol{U} + \boldsymbol{U}\cdot\nabla\boldsymbol{\omega}\,.\end{aligned}$$

Then the 3-D vorticity transport equation can be expressed as

$$\boldsymbol{\omega}_t + \boldsymbol{U}\cdot\nabla\boldsymbol{\omega} = -\boldsymbol{\omega}\cdot\nabla\boldsymbol{U} + \nu\Delta\boldsymbol{\omega}\,, \tag{2.21}$$

which is of the same general form found for the 2-D case, Eq. (2.20), except for the term $-\boldsymbol{\omega}\cdot\nabla\boldsymbol{U}$ on the right-hand side. Now recall, as we have shown in Chap. 1, that $\nabla\boldsymbol{U}$ can be decomposed as the sum of strain rate and rotation tensors. In particular, this extra term appearing in the 3-D case

is associated with the interaction of vorticity and and the velocity gradient tensor (which contains strain rate—but also vorticity), and is called the "vortex-stretching" term.

It is important to recognize, however, that while vorticity can be a useful tool in studying fluid flow, its overemphasis can be misleading, as we have already suggested. In the present context there has been a tendency to identify existence of vorticity with actual vortices, or "eddies," and arrive at a cartoon of turbulence in which these hypothesized eddies and their behaviors are used to explain the nature of turbulence. The cascade of turbulent energy from large to small scales is a prime example of this. In particular, we have contrived a theory of the 3-D vorticity equation that supports such a cascade when it is now known that the simple idea of vortex stretching and subsequent breaking into yet smaller vortices (eddies) is not an accurate picture of actual physics. Moreover, especially in wall-bounded shear flows it has been found that as much as one third of the energy cascaded to small scales is "back scattered" up to the large scales. But in addition to this, the simple eddy cartoon of a turbulent flow is hardly an accurate representation in any circumstance. At any given instant the fraction of the volume of a turbulent flow actually occupied by eddies can be rather small. This, of course, does not imply that the vorticity is zero nearly everywhere but rather serves to emphasize that constructing a physical theory based on this cartoon of vortical eddies may be ill advised. Finally, we observe that there is a logical inconsistency in attempting to view strain rate as the cause of vortex stretching. Namely, vorticity and strain rate are somewhat artificially contrived contributions to the velocity gradient tensor. Hence, they occur simultaneously, thus removing the ability of one of them the "cause" the other.

2.1.4 Some general problems with RANS formulations

In this section we discuss several specific problems that are inherent in RANS formulations independent of details of any specific model. These are: *i*) the time-averaging process, *ii*) the Reynolds decomposition, *iii*) the generic form of the RANS equations and *iv*) general inequality of time-averaged N.–S. solutions and RANS solutions.

Effects of Time Averaging

We will here restrict attention to time averaging since this is essentially always employed in practical RANS formulations, and we begin by noting that averaging intrinsically results in a loss of information. Indeed, this is the purpose of averaging—to strip away "non-essential" details leaving only the core information. Figure 2.2 displays this idea in a pictorial way, and with respect to this it is worthwhile to recall Reynolds' [16] proposal that it would be impossible to understand turbulence in detail, and it would thus be more useful to consider only statistical properties of a turbulent flow.

There are at least two fundamental flaws with this viewpoint in the context of fluid physics. The first is that in many flow situations details are, in fact, quite essential. An important example of this, which we will consider again later, is turbulent combustion: the small-scale flow field fluctuations can interact with the chemical kinetics leading to significant changes in reaction rates, temperature and species concentrations. These effects cannot be ignored if accurate predictions are desired.

But there is a second difficulty that arises from averaging, or from use of statistical characterizations in general. It is that the mapping

$$physics \longrightarrow statistics$$

is many to one. That is, the same statistics (at least at the level of only the first few moments, as considered in RANS formulations) can be generated from many different physical flows. This is easy to imagine simply be replacing the erratic signal shown in Fig. 2.2(a) with any periodic signal

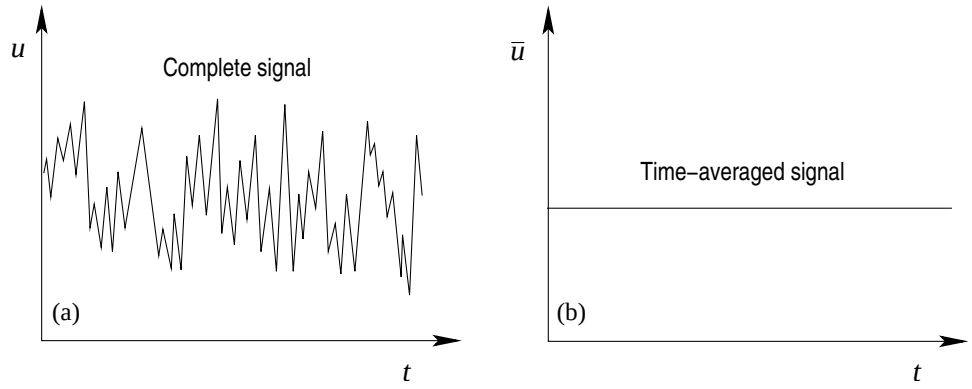

Figure 2.2: Loss of information due to averaging; (a) the complete signal, and (b) the time-averaged signal.

having the same mean—and this can be done completely independent of the frequency content. The result of averaging of any member of this uncountable set of signals will always be that shown in Fig. 2.2(b). Furthermore, second-order statistical quantities (auto and cross correlations) that comprise the Reynolds stress tensor can even be independent of flow direction; *e.g.*, $\overline{u'^2} = \overline{(-u')^2}$, *etc.*

From a formal mathematics perspective any many-to-one mapping is generally noninvertible. But it is precisely inversion that is being attempted when we compute with RANS formulations. In particular, we are attempting to recover a specific realization of the flow field (expressed in terms of its time mean) usually by employing this (unknown) mean to predict higher-order statistics (the Reynolds stresses) which, themselves, affect the mean that is being sought. This is a quite convoluted and complicated approach, and again from the standpoint of pure mathematics, one would never expect this to succeed—and with respect to true predictability, it does not! Indeed, we remark that, except in cases of very simple flow situations, RANS models are never able to accurately reproduce an entire flow field. They may be quite accurate in some locations while being very inaccurate in others, even when model constants have been carefully "tuned." An important question to consider here is what if this were observed for the N.–S. equations—would we then accept them as the equations of fluid motion as the RANS equations are accepted as a proper description of turbulence? One should hope not!

Finally, as we have already seen in Sec. 2.1, averaging the equations of motion (in contrast to simply averaging their solutions) leads directly to the closure problem alluded to above. It is clear, and this will be more evident when we discuss specific RANS models below, that averaging of both solutions and equations as required in constructing the RANS equations, results in many difficulties and shortcomings. Some of these are essentially insurmountable as witnessed by more than a century of effort with little in the way of predictive capability resulting.

The Reynolds Decomposition

We have already considered the details of Reynolds decomposition in Chap. 1, and we have employed it in Sec. 2.1; but it is worthwhile to recall some of these details viewed from a slightly different perspective than that taken earlier. Recall that Reynolds decomposition of any flow variable, say $u(\boldsymbol{x}, t)$ the first component of the velocity vector, for definiteness, can be expressed as

$$u(\boldsymbol{x}, t) = \overline{u}(\boldsymbol{x}) + u'(\boldsymbol{x}, t)\,. \tag{2.22}$$

It is interesting to consider the mathematics of this decomposition as we did earlier for the LES decomposition. In particular, we suppose $u(\boldsymbol{x}, t) \in L^2(\Omega)\times C^1(0, t_f)$ for some spatial domain Ω and

time interval $(0, t_f]$, possibly with $t_f \to \infty$. Then, as done in Chap. 1, we can express $u(\boldsymbol{x}, t)$ as the Fourier series

$$u(\boldsymbol{x}, t) = \sum_{|\boldsymbol{k}|\geq 0}^{\infty} a_{\boldsymbol{k}}(t)\varphi_{\boldsymbol{k}}(\boldsymbol{x})\,.$$

We now time average this representation:

$$\begin{aligned}\lim_{T\to\infty}\frac{1}{T}\int_0^T u(\boldsymbol{x},t)\,dt &= \lim_{T\to\infty}\frac{1}{T}\int_0^T \sum_{\boldsymbol{k}} a_{\boldsymbol{k}}(t)\varphi_{\boldsymbol{k}}(\boldsymbol{x})\,dt\\ &= \sum_{\boldsymbol{k}}\left(\lim_{T\to\infty}\frac{1}{T}\int_0^T a_{\boldsymbol{k}}(t)\,dt\right)\varphi_{\boldsymbol{k}}\\ &= \sum_{\boldsymbol{k}}\overline{a_{\boldsymbol{k}}}\varphi_{\boldsymbol{k}}(\boldsymbol{x})\,.\end{aligned}$$

It then follows from (2.22) that

$$u'(\boldsymbol{x}, t) = \sum_{\boldsymbol{k}} \left(a_{\boldsymbol{k}}(t) - \overline{a_{\boldsymbol{k}}}\right)\varphi_{\boldsymbol{k}}(\boldsymbol{x})\,. \tag{2.23}$$

Now if we associate $\overline{a_{\boldsymbol{k}}}$ with the "dc" part of the original signal $u(\boldsymbol{x}, t)$, as would commonly be done in signal processing analyses, we see that $u'(\boldsymbol{x}, t)$ contains all other (temporal) modes of the Fourier representation and, moreover, contains all spatial wavenumbers. That is, in general, u' represents all of the solution except the zeroth temporal modes, and as a consequence we should expect that it would be quite difficult to model. But, at least in principle, this is precisely what must be modeled in RANS formulations; that is, essentially the entire solution must be modeled!

Generic Form of RANS Equations—Absence of Interactions on Small Scales

The above implied difficulties led, from the beginning, to modeling of flow statistics (the Reynolds stresses, *e.g.*, $\overline{u'v'}$) rather than flow physics (u', v', *etc.*), and this essentially precludes any possibility for resolving short-time/small-length scale interactions between turbulence and other physical phenomena. To begin with, the quantities actually modeled (*e.g.*, $\overline{u'v'}$) have already been averaged over a time considered long compared with time scales of the turbulent fluctuating quantities themselves. In fact, as we have discussed earlier, for this averaging to be well defined it is necessary that the averaging time approach infinity. Thus, $\overline{u'v'}$ for example, is not even time dependent. At the same time, in typical computations the spatial scales are chosen to be small enough for resolution of large-scale behavior, but little more. Thus, in RANS formalisms, both spatial and temporal scales on which interactions would occur are absent from the models.

There are numerous physical situations in which such an approach cannot succeed. We will describe two of these here. It is well known (see Warhaft [111] and references therein) that passive scalars tend to exhibit a greater degree of intermittency (in the sense related to flatness, as discussed by Frisch [80]) than does the turbulent flow in which they are being carried. So consider a fluctuating temperature $\theta'(\boldsymbol{x}, t)$ at a fixed location $\boldsymbol{x}$ and its interaction with a component of fluctuating velocity, say $u'(\boldsymbol{x}, t)$. Now recall that, by definition, both θ' and u' have zero mean; but in the case of θ', the fluctuations will remain close to this mean value for long periods with an ocassional brief departure that may be of quite large magnitude. The consequence of this will be a temperature-velocity correlation such that $\overline{u'\theta'} \ll \mathcal{O}(1)$, implying that the turbulent heat fluxes have essentially no effect on the mean temperature. But in reality, the turbulent fluctuations could produce very significant local in space and instantaneous in time effects. RANS models are unable, in general, to reproduce

such phenomena, and in fact do not attempt to do so. Rather, the goal of RANS modeling in such situations is simply to produce averaged scalar fluxes whose overall effect is close to a "smearing" over time of the actual physics. This is sometimes acceptable, but often it is not.

A second example comes from combustion chemistry. It is well known that typical reaction rates can be expressed in terms of various physical parameters by a formula of the form

$$k(T) = AT^n \exp\left(\frac{-E_a}{R_0 T}\right) ,$$

where E_a is activation energy which, along with A and n, is an empirical constant; R_0 is the universal gas constant, and T is (absolute) temperature (see, *e.g.*, Williams [112]). Species production rates (needed to compute species concentrations) are obtained as the reaction rate times a product of species concentrations raised to various powers—an extremely nonlinear form—and these must be averaged in the context of the RANS formalism. We will consider details of averaging only the above formula since, alone, it already poses extreme difficulties. It is clear that

$$\begin{aligned} \overline{k}(T) &= \overline{AT^n \exp\left(\frac{-E_a}{R_0 T}\right)} \\ &\neq A\overline{T}^n \exp\left(\frac{-E_a}{R_0 \overline{T}}\right) = k(\overline{T}) , \end{aligned}$$

and lack of equality is so severe that the second formula on the right simply cannot be used. In particular, reaction rates cannot be accurately predicted on an instantaneous basis, and this would be true even if $\overline{k}(T) = k(\overline{T})$ because only the time-mean temperature is used in the second formula; that is, no effects of instantaneous fluctuating temperatures are included. Hence, species production rates will be inaccurate.

Beyond this, the species transport equations have been time averaged and contain terms of the form, for example, $\overline{u'c_i'}$, where c_i' represents temporal fluctuations in the i^{th} species concentration. But this scalar flux has also been time averaged, as the notation indicates, and carries no temporal information. Thus, there is no instantaneous interaction between the fluctuating flow field and chemical kinetics. So it is clear that no useful details of a chemically-reacting turbulent flow can be obtained from a RANS calculation unless the overall physics is such as to remain very close to the mean values of all variables at all times—which is rare in reaction chemistry, and even in this case an extreme amount of difficult modeling is necessary.

Inequality of RANS Solutions and Time-Averaged Solutions of N.–S. Equations

In this final subsection we consider what is possibly the most important question related to RANS solutions. To put this into the proper setting we first think in terms of time-averaged experimental results. Since it is now almost universally accepted that the N.–S. equations embody all the physics of turbulence, it is reasonable to view experimental time series as solutions (albeit, analog) to the N.–S. equations, subject, of course, to measurement errors. Hence, averaging of any such time series yields a time-averaged solution to the N.–S. equations. Then, with respect to the RANS equations, the natural question to consider is whether solutions to these equations equal time-averaged solutions to the N.–S. equations. Failure to demonstrate such equality would obviously raise serious questions regarding use of RANS formulations in general, and it is well known that comparisons of RANS solutions with experimental data have, from the earliest calculations to the present, always shown more than minor discrepancies.

This fact motivated an analytical study by McDonough [113], carried out in terms of ensemble averaging, which demonstrated that the desired equality could be obtained only if exact Reynolds

stresses could be constructed and used in the RANS equations. In the current section we will present an analogous result derived in the more useful context of time averaging, and then discuss some of its consequences.

We begin by observing that from the standpoint of mathematics there are two operators to consider in making the required comparisons. The first, and probably the more familiar one, is the *averaging operator* that we have already introduced; *viz.*, we can write the time-averaging procedure in the general form

$$\overline{(\,\cdot\,)} \equiv \lim_{T\to\infty} \frac{1}{T}\int_0^T (\,\cdot\,)\,dt\,, \tag{2.24}$$

with $(\,\cdot\,)$ representing any quantity of interest for which the combination of integration and limit processes exists. The second is a construct that often appears in modern PDE theory and is called the *solution operator*. This is simply a mapping from initial and/or boundary data for a PDE (or ODE) system to the solution of that system, evaluated at any prescribed time and/or spatial location. Such operators appear in many different forms (even for the same problem), and we will initially employ the generic notation

$$\boldsymbol{u}(\boldsymbol{x},t) = \boldsymbol{\mathcal{S}}(t)\boldsymbol{u}_0(\boldsymbol{x}) \tag{2.25}$$

used for initial-boundary value problems. But we note that little change is needed in the symbolism to accomodate steady-state boundary value problems. Here, $\boldsymbol{\mathcal{S}}(t)$ is the solution operator being applied to initial data $\boldsymbol{u}_0(\boldsymbol{x})$ to produce the solution $\boldsymbol{u}$ at time t and location $\boldsymbol{x}$.

The notation provided by Eqs. (2.24) and (2.25) permits us to examine the basic question being considered here in a more formal and abstract way. In particular, the question of whether solutions to RANS equations equal time-averaged solutions to the N.–S. equations takes the form

$$\boldsymbol{\mathcal{S}}_{RANS}\overline{\boldsymbol{u}}_0 \stackrel{?}{=} \overline{\boldsymbol{\mathcal{S}}_{NS}(t)\boldsymbol{u}_0}\,.$$

Thus, we are simply inquiring into the commutativity of the averaging and solution operators, but we note that the solution operators in the above expression are identical only with respect to space since the one the left-hand side is independent of time. But we wish to investigate the overall effect of the combination of both operators on each side of the equation to be the same. In particular, we can formally view $\boldsymbol{\mathcal{S}}_{RANS}$ on the left-hand side as containing a temporal operator employed in a pseudo-time integration to steady state, as would often be carried out in a numerical computation. In this context, $\overline{\boldsymbol{u}}_0$ could be the same on both sides of the above equation (*i.e.*, an initial guess), but it would not necessarily have to be if uniqueness of steady solutions could be shown. Then we can define the *commutator* of these two operators (solution operator and averaging) as

$$[\boldsymbol{\mathcal{S}}, \overline{(\,\cdot\,)}] \equiv \overline{\boldsymbol{\mathcal{S}}_{NS}\boldsymbol{u}_0} - \boldsymbol{\mathcal{S}}_{RANS}\overline{\boldsymbol{u}_0}\,, \tag{2.26}$$

and if this commutator is identically zero we are guaranteed that interchanging the order of averaging and solving the equations is valid; otherwise, it is not.

We can now prove the following theorem.

Theorem 2.1 *Suppose $\boldsymbol{U}(\boldsymbol{x},t)$ is a stationary solution to the Navier–Stokes equations on a domain $\Omega \subseteq \mathbb{R}^3$, and let averaging and solution operators be defined as above. Then the RANS equations are equivalent to the time-averaged N.–S. equations if, and only if, the Reynolds stress tensor employed in the RANS equations is exact.*

Proof. We express the N.–S. equations as

$$\boldsymbol{U}_t = \nu\Delta\boldsymbol{U} - \nabla p - \boldsymbol{U}\cdot\nabla\boldsymbol{U}$$

with initial data $\boldsymbol{U}_0(\boldsymbol{x}) = \boldsymbol{U}(\boldsymbol{x}, 0)$ and formally integrate in time:

$$\boldsymbol{U}(\boldsymbol{x}, t) = \boldsymbol{U}(\boldsymbol{x}, 0) + \int_0^t \nu\Delta\boldsymbol{U} - \nabla p - \boldsymbol{U}\cdot\nabla\boldsymbol{U}\, d\tau \quad \left(\equiv \boldsymbol{\mathcal{S}}_{NS}(t)\boldsymbol{U}_0\right), \tag{2.27}$$

where we assume p is such that the divergence-free constraint is satisfied by $\boldsymbol{U}$. Observe that Eq. (2.27) is a specific solution operator that could be readily implemented in a numerical algorithm even though the unknown solution appears on both sides of the equation—and, even nonlinearly on the right-hand side.

We next rearrange the above N.–S. solution and formally average the result:

$$\lim_{T\to\infty}\frac{1}{T}\int_0^T \overline{\boldsymbol{U}}(\boldsymbol{x}, t) - \overline{\boldsymbol{U}}(\boldsymbol{x}, 0)\, dt = \lim_{T\to\infty}\frac{1}{T}\int_0^T\int_0^t \nu\Delta\boldsymbol{U} - \nabla p - \boldsymbol{U}\cdot\nabla\boldsymbol{U}\, d\tau dt\,.$$

Then we observe that for each value of T occurring as the indicated limit is taken, the upper limit in the inner integration with respect to τ must be T. Thus, we express the above as

$$\overline{\boldsymbol{U}}(\boldsymbol{x}) - \overline{\boldsymbol{U}_0}(\boldsymbol{x}) = \lim_{T\to\infty}\int_0^T \frac{1}{T}\int_0^T \nu\Delta\boldsymbol{U} - \nabla p - \boldsymbol{U}\cdot\nabla\boldsymbol{U}\, d\tau dt\,.$$

Now as the limit is taken, the entire integrand of the outer integral approaches the average of the N.–S. equations, but at the same time the limit on this outer integral approaches infinity. Thus, some care is required in treating this limit process. We first apply the Fubini–Tonelli theorem (see, *e.g.*, Royden [114]) to the double integral to obtain

$$\lim_{T\to\infty}\int_0^T \frac{1}{T}\int_0^T \nu\Delta\boldsymbol{U} - \nabla p - \boldsymbol{U}\cdot\nabla\boldsymbol{U}\, d\tau dt = \lim_{T\to\infty}\int_0^T dt \lim_{T\to\infty}\frac{1}{T}\int_0^T \nu\Delta\boldsymbol{U} - \nabla p - \boldsymbol{U}\cdot\nabla\boldsymbol{U}\, d\tau\,,$$

which is valid provided the terms in the N.–S. equations are integrable (an already implicit assumption in constructing the solution operator). Hence,

$$\frac{\overline{\boldsymbol{U}}(\boldsymbol{x}) - \overline{\boldsymbol{U}_0}(\boldsymbol{x})}{\lim_{T\to\infty}\int_0^T dt} = \lim_{T\to\infty}\frac{1}{T}\int_0^T \nu\Delta\boldsymbol{U} - \nabla p - \boldsymbol{U}\cdot\nabla\boldsymbol{U}\, d\tau\,,$$

and it follows, under the assumption $\overline{\boldsymbol{U}}$ is bounded, that

$$\overline{\nu\Delta\boldsymbol{U} - \nabla p - \boldsymbol{U}\cdot\nabla\boldsymbol{U}} = 0\,.$$

Thus,

$$\nu\Delta\overline{\boldsymbol{U}} - \nabla\overline{p} - \overline{\boldsymbol{U}\cdot\nabla\boldsymbol{U}} = 0\,,$$

and if we introduce the Reynolds decomposition we obtain, after rearrangement,

$$\nabla\cdot\overline{\boldsymbol{u}}^2 = -\nabla\overline{p} + \nu\Delta\overline{\boldsymbol{u}} - \boldsymbol{R}(\boldsymbol{u'}, \boldsymbol{u'})\,, \tag{2.28}$$

which is the time-averaged N.-S. equation. Here, as earlier, $\boldsymbol{R}(\boldsymbol{u'}, \boldsymbol{u'})$ is the divergence of the Reynolds stress tensor. We see from this that if $\boldsymbol{R}(\boldsymbol{u'}, \boldsymbol{u'})$ is exact in the RANS equations (recall Eq. (2.12a)), then these equations are identical to what is obtained here by time averaging the N.–S. equations; and conversely, if $\boldsymbol{R}(\boldsymbol{u'}, \boldsymbol{u'})$ is not exact the identity does not hold. This completes the proof.

In the more abstract terminology introduced earlier, this shows that the commutator $[\boldsymbol{\mathcal{S}}, \overline{(\,\cdot\,)}]$ in Eq. (2.26) is zero if, and only if, exact Reynolds stresses are employed in the RANS equations. But we know that, in fact, we do not even have equations for the statistical correlations—other

than their definitions—that comprise the elements of this tensor. Consequently, there is essentially no way in which $\boldsymbol{R}(\boldsymbol{u'},\boldsymbol{u'})$ can be exact, and we should not expect to obtain very accurate results from the RANS equations. Indeed, the only way in which $\boldsymbol{R}(\boldsymbol{u'},\boldsymbol{u'})$ could be exact is if it were constructed with results from a DNS, and if such results were available there would be no need for a RANS calculation. Thus, it is important to investigate the error likely to occur when solving RANS equations if the Reynolds stress tensor is not exact.

To do this we need to first introduce some additional notation. We now write the RANS equations in terms of the dependent variable $\overline{\boldsymbol{v}}$ as

$$\nabla\cdot\overline{\boldsymbol{v}}^2 = -\nabla\overline{p}_M + \nu\Delta\overline{\boldsymbol{v}} - \boldsymbol{R}_M(\boldsymbol{v'},\boldsymbol{v'})\,, \tag{2.29}$$

where $\boldsymbol{R}_M$ and $\overline{p}_M$ denote modeled Reynolds stress and the corresponding pressure field required to maintain the divergence-free condition on $\overline{\boldsymbol{v}}$, respectively. We further define the error in a computed solution as the difference between the time-averaged N.–S. result and the RANS result:

$$\boldsymbol{e}(\boldsymbol{x}) = \overline{\boldsymbol{u}}(\boldsymbol{x}) - \overline{\boldsymbol{v}}(\boldsymbol{x})\,, \tag{2.30}$$

and similarly for the error in the pressure,

$$e_p(\boldsymbol{x}) = \overline{p}(\boldsymbol{x}) - \overline{p}_M(\boldsymbol{x})\,. \tag{2.31}$$

We recognize the first of these as the commutator given in Eq. (2.26) since $\overline{\boldsymbol{u}}$ results from $\overline{\boldsymbol{\mathcal{S}}_{\boldsymbol{NS}}\boldsymbol{u}_0}$, and $\overline{\boldsymbol{v}}$ is computed from $\boldsymbol{\mathcal{S}}_{RANS}\overline{\boldsymbol{v}}_0$ with $\overline{\boldsymbol{v}}_{\boldsymbol{0}} = \boldsymbol{u}_0$. Finally, we denote the difference in divergences of the Reynolds stress tensor as

$$\begin{aligned}\boldsymbol{E}(\boldsymbol{u'},\boldsymbol{v'}) &= \boldsymbol{R}(\boldsymbol{u'},\boldsymbol{u'}) - \boldsymbol{R}_M(\boldsymbol{v'},\boldsymbol{v'})\\ &= \nabla\cdot\begin{pmatrix} \overline{u_1'u_1'} - \overline{v_1'v_1'} & \overline{u_1'u_2'} - \overline{v_1'v_2'} & \overline{u_1'u_3'} - \overline{v_1'v_3'}\\ \overline{u_1'u_2'} - \overline{v_1'v_2'} & \overline{u_2'u_2'} - \overline{v_2'v_2'} & \overline{u_2'u_3'} - \overline{v_2'v_3'}\\ \overline{u_1'u_3'} - \overline{v_1'v_3'} & \overline{u_2'u_3'} - \overline{v_2'v_3'} & \overline{u_3'u_3'} - \overline{v_3'v_3'}\end{pmatrix}.\end{aligned} \tag{2.32}$$

We can now prove the following result that provides an estimate of the error incurred by using the RANS equations.

Theorem 2.2 *Suppose $\boldsymbol{R}_M$ is the divergence of the Reynolds stress tensor corresponding to any arbitrary RANS model, and $\boldsymbol{E}$ is the corresponding Reynolds stress error given in Eq.* (2.32)*. Then a bound for the error $\boldsymbol{e}$ of Eq.* (2.30)*, i.e., the commutator of averaging and solution operators, Eq.* (2.26)*, on a domain $\Omega\subseteq\mathbb{R}^3$ is of order $\mathcal{O}\big((\|\boldsymbol{E}\| + \|\boldsymbol{R_M}\|)^{1/2}\big)$ and satisfies*

$$\begin{aligned}\|\boldsymbol{e}\| \sim \frac{1}{2}\Bigg\{ 2\nu^*\frac{C_2}{C_1} \;+\; \sqrt{\left(\nu^*\frac{C_2}{C_1}\right)^2 + \frac{2}{C_1}\big(2\|\boldsymbol{E}\| + \|\boldsymbol{R}_M\|\big)} \;+\\ \sqrt{\left[2\nu^*\frac{C_2}{C_1} + \sqrt{\left(\nu^*\frac{C_2}{C_1}\right)^2 + \frac{2}{C_1}\big(2\|\boldsymbol{E}\| + \|\boldsymbol{R}_M\|\big)}\,\right]^2 + \frac{2}{C_1}\|\boldsymbol{E}\|}\Bigg\},\end{aligned} \tag{2.33}$$

where C_1 and C_2 are $\mathcal{O}(1)$ positive constants with values depending on the domain Ω, ν^* is one half the usual kinematic viscosity, and $\|\cdot\|$ is the L^2 norm.

Proof. We first obtain a differential equation for the commutator by taking the difference between Eqs. (2.28) and (2.29). This yields

$$\nabla\cdot\overline{\boldsymbol{u}}^2-\nabla\cdot\overline{\boldsymbol{v}}^2=\nu\Delta(\overline{\boldsymbol{u}}-\overline{\boldsymbol{v}})-\nabla(\overline{p}-\overline{p}_M)-(\boldsymbol{R}(\boldsymbol{u}',\boldsymbol{u}')-\boldsymbol{R}_M(\boldsymbol{v}',\boldsymbol{v}'))\,,$$

or in the notation introduced above,

$$\nabla\cdot\overline{\boldsymbol{u}}^2-\nabla\cdot\overline{\boldsymbol{v}}^2=\nu\Delta\boldsymbol{e}-\nabla e_p-\boldsymbol{E}\,. \tag{2.34}$$

We begin by rewriting the left-hand side of (2.34) as

$$\begin{aligned}\nabla\cdot\overline{\boldsymbol{u}}^2-\nabla\cdot\overline{\boldsymbol{v}}^2 &= \nabla\cdot\big(\overline{\boldsymbol{u}}^2-\overline{\boldsymbol{v}}^2\big)=\nabla\cdot\big(\overline{\boldsymbol{u}}+\overline{\boldsymbol{v}}\big)\big(\overline{\boldsymbol{u}}-\overline{\boldsymbol{v}}\big)\\ &= \nabla\cdot\big(\overline{\boldsymbol{e}}+2\overline{\boldsymbol{v}}\big)\boldsymbol{e})\\ &= 2\boldsymbol{e}\cdot\nabla\boldsymbol{e}+2\big(\boldsymbol{e}\cdot\nabla\overline{\boldsymbol{v}}+\overline{\boldsymbol{v}}\cdot\nabla\boldsymbol{e}\big)\,.\end{aligned}$$

Then, by using this along with Eqs. (2.30)–(2.32), we can express (2.34) as

$$\boldsymbol{e}\cdot\nabla\boldsymbol{e}-\frac{\nu}{2}\Delta\boldsymbol{e}+\frac{1}{2}\nabla e_p=-\frac{1}{2}\boldsymbol{E}-\big(\boldsymbol{e}\cdot\nabla\overline{\boldsymbol{v}}+\overline{\boldsymbol{v}}\cdot\nabla\boldsymbol{e}\big)\,. \tag{2.35}$$

We immediately observe that, except for the factor 1/2, the left-hand side of this equation is in precisely the same form as occurs for the steady N.–S. equations. Thus, it is fairly easy to prove existence of (strong) solutions (in both 2D and 3D)—although the latter may not be unique—provided the right-hand side is at least in $L^2(\Omega)$ and is divergence free. The reader should consult [2] and [89] for more details, especially the former. We remark that, as noted in these references, the situation for steady N.–S. equations is not nearly so delicate as described in Chap. 1 for the time-dependent equations. On the other hand, the right-hand side of (2.35) is not, in general, divergence free.

The simplest treatment of Eq. (2.35) is via Leray projection to remove the term corresponding to error (due indirectly to modeling) in the pressure gradient and at the same time to force the right-hand side to be divergence free. This results in

$$\boldsymbol{e}\cdot\nabla\boldsymbol{e}-\nu^*\Delta\boldsymbol{e}=-\boldsymbol{E}^*\,, \tag{2.36}$$

where we have not introduced any specific notation to signify Leray projection. Furthermore, we have set $\nu^*=\nu/2$ and defined

$$\boldsymbol{E}^*\equiv\frac{1}{2}\boldsymbol{E}+\boldsymbol{e}\cdot\nabla\overline{\boldsymbol{v}}+\overline{\boldsymbol{v}}\cdot\nabla\boldsymbol{e}\,, \tag{2.37}$$

which now can be assumed to be divergence free.

It is worthwhile at this point to note that Eq. (2.36) is expressed entirely in terms of modeling errors and the RANS solution $\overline{\boldsymbol{v}}$ that results from these. At the same time, due to the similarity of this equation with the steady N.–S. equation we know that a solution $\boldsymbol{e}$ exists, and is in $H^2(\Omega)$. It is then clear that we can write

$$\|\boldsymbol{e}\cdot\nabla\boldsymbol{e}-\nu^*\Delta\boldsymbol{e}\|=\|\boldsymbol{E}^*\|\,,$$

from which it follows that

$$\|\boldsymbol{e}\cdot\nabla\boldsymbol{e}\|-\nu^*\|\Delta\boldsymbol{e}\|\le\|\boldsymbol{E}^*\|\,,$$

and from the Poincaré inequality (see, *e.g.*, [89]), $C_1\|\boldsymbol{e}\|\le\|\nabla\boldsymbol{e}\|$ with C_1 depending on the geometry of Ω, it follows that

$$C_1\|\boldsymbol{e}\|^2-\nu^*\|\Delta\boldsymbol{e}\|\le\|\boldsymbol{E}^*\|\,. \tag{2.38}$$

Now if $\boldsymbol{e}$ is sufficiently regular ($\boldsymbol{e} \in H^2(\Omega)$ is need, but this is required for $\boldsymbol{e}$ to be a strong solution to (2.36) in any case), we can "boot strap" the Poincaré inequality to obtain a Sobolev inequality (see [89])

$$\|\nabla \boldsymbol{e}\| \leq \frac{1}{C}\|\Delta \boldsymbol{e}\| \qquad \Rightarrow \qquad C_1\|\boldsymbol{e}\| \leq \|\nabla \boldsymbol{e}\| \leq \frac{1}{C}\|\Delta \boldsymbol{e}\|\,,$$

and hence,

$$C_2\|\boldsymbol{e}\| \leq \|\Delta \boldsymbol{e}\|\,.$$

But we are now unable to maintain the direction of the inequality in (2.38), so we must simply express it as an order relation:

$$C_1\|\boldsymbol{e}\|^2 - \nu^* C_2\|\boldsymbol{e}\| \sim \|\boldsymbol{E}^*\|\,.$$

Clearly, estimates based on this can no longer be sharp, but they should nevertheless be informative.

We observe that the above expression is a quadratic in $\|\boldsymbol{e}\|$ with $\boldsymbol{E}^*$ presumed known. We rearrange this to obtain

$$\|\boldsymbol{e}\|^2 - \nu^*\frac{C_2}{C_1}\|\boldsymbol{e}\| - \|\boldsymbol{E}^*\| \sim 0\,.$$

At this point we reintroduce the definition (2.37) for $\boldsymbol{E}^*$. Then we have

$$\begin{aligned}
\|\boldsymbol{E}^*\| &= \frac{1}{2}\|\boldsymbol{E} + \boldsymbol{e}\cdot\nabla\overline{\boldsymbol{v}} + \overline{\boldsymbol{v}}\cdot\nabla\boldsymbol{e}\| \\
&\leq \frac{1}{2}\|\boldsymbol{E}\| + \|\boldsymbol{e}\|\|\nabla\overline{\boldsymbol{v}}\| + \|\overline{\boldsymbol{v}}\|\|\nabla\boldsymbol{e}\| \\
&\sim \frac{1}{2}\|\boldsymbol{E}\| + 2C_1\|\boldsymbol{e}\|\|\overline{\boldsymbol{v}}\|\,,
\end{aligned}$$

with the last line following from the Poincaré inequality, and again leading to inability to maintain direction of the inequality. Substitution of this expression into the above quadratic expression results in

$$\|\boldsymbol{e}\|^2 - \left(\nu^*\frac{C_2}{C_1} + 2\|\overline{\boldsymbol{v}}\|\right)\|\boldsymbol{e}\| - \frac{1}{2C_1}\|\boldsymbol{E}\| \sim 0\,, \tag{2.39}$$

the solution of which is

$$\|\boldsymbol{e}\| \sim \frac{1}{2}\left[\nu^*\frac{C_2}{C_1} + 2\|\overline{\boldsymbol{v}}\| + \sqrt{\left(\nu^*\frac{C_2}{C_1} + 2\|\overline{\boldsymbol{v}}\|\right)^2 + \frac{2}{C_1}\|\boldsymbol{E}\|}\right]\,, \tag{2.40}$$

where we have taken the larger of the two possible solutions to obtain an upper bound.

To complete the proof of Theorem 2.2 we need an estimate of $\|\overline{\boldsymbol{v}}\|$. This will be provided by the following lemma.

Lemma 2.1 *Suppose $\overline{\boldsymbol{v}}$ is a solution to the steady-state RANS equation* (2.29)*. The $\|\overline{\boldsymbol{v}}\|$ satisfies the order relation*

$$\|\overline{\boldsymbol{v}}\| \sim \frac{\nu^*}{2}\frac{C_2}{C_1} + \sqrt{\left(\frac{\nu^*}{2}\frac{C_2}{C_1}\right)^2 + \frac{1}{2C_1}\|R_M\|}\,,$$

where $\|\cdot\|$ is the L^2 norm on Ω.

Proof (of lemma). The proof of this estimate is quite similar to that already carried out to estimate $\|\boldsymbol{e}\|$. In particular, since (2.29) is of the same form a steady-state N.–S. equation, we know strong solutions exist in both 2D and 3D provided, in this case, R_M is divergence free. Of course, this will not be true, so Leray projection must be applied to arrive at the form

$$\nabla \cdot \left(\overline{\boldsymbol{v}}^2\right) - \nu \Delta \overline{\boldsymbol{v}} = -\boldsymbol{R}_M$$

where, as before, we have not specifically denoted the projection operator. Observe that this is now of precisely the form (2.36), the starting point for estimating $\|\boldsymbol{e}\|$, but the procedure is now somewhat easier because $\|\boldsymbol{E}^*\|$ of (2.36) explicitly contained $\|\overline{\boldsymbol{v}}\|$ (the reason the current lemma is needed); but here $\boldsymbol{R}_M$ is simply a RANS model. Thus, calculations analogous to those performed above lead directly to the stated result, and the proof of the lemma is complete. A somewhat more useful form of the result given above is contained in the following:

Corollary 2.1 *Suppose $\|\overline{\boldsymbol{v}}\|$ satisfies the estimate of Lemma 2.1. Then it can be expressed as*

$$\|\overline{\boldsymbol{v}}\| \sim \frac{1}{2}\left[\nu^* \frac{C_2}{C_1} + \sqrt{\left(\nu^* \frac{C_2}{C_1}\right)^2 + \frac{2}{C_1}\left(2\|\boldsymbol{E}\| + \|\boldsymbol{R}_M\|\right)}\right] . \tag{2.41}$$

Proof (of corollary). We have the following estimates.

$$\|\boldsymbol{R}_M\| = \|\boldsymbol{R} - \boldsymbol{E}\| = \|\boldsymbol{E} - \boldsymbol{R}\|$$

by the definition of $\boldsymbol{R}_M$. But

$$\|\boldsymbol{E} - \boldsymbol{R}\| < \|\boldsymbol{E} + \boldsymbol{R}\| < \|\boldsymbol{E}\| + \|\boldsymbol{R}\| < 2\|\boldsymbol{E}\| + \|\boldsymbol{R}_M\| .$$

This proves the corollary.

Proof of Theorem 2.2 is now completed by substituting (2.41) into the occurences of $\|\overline{\boldsymbol{v}}\|$ in Eq. (2.40).

There are several interesting observations to be made regarding this result. The first is that, in some sense to leading order, the difference between time-averaged solutions and solutions to time-averaged equations scales as the square root of the error in the Reynolds stress tensor. That is

$$\|\boldsymbol{e}\| \sim \|\boldsymbol{E}\|^{1/2} , \tag{2.42}$$

and this at least hints at an explanation for why RANS models are able to produce at least reasonable results in portions of many flow fields (after constants have been tuned) despite the fact that they have many fundamental flaws. That is, despite all the flaws, the end result is not as bad as might have been expected. Second, we see that effects of molecular viscosity are not likely to be an important contribution to $\|\boldsymbol{e}\|$ for high-Re flows. In addition, because of the way in which the Poincaré inequality was applied, it can be checked that the constants C_1 and C_2 become small for large domains Ω, and this will tend to increase $\|\boldsymbol{e}\|$. At the same time, it must be recognized that Eq. (2.33) is a global, order-of-magnitude estimate; it tells us little about what occurs on a point-to-point basis unless the exact turbulent solution is homogeneous on Ω, and this is also correctly reflected in $\boldsymbol{R}_M$. Furthermore, we remark that although these results have formally been derived for the 3-D situation, there is nothing that intrinsically restricts them to this case. Thus, we could expect the relation (2.42) to hold in 2D, and even for a 1-D Burgers' equation problem.

Finally, it is important to recognize that the result given in this theorem is completely analytical and independent of any numerical computational technique. This implies that refining discretization step sizes will not, in general, decrease $\|\boldsymbol{e}\|$ and, in fact, could even increase it—something that is often observed in computations. As a consequence, we see that solutions to RANS equations cannot converge to solutions to the time-averaged Navier–Stokes equations, no matter what RANS model is used, unless the modeling error $\boldsymbol{E}$ is identically zero.

2.1.5 Reynolds-averaged Navier–Stokes Models

This section will be devoted to treatment of some of the best-known and most widely used of the classical turbulence models. One might wonder, in light of the theorems of the preceding section, what is the point of considering this modeling approach at all. But it is rather clear that studies associated with such models have so dominated turbulence "theory" for so long that no treatment of the subject could be complete without a fairly detailed presentation of these. Moreover, these models still are very often used, especially for industrial-scale computations performed in the context of commercial CFD software. We will in a first course, however, restrict discussions to the most fundamental versions of these methods.

We begin with a few rather general remarks on modeling itself. From the viewpoint of mathematics, and probably also from that of engineering, other than turbulence, models are usually taken to be simplifications of problems (or of the equations representing them) that allow analyses that might not otherwise be possible. Thus, we would expect that usually the equations of a model problem should be simpler than those of the actual complete situation being modeled. A familiar example of this in fluid dynamics is potential flow. Here, the Navier–Stokes equations are replaced with Laplace's equation which often (at least formally, always) has exact solutions. But of course we give up the ability to predict skin friction drag with this model. If we upgrade this slightly with boundary-layer theory, then we are able to readily solve a wide range of flow problems because the equations involved are still much simpler than the full N.–S. equations. Rather generally, the goal of modeling is to replace a complicated system of equations with a much simpler one that can be solved analytically, or at worst, with minimal numerical analytic effort—and, hopefully, be able to assess, *a priori*, the range of validity of any such solution.

We will see, however, that this does not occur with classical (or any other) turbulence models. In fact, in all cases known to date, the model of turbulence presents equations that are more complicated than the N.–S. equations themselves. On the other hand, these equations are in a sense "more easily solved" when viewed in the context of required computer CPU time because their run times usually scale as $\mathcal{O}(Re^{\alpha})$ with $\alpha < 3$. Indeed, the methods we will consider herein have run times generally independent of Re, but in some cases, nevertheless, not very short as will be obvious. Thus, turbulence models accomplish what a model should in a rather general sense. But at the same time it is usually the case that ranges of validity of such models cannot be generally predicted.

We begin this section with presentation of some widely-used terminology associated with classification of turbulence models followed by a fairly detailed treatment of the Boussinesq hypothesis. We then consider, in turn, Prandtl's mixing-length theory, the basic k–ε model, and finally the second-moment closure methods.

RANS Turbulence Model Terminology

There are several different ways by means of which RANS models are classified. One of the more common is in terms of the number of additional PDEs one must solve beyond those of the Navier–Stokes equations. Thus, as described earlier in Def. 1.119, one considers zero-equation, one-equation and two-equation models, and as noted by Wilcox [98] there are also $\frac{1}{2}$-equation models which include a single ordinary differential equation in their formulation. But this does not exhaust the possibilties. Indeed, the "second-moment closures," so named because they include equations for each of the second-moment statistical quantities comprising the Reynolds stress tensor, employ at least five additional PDEs, and in some forms might include seven. In principle, one could construct third- and higher-moment closures, although this is almost never done, even theoretically. In part because of all these various possible formulations and associated terminologies, turbulence models

are often simply called "algebraic" if they include no additional ODEs or PDEs, and otherwise they are termed "differential."

There is yet a different way to view modeling and, correspondingly, construct models which we have not previously considered. It involves two-point (or, in general, multi-point) closures. All of the models to be considered herein will be based on statistics constructed at a single point (*one-point closures*). For example, $\overline{u'v'}$ is a time average of data collected at a single point. Of course, this correlation must be constructed at every point in a computational flow field, but at each such point data only from that point are used to construct $\overline{u'v'}$. There is no particular reason to do this (beyond its relative simplicity, of course), and indeed it can be argued that this is completely inadequate due to the fact that turbulence involves erratic spatial fluctuations as well as temporal ones; no account of the effects of these can be included in a one-point closure. Thus, one might consider a *two-point correlation* such as $\overline{u'(\boldsymbol{x})v'(\boldsymbol{x}+\boldsymbol{r})}$ where $\boldsymbol{x}$ and $\boldsymbol{x}+\boldsymbol{r}$ are two distinct points where data might be collected. One could generalize this even further by first averaging in space over all points a distance r from the chosen point $\boldsymbol{x}$, and then perform the indicated time average. We remark, however, that with the exception of structure function based sub-grid scale LES models (see Métais and Lesieur [115] and Sagaut [73]) there presently are no multi-point closure models in use.

The Boussinesq Hypothesis

The Boussinesq hypothesis, often called the "Boussinesq approximation"—which leads to confusion with a completely different concept associated with natural convection—was described in Def. 1.56 as a means to relate turbulent shear stress to the mean flow strain rate. (This is also termed "gradient transport," especially in the context of passive scalars; see [91] and also [7].) Thus, for example, we might hypothesize that

$$\overline{u'v'} \sim \frac{1}{2}\left(\frac{\partial \overline{u}}{\partial y} + \frac{\partial \overline{v}}{\partial x}\right) . \tag{2.43}$$

The first question one should ask is "Why would we suppose such a relationship might be valid?" To answer this, we need only recall Newton's law of viscosity which might be paraphrased as "shear stress is proportional to strain rate, with viscosity being the constant of proportionality." That is,

$$\tau = \mu \frac{\partial u}{\partial y} , \tag{2.44}$$

for example. But it is important to bear in mind that the physical situation is very different for the two cases corresponding to the above respective relations. In the laminar flow case, Eq. (2.44) is an empirical relationship that is extremely well supported by experimental results. Moreover, the (dynamic) viscosity μ is a property of the fluid, and at least in the case of perfect gases, a first-principles formula for it can be derived from the kinetic theory of gases (see, *e.g.*, Vincenti and Kruger [116]).

The turbulent flow case, which might be expressed as

$$-\overline{u'v'} = \nu_T \left(\frac{\partial \overline{u}}{\partial y} + \frac{\partial \overline{v}}{\partial x}\right) , \tag{2.45}$$

where ν_T is the *turbulent eddy viscosity*, is not an empirically-supported physical result—in fact, just the opposite is true; see Chen *et al.* [117]. (Here, the factor $\frac{1}{2}$ from Eq. (2.43) has been absorbed into ν_T, and the negative sign has been introduced on the left-hand side for reasons that will be apparent

as we proceed.) Indeed, Boussinesq himself warned of its shortcomings, not the least of which would be actually determining values of ν_T. Moreover, DNS results, from which all factors in Eq. (2.45) (except for ν_T—and actually even this by inference) can be directly computed, indicate a rather poor correlation between $\overline{u'v'}$ and $\partial\overline{u}/\partial y + \partial\overline{v}/\partial x$, or any other similar relationship. Finally, it should be observed that the Boussinesq hypothesis attempts to relate presumably small-scale statistical behavior characterized by the Reynolds stress to large-scale, mean (again, statistical) behavior (the mean strain rates). Of course, we should recall that due to the nature of the Reynolds decomposition, u' and v' are not necessarily small; but independent of this, Eq. (2.45) relates statistical, and not physical, properties of the flow.

It is thus clear that Eq. (2.45) has no basis in physics beyond a simple analogy from laminar flow, and ν_T cannot be a physical fluid property. Indeed, it changes with each flow, and if one takes the view that Eq. (2.45) is nothing but a definition for ν_T, *viz.*,

$$\nu_T \equiv -\frac{\overline{u'v'}}{\frac{1}{2}\left(\frac{\partial\overline{u}}{\partial y} + \frac{\partial\overline{v}}{\partial x}\right)}, \tag{2.46}$$

we see that ν_T cannot be a constant, except for extremely simple flows. Moreover, it is also easy to see that ν_T should actually be a tensor if we are to employ this in 3D. This is all very different from the situation for the physical viscosity ν, as should be obvious.

Wilcox [98] provides a fairly detailed discussion of the Boussinesq hypothesis, emphasizing its shortcomings, in order to justify consideration of more sophisticated approximations to the Reynolds stress tensor. He lists the following set of flow situations in which the Boussinesq hypothesis has proven to be inadequate:

i) flows with sudden changes in mean strain rate,

ii) flows over curved surfaces,

iii) flow in ducts or, in general, those containing secondary fluid motions, including boundary-layer separation,

iv) flow of rotating and/or stratified fluids,

v) three-dimensional flows.

We will not here provide an exhaustive discussion regarding exactly what goes wrong with the Boussinesq hypothesis in each of these separate situations but rather give some general comments that apply to several of these. The reader is referred to [98] for more details.

It should first be observed that turbulence tends to "remember" its past history, at least for short times. This is, of course, one of the strongest arguments against randomness (but not necessarily against stochasticity) of turbulence, and indeed, the Taylor hypothesis would never work if turbulence did not, however briefly, remember its past. Equation (2.45) does not reflect this. It is clear that in a flow field exhibiting mean strain rate this equation predicts nonzero turbulent stresses (which is qualitatively correct); but as soon as the strain is removed, this same relationship instantaneously predicts zero turbulent shear stress. This does not at all coincide with experimental observations, and related to something we mentioned earlier in our discussion of isotropic turbulence, the rate of decay of turbulent stresses is a physical observable (albeit, statistical) that can be used to calibrate turbulence models that are more sophisticated than the Boussinesq hypothesis-based models, and which are to some extent able to correctly account for this. We also remark that the instantaneous decay to zero of turbulent stress predicted by Eq. (2.45) violates the basic notion of causality—an effect can neither precede, nor be simultaneous with, its cause.

Somewhat as an aside, we mention that precisely this same argument (lack of causality) might be made against using Newton's law of viscosity for laminar flows. Indeed, this is a valid criticism, and such use might conceivably have some influence on detailed accuracy of Stokes flow ($Re \to 0$ and hence viscously dominated). But in the case $Re \to \infty$ as in turbulent flows, the effects of the viscous terms are balanced (at some point even dominated) by the nonlinear advective terms which lead to finite-time responses. This, in turn, leads to ostensibly correct behaviors regarding the ability of flows computed via DNS and LES to remember their past for short times, and at least in the case of DNS such results agree extremely well with experiment even in the relatively low Re regimes currently accessible via DNS.

This very basic and fundamental flaw in the form of the Boussinesq hypothesis lies at the heart of essentially all of the items in the preceding list, influencing some more than others. But there are additional problems with the hypothesis, and we will touch on just two of these. First, regarding failure in three-dimensional flows, we remark that this is not as serious as might be thought because, as already mentioned, Eq. (2.46) implies that ν_T should actually be a tensor; and if we allow this, then use of the Boussinesq hypothesis can, at least formally, be extended to 3-D flows—modulo its more fundamental flaws, of course. Second, in association with rotating and/or stratified fluid flows, the form (2.43) is, from the start, lacking the physics of these flow situations. In principle, one might consider attempting to include this physics, and at least in a very informal way, this is often done in the context of stratification resulting from thermal effects. In particular, a turbulent Prandtl number, Pr_T, is often used with the Reynolds analogy (see Def. 1.100) as the justification, to account for fluctuating turbulent fluxes (of thermal energy). Similar arguments are also used to obtain turbulent Schmidt numbers, Sc_T, for application in models of turbulent transport of chemical species. But in the end, the basic flaws of the Boussinesq hypothesis persist, and computed results always reflect this except in the simplest flow situations.

Finally, we mention a difficulty arising from the Boussinesq approximation that is of a more mathematical nature and is seldom cited (or even recognized) in the engineering literature. To understand this particular shortcoming we must first recall that the source of the Reynolds stresses was averaging the nonlinear advective terms of the N.–S. equations. But, as will be apparent in the sequel, the Boussinesq hypothesis leads to replacing these with linear diffusive terms. The mathematical consequences of this are far reaching. To see this we should recall the balance of nonlinear advection and linear diffusion (dissipation) exhibited in the Galerkin form of the N.–S. equations (1.44) and (1.45). Clearly, adding a diffusive term of any size (and in practice it turns out to be large) will significantly upset this balance; the resulting equations will be far more dissipative. While this may be good for numerical computation, it is bad physics, and in the context of the modern view of the N.–S. equations it is guaranteed to alter bifurcation sequences. This essentially precludes any chance of predicting transition to turbulence with RANS models, and it is well know that this is true. Separate, *ad hoc*, models of transition must be implemented.

Some Specific RANS Models

In this section we will somewhat briefly treat three specific RANS turbulence models. The first of these will be the "mixing-length" model of Prandtl [19] which, in the terminology introduced earlier, is an example of a zero-equation, or algebraic, model. We will follow this with a description of the widely-used k–ε models that in some sense were already foreseen by Kolmogorov [106] and Chou [118], and introduced in their present form by Jones and Launder [108] and Launder and Sharma [119]. These are the so-called two-equation models. We will then present the equations corresponding to a second-moment closure. These include partial differential equations for each of the components of the Reynolds stress tensor. We finally close the section with a few remarks regarding methods we have neglected to discuss in any detail, including some quite recent developments, and

an overall assessment of results from RANS models. We comment at the outset that much of the material presented here can be found in far more detail in the book by Wilcox [98].

Mixing-length theory. As we have noted earlier, the "mixing-length theory" introduced by Prandtl [19] in 1925 was the first attempt to formally derive the turbulent eddy viscosity appearing in the Boussinesq hypothesis. Our treatment of this strongly reflects that of [98]; there is a yet more detailed analysis to be found in Tennekes and Lumley [7] which is highly-recommended reading for the serious student.

We begin our discussions by recalling the form of the RANS equations presented earlier in Eqs. (2.12) as

$$\nabla \cdot \overline{\boldsymbol{u}}^2 = -\nabla \overline{p} + \nu \Delta \overline{\boldsymbol{u}} - \boldsymbol{R}(\boldsymbol{u'}, \boldsymbol{u'}) \,,$$

or now restricted to 2D:

$$\frac{\partial \overline{u}^2}{\partial x} + \frac{\partial \overline{u}\,\overline{v}}{\partial y} = -\frac{\partial \overline{p}}{\partial x} + \nu \Delta \overline{u} - \frac{\partial \overline{u'^2}}{\partial x} - \frac{\partial \overline{u'v'}}{\partial y} \,, \tag{2.47a}$$

$$\frac{\partial \overline{u}\,\overline{v}}{\partial x} + \frac{\partial \overline{v}^2}{\partial y} = -\frac{\partial \overline{p}}{\partial y} + \nu \Delta \overline{v} - \frac{\partial \overline{u'v'}}{\partial x} - \frac{\partial \overline{v'^2}}{\partial y} \,, \tag{2.47b}$$

and

$$\frac{\partial \overline{u}}{\partial x} + \frac{\partial \overline{v}}{\partial y} = 0 \,. \tag{2.48}$$

We now observe that, for example,

$$\nu \Delta \overline{u} = \nu \left(\frac{\partial^2 \overline{u}}{\partial x^2} + \frac{\partial^2 \overline{u}}{\partial y^2} \right) = \frac{\partial}{\partial x} \left(\nu \frac{\partial \overline{u}}{\partial x} \right) + \frac{\partial}{\partial y} \left(\nu \frac{\partial \overline{u}}{\partial y} \right) \,, \tag{2.49}$$

thus showing the well-known relationship between the viscous terms of the N.–S. equations and Newton's law of viscosity. But even more important for our purposes is the suggestive relationship of the last two terms on the right-hand sides of Eqs. (2.47) to the terms on the far right-hand side of Eq. (2.49) via the Boussinesq hypothesis. Namely, we have

$$-\overline{u'^2} = \nu_T \frac{\partial \overline{u}}{\partial x} \,, \qquad -\overline{u'v'} = \nu_T \frac{\partial \overline{u}}{\partial y} \,, \qquad -\overline{v'^2} = \nu_T \frac{\partial \overline{v}}{\partial y} \,. \tag{2.50}$$

We note that within the framework of 2-D free shear flows and boundary layers, where use of a mixing-length model has been most successful, only the second of the above relations actually appears in the RANS equations, so we will concentrate only on this term.

We now observe that if ν_T were known, we could immediately compute $\overline{u'v'}$ (and $\overline{u'^2}$, $\overline{v'^2}$, as appropriate) because $\overline{u}$, itself, is the solution to Eq. (2.47a), and we could express this equation as

$$\frac{\partial \overline{u}^2}{\partial x} + \frac{\partial \overline{u}\,\overline{v}}{\partial y} = -\frac{\partial \overline{p}}{\partial x} + \frac{\partial}{\partial x} \left((\nu + \nu_T) \frac{\partial \overline{u}}{\partial x} \right) + \frac{\partial}{\partial y} \left((\nu + \nu_T) \frac{\partial \overline{u}}{\partial y} \right) \,. \tag{2.51}$$

Clearly, an analogous equation could be developed from Eq. (2.47b) to complete the construction in 2D, and a similar derivation can be done for 3-D problems.

Then to complete the modeling effort and thus "close" the equations, we must determine ν_T—precisely what Boussinesq warned "might be very difficult, or even impossible." It is worthwhile to follow the line of reasoning employed by Prandtl, as is carried out by Wilcox [98], not because it is correct, but more because such a presentation makes it easier to highlight the fallacies. Thus, we first recall that the kinetic theory derivation of the molecular coefficient of viscosity μ (here replaced

by $\nu = \mu/\rho$) requires, among other things, establishing length and velocity scales associated with transport of momentum in the gas under consideration. Indeed, we know that $\nu \sim \mathrm{L}^2/\mathrm{T} \sim \mathrm{L}\cdot\mathrm{L}/\mathrm{T} \sim \mathrm{L}\cdot\mathrm{U}$ where L, T and U are general length, time and velocity dimensions, respectively. Also recall that within the framework of kinetic theory, the length scale is the mean free path between molecular collisions (assumed to be elastic in the simplest analyses), and the velocity scale comes from an average over the Maxwellian distribution for a fixed volume of gas in thermodynamic equilibrium. A direct, but tedious, derivation results in a closed-form formula for μ (or ν) that agrees fairly well with experimental measurements, and which can be brought into very good agreement by replacing the elastic "billiard-ball" collisions with use of more realistic molecular potential energy distribution functions and collision cross sections (see, *e.g.*, Hirschfelder *et al.* [120]). It was Prandtl's goal to construct an analogous treatment leading to ν_T—despite the fact that it is not a physical property of the fluid.

Prandtl first hypothesized that one might envision the fluid flow as consisting of collections of fluid parcels moving about randomly with some characteristic speed, which here we denote as v_{mix} consistent with [98], and which over some characteristic length scale ℓ_{mix} they would essentially retain their momentum. We remark that one often encounters a similar description based on turbulent eddies (*á la* Richardson [4]) instead of fluid parcels, and with ℓ_{mix} corresponding to the distance over which these eddies retain there identities.

Independent of the details of these descriptions, it is clear that the behavior of a fluid parcel (or eddy) is rather different from that of molecules of a gas; so attempting to apply anything resembling the formalism from the kinetic theory of gases is dubious at best. Nevertheless, we will proceed in an attempt to establish formulas for v_{mix} and ℓ_{mix}. Prandtl [19] chose to employ dimensional analysis to find v_{mix}, and since it is obviously desirable to express this in terms of mean flow quantities (we have equations for these), he hypothesized that

$$v_{mix} \sim \ell_{mix} \left| \frac{d\overline{u}}{dy} \right| . \tag{2.52}$$

Now since, again simply on dimensional grounds, we should have

$$\nu_T \sim \ell_{mix} v_{mix} ,$$

we set

$$\nu_T = \ell_{mix}^2 \left| \frac{d\overline{u}}{dy} \right| , \tag{2.53}$$

where any required constants will be absorbed into ℓ_{mix}. We observe that this formulation preserves the positivity of ν_T as must be done, if for no other reason than mathematical well posedness. Also, if we use the above in the Boussinesq hypothesis, we obtain an estimate of the magnitude of $\overline{u'v'}$ component of the Reynolds stress, namely,

$$\left| -\overline{u'v'} \right| = \left| \ell_{mix} \frac{d\overline{u}}{dy} \right|^2 . \tag{2.54}$$

It is now worthwhile, again following [98], to examine the validity of the mixing-length analysis. We first observe that in the turbulent flow situation there is, in general, no true analog of thermodynamic equilibrium, and unlike the kinetic theory situation the fluid parcels are constantly changing. Thus, on very fundamental grounds, attempting to apply anything like kinetic theory makes essentially no sense. But beyond this (or, maybe, because of it) is the fact that we have not been able to obtain a formula for ℓ_{mix}. In this sense the mixing-length theory is an "incomplete" one in that a major part of the corresponding model must be determined entirely from experiment

(or, possibly, from DNS). Despite this shortcoming, ℓ_{mix} has been quite successfully determined for a number of free shear flows possessing only a single length scale. But the value of ℓ_{mix} is different for each type of flow, as should be expected. In particular, it has been typical to express ℓ_{mix} as

$$\ell_{mix} = C_1 \delta(x) \,, \tag{2.55}$$

where $\delta(x)$ is the single characteristic length (*e.g.*, the diameter of a jet), and C_1 is the closure constant to be determined such that computed mean velocity profiles match corresponding measurements. Flows exhibiting a single length scale include the free shear flows, far wake, mixing layer and jet (both plane and round). The values obtained for C_1 for these cases are given in [98] and repeated here as Table 2.1.

Table 2.1: Values of mixing-length closure constant for various flows

Flow Type	Far Wake	Mixing Layer	Plane Jet	Round Jet
C_1	0.180	0.071	0.098	0.080

We observe that there is a nearly 20% discrepancy even between the two types of jets, although with these constants the match of $\overline{u}$ is quite good. But since C_1 is associated with a single flow feature, ℓ_{mix}, this wide range of values required for very similar types of flows strongly suggests basic flaws in the theory.

Finally, we note that in the context of wall-bounded shear flows (*e.g.*, boundary layers), significant modifications must be made not only to the closure coefficient (recall that we absorbed an undetermined constant into ℓ_{mix} in Eq. (2.53)), but also to the structure of the model itself. Nevertheless, this has been carried out with a fair degree of success for unseparated boundary layers. We will discuss only the simplest instance of this here; the interested reader is encouraged to consult [98] and references therein for further details.

We first recall our discussion of the law of the wall in Chap. 1. There it was shown that four distinct regions exist within a boundary layer: *i*) viscous sublayer, *ii*) buffer layer, *iii*) log layer and *iv)* defect layer. Thus it appears likely that there must be more than a single length scale needed to adequately describe a boundary layer, and indeed, this observation was already made and utilized in producing the formulas for the velocity profile corresponding to the law of the wall. Wilcox [98] notes that Prandtl [19] had proposed that the mixing length might be given by

$$\ell_{mix} = \kappa y \,, \tag{2.56}$$

where κ is the Kármán constant having the experimentally determined value $\kappa = 0.41$, and y is the distance normal to the surface bounding the flow. But apparently Prandtl did not expect this single form of ℓ_{mix} to work throughout the boundary layer—and it does not. It was not until introduction of the Van Driest *damping function* [121] in 1956, however, that a workable result was available. This empirical modification brings the mixing-length formula to

$$\ell_{mix} = \kappa y \left[1 - e^{-y^+/A_0^+} \right] \,, \tag{2.57}$$

where, as in Chap. 1, $y^+ = u_\tau y/\nu$ with $u_\tau = \sqrt{\tau_w/\rho}$, and $A_0^+ = 26$. This result works reasonably well from the bounding surface out through the log layer, but it does not adequately describe the defect layer.

There have been numerous other modifications to the basic mixing-length model through the years. Here we simply mention two of the most widely used: the Cebeci–Smith model [122] and

the Baldwin–Lomax model [123]. The former is valid only for 2-D calculations, and while the latter is used in 3-D, it already contains a total of six adjustable closure constants. There are a couple items deserving mention here. First, both the Cebeci–Smith and Baldwin–Lomax models provide treatment of the defect layer. But because the precise value of y at which this layer begins is not known, *a priori*, it is necessary to construct testable "switching functions" to determine where to begin using the appropriate formulas for the defect layer. In the case of the Cebeci–Smith model (but not for Baldwin–Lomax) this results in a formula for ℓ_{mix} that depends on velocity. In turn, this implies that ν_T depends on velocity, and it is easy to check that the resulting RANS equations are no longer Galilean invariant. Proof of this is left as an exercise for the student. The reader will find further details of these models in [98].

***k*–ε models.** In this section we provide an introductory treatment of the widely-used k–ε models. As might be inferred from the terminology, these are two-equation models. In their basic form they consist of a PDE for each of turbulence kinetic energy, k, and turbulence kinetic energy dissipation rate ε. Together, these two quantities provide velocity and length scales needed to directly construct eddy viscosity at each point in a computational domain. In this sense the k–ε models can be viewed as being "closed" because unlike zero- and one-equation models (with the possible exception of the Spalart–Allmaras model [124], and other similar ones, for eddy viscosity) these models possess sufficient equations for constructing eddy viscosity with no direct appeal to experimental results needed. We also note that k and ε are not the only variables employed in two-equation models. For example, there are also k–ℓ, k–ω and k–ω^2 models, among others. The latter two of these are quite well known, but we leave it to the interested reader to seek more information on these; a good starting point is Wilcox [98] and references therein.

We will begin this section with a derivation of a general total kinetic energy equation for incompressible flow expressed in its simplest form. We follow this with an outline of the derivation for kinetic energy of the mean flow, again expressed in the simplest manner. Then, following Tennekes and Lumley [7], we present a detailed analysis of the equation for turbulence kinetic energy that permits us to analyze the physics represented by each term. But this is not the equation employed in practical calculations, so we next present this with a brief comparison to the theoretical form. We then very briefly treat construction of a PDE for dissipation rate and collect results to obtain a complete k–ε model. We close the section with presentation of some well-known results and discussion of the consequences these imply.

To obtain a general energy equation for incompressible flow we simply form the dot product of the N.–S. equations with the velocity vector that is their solution. Thus,

$$\boldsymbol{U}\cdot(\boldsymbol{U}_t + \boldsymbol{U}\cdot\nabla\boldsymbol{U}) = \boldsymbol{U}\cdot(-\nabla p + \nu\Delta\boldsymbol{U})\,. \tag{2.58}$$

A straightforward calculation shows that

$$\frac{\partial}{\partial t}\left(\frac{1}{2}|\boldsymbol{U}|^2\right) + \boldsymbol{U}\cdot\nabla\left(\frac{1}{2}|\boldsymbol{U}|^2\right) = -\nabla\cdot(p\boldsymbol{U}) + \nu\Delta\left(\frac{1}{2}|\boldsymbol{U}|^2\right)\,,$$

where

$$\frac{1}{2}|\boldsymbol{U}|^2 = \frac{1}{2}(u^2+v^2+w^2) \equiv K\,, \tag{2.59}$$

the total kinetic energy per unit mass. Thus, the above can be expressed as

$$K_t + \boldsymbol{U}\cdot\nabla K = -\nabla\cdot(p\boldsymbol{U}) + \nu\Delta K\,. \tag{2.60}$$

This equation is a very general result; but at the same time, in the context of incompressible flows, it carries no new information not already found in the N.–S. equations themselves. Nevertheless,

we will see that various manipulations of it can lead to useful insights. In this regard we stress that in the current chapter only the Reynolds decomposition will be studied. But, in fact, essentially any decomposition could be used, and as will be apparent as we proceed, details of any associated analysis will depend on the form of decomposition employed. Thus, we emphasize that for a LES or Hilbert space decomposition, results we will obtain here would not be valid in detail.

We now introduce the Reynolds decomposition into Eq. (2.59) to obtain

$$\begin{aligned} K &= \frac{1}{2}\left[(\overline{u}+u')^2+(\overline{v}+v')^2+(\overline{w}+w')^2\right] \\ &= \frac{1}{2}\left(\overline{u}^2+\overline{v}^2+\overline{w}^2\right)+\frac{1}{2}\left(u'^2+v'^2+w'^2\right)+\overline{u}u'+\overline{v}v'+\overline{w}w' \\ &\equiv \overline{k}+k'+\overline{\boldsymbol{U}}\cdot\boldsymbol{U}'\,. \end{aligned} \tag{2.61}$$

Finally, we note that the definition of *turbulence kinetic energy* is

$$k \equiv \frac{1}{2}\left(\overline{u'^2}+\overline{v'^2}+\overline{w'^2}\right)=\overline{k'}\,. \tag{2.62}$$

We can also construct an energy equation analogous to Eq. (2.60) for the mean flow in exactly the same way, *viz.*, form the dot product of the mean velocity vector with the RANS equations. This results in

$$\overline{\boldsymbol{u}}\cdot\nabla\overline{k}=-\nabla\cdot(\overline{p}\,\overline{\boldsymbol{u}})+\nu\Delta\overline{k}-\overline{\boldsymbol{u}}\cdot\boldsymbol{R}(\boldsymbol{u'},\boldsymbol{u'})\,. \tag{2.63}$$

There are two key observations to make in comparing Eq. (2.63) for the mean flow energy with Eq. (2.60) for total energy. First, Eq. (2.63) is a steady-state equation, as it must be, while Eq. (2.60) is time dependent. Second is the appearance of terms involving the Reynolds stress tensor. Particularly the second of these will ultimately play a major role in interpretations given to terms in the equation for turbulence kinetic energy.

We should now observe that we still do not have an equation for what we actually want—the turbulence kinetic energy. Derivation of this is extremely tedious and is essentially always left as "an exercise to the reader." Here too, we will mainly follow this approach, but we will provide a few intermediate steps. As noted in [7], the starting point is to time average the total energy equation (2.60), and then subtract the mean flow energy equation (2.63) from this result.

Formal time averaging of (2.60) after substituting the decomposition obtained in Eq. (2.61) yields

$$\overline{\boldsymbol{U}\cdot\nabla(\overline{k}+k')}+\overline{\boldsymbol{U}\cdot\nabla(\overline{\boldsymbol{U}}\cdot\boldsymbol{U}')}=-\overline{\nabla\cdot(p\,\boldsymbol{U})}+\nu\overline{\Delta(\overline{k}+k')}\,,$$

and after considerable manipulation this can be expressed as

$$\overline{\boldsymbol{u}}\cdot\nabla(\overline{k}+k)+\nabla\cdot\left(\overline{\boldsymbol{u}'k'}\right)=-\nabla(\overline{p}\,\overline{\boldsymbol{u}})-\nabla\cdot\left(\overline{p'\boldsymbol{u}'}\right)+\nu\Delta\overline{k}+\nu\Delta k\,, \tag{2.64}$$

where p' is fluctuating pressure divided by density.

We observe that Eq. (2.64) is now steady state, and that it is expressed in terms of both mean and fluctuating quantities. Subtracting Eq. (2.63) from this leads to an equation for the turbulence kinetic energy:

$$\overline{\boldsymbol{u}}\cdot\nabla k+\nabla\cdot\left(\overline{\boldsymbol{u}'k'}\right)=-\nabla\cdot\left(\overline{p'\boldsymbol{u}'}\right)+\nu\Delta k+\overline{\boldsymbol{u}}\cdot\boldsymbol{R}(\boldsymbol{u'},\boldsymbol{u'})\,. \tag{2.65}$$

But this is <u>not</u> the form found in either [7] or [98], so we need to perform considerable further manipulations to arrive at either of these. We will not carry out the details here, but instead merely present the results with a brief indication of how they are obtained.

We begin by presenting the results given in each of the two references being used. From Tennekes and Lumley [7] we have

$$\overline{u}_j\frac{\partial k}{\partial x_j}=-\frac{\partial}{\partial x_j}\left(\overline{p'u'_j}+\frac{1}{2}\overline{u'_iu'_iu'_j}-2\nu\overline{u'_is'_{ij}}\right)-\overline{u'_iu'_j}\,\overline{s_{ij}}-2\nu\overline{s'_{ij}s'_{ij}}\,. \tag{2.66}$$

The $\overline{s_{ij}}$ and s'_{ij} represent components of mean and fluctuating strain rate tensors, respectively:

$$\overline{s_{ij}} = \frac{1}{2}\left(\frac{\partial \overline{u}_i}{\partial x_j} + \frac{\partial \overline{u}_j}{\partial x_i}\right), \qquad \text{and} \qquad s'_{ij} = \frac{1}{2}\left(\frac{\partial u'_i}{\partial x_j} + \frac{\partial u'_j}{\partial x_i}\right). \tag{2.67}$$

Similarly, from Wilcox [98] we obtain

$$k_t + \overline{u}_j \frac{\partial k}{\partial x_j} = -\overline{u'_i u'_j}\frac{\partial \overline{u}_i}{\partial x_j} - \varepsilon + \frac{\partial}{\partial x_j}\left(\nu \frac{\partial k}{\partial x_j} - \overline{p' u'_j} - \frac{1}{2}\overline{u'_i u'_i u'_j}\right), \tag{2.68}$$

with ε expressed in Cartesian tensor form as

$$\varepsilon = 2\nu \overline{s'_{ij} s'_{ij}}. \tag{2.69}$$

At first glance Eqs. (2.65), (2.66) and (2.68) all appear to be rather different; but with the exception of the time-derivative term (incorrectly) retained in (2.68), these equations are actually equivalent. To see this we first note the similarities and then briefly explain the apparent discrepancies.

We immediately recognize the term corresponding to advection of turbulence kinetic energy appearing in the left-hand side of all three equations. The only difference between (2.65) and the other two equations is the Cartesian tensor notation employed in the latter two; in particular, we have $\overline{\boldsymbol{u}} \cdot \nabla k = \overline{u}_j \partial k / \partial x_j$. Similarly, the first term on the right-hand side of (2.65) equals the corresponding term in (2.66) and the fourth term on the right-hand side of (2.68):

$$\nabla \cdot \left(\overline{p' \boldsymbol{u}'}\right) = \frac{\partial}{\partial x_j}\left(\overline{p' u'_j}\right),$$

showing the equivalence of these terms. We next observe that the remaining term on the left-hand side of Eq. (2.65) is easily shown to be identical to the velocity "triple correlations" appearing in the right-hand sides of Eqs. (2.66) and (2.68). That is,

$$\nabla \cdot \left(\overline{\boldsymbol{u}' k'}\right) = \frac{1}{2}\frac{\partial}{\partial x_j}\overline{u'_i u'_i u'_j}.$$

Finally, we easily see that the diffusion terms on the right-hand sides of Eqs. (2.65) and (2.68) exactly match; namely,

$$\nu \Delta k = \frac{\partial}{\partial x_j}\left(\nu \frac{\partial k}{\partial x_j}\right),$$

but there appears to be no corresponding term in Eq. (2.66). At the same time, there seem to be no strain-rate terms in either of Eqs. (2.65) and (2.68) while there are three such terms in Eq. (2.66). However, using the definition of ε from Eq. (2.69), we see that the last term on the right-hand side of (2.66) is precisely the ε term of (2.68). Finally, we observe that by making use of the properties of the Cartesian tensor notation it can be checked that to within a factor of two, we have

$$\overline{u'_i u'_j}\,\overline{s_{ij}} \sim \overline{u'_i u'_j}\frac{\partial \overline{u}_i}{\partial x_j}.$$

Thus, to reconcile the forms of Eqs. (2.66) and (2.68) we must account for the factor of two, and for the lack of a term in (2.66) corresponding to diffusion of turbulence kinetic energy.

The latter of these is easily explained, and this in turn leads to resolution of the factor of two as well as all the apparent discrepancies between Eqs. (2.65) and (2.66). In particular, if one begins derivation of the equation for turbulence kinetic energy with the N.–S. equations expressed in terms

of strain rate and ignores the cancellations resulting from the divergence-free condition (as is done in [7]), then one arrives at Eq. (2.66). On the other hand, if this is not done (following the approach we have taken in the present notes), no explicit strain rate terms appear. This essentially resolves all differences between Eq. (2.66) and Eq. (2.68).

We still need to provide arguments to show that Eqs. (2.65) and (2.66) are equivalent. The basic one has already been noted, namely, retaining strain rate in lieu of invoking incompressibility to remove most of the terms arising in this representation. But Eq. (2.65) contains a Reynolds stress term that has not been matched with any terms in either of Eqs. (2.66) and (2.68). It is easily checked that use of product-rule differentiation permits one to recover the terms $\overline{u_i'u_j'}\,\overline{s_{ij}}$ of (2.66) in Eq. (2.65) (recall that $\boldsymbol{R}$ includes a divergence operator), but additional terms also arise. Similar product-rule differentiations in the strain rate terms (along with judicious combinations of these with terms from the Reynolds stress tensor) lead to the remaining terms appearing in (2.66). As indicated in [7], these are very tedious calculations, and they will not be repeated here.

The point of this rather lengthy development is the following. The derivation we have provided for Eq. (2.65) is fairly easy and straightforward and corresponds to a purely math-oriented approach. The form corresponding to Eq. (2.66) leads to a number of significant physical insights as we will briefly discuss below, and Eq. (2.68) is directly related to the turbulence energy equation actually used in RANS models. Thus, each form has an important application, and demonstration of their equivalence is important.

The physical interpretation of terms in Eq. (2.66), as presented in [7], can be summarized as follows. As already mentioned, the term on the left-hand side is simply advection of turbulence kinetic energy, and the first term on the right-hand side represents pressure work due only to turbulence. The second term on the right-hand side corresponds to transport of turbulence kinetic energy by turbulent fluctuations, which is more clearly seen in the equivalent term appearing on the left-hand side of Eq. (2.65). Tennekes and Lumley [7] describe the next term on the right-hand side, $2\nu\overline{u_i's_{ij}'}$, as "transport by viscous stresses," but they do not make clear what is being transported. Indeed, Eq. (2.66) is a transport equation for turbulence kinetic energy; but if the term in question is accomplishing transport by viscous stresses, then it is clear that it is not kinetic energy that is being transported, but rather simply the turbulent fluctuations. In fact, from our earlier discussions it is evident that this term is actually associated with the diffusive transport of turbulence kinetic energy which is the same as transport of viscous stresses, but is an intuitively clearer concept.

Tennekes and Lumley [7] associate the final two terms on the right-hand side of Eq. (2.66) with "two kinds of deformation work." They term the first of these "turbulence production," and use the notation

$$\mathcal{P} \equiv -\overline{u_i'u_j'}\,\overline{s_{ij}} \tag{2.70}$$

to identify it. We see from this that what is actually represented is amplification of the Reynolds stress tensor components by the mean strain rate. We have already argued that this is equivalent to the first term on the right-hand side of Eq. (2.68), and that it arises from rearrangement of terms involving the Reynolds stress tensor as appears on the right-hand side of (2.65). We remark that if true "scale separation" occurs, as is usually supposed in the context of model construction, we must question whether the mean strain rate would actually be capable of interaction with Reynolds stress components.

The final term on the right-hand side of Eq. (2.66) is described in [7] as the "rate at which fluctuating viscous stresses ($\nu s_{ij}'$) perform deformation work against the fluctuating strain rate (s_{ij}')." But we have already seen that the term $\nu\overline{s_{ij}'s_{ij}'}$ is (up to a constant factor) simply the dissipation rate of turbulence kinetic energy given in Eq. (2.69). A physical interpretation based on the latter description seems far more intuitive, especially with regard to the purported consequences.

At this point it is worthwhile to observe, as noted in [7], that $\varepsilon \simeq \mathcal{P}$ often holds, at least

in an order-of-magnitude sense; that is, dissipation of turbulence kinetic energy by viscosity is approximately equal to its production via strain rate amplification of Reynolds stress. Recall that we earlier, in Def. 1.70, termed this type of situation "equilibrium turbulence." This corresponds to

$$-\overline{u_i' u_j'}\,\overline{s_{ij}} = 2\nu\overline{s_{ij}' s_{ij}'}\,. \tag{2.71}$$

We immediately note that the right-hand side of this expression is nonnegative, which implies that velocity correlations and corresponding mean strain rates would be expected to usually have opposite signs. On the other hand, we must be careful to recognize that Eq. (2.71) does not often hold exactly, and as also observed in [7], it will seldom be satisfied for shear flows. If we recall the Galerkin form of the N.–S. equations presented in Chap. 1, we see that "production" actually arises simply from the nonlinear terms, and dissipation is a result of only the diffusive terms. On small scales (high wavenumbers) it is indeed a "balance"—not necessarily equality—of these that produces the strange attractor of the dynamical system. At the same time, if equality holds over essentially all scales (implied by mathematics of the Reynolds decomposition) as required in equilibrium turbulence, the Galerkin form of the equations of motion collapses to triviality. This all reflects the difficulties that can arise when one attempts to associate physics with quantities that are merely statistical. Nevertheless, the authors of [7] proceed to derive the estimate

$$\overline{s_{ij}' s_{ij}'} \gg \overline{s_{ij}}\,\overline{s_{ij}} \tag{2.72}$$

using Eq. (2.71). Indeed, the magnitude of the difference between these two parts of the strain rate tensor is of the order of the integral scale Reynolds number, Re_ℓ. Despite the rather unrealistic assumptions made in [7] in the course of obtaining the inequality (2.72) (equilibrium turbulence with single length and velocity scales), it is an expected result from the standpoint of our understanding of the mathematical nature of solutions to the N.–S. equations. We recall that a feature of turbulent solutions is their intrinsic nonsmoothness, particularly in their high-wavenumber/high-frequency components; that is, spatial derivatives on small scales can be expected to be much larger than those on larger scales. Since strain rates are simply combinations of spatial derivatives, inequality (2.72) is completely consistent with this.

We are now prepared to return to Eq. (2.68), the form of the turbulence kinetic energy equation employed for computations, and indicate how it is converted to the form usually taken in k–ε models. We begin by observing that there are three terms that must be modeled, a remaining one that will be computed from a PDE, which itself requires very significant modeling, and then construction of eddy viscosity from the results of all this. These terms correspond to production, pressure work, velocity triple correlation and dissipation. Following Wilcox [98] we employ the Boussinesq hypothesis in the form

$$-\overline{u_i' u_j'} = 2\nu_T\overline{s_{ij}} - \frac{2}{3}k\delta_{ij} \tag{2.73}$$

to model the Reynolds stress factor in the production term. The second term in this expression accounts for the fact that the trace of the strain rate tensor is zero for an incompressible flow, and is thus necessary to provide consistency with the definition of turbulence kinetic energy k from Eq. (2.62) when $i = j$. We remark, however, that this results in employing normal viscous stress terms in the absence of eddy viscosity which is logically inconsistent with the Boussinesq hypothesis. Despite this, it is useful to recognize that Eq. (2.73) provides a means to calculate each term of the Reynolds stress tensor once a calculation of the mean velocity has been completed. Thus, second-moment information can be retrieved even though we have not directly calculated it as we would in a second-moment closure approach. In principle, this is true for any model employing a Boussinesq hypothesis; in particular, once the mean velocity field has been obtained, if ν_T is available, we can directly compute the $\overline{u_i' u_j'}$s.

We next consider approximation of the pressure work term $\overline{p'u'_j}$ appearing in the right-hand side of Eq. (2.68). To accomplish this we first observe that a slight modification of the Boussinesq hypothesis leads to an approximation of scalar turbulent fluxes analogous to the Reynolds stress component fluctuations. For example, suppose we decompose temperature as $T = \overline{\theta} + \theta'$. Then the Reynolds-averaged thermal energy equation will contain terms of the form $\overline{u'_j\theta'}$ known as *scalar fluxes*, and these are approximated as $(\nu_T/\sigma_T)\partial\overline{\theta}/\partial x_j$ where σ_T is the turbulent Prandtl number. So one might suppose that we could represent the pressure work term (usually called *pressure diffusion* in the modeling literature) as $\overline{p'u'_j} = \nu_T\partial\overline{p}/\partial x_j$. But this does not work. Indeed, it is clear that pressure does not satisfy a transport equation, so the notion of pressure diffusion is completely fallacious. Much effort has been devoted to modeling this term, and rather complicated results have been obtained and used in the context of second-moment closures (see [98]). There is little physical or mathematical justification for any of these, so we will not provide any further descriptions.

Within the confines of k–ε models a simpler approach is usually taken. This begins by also recognizing that we have little in the way of sound theory for modeling the velocity triple correlation. So we (arbitrarily) combine this with the pressure diffusion term and model these together as "diffusion of kinetic energy:"

$$-\overline{p'u'_j} - \frac{1}{2}\overline{u'_iu'_iu'_j} = \frac{\nu_T}{\sigma_k}\frac{\partial k}{\partial x_j}\,. \tag{2.74}$$

Here, σ_k is an additional closure constant. We emphasize that there is no physical justification for use of Eq. (2.74). It is merely a convenience that permits us to express Eq. (2.68) as

$$\frac{\partial k}{\partial t} + \overline{u}_j\frac{\partial k}{\partial x_j} = -\overline{u'_iu'_j}\frac{\partial\overline{u}_i}{\partial x_j} - \varepsilon + \frac{\partial}{\partial x_j}\left[(\nu + \nu_T/\sigma_k)\frac{\partial k}{\partial x_j}\right]\,. \tag{2.75}$$

It is worthwhile to point out the types of terms appearing in this equation because it is common practice in RANS modeling to presume all turbulence quantities, physical or otherwise, satisfy transport equations of this basic form. Thus, we observe that in addition to the time-dependent term on the left-hand side there is also advection (macroscopic transport) of kinetic energy. On the right-hand side we find production, dissipation and diffusion of kinetic energy. This is almost always the point of departure for construction of essentially all such models. The fact that there is no *a priori* reason to expect that many such quantities (*e.g.*, eddy viscosity, as in the Spalart–Allmaras models) should satisfy a transport equation is never considered.

We have yet to construct a model for ε. We begin this by recalling the definition from Eqs. (2.67) and (2.69). Clearly, a model is needed since we do not know u'_i, and we therefore cannot compute the required derivatives. We also observe that when we present the equation for second-moment closure in the next section we will see that within this context dissipation of turbulence kinetic energy is a second-rank tensor—a matrix:

$$\varepsilon_{ij} = 2\nu\overline{\frac{\partial u'_i}{\partial x_k}\frac{\partial u'_j}{\partial x_k}}\,.$$

So the first approximation we make here in our effort to model dissipation as it appears in (2.75) is the usual one; that is, dissipation is locally isotropic, meaning that small-scale (high wavenumber) behavior is independent of direction and satisfies required rotation and reflection properties discussed earlier. Then we express the above as

$$\varepsilon_{ij} = \frac{2}{3}\varepsilon\delta_{ij}\,,$$

with

$$\varepsilon \equiv \nu\overline{\frac{\partial u'_i}{\partial x_k}\frac{\partial u'_i}{\partial x_k}}\,. \tag{2.76}$$

Effects of this, and modifications, have been widely studied (see [98] and references therein), and we will not dwell on this in these lectures. We mention, however, that in the context of a Reynolds decomposition it is much more difficult to justify invocation of a local isotropy hypothesis than is the case for a LES decomposition for the simple reason that the u_i' represent far more than high-wavenumber behavior, in general, in the present case.

Even within the confines of local isotropy, the equation for ε is difficult to derive, and it introduces a total of six new higher-order correlations that must be modeled. These are all statistical, and as noted by Wilcox [98], the result is not a model of physics but rather a model of the original PDE(s) associated with the tensor components ε_{ij}. Nevertheless, it is widely used, and we present it here:

$$\frac{\partial\varepsilon}{\partial t}+\overline{u}_j\frac{\partial\varepsilon}{\partial x_j}=-C_{\varepsilon1}\frac{\varepsilon}{k}\overline{u_i'u_j'}\frac{\partial\overline{u}_i}{\partial x_j}-C_{\varepsilon2}\frac{\varepsilon^2}{k}+\frac{\partial}{\partial x_j}\left[(\nu+\nu_T/\sigma_\varepsilon)\frac{\partial\varepsilon}{\partial x_j}\right].$$

This equation contains three new closure constants, $C_{\varepsilon1}$, $C_{\varepsilon2}$ and σ_ε. But assuming values of these can be found, it can be seen that we now have a complete system of equations. In particular, we observe that the required length and velocity scales needed for construction of ν_T now are available at each point of a computational grid. To see this, we note that $k\sim(\mathrm{L/T})^2$, and $\varepsilon\sim\mathrm{L}^2/\mathrm{T}^3$. Thus, we immediately obtain a velocity scale from $k^{1/2}$, as is obvious; moreover, it is clear that $k/\varepsilon\sim\mathrm{T}$. It then follows that $k^{3/2}/\varepsilon\sim\mathrm{L}$, and since $\nu_T\sim\mathrm{L}^2/\mathrm{T}$, we obtain

$$\nu_T=C_\nu\frac{k^2}{\varepsilon},$$

where C_ν is a constant needed to account for the fact that this has been derived entirely via dimensional analysis. It is important to recognize that with ν_T given in this form the diffusion terms of the kinetic energy and dissipation equations become highly nonlinear, and while fairly elaborate numerical techniques probably should be employed to handle this, in practice this is seldom done.

We now collect the complete set of equations comprising the "standard" k–ε RANS model. These consist of the mean flow continuity and momentum equations, with the latter containing the eddy viscosity, the equations for turbulence kinetic energy and its dissipation rate, and those corresponding to the Boussinesq hypothesis and eddy viscosity in this context. Thus, we have the following:

$$\nabla\cdot\overline{\boldsymbol{u}}=0, \tag{2.77a}$$

$$\overline{\boldsymbol{u}}_t+\overline{\boldsymbol{u}}\cdot\nabla\overline{\boldsymbol{u}}=-\nabla\overline{p}+\nabla\cdot[(\nu+\nu_T)\nabla\overline{\boldsymbol{u}}], \tag{2.77b}$$

$$k_t+\overline{\boldsymbol{u}}\cdot\nabla k=\mathcal{P}-\varepsilon+\nabla\cdot[(\nu+\nu_T/\sigma_k)\nabla k], \tag{2.77c}$$

$$\varepsilon_t+\overline{\boldsymbol{u}}\cdot\nabla\varepsilon=C_{\varepsilon1}\frac{\varepsilon}{k}\mathcal{P}-C_{\varepsilon2}\frac{\varepsilon^2}{k}+\nabla\cdot[(\nu+\nu_T/\sigma_\varepsilon)\nabla\varepsilon], \tag{2.77d}$$

with production $\mathcal{P}$ given by

$$\mathcal{P}=-\overline{u_i'u_j'}\frac{\partial\overline{u}_i}{\partial x_j}, \tag{2.78a}$$

$$-\overline{u_i'u_j'}=2\nu_T\overline{s_{ij}}-\frac{2}{3}k\delta_{ij}, \tag{2.78b}$$

and

$$\nu_T=C_\nu\frac{k^2}{\varepsilon}. \tag{2.79}$$

The closure constants corresponding to the so-called standard k–ε model are:

$$C_\nu=0.09,\quad C_{\varepsilon1}=1.44,\quad C_{\varepsilon2}=1.92,\quad \sigma_k=1.0,\quad \sigma_\varepsilon=1.3. \tag{2.80}$$

It is clear that this provides a completely closed system for computing $\overline{\boldsymbol{u}}$, $\overline{p}$, k and ε, and from (2.78b) all components of the Reynolds stress tensor.

Several remarks are in order regarding Eqs. (2.77)–(2.80). The first, and possibly most important, is that they cannot be integrated all the way to a solid boundary. Instead, the law of the wall must be employed to provide velocity "boundary conditions" away from solid boundaries. Numerous physical arguments have been proposed to explain this (*e.g.*, approximations made to obtain equations for k and ε are not strictly valid in the viscous sublayer), but the simplest follows directly from the mathematics of the energy dissipation equation Eq. (2.77d). It is clear from the definition of ε, Eq. (2.76), that ε will not generally go to zero approaching a solid boundary. At the same time we see from Eq. (2.62) that $k \to 0$ as a solid boundary is approached. Now the term involving $\mathcal{P}$ in Eq. (2.77d) approaches zero near solid surfaces, but the term containing ε^2/k is singular. Clearly, if we can assign Dirichlet conditions for ε on solid boundaries this formally removes the singularity since in this case the dissipation rate equation will not be solved on the boundary. But despite this, as gridding is refined, the ε^2/k term can become large at the first few grid points near the boundary, and thus difficult to handle from a computational aspect. In particular, this is often associated with "stiffness" of the numerical problem.

The simplest way to avoid this difficulty is to employ the law of the wall to permit assigning velocity (and all other) conditions at a distance from the wall corresponding to somewhere in the log layer. One must be somewhat careful with this, however. It is well known that turbulence kinetic energy attains its maximum fairly close to the wall (but somewhere in the log layer) for wall-bounded shear flows. Thus, if one is to have any hope of predicting this correctly it is necessary to evaluate the log law fairly close to the wall to find the required velocity boundary condition. We also note that the specific difficulty we have analyzed here is not present if we employ Neumann conditions on solid walls for the dissipation rate equation. In some respects, this might be preferable in any case since there is actually no way to accurately assign Dirichlet conditions for dissipation rate. Another, somewhat more complicated, approach is to employ the so-called "low-Re" k–ε formalisms in the vicinity of solid boundaries. Treatment of these can be found in [98] and references therein.

The second observation is that the equations for kinetic energy and dissipation rate are strongly coupled with each other, but only weakly so with the momentum equations. This suggests that significant difficulties can arise in solution procedures for the k and ε equations which are not present for the momentum equations alone. Moreover, little mathematical analysis has been applied to these equations beyond the rather naïve treatment provided by Mohammadi and Pironneau [125]. In particular, these equations exhibit, in addition to their stronger coupling, forms of nonlinearity not generally encountered in transport equations raising numerous questions regarding existence and uniqueness of their solutions, and their regularity. Even in the absence of these questions there clearly is significant extra computer arithmetic because there are two additional nonlinear PDEs to be solved.

In general, assignment of boundary conditions, for k and ε is difficult. We have already mentioned the problems with ε on solid walls. But at inflows and outflows values are not known for either k or ε. For outflows it is common practice (although not necessarily accurate) to assume fully-developed flow and employ homogeneous Neumann conditions for both of these. But at inflows this is not valid, and it is difficult to obtain accurate conditions, especially for ε. Indeed, both k and ε are statistical, rather than physical, so it is not easy to deduce their values simply from physical considerations alone.

Finally, we note that the definition of ν_T in terms of k and ε renders the system of equations no longer Galilean invariant. (We leave demonstration of this as an exercise for the reader; consultation of [99] may be helpful.) As we have emphasized earlier, Galilean invariabnce is a crucial physical property of solutions to the N.–S. equations, and its lose during construction of any turbulence model must cast strong doubts on validity of such a model.

We remark that none of these issues is new; they are all well known. So poor performance of k–ε methods applied to realistic flow situations is not surprising. There have been many modifications through the years to improve this situation including those associated with guaranteeing realizability and Galilean invariance (see Durbin and Pettersson Reif [126] for a brief treatment of these), and the modern renormalization group (RNG) approaches begun by Yahot and Orszag [93]. We will not provide treatment of these methods herein, leaving this material to the cited references. One of the main reasons for this is suggested by Fig. 2.3, a replot of one given by Freitas [94]. This

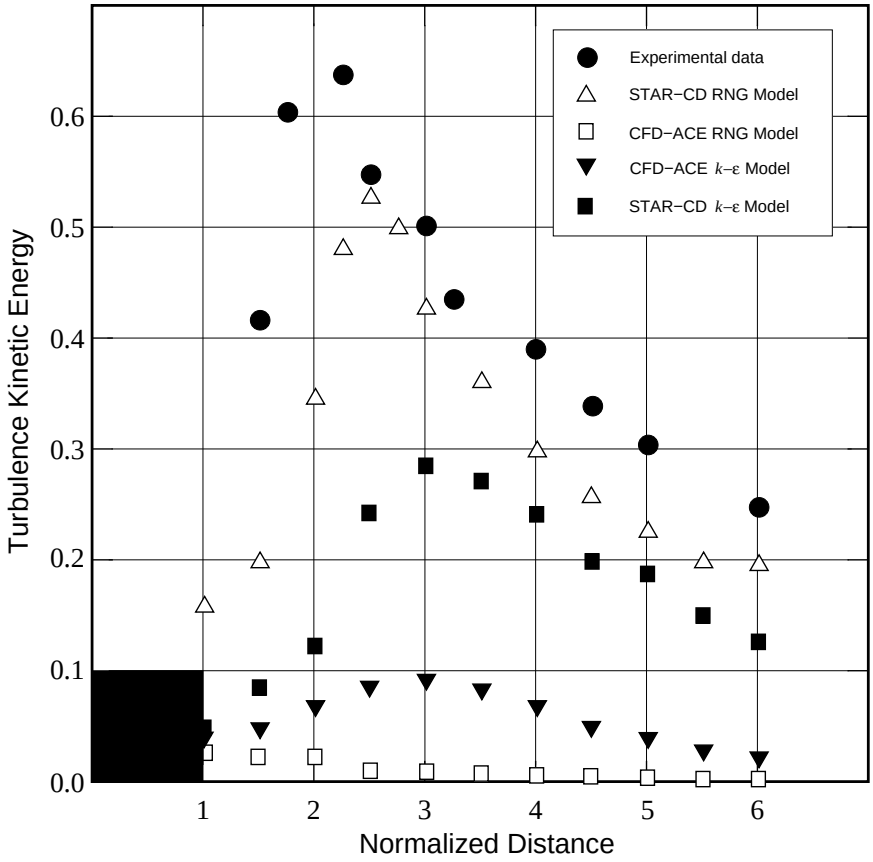

Figure 2.3: Comparison of various k–ε models for flow over a rib of square cross section.

figure displays computed results obtained from several different commercial flow codes/k–ε models for a very straightforward problem, 3-D flow over a rib located in the center of a wind tunnel, and compared with experimental data. The figure shows measured (normalized) turbulence kinetic energy and calculations along the centerline of the wind tunnel at a height corresponding to the center of the rib. The Reynolds number based on the wind tunnel upstream flow speed and the rib height was approximately 14000.

We observe from these results that no version of the k–ε model performed extremely well over the entire flow field despite the fact that all calculations were performed with professionally-developed commercial CFD software set up and run by professional staff from each code vendor. As can be seen from the figure, both the best and worst results were obtained by the same (RNG) version of the k–ε model, but implemented by different vendors. When it is further considered that the results presented here were ultimately computed with knowledge of the experimental data, and represent the best that could be done *circa* 1994, it is difficult to have much confidence in this modeling approach. Indeed, it appears that success of RNG k–ε depends very much on the particular implementation, but in addition no two k–ε models produced the same results. As we have already

emphasized, this poor performance is not unexpected; the Boussinseq hypothesis lacks any rigorous physical justification, and as a consequence, for anything but the simplest flows it cannot be expected to provide adequate accuracy, almost independent of the implementation. But as we have already observed, there are further (unjustified) approximations made in the derivations of both the k and ε equations, and beyond this the fundamental flaws inherent in all Reynolds-averaged approaches.

Second-moment closure (SMC) models. The poor results produced by k–ε models have long suggested attempting use of models "containing more physics." Recall that the Reynolds stresses arising in the RANS formulation are modeled entirely in terms of the Boussinesq hypothesis in k–ε models, and as we have already seen, reasons can be cited for probable failure of this for most nontrivial physical flows. Launder *et al.* [63] first proposed use of second-moment closure (also known as Reynolds stress models, RSMs) in 1975 shortly after introduction of the modern k–ε models. Such methods provide PDEs for each component of the Reynolds stress tensor rather than depending on the Boussinesq hypothesis. But as we will show, these equations contain so many unclosed terms which themselves must be modeled, often in nearly arbitrary ways (including analogs of the Boussinesq hypothesis), that it is rather doubtful much has been gained. Moreover, in 3-D calculations typically nine PDEs must be solved, compared with four for the N.–S. equations alone, and six for k–ε models. Thus, the computational expense is extremely high, approaching that of DNS for low-Reynolds number calculations. On the other hand, the total arithmetic required by SMCs is not strongly dependent on Re, so at least in principle these might be used effectively for high-Re flow calculations not currently accessible by DNS. Indeed, commercial CFD codes are beginning to contain SMC options.

Herein we will provide a simple derivation of the equations corresponding to typical SMCs, for the most part following Wilcox [98] but employing somewhat different notation more consistent with that used throughout these lectures.

The RANS equations used in conjunction with SMC models are the same as originally derived, but they are typically expressed in a form containing the temporal derivative operators. Thus, in the context of Cartesian tensor notation these are

$$\frac{\partial \overline{u}_i}{\partial t} + \overline{u}_j \frac{\partial \overline{u}_i}{\partial x_j} = -\frac{\partial \overline{p}}{\partial x_i} + \nu \frac{\partial^2 \overline{u}_i}{\partial x_j \partial x_j} - \frac{\partial \left(\overline{u'_i u'_j} \right)}{\partial x_j} , \qquad i = 1, 2, 3 , \tag{2.81}$$

along with the divergence-free condition $\partial \overline{u}_i / \partial x_i = 0$.

Now let $\mathcal{NS}(U_i)$ denote the *Navier–Stokes operator* for the i^{th} component of velocity; *i.e.*, we have

$$\mathcal{NS}(U_i) \equiv \left(\frac{\partial}{\partial t} + \boldsymbol{U} \cdot \nabla - \nu \Delta \right) U_i + \frac{\partial P}{\partial x_i} = 0 , \qquad i = 1, 2, 3 . \tag{2.82}$$

Then to derive the equation for the Reynolds stress component $\overline{u'_i u'_j}$ we form

$$\overline{u'_i \mathcal{NS}(U_j) + u'_j \mathcal{NS}(U_i)} = 0 , \qquad i, j = 1, 2, 3 , \quad i \neq j . \tag{2.83}$$

We now carry out the indicated operations in a term-by-term fashion.

We first consider the time derivative terms. These can be expressed as

$$\begin{aligned}
\overline{u_i'\frac{\partial U_j}{\partial t}+u_j'\frac{\partial U_i}{\partial t}} &= \overline{u_i'\frac{\partial}{\partial t}\left(\overline{u}_j+u_j'\right)+u_j'\frac{\partial}{\partial t}\left(\overline{u}_i+u_i'\right)} \\
&= \overline{u_i'\frac{\partial \overline{u}_j}{\partial t}}+\overline{u_i'\frac{\partial u_j'}{\partial t}}+\overline{u_j'\frac{\partial \overline{u}_i}{\partial t}}+\overline{u_j'\frac{\partial u_i'}{\partial t}} \\
&= \overline{u_i'\frac{\partial u_j'}{\partial t}}+\overline{u_j'\frac{\partial u_i'}{\partial t}} \\
&= \frac{\partial\left(\overline{u_i'u_j'}\right)}{\partial t}\equiv\frac{\partial r_{ij}}{\partial t}\,, \qquad (2.84)
\end{aligned}$$

where we have introduced the notation $r_{ij}\equiv\overline{u_i'u_j'}$. From the advective terms of the $\mathcal{NS}$ operators appearing in Eq. (2.83) we obtain

$$\begin{aligned}
\overline{u_i'U_k\frac{\partial U_j}{\partial x_k}+u_j'U_k\frac{\partial U_i}{\partial x_k}} &= \overline{u_i'\left(\overline{u}_k+u_k'\right)\frac{\partial}{\partial x_k}\left(\overline{u}_j+u_j'\right)}+\overline{u_j'\left(\overline{u}_k+u_k'\right)\frac{\partial}{\partial x_k}\left(\overline{u}_i+u_i'\right)} \\
&= \overline{u_i'u_k'\frac{\partial \overline{u}_j}{\partial x_k}}+\overline{u_i'\overline{u}_k\frac{\partial u_j'}{\partial x_k}}+\overline{u_i'u_k'\frac{\partial u_j'}{\partial x_k}}+\overline{u_j'u_k'\frac{\partial \overline{u}_i}{\partial x_k}}+\overline{u_j'\overline{u}_k\frac{\partial u_i'}{\partial x_k}}+\overline{u_j'u_k'\frac{\partial u_i'}{\partial x_k}} \\
&= \overline{u}_k\frac{\partial r_{ij}}{\partial x_k}+r_{ik}\frac{\partial \overline{u}_j}{\partial x_k}+r_{jk}\frac{\partial \overline{u}_i}{\partial x_k}+\frac{\partial}{\partial x_k}\left(\overline{u_i'u_j'u_k'}\right)\,. \qquad (2.85)
\end{aligned}$$

We note that the last term has been simplified by invoking the divergence-free condition shown to hold for fluctuating quantities in Eq. (2.5), which in the Cartesian tensor form used here is $\partial u_k'/\partial x_k=0$.

Similarly, the diffusive terms from Eq. (2.83) can be treated as follows:

$$\begin{aligned}
\nu\left(\overline{u_i'\frac{\partial^2 U_j}{\partial x_k\partial x_k}+u_j'\frac{\partial^2 U_i}{\partial x_k\partial x_k}}\right) &= \nu\overline{u_i'\frac{\partial^2\left(\overline{u}_j+u_j'\right)}{\partial x_k\partial x_k}}+\nu\overline{u_j'\frac{\partial^2\left(\overline{u}_i+u_i'\right)}{\partial x_k\partial x_k}} \\
&= \nu\overline{u_i'\frac{\partial^2 u_j'}{\partial x_k\partial x_k}}+\nu\overline{u_j'\frac{\partial^2 u_i'}{\partial x_k\partial x_k}} \\
&= \nu\frac{\partial^2 r_{ij}}{\partial x_k\partial x_k}-2\nu\overline{\frac{\partial u_i'}{\partial x_k}\frac{\partial u_j'}{\partial x_k}}\,. \qquad (2.86)
\end{aligned}$$

Finally, terms arising from the pressure gradient in the $\mathcal{NS}$ operator are

$$\begin{aligned}
\overline{u_i'\frac{\partial P}{\partial x_j}+u_j'\frac{\partial P}{\partial x_i}} &= \overline{u_i'\frac{\partial(\overline{p}+p')}{\partial x_j}+u_j'\frac{\partial(\overline{p}+p')}{\partial x_i}} \\
&= \overline{u_i'\frac{\partial p'}{\partial x_j}+u_j'\frac{\partial p'}{\partial x_i}}\,. \qquad (2.87)
\end{aligned}$$

We now collect the results contained in Eqs. (2.84)–(2.87) to complete the construction corresponding to (2.83) and thus obtain our equations for the Reynolds stresses.

$$\begin{aligned}
\frac{\partial r_{ij}}{\partial t}+\overline{u}_k\frac{\partial r_{ij}}{\partial x_k} = &- r_{ik}\frac{\partial \overline{u}_j}{\partial x_k}-r_{jk}\frac{\partial \overline{u}_i}{\partial x_k}+\nu\frac{\partial^2 r_{ij}}{\partial x_k\partial x_k} \\
&- \frac{\partial}{\partial x_k}\left(\overline{u_i'u_j'u_k'}\right)-2\nu\overline{\frac{\partial u_i'}{\partial x_k}\frac{\partial u_j'}{\partial x_k}}-\overline{u_i'\frac{\partial p'}{\partial x_j}+u_j'\frac{\partial p'}{\partial x_i}}\,. \qquad (2.88)
\end{aligned}$$

We observe that for each $i, j = 1, 2, 3$ the first line of this tensor equation could be directly solved for the r_{ij}s. This comprises a system of six independent equations coupled with Eqs. (2.81) for the mean flow quantities. All of the terms in the second line, however, must be modeled in order to accomplish this and thus close the SMC formulation.

Following [98] we first note the extent of this task. The first term in the second line of Eq. (2.88) is divergence of the velocity triple correlations $\overline{u_i' u_j' u_k'}$, which when expanded taking into account the summation convention and the i-j, *etc.* symmetries results in a total of 10 independent terms. There are six further terms from expansion of the second term on line 2 of (2.88), which we recognize as the dissipation rate tensor also defined earlier in connection with the k–ε method. Finally, there are six more terms arising in the pressure-velocity correlations, the last terms on the second line of (2.88). Thus, there are now a total of 22 additional unknowns. (The reader will recall that there were only eight such terms appearing in the k–ε formulation once the Boussinesq hypothesis was invoked.) These are all statistical quantities with little verifiable physical relevance although much effort has been devoted to attempts to ascribe such relevance to them. (See [98] for a slightly different arrangement of (2.88) which is cast in terms of such physical interpretations.) So, despite the fact that SMCs are claimed to "contain more physics" than do k–ε models, in reality if this were actually true to begin with (which requires contrived interpretations to argue), any benefits of this additional physics are likely to be lost in the modeling of these terms. In the end, what is actually accomplished is not a model containing more physics but rather one containing far more closure "constants" that can be adjusted at will in efforts to match observed flow field behaviors. But there is no significantly improved predictability.

Summary of RANS Modeling

We summarize the preceding discussions of RANS methods as follows. After introducing commonly-used terminology we provided a fairly detailed treatment of the Boussinesq hypothesis, emphasizing its fundamental shortcomings. The main ones of these are its lack of physical basis and the need to compute eddy viscosity (which is actually a tensor) in order to use it. We then discussed in considerable detail two of the main classes of RANS methods still in wide use, namely Prandtl's mixing-length theory and the k–ε models. The first of these is an algebraic model while the second is of the two-equation type. Both of these explicitly employ the Boussinesq hypothesis to convert fluctuating (nonlinear) advective terms of the N.–S. equations to mean (linear) diffusive ones. In addition, the k–ε methods require completely unsubstantiated modeling of various other statistical terms to arrive at the final forms used in computational procedures. We then presented specific results from the extant literature that demonstrate the inadequacies of the k–ε models when applied to nontrivial flows.

Following this we briefly described second-moment closures which were introduced specifically to include "more physics" in turbulence models. These approaches utilize PDEs to directly compute Reynolds stress tensor components rather than relying in the Boussinesq-hypothesis model. Unfortunately, the process of obtaining these equations leads to a total of 22 unclosed statistical terms that must be modeled. Most of these are sufficiently removed from actual physics that there is little guidance in constructing their models, and the end result is that SMCs are "models of differential equations" rather than models containing more physics. Experience with these to date has been only slightly better than that with k–ε models, but with an extremely increased arithmetic penalty. As a consequence, they are not widely used.

2.2 The Kolmogorov Theory of Turbulence

The Kolmogorov theory of turbulence published in a series of three brief papers [26] in 1941 is one of the most quoted, and yet least well understood of turbulence theories. It is not well understood,at least in part, because the original papers are difficult to read: the assumptions made are not clearly stated, and the purportedly proven results do not clearly follow from the assumptions—because they are highly nontrivial. On the other hand, these results have withstood the "test of time," and in general should be viewed as correct within the confines of the assumptions made to obtain them. The monograph by Frisch [80] and references therein provide a very good but still not easily comprehended modern treatment and interpretation of this seminal work. In the present lectures we will only briefly summarize Frisch's treatment by stating main hypotheses and results, but not attempting the detailed proofs provided in [80].

We will begin by stating Kolmogorov's "universality assumptions," and then introduce three related hypotheses employed by Frisch [80] to prove the celebrated 4/5 law. We then present the three main results provided by Kolmogorov in 1941 (known as the K41 theory), and discuss some of the consequences of these.

2.2.1 Kolmogorov's "universality" assumptions

Before presenting specific statements of these assumptions it is worthwhile to briefly consider the overall setting in which they were made. In particular, we should recall the general nature of solutions to the Navier–Stokes equations, *viz.*, that they possess certain symmetries (discussed in Chaps. 1 and 2 of [80] and in Sec. 1.4.3 of these notes) at very low Reynolds number which are successively broken as Re is increased, and which are ultimately recovered in a statistical sense as Re becomes very large and the flow becomes turbulent. Within this framework Kolmogorov's first universality assumption (also called his "first similarity hypothesis") is the following:

Kolmogorov's first universality assumption. *At very high, but not infinite, Reynolds number, all of the small-scale statistical properties are uniquely and universally determined by the length scale ℓ, the mean dissipation rate (per unit mass) ε and the viscosity ν.*

Similarly, the second universality assumption is as follows:

Kolmogorov's second universality assumption. *In the limit of infinite Reynolds number, all small-scale statistical properties are uniquely and universally determined by the length scale ℓ and the mean dissipation rate ε.*

It is worthwhile at this point, before proceeding further, to clarify terminology and notation. First, it is important to note that both of these statements concern "small-scale" statistical properties. Thus, we need first to indicate precisely what is meant by small scale. Indeed, this is not the same as the fluctuating quantities in a Reynolds decomposition (although classical turbulence theorists often equate these). In the first place, we should recall that the lowest moments of the fluctuating quantities in a Reynolds decomposition are null; for example, $\overline{u'} = 0$. In the present case we should associate small-scale quantities with the high-pass filtered part of a N.–S. solution, *i.e.*, the high-wavenumber/high-frequency components. Thus, in the context of our Hilbert-space decomposition discussed earlier, the small-scale part would be, *e.g.*,

$$u'(\boldsymbol{x},t) = \sum_{|\boldsymbol{k}|>k_c} a_{\boldsymbol{k}}(t)\varphi_{\boldsymbol{k}}(\boldsymbol{x})\,. \tag{2.89}$$

Clearly, in this case we would generally expect $\overline{u'} \neq 0$. A slightly different, but related (essentially equivalent) interpretation of small scale is to associate this with length scales that are very much smaller than the integral scale.

We can now compare details of Kolmogorov's two similarity hypotheses. The first of these concerns finite-Re behavior and as such implies that statistics will be uniquely (and universally) set by length and dissipation scales, and viscosity. The second hypothesis, on the other hand, concerns the $Re \to \infty$ limit. In this case we expect $\nu \to 0$, so the universal statistical behaviors should no longer depend on ν. We remark that it is this hypothesized universality that causes most problems (and criticisms) regarding these assumptions. It should also be mentioned that the flow properties homogeneity and (local) isotropy are generally regarded as being essential to applicability of the Kolmogorov theory, but we see neither of these explicitly required in the similarity hypotheses. We will see below, however, that these will be needed in order to prove the 4/5 law.

2.2.2 Hypotheses employed by Frisch [80]

Treatment of the Kolmogorov theory provided by Frisch is the outgrowth of many careful and thorough examinations and analyses by a number of researchers during the 1980s and early 1990s, and it has culminated in a much more easily understood representation of this theory. We will provide an abbreviated, and slightly reorganized, version of this in the present lectures. The starting point is the set of three hypotheses put forth by Frisch in [80]. These can be stated as follows.

Hypothesis 1. *In the $Re \to \infty$ limit, all possible symmetries of the N.–S. equations, usually broken by the (physical) mechanisms producing turbulence, are restored in a statistical sense at small scales and away from boundaries.*

Hypothesis 2. *Under the same assumptions as above, turbulent flow is self similar at small scales; i.e., it possesses a unique scaling exponent h such that*

$$\delta\boldsymbol{u}(\boldsymbol{x}, \lambda\boldsymbol{\ell}) = \lambda^h \delta\boldsymbol{u}(\boldsymbol{x}, \boldsymbol{\ell}) \,, \qquad \forall \ \lambda \in \mathbb{R}_+ \,, \quad \boldsymbol{x} \in \mathbb{R}^3 \,, \tag{2.90}$$

with increments $\boldsymbol{\ell}$ and $\lambda\boldsymbol{\ell}$ small compared with the integral scale.

Hypothesis 3. *Again, under the same assumptions as in Hypothesis 1, turbulent flow has a finite, nonvanishing mean rate of dissipation per unit mass, ε.*

It is important to consider some of the details and implications of these hypotheses and, further, to compare them with Kolmogorov's two original universality assumptions. We should first note, again, that small scales are associated with length scales very much smaller than the integral scale. As explained in [80], we can interpret small-scale homogeneity, for example (the requirement for which is not explicitly stated), in the context of *velocity increments*

$$\delta\boldsymbol{u}(\boldsymbol{x}, \boldsymbol{\ell}) \equiv \boldsymbol{u}(\boldsymbol{x} + \boldsymbol{\ell}) - \boldsymbol{u}(\boldsymbol{x}) \,, \tag{2.91}$$

and in particular require that statistics of these increments be invariant under arbitrary translations $\boldsymbol{r}$ for stationary flows. That is, *e.g.*,

$$\langle \delta\boldsymbol{u}(\boldsymbol{x} + \boldsymbol{r}, \boldsymbol{\ell}) \rangle = \langle \delta\boldsymbol{u}(\boldsymbol{x}, \boldsymbol{\ell}) \rangle \,,$$

where $\langle \cdot \rangle$ denotes any generic averaging procedure, as described in Chap. 1. A similar interpretation applies to isotropy of velocity increments in which invariance must hold under arbitrary rotations of $\boldsymbol{\ell}$ and $\delta\boldsymbol{u}$.

We now provide a few clarifying remarks associated with each of these hypotheses. First, taken as a whole, they are used to obtain consequences of the 2/3 law and proof of the 4/5 law. Also, we note that they are not identical with the Kolmogorov universality assumptions. In this regard, we observe (as does Frisch, himself) that Frisch does not employ the first of the Kolmogorov assumptions; in particular, all of the hypotheses used in [80] involve the $Re \to \infty$ limit. Second, all three of Frisch's

hypotheses require the assumption of recovery of N.–S. symmetries in the statistical sense, while the Kolmogorov assumptions do not explicitly state this. It is, however, needed in the proofs of the Kolmogorov results, but it is not clear that Kolmogorov actually intended this simply because experimental evidence of this property was not yet particularly strong at the time of his K41 analyses. The final aspect of Frisch's first hypothesis is the requirement of being "away from boundaries." Again, this did not explicitly appear in the Kolmogorov assumptions, and it is the first hint that homogeneity and or isotropy will be invoked—neither of these can be expected to hold in the vicinity of (solid) boundaries.

It is important to consider the details of Frisch's Hypothesis 2. It proposes self similarity on small scales and expresses this in terms of a typical power law representation. We earlier provided an heuristic description of self similarity in Def. 1.45; here, Eq. (2.90) establishes a more concrete characterization in terms of velocity increments. It is worth noting that this is a vector expression; but λ and h are scalars, and this is hypothesized to hold for all spatial increments $\boldsymbol{\ell}$ and positive multipliers λ provided these are small compared with the integral scale. We remark that this seems basically equivalent to small-scale homogeneity and isotropy assumptions, the latter due to the vector nature of Eq. (2.90). We also mention that this equation expresses a "pure math" version of homogeneity of velocity increments; *viz.*, it is precisely a statement that these are "homogeneous of degree h," a concept often encountered in analysis of differential equations—*e.g.*, dynamical systems. At the same time, the power law representation is typical of characterizations associated with fractal attractors of differential (and discrete) dynamical systems, and often phase portraits of their trajectories appear the same on all scales (see, for example, Alligood *et al.* [85]).

Finally, Hypothesis 3 coincides with experimental observations, and it also embodies important mathematical consequences. To understand the nature of these we again recall the definition of dissipation rate, given originally in Eq. (1.53):

$$\varepsilon = 2\nu\|\boldsymbol{S}\|^2 ,$$

where $\boldsymbol{S}$ is the strain rate tensor whose elements are of the form

$$S_{ij} = \frac{1}{2}\left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i}\right) , \qquad i,j = 1,2,3 .$$

Clearly, as $Re \to \infty$ as required in all of Frisch's hypotheses (and in Kolmogorov's second assumption), $\nu \to 0$. Then if ε is to remain finite, it must be the case that some (first) derivatives of $\boldsymbol{U}$ become unbounded. But we emphasize that ε is a mean dissipation rate, with the average being constructed spatially in the context of Frisch's proofs. We see from this that finite ε in the $Re \to \infty$ limit is not at all inconsistent with modern mathematical theories of the N.–S. equations, despite the the form of the definition of ε. In particular, such theories permit unboundedness of $\boldsymbol{U}$ on sets of zero measure, and an averaging process (which is basically integration) can ignore such sets without affecting the value of the average. We also emphasize that ε, itself, is just a definition, and it does not naturally appear in the N.–S. equations. In particular, within the context of the Galerkin representation of the N.–S. equations provided in Eqs. (1.42)–(1.45), viscous dissipation occurs due to linear diffusion terms of the form, *e.g.*, $|\boldsymbol{k}|^2 a_{\boldsymbol{k}}/Re$, and only from such terms. This merely implies that as $Re \to \infty$ larger values of $|\boldsymbol{k}|$ must be retained in Fourier representations of N.–S. solutions to maintain diffusive (dissipative) physical effects. Thus, finite dissipation as $Re \to \infty$ is not inconsistent with the N.–S. equations themselves despite a possible inconsistency with the definition of ε.

2.2.3 Principal results of the K41 theory

As mentioned earlier, there are three main results of the K41 theory, namely, the finite dissipation rate just discussed, the 2/3 law and the 4/5 law. The first two of these follow from dimensional

analysis and experimental observations (to set undetermined constants), with the second modified somewhat in Kolmogorov's K62 theory [27] to bring it into better agreement with more recent experimental findings. The third one, on the other hand, is a mathematically-provable outcome of the N.–S. equations. We will treat each of these here, beginning with the result concerning the dissipation rate.

Finite Dissipation Rate as $Re \to \infty$

We begin by noting that considerable discussion of this appears already in the previous subsection. Indeed, this is simply one of the three hypotheses employed by Frisch [80] to prove the 4/5 law. While our earlier discussions were of a more mathematical (but not rigorous) nature, in the present treatment we focus on physical confirmation (or, at least suggestion) that dissipation rate of turbulence kinetic energy remains finite as $Re \to \infty$, despite the fact that $\nu \to 0$. We will mainly follow the treatment provided in Chap. 5 of [80], with some minor changes.

Frisch employs the physical example of drag created by flow past a bluff object to arrive at a physical argument for finite dissipation rate as $Re \to \infty$. It is first noted that wind tunnel experiments, although differing in detail from one geometric shape to another, rather generally produce plots of drag coefficient, C_D (defined as $F_D/(\frac{1}{2}\rho U^2 A)$, where F_D is (measured) drag force, and A is area over which it acts), versus Reynolds number of the form shown in Fig. 2.4. Plots of this

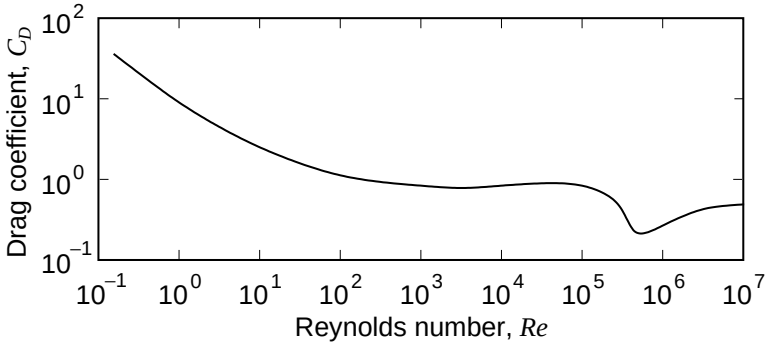

Figure 2.4: Drag coefficient *vs.* Re for flow over a circular cylinder, from experimental data referenced in Tritton [127].

nature underscore several main ideas. The first is that for low Reynolds number C_D depends fairly strongly on Re, and up to $Re \sim 10$ is approximately inversely proportional to Re. It should be noted that such flows are laminar and steady, but at least some of the broken symmetries occurring on the route to turbulence will have already taken place. Next, it should be observed that for $Re \gtrsim 10^2$ there is essentially no Re dependence of C_D until onset of the "drag crisis" which occurs for Re somewhere between 10^5 and 10^6. But with continued increases in Re beyond this, C_D is essentially constant. The reader may recall an analogous behavior represented in the Moody diagram of skin friction coefficient *vs.* Re for pipe and duct flows.

At this point it is interesting to relate the drag coefficient to the energy dissipation rate because the drag coefficient in general depends on Re, and this will lead to the result that dissipation remains nonzero as $Re \to \infty$. This will be done simply via dimensional analysis. We observe that the power $\dot{W}$ needed to move flow past an object should be related to the speed of the flow U and the size of the object, here characterized by a length scale L. In particular,

$$\dot{W} \sim W/T \sim F \cdot L/T \sim M \cdot L/T^2 \cdot L/T \sim M \cdot L^2/T^3 \,,$$

where F is force; M is mass, and T is time. We observe that $\dot{W}$ has units mass times rate of viscous dissipation and also can be related to the drag coefficient as

$$\dot{W} \sim C_D \rho L^2 U^3 \,.$$

To see this, observe that by definition

$$C_D = \frac{F_D}{\frac{1}{2}\rho U^2} \,.$$

At the same time

$$W = F_D \, L = \frac{1}{2}\rho U^2 A L C_D \,, \quad \Rightarrow \dot{W} = W/T = \frac{1}{2}\rho U^3 A C_D \sim C_D \rho L^2 U^3 \,.$$

The dissipation rate is generally written per unit mass; in particular, we recall that $\varepsilon \sim L^2/T^3$. Thus, we write

$$\varepsilon \sim \frac{\dot{W}}{\rho L^3} \sim C_D \frac{U^3}{L} \,,$$

by equating ε to power per unit mass. Now recall from Fig. 2.4 that at high Re when a flow should be turbulent, C_D is independent of Re, and hence ε is independent of viscosity for $Re \to \infty$. This provides an heuristic demonstration of the following (as appears in [80]):

Finite dissipation. *If, in an experiment on turbulent flow, all control parameters are held constant except for viscosity, which is lowered as much as possible, the energy dissipation per unit mass $de/dt \sim \varepsilon$ behaves in a way consistent with a finite positive limit.*

The 2/3 Law

We begin by stating Kolmogorov's 2/3 law as given by Frisch [80]. We follow this with a specific formula corresponding to the statement, and finally show how the well-known $k^{-5/3}$ energy spectrum follows directly from this formula. We emphasize from the start that these are mainly empirical results, and the reader is encouraged to consult Chap. 5 of [80] to see specific comparisons with experimental data. The 2/3 law can be stated as follows (Frisch [80]):

Kolmogorov's 2/3 law. *In a turbulent flow at very high Reynolds number, the mean-square velocity increment $\langle(\delta \boldsymbol{u}(\boldsymbol{\ell}))^2\rangle$ between two points separated by a distance ℓ behaves approximately as the two-thirds power of the distance.*

Several remarks are in order for interpretation of this result. First, we should observe that $(\delta \boldsymbol{u})^2$ is a (squared) magnitude of the vector quantity $\delta \boldsymbol{u}$, so one might question whether any sense can be made of this unless the turbulence is isotropic. But it turns out that this same scaling seems to hold for individual components of velocity increments, independent of isotropy (see experimental results presented in Chap. 5 of [80]). Of more importance is the length ℓ. We would expect from hypotheses discussed in preceding subsections that ℓ should be very much smaller that the integral scale. Indeed, the 2/3 law holds specifically in the inertial subrange of the energy spectrum corresponding, roughly, to Taylor microscale lengths as we have previously discussed in connection with Fig. 1.5. Finally, the "very-high Reynolds number" aspect of the statement of this law is very important. For low Re (but still high enough to be turbulent) the range of ℓ over which the 2/3 law holds is very small; but as Re increases this range increases. We again remind the reader of discussions associated with Fig. 1.5 where it was argued that this type of behavior is predicted, at least qualitatively, from the Galerkin form of the N.–S. equations.

We can now provide a formula corresponding to the 2/3 law. To do this we first recognize that $\langle(\delta\boldsymbol{u}(\boldsymbol{\ell}))^2\rangle$ is actually the second-order structure function of $\boldsymbol{u}$ defined in general in Eq. (1.22). That is,

$$\begin{aligned} S_2(\ell) &= \langle(\delta\boldsymbol{u}(\boldsymbol{\ell}))^2\rangle \\ &= \langle(\boldsymbol{u}(\boldsymbol{x}+\boldsymbol{\ell}) - \boldsymbol{u}(\boldsymbol{x}))^2\rangle\,. \end{aligned}$$

It is useful to note, as we have earlier, that in the context of velocity components S_2 is clearly related to kinetic energy, and that in general it is associated with "energy" in the sense of L^2 functions.

We next recall the meaning of the inertial subrange as given in Def. 1.77; that is, this is the range of scales over which viscosity is unimportant—viscous forces are dominated by inertial forces. In particular, this corresponds to Re sufficiently high that at low to moderate wavenumbers diffusive effects are completely dominated by advective effects in the N.–S. equations. We should recall that this would correspond to Kolmogorov's second universality assumption for which turbulence statistics depend only on the length scale ℓ and the dissipation rate ε. Now because S_2 is associated with kinetic energy, it must have general dimensions $\mathrm{L}^2/\mathrm{T}^2$. Also, we know that $\varepsilon \sim \mathrm{L}^2/\mathrm{T}^3$, and we want to find a combination of ℓ and ε that has the dimensions of energy. It is thus easily seen that the only possible combination of these variables is

$$S_2(\ell) = C\varepsilon^{2/3}\ell^{2/3}\,, \tag{2.92}$$

where C is a universal constant.

There are two main consequences that can be readily deduced from the 2/3 law. The first is the form of the turbulence kinetic energy spectrum, and a second lesser-known one is the exponent h in the Frisch's Hypothesis 2 regarding self similarity of N.–S. turbulence. We will treat each of these in the following two subsections.

The $k^{-5/3}$ energy spectrum. We can now use Eq. (2.92) to derive the widely-quoted Kolmogorov $k^{-5/3}$ inertial-range scaling of the turbulent energy spectrum. To do this we first observe that S_2 in Eq. (2.92) is expressed in physical space, and we will denote this as $\mathcal{E}$ as a reminder that it is energy. Thus, we write

$$\mathcal{E} = C\varepsilon^{2/3}\ell^{2/3}\,, \tag{2.93}$$

But what is needed is a function of the wavenumber k. Clearly, we should be able to express an incremental part of $\mathcal{E}$ in Fourier space as $d\mathcal{E} = E(k)dk$. From this it follows that the small-scale total (cumulative) energy corresponding to all wavenumbers higher than an arbitrary (but evidently chosen in the inertial subrange) one k must be

$$\mathcal{E}(k) = \int_k^\infty d\mathcal{E} = \int_k^\infty E(k)dk\,.$$

It is shown in [80], under rather general circumstances, that $E(k)$ must satisfy

$$E(k) \propto k^{-n}\,, \qquad 1 < n < 3\,, \tag{2.94}$$

and substitution of this into the integral on the right-hand side above leads to

$$\begin{aligned} \mathcal{E}(k) &\simeq \int_k^\infty k^{-n}\,dk \\ &= -\frac{1}{1-n}k^{-n+1} \\ &\simeq -\frac{1}{1-n}E(k)k\,. \end{aligned}$$

We remark that the estimate of $E(k)$ given in the relation (2.94) is consistent with velocity being at least in L^2, so it is a mathematically-reasonable assertion.

Now observe that the range of n given above guarantees that

$$C_n \equiv -\frac{1}{1-n} > 0\,,$$

and since k is arbitrary we can write $\mathcal{E}(k) = C_n E(k) k$ for any k. Finally, we note that up to a scaling constant (that depends on details of basis functions chosen for a Fourier representation) we have $k = 1/\ell$, so from Eq. (2.93) we obtain

$$C_n E(k) k = C\varepsilon^{2/3} k^{-2/3}\,,$$

and rearrangement leads to Kolmogorov's well-known result for the inertial range energy spectrum,

$$E(k) = C_K \varepsilon^{2/3} k^{-5/3}\,, \tag{2.95}$$

where C_K is the *Kolmogorov constant*. This constant must be determined from experimental measurements, and it appears not to be truly constant with observed values generally in the range from unity to two (see, *e.g.*, Chasnov [128]).

We remark, as noted in [80], that Kolmogorov did not actually present this result in his 1941 papers. Rather, it was first given by Obukhov [129] in 1941. Furthermore, because the Russian technical literature was not widely available in the West until considerably later, this same result was independently discovered by several other researchers during the mid to late 1940s.

The scaling exponent *h*. There is yet another result that can be easily deduced from the 2/3 law. Recall that in Frisch's second hypothesis it was assumed that velocity increments are self similar on small scales, and that there is a unique scaling exponent h which was not specified at that time. The 2/3 law permits us to determine the value of this exponent.

From Eq. (2.90) we have

$$\delta\boldsymbol{u}(\boldsymbol{x}, \lambda\boldsymbol{\ell}) = \lambda^h \delta\boldsymbol{u}(\boldsymbol{x}, \boldsymbol{\ell})\,,$$

which we write in terms of the second-order structure function as

$$S_2(\lambda\ell) = \langle (\delta\boldsymbol{u}(\boldsymbol{x}, \lambda\boldsymbol{\ell}))^2 \rangle = C\varepsilon^{2/3} (\lambda\ell)^{2/3}\,.$$

But by (2.90) this must be

$$\lambda^{2h} \langle (\delta\boldsymbol{u}(\ell))^2 \rangle = C\varepsilon^{2/3} \lambda^{2/3} \ell^{2/3}\,.$$

Thus, it follows that the unique scaling exponent has the value $h = 1/3$. We remark that while this derivation would appear to require isotropy, Frisch [80] notes that there is an alternative derivation that does not require this.

The 4/5 Law

The 4/5 law, although probably less quoted and possibly less directly useful than the 2/3 law, is in many respects the most important of the K41 results. This is true for two main reasons. First, it is derived directly from the N.–S. equations (albeit, with the use of a number of restrictive assumptions), and unlike the 2/3 law it contains no adjustable constants—it is the only exact result for the Navier–Stokes equations at high *Re*. Second, and equally important, it has been validated in numerous laboratory experiments. The 4/5 law proposes that third-order structure functions of velocity increments scale linearly with separation distance ℓ. More precisely, we have the following:

Kolmogorov's 4/5 law. *In the limit of infinite Reynolds number, the third-order longitudinal structure function of homogeneous isotropic turbulence, evaluated for increments ℓ small compared with the integral scale, is given in terms of the mean energy dissipation rate per unit mass as*

$$S_{3\parallel}(\ell) = -\frac{4}{5}\varepsilon\ell\,. \tag{2.96}$$

Derivation of this exact result is lengthy and nontrivial, and we will not present it here. (The interested reader should consult Chap. 6 of [80].) We simply observe that the derivation provided in [80] requires three conditions: homogeneity, isotropy and fully-developed turbulence. We have not previously emphasized the last of these, so we will provide a few additional remarks here. We first recall from Def. 1.72 that this terminology implies high-Re turbulence in which statistics are no longer changing in the flow direction. Thus, we see that within the context of the other two requirements, the only thing added by requiring fully-developed turbulence is the specific statement of high Re (as given in the 4/5 law). Hence, this has in part been implied by all of our preceding hypotheses and is basically subsumed by the homogeneity and isotropy requirements. The specific information employed by Frisch [80] to prove Eq. (2.96) is the following:

i) energy is input only on large scales;

ii) the N.–S. solutions tend to a stationary state at large times, and

iii) mean dissipation rate ε remains finite as $Re \to \infty$.

Moreover, it is emphasized that the self similarity of Frisch's Hypothesis 2 stated earlier is not used. With this combination of hypotheses and their consequences, it is possible to derive the following third-order ODE for S_3 forced by ε:

$$\left(1+\ell\frac{d}{d\ell}\right)\left(3+\ell\frac{d}{d\ell}\right)\left(5+\ell\frac{d}{d\ell}\right)\frac{S_3(\ell)}{6\pi\ell} = -12\varepsilon\,. \tag{2.97}$$

The unique, exact solution to this equation is given in Eq. (2.96) for the longitudinal case.

Higher-Order Structure Functions

Finally we note that Hypothesis 2 suggests that structure functions of arbitrary (but finite) order p should scale as

$$S_p(\ell) = C_p\varepsilon^{p/3}\ell^{p/3}\,, \tag{2.98}$$

where the C_ps are dimensionless and independent of Re. It is clear that when $p = 2$ we recover the 2/3 law with $C_2 = C_K$, and when $p = 3$ we obtain the 4/5 law with $C_3 = -4/5$. We remark that experimental evidence for the validity of Eq. (2.98) for large p is not strong. It is believed that this discrepancy is related to intermittency, effects of which have not been included in the K41 theory (but attempts at this were made in the K62 theory), and there is still considerable ongoing research related to this. We refer the reader to [80] and references therein for more details on this important and interesting topic.

2.3 Summary

This chapter has been devoted to statistical analysis and modeling of Navier–Stokes turbulence. We began with a fairly detailed treatment of the Reynolds-averaged Navier–Stokes equations. This

included analysis of the Reynolds decomposition and of the RANS equations, independent of any specific modeling procedures. The key results of this were *i*) the Reynolds decomposition imposes an extremely difficult modeling task because essentially all flow details most be modeled—using statistical quantifications; and *ii*) the form of the RANS equations is such that their solutions can never converge to time-averaged solutions to the N.–S. equations except in the case of availability of exact Reynolds stresses—which can be obtained only via DNS.

We next described three main widely-used classes of RANS models: *i*) mixing-length theory, *ii*) k–ε models and *iii*) second-moment closures. It was clear that none of these is well founded in either physics or mathematics; so the generally poor results they produce comes as no surprise, and as a consequence they cannot be used as predictive tools. On the other hand, in the presence of sufficient data required for detailed calibration of closure constants and validation of results, it is possible to use such models in an interpolatory fashion—of course, with caution. They are sometimes able to at least mimic observed physical trends, but they are essentially never able to accurately reproduce entire flow fields. This, again, is a completely expected outcome—too much information has been lost in the averaging process to be able to recover it by anything short of exact statistical correlations (the Reynolds stresses).

Finally, we presented a brief treatment of the Kolmogorov K41 theory. We emphasized that this is also a statistical treatment of the N.–S. equations, but it is intrinsically different from Reynolds averaging in that the (implied) decomposition of flow variables is a Hilbert-space decomposition. Furthermore, no modeling, *per se*, is contained in the Kolmogorov results; on the other hand, rather restrictive assumptions must be imposed to obtain them. There are three main results: *i*) finite dissipation rate as $Re \to \infty$, *ii*) the 2/3 law, and *iii*) the 4/5 law. The first two of these are mainly empirical but can be derived on the basis of dimensional analysis. The second of these leads to two further results—the $k^{-5/3}$ energy spectrum of the inertial subrange, and a value $h = 1/3$ for a universal scaling constant associated with self similarity on small scales. The last of the key results, the 4/5 law, is directly derivable from the N.–S. equations under the hypotheses of high-Re homogeneous isotropic turbulence. It is exact (no adjustable constants), nontrivial and in good agreement with experimental data. It is the only such result known up to the present time.

Chapter 3

Large-Eddy Simulation and Multi-Scale Methods

In the preceding chapter we provided an overview of RANS methods that demonstrated their inability to produce solutions to the N.–S. equations. This is obviously a very serious shortcoming of any turbulence modeling procedure, and although it has certainly been recognized for a long time by theorists, especially mathematicians, it has had little, if any, impact on engineering analyses of turbulence. On the other hand, failures of RANS methods are so widespread that engineers have sought other techniques from the earliest that computers have been available. Large-eddy simulation (LES) was the first of these, proposed by Smagorinsky [130] in 1963 and by Deardorff [10] in 1970, and actually preceded both modern k–ε models and SMCs.

In this chapter we will present a fairly thorough overview of the main aspects of LES, particularly with regard to subgrid-scale (SGS) models, and to a lesser extent the problem of filtering the governing equations. But in addition we will propose viewing LES as a subset of the much wider class of multi-scale methods (see *e.g.*, E and Engquist [131], and references therein) and show that it, in a sense, represents a link between classical turbulence modeling procedures and new approaches (such as DNS) that emphasize the dynamical systems aspects of the N.–S. equations. We will establish this link by first providing a fairly detailed treatment of LES, then presenting details of the dynamical systems view of the N.–S. equations, *á la* Ruelle and Takens [8], and finally showing how this relates to multi-scale methods in general and LES in particular. We will then conclude the chapter with a somewhat altered form of LES that explicitly utilizes this dynamical systems viewpoint for construction of SGS models and at the same time employs multi-scale ideas to combine the large and small scales.

3.1 Large-Eddy Simulation

In Chap. 1 we briefly discussed large-eddy simulation decomposition of dependent variables and further described LES as a turbulence computation method lying somewhere between RANS and DNS. But these discussions were mainly qualitative; here we will provide considerably more detail. We begin with a more thorough discussion of just where LES lies on the spectrum of turbulence models. Following this we review material presented earlier on the LES decomposition, and we carry out the formal procedure for applying LES to the N.–S. equations, just as we have previously done in the RANS case. This process will include two specific aspects of crucial importance to LES: filtering the governing equations, and construction of SGS models. It is the specific approach used for each of these that allows distinctions to be made amongst the various forms of LES, and we will consider several of the currently most widely used of these.

3.1.1 Comparison of DNS, LES and RANS methods

We have already emphasized in Chap. 1 that in the context of turbulent flow simulations DNS requires no modeling, and RANS approaches entail extensive modeling, as we have seen in Chap. 2. The former simply exploits numerical analysis to construct an efficient implementation of an appropriate N.–S. solution procedure. No physical assumptions need be made beyond those already embodied in the N.–S. equations themselves; but because of the extremely wide range of length and time scales, reflected in the magnitude of the Reynolds number, it is not yet possible to perform such calculations if Re is at all large. The reader will recall in Chap. 1 we argued that the total required arithmetic scales as Re^3 for DNS calculations, and this is an optimistic estimate. It is often pointed out that at the current rate of improvement in computing hardware, this translates to being able to double the size of Re in a calculation only once in a five to ten year period. Clearly, this is completely unacceptable from a practical standpoint as is especially obvious when it is considered that $Re \sim \mathcal{O}(10^4)$ is the current state of the art.

Until recently, the only alternatives have been RANS methods. In these required arithmetic is nearly independent of Re, and in fact many such methods work best at very high Re where scale separation is more likely to be a reasonable approximation. But as we have seen, no RANS approach can be predictive due to the extensive modeling requirements; so this too is unacceptable.

The same type of analysis leading to the Re^3 arithmetic requirement for DNS shows that total arithmetic required for LES should not exceed Re^2. In particular, the goal in LES is to directly simulate only into the inertial subrange, and not all the way to the dissipation scales as must be done in DNS. At the same time we see that the LES models need represent only the high-wavenumber part of the inertial subrange, and the dissipation scales—in sharp contrast with RANS models which are formally required to model everything from (at least) the integral scales through the dissipation range.

Furthermore, we know the Taylor microscales lie fairly deep within the inertial subrange, and it was shown in Chap. 1 that $\ell/\lambda \sim Re_\ell^{1/2}$, where ℓ is the integral scale length, and λ is the Taylor microscale length. Thus, in 3D the ratio of largest to smallest scales that must be resolved is proportional to $Re_\ell^{3/2}$, and as we did in the context of DNS estimates we assume the numerical time step sizes should satisfy a Courant condition; so the total arithmetic is $\sim Re^2$ for LES provide SGS model evaluation is independent of Re, which usually is the case. Unlike the DNS case for which Re^3 is actually somewhat optimistic, the Re^2 estimate for LES is likely to be somewhat conservative, and one sometimes finds predictions of total arithmetic as low as $Re^{3/2}$. But even the Re^2 estimate is beginning to be an acceptable amount of arithmetic on modern computing machinery, and LES is now an available option in many commercial CFD software suites intended for practical engineering applications.

We can summarize these comparisons as follows:

i) DNS requires no modeling, but it demands resolution from the large scales all the way through at least the beginning of the dissipation scales. This results in total arithmetic scaling at least as Re^3, or worse.

ii) LES requires modeling of part of the inertial subrange and into the beginning of the dissipation scales. The amount of required modeling is set by the amount of resolution that can be afforded, but it is unlikely that total arithmetic will scale worse than Re^2.

iii) RANS requires modeling of everything from the integral scales into the dissipation range—only mean (zeroth-mode) quantities are directly computed. As a consequence, total arithmetic is at most a weak function of Re.

In addition, we remark here (and we will demonstrate this in detail later) that in contrast to RANS, it can be shown that LES procedures generally converge to DNS (and thus their solutions can be expected to converge to N.–S. solutions) as discretization step sizes (and filter widths) are refined.

Finally, before proceeding to details, it is worthwhile to mention that just as is true for RANS methods, there are by now many different forms of LES. We will not here attempt a detailed treatment of all of these but instead concentrate on a few of the better-known and more widely-used approaches. The oldest, but yet still widely used, is the Smagorinsky model [130], with "model" referring to treatment of the unresolved stresses. We will later see that this approach in not very different from mixing-length models used in RANS methods, and this is not surprising since early work with LES was actually done in the shadow of RANS theory. Like mixing-length models, the Smagorinsky model works well only for a very limited number of flow situations, and the first attempts to correct this resulted in what are now termed "scale-similarity" models, first proposed by Bardina *et al.* [132]. These were not successful in general, and the next "bandaid" consisted of the so-called "dynamic" models introduced by Germano *et al.* [68]. Both of these approaches tend to be too little dissipative (although this can be corrected in the latter), in contrast to the Smagorinsky model which is essentially always overly dissipative. The natural response to this conundrum was to effectively "average" scale-similarity and more dissipative models such as Smagorinsky to produce "mixed" models. All of these methods have been constructed via the "classical" approach to LES consisting of the following steps:

i) decompose flow variables into large- and small-scale parts, with the large-scale part purportedly defined by a filtering process;

ii) filter the governing equations, and substitute the decomposition from part *i*) into the nonlinear terms to construct the unclosed terms to be modeled and obtain a system of equations for resolved-scale variables that is as close as possible to the N.–S. equations;

iii) model the unresolved stresses;

iv) solve equations for the large-scale contribution (while essentially ignoring the small-scale part).

With the exception of the last, we will treat these steps in detail in the sequel; but before doing this we wish to emphasize that there are beginning to appear forms of LES that do not follow this standard formulation, and we will mention a few of these at this time and provide further details in the section devoted to multi-scale methods.

There are many different approaches being attempted, ranging from simply solving the governing equations with numerical methods that are strongly dissipative and thus replacing physical (actually, model) dissipation with numerical dissipation, to quite sophisticated procedures which attempt to directly model the subgrid-scale variables. The former approaches are usually termed "implicit" LES (ILES), due to Fureby and Grinstein [133], an example of which is monotonically integrated LES (MILES), as first proposed by Boris *et al.* [134], while the latter comprise a wide class of methods reviewed in some detail by Sagaut [73] and often called "synthetic velocity" approaches. Specific ones of these include the linear-eddy models (LEMs) and one-dimensional turbulence (ODT) models of Kerstein and coworkers (see, *e.g.*, Kerstein [135] and Echekki [71], respectively), the "estimation" models of Domaradzki and coworkers (*e.g.*, Domaradzki and Saiki [70]) and the chaotic-map models of McDonough and coworkers (*e.g.*, Hylin and McDonough [136], and the summary in [73]).

3.1.2 The LES decomposition

We earlier presented the LES decomposition in Eq. (1.27) for a generic scalar variable. Here, we consider this in more detail for the velocity vector $\boldsymbol{U} = (u, v, w)^T$ and begin by writing

$$\boldsymbol{U}(\boldsymbol{x}, t) = \widetilde{\boldsymbol{u}}(\boldsymbol{x}, t) + \boldsymbol{u}'(\boldsymbol{x}, t)\,. \tag{3.1}$$

We remind the reader of two key points associated with this form of decomposition. The first is that '$\widetilde{\ }$' represents a spatial filter that should be considered a low-pass filter which, in principle, removes all wavenumbers in the Fourier representation of $\boldsymbol{U}$ above those supported by the chosen discretization of the governing equations. Formally, we have

$$\widetilde{\boldsymbol{U}}(\boldsymbol{x}, t) = \int_{\Omega_i} G(\boldsymbol{x}|\boldsymbol{\xi})\, \boldsymbol{U}(\boldsymbol{\xi}, t)\, d\boldsymbol{\xi} \equiv \widetilde{\boldsymbol{u}}(\boldsymbol{x}, t)\,, \tag{3.2}$$

where the *filter kernel* G is often taken to be a Gaussian, and Ω_i is a subdomain of the solution domain Ω such that the volume of Ω_i is approximately h^3 with h being the discrete step size of the numerical approximation. We remark that there is an alternative interpretation of this. Namely, we might instead take the domain of integration to be the entire problem solution domain Ω and then demand that G have compact support over just a few multiples of h (in each direction). Clearly, this is equivalent to our notation. In addition, it is probably more precise to measure this support in terms of the "filter width," usually denoted Δ, which itself is a few multiples of h in length.

The second point, which follows from the first, is that $\boldsymbol{u}'(\boldsymbol{x}, t)$ is a very different mathematical object from that possessing the same notation in RANS formalisms. Indeed, we see that $\boldsymbol{u}'$ is the high-pass filtered part (or, remainder in the Fourier series context) of the solution $\boldsymbol{U}(\boldsymbol{x}, t)$ after the low-pass filtering of Eq. (3.2) has been applied. That is,

$$\boldsymbol{u}'(\boldsymbol{x}, t) = \boldsymbol{U}(\boldsymbol{x}, t) - \widetilde{\boldsymbol{u}}(\boldsymbol{x}, t)\,, \tag{3.3}$$

which formally has exactly the same appearance as the RANS result. But due to the nature of the LES decomposition the are obvious consequences of this formalism that differ from the RANS case, namely,

$$\widetilde{\widetilde{\boldsymbol{u}}}(\boldsymbol{x}, t) \neq \widetilde{\boldsymbol{u}}(\boldsymbol{x}, t) \qquad \text{and} \qquad \widetilde{\boldsymbol{u}'}\,(\boldsymbol{x}, t) \neq 0\,, \tag{3.4}$$

except in the special case of a projective filter. In particular, from Eq. (3.3) we have

$$\begin{aligned} \widetilde{\boldsymbol{u}'}\,(\boldsymbol{x}, t) &= \widetilde{\boldsymbol{U}}(\boldsymbol{x}, t) - \widetilde{\widetilde{\boldsymbol{u}}}(\boldsymbol{x}, t) \\ &= \widetilde{\boldsymbol{u}}(\boldsymbol{x}, t) - \widetilde{\widetilde{\boldsymbol{u}}}(\boldsymbol{x}, t) \\ &\neq 0\,, \end{aligned}$$

since in general,

$$\widetilde{\widetilde{\boldsymbol{u}}}(\boldsymbol{x}, t) = \int_{\Omega_i} G(\boldsymbol{x}|\boldsymbol{\xi})\, \widetilde{\boldsymbol{u}}(\boldsymbol{\xi}, t)\, d\boldsymbol{\xi} \neq \int_{\Omega_i} G(\boldsymbol{x}|\boldsymbol{\xi})\, \boldsymbol{U}(\boldsymbol{\xi}, t)\, d\boldsymbol{\xi} = \widetilde{\boldsymbol{u}}(\boldsymbol{x}, t)\,,$$

for typical (non-projective) filter kernels. We remark that a specific exception to this is the sharp-cutoff Fourier-space filter which is projective.

As we have previously noted, the form of Eq. (3.1) is suggestive of a Hilbert-space decomposition except, as already indicated, the subspaces are not perfectly orthogonal. Despite this technicality, it is useful to identify (3.1) with the Fourier representation

$$\boldsymbol{U}(\boldsymbol{x}, t) = \sum_{|\boldsymbol{k}|\geq 0}^{\boldsymbol{k}_c} \boldsymbol{a}_{\boldsymbol{k}}(t)\boldsymbol{\varphi}_{\boldsymbol{k}}(\boldsymbol{x}) \;+ \sum_{|\boldsymbol{k}|=\boldsymbol{k}_c+1}^{\infty} \boldsymbol{a}_{\boldsymbol{k}}(t)\boldsymbol{\varphi}_{\boldsymbol{k}}(\boldsymbol{x})\,. \tag{3.5}$$

In this manner it is easy to provide an heuristic (but not rigorous) argument for convergence of LES solutions to solutions to the N.–S. equations. In particular, in contrast to a RANS decomposition in which significant information may be lost in a nonrecoverable way due to the simple averaging procedure employed, Eq. (3.5) representing LES shows that as discrete resolution is increased ($\boldsymbol{k}_c \to \infty$), the computed result should converge to the true function being represented (assuming, of course, that the function is at least in $L^2(\Omega)$).

3.1.3 Derivation of the LES filtered equations

In this section we will derive the filtered equations employed in the LES formalism. We begin with conservation of mass, which in the incompressible case treated here is extremely simple, and we then proceed to the more tedious momentum equations; *viz.*, we follow the same process introduced in treatment of the RANS equations in the preceding chapter.

The Continuity Equation

Conservation of mass, or the continuity equation, for an incompressible flow is just the divergence-free condition $\nabla\cdot\boldsymbol{U} = 0$, or

$$u_x + v_y + w_z = 0 \tag{3.6}$$

in 3D. Formal application of the filter results in

$$\widetilde{u_x} + \widetilde{v_y} + \widetilde{w_z} = 0\,.$$

We see that to obtain the desired form will require commutativity of the spatial filter and spatial differentiation. For example, the first term in the above is of the form

$$\int_{\Omega_i} G(\boldsymbol{x}|\boldsymbol{\xi})\,\frac{\partial u}{\partial \xi}(\boldsymbol{\xi},t)\,d\boldsymbol{\xi}\,.$$

It should be clear that commutativity will not hold for all possible kernels G, but the typical ones used for LES filters usually possess arguments of the form $\boldsymbol{x}-\boldsymbol{\xi}$, and they generally have (nearly) compact support on Ω. In this case it can be shown that $\widetilde{u_x} = \widetilde{u}_x$, *etc.*, at least in Cartesian coordinate systems. We leave demonstration of this for the reader. Thus, Eq. (3.6) can be replaced with

$$\widetilde{u}_x + \widetilde{v}_y + \widetilde{w}_z = 0\,, \tag{3.7}$$

or $\nabla\cdot\widetilde{\boldsymbol{u}} = 0$.

At this point it is worthwhile to again recall that $\widetilde{\boldsymbol{u}}$ is a low-pass filtered quantity associated with the first few terms of a Fourier representation of $\boldsymbol{U}$ as given in Eq. (3.5). Thus, $\nabla\cdot\widetilde{\boldsymbol{U}} = \nabla\cdot\widetilde{\boldsymbol{u}}$ by definition. Then it follows that

$$0 = \nabla\cdot\boldsymbol{U} = \nabla\cdot(\widetilde{\boldsymbol{u}} + \boldsymbol{u}')\ .$$

But $\nabla\cdot\widetilde{\boldsymbol{u}} = 0$ also holds as seen from Eq. (3.7). Hence, just as was the case for the Reynolds decomposition, we see that

$$\nabla\cdot\boldsymbol{u}' = 0\,. \tag{3.8}$$

We again recall from our Galerkin representation of the Navier–Stokes equations in Chap. 1 that in Fourier space this must hold on a mode-by-mode basis; but we see that even in the absence of a projective filter the divergence-free condition must hold separately on large and small scales, and in synthetic-velocity SGS models this requirement should be explicitly enforced.

The Momentum Equations

From the preceding straightforward exercise in filtering the continuity equation we can easily deduce the filtered form of the momentum equations. These can be expressed as

$$\frac{\partial \widetilde{\boldsymbol{u}}}{\partial t} + \nabla\cdot\left(\widetilde{\boldsymbol{U}\boldsymbol{U}}\right) = -\nabla\widetilde{p} + \nu\Delta\widetilde{\boldsymbol{u}}\,. \tag{3.9}$$

Obtaining the first term on the left-hand side involves a trivial (assuming time-independent coordinates) commutativity of the filter with the temporal derivative, while both terms on the right-hand side follow in the same way as in the continuity equation. Thus, just as was true with Reynolds averaging, filtering of the momentum equations leads to significant difficulty only in the nonlinear terms (provided simple filters with properties as we have been supposing here are employed). As will be clear as we proceed, the difficulties associated with the nonlinear terms are similar to (but more complicated than) those arising in the RANS case, and in generalized coordinates they occur for linear terms as well.

For the nonlinear terms we introduce the LES decomposition Eq. (3.1) and write

$$\nabla\cdot\left(\widetilde{\boldsymbol{U}\boldsymbol{U}}\right) = \nabla\cdot\left(\widetilde{(\widetilde{\boldsymbol{u}}+\boldsymbol{u}')\,(\widetilde{\boldsymbol{u}}+\boldsymbol{u}')}\right)\,.$$

$$\widetilde{(\widetilde{\boldsymbol{u}}+\boldsymbol{u}')\,(\widetilde{\boldsymbol{u}}+\boldsymbol{u}')} = \begin{pmatrix} \widetilde{(\widetilde{u}+u')\,(\widetilde{u}+u')} & \widetilde{(\widetilde{u}+u')\,(\widetilde{v}+v')} & \widetilde{(\widetilde{u}+u')\,(\widetilde{w}+w')} \\ \widetilde{(\widetilde{u}+u')\,(\widetilde{v}+v')} & \widetilde{(\widetilde{v}+v')\,(\widetilde{v}+v')} & \widetilde{(\widetilde{v}+v')\,(\widetilde{w}+w')} \\ \widetilde{(\widetilde{u}+u')\,(\widetilde{w}+w')} & \widetilde{(\widetilde{v}+v')\,(\widetilde{w}+w')} & \widetilde{(\widetilde{w}+w')\,(\widetilde{w}+w')} \end{pmatrix}. \tag{3.10}$$

We examine one component of this tensor in detail to deduce the structure induced by decomposition and filtering. Consider

$$\widetilde{(\widetilde{u}+u')\,(\widetilde{v}+v')} = \widetilde{\widetilde{u}\widetilde{v}} + \widetilde{\widetilde{u}v'} + \widetilde{\widetilde{v}u'} + \widetilde{u'v'}\,. \tag{3.11}$$

We observe that there are three specific types of terms in this expression, and in all of the analogous ones corresponding to the remaining components of Eq. (3.10). The first of these, $\widetilde{\widetilde{u}\widetilde{v}}$ (a part of what is known as the *Leonard stress*), would have simply been of the form $\overline{u}\,\overline{v}$ in a Reynolds-averaging procedure. On the other hand, despite its greater complexity, $\widetilde{\widetilde{u}\widetilde{v}}$ still can be computed directly (without any modeling required) since we have equations for $\widetilde{u}$ and $\widetilde{v}$.

The second pair of terms, $\widetilde{\widetilde{u}v'} + \widetilde{\widetilde{v}u'}$, is known as the *cross stress*, and this would be identically zero in a Reynolds-averaging formalism. In the present context, however, these terms are not only nonzero, but they must be modeled because they contain small-scale factors.

Finally, terms of the form $\widetilde{u'v'}$ are analogous to components of the Reynolds stress tensor that occurs in RANS procedures, and they are usually called *Reynolds stresses* even in the LES context. In usual treatments of LES, the terms on the right-hand side of Eq. (3.11) are expressed in Cartesian tensor notation as given here parenthetically:

$$L_{ij} \equiv \widetilde{\widetilde{u}\widetilde{v}} - \widetilde{u}\widetilde{v} \qquad \left(= \widetilde{\widetilde{u}_i\widetilde{u}_j} - \widetilde{u}_i\widetilde{u}_j\right)\,, \qquad \text{(Leonard stress)}$$

$$C_{ij} \equiv \widetilde{\widetilde{u}_i u'_j} + \widetilde{\widetilde{u}_i u'_j} \quad \left(= \widetilde{\widetilde{u}_i u'_j} + \widetilde{\widetilde{u}_i u'_j}\right)\,, \qquad \text{(cross stress)}$$

$$R_{ij} \equiv \widetilde{u'v'} \qquad \left(= \widetilde{u'_i u'_j}\right)\,, \qquad \text{(Reynolds stress)}$$

where the origin of the second term in the Leonard stress will be made clear below.

It is important to recognize that the last of these terms are fundamentally different from the Reynolds stress components arising in RANS formalisms, despite their common name. In particular we should recall that the individual fluctuating quantities appearing in the usual Reynolds stress tensor contain information from all but the zeroth mode of their Fourier representations (recall Eq. (2.23)), while the subgrid part of a LES representation is associated with a high-pass filtering of the solution, thus carrying information only from the modes above some cut-off wavenumber k_c corresponding to what can be supported by the discretization employed in treating the filtered, large-scale equations. Furthermore, it should again be observed that, for example, $\overline{u'} = 0$ in RANS formulations; but $\widetilde{u'} \neq 0$ in the LES case, as we have already stressed. In short, at least from a fundamental mathematical viewpoint we must expect that $\widetilde{u'v'}$ and $\overline{u'v'}$ are quite different, and in particular we should not expect $\widetilde{u'v'} = \overline{u'v'}$ in general, although they have been given the same nomenclature and notation. On the other hand, due to presence of the Leonard stress, the overall SGS stresses may not differ significantly from the corresponding Reynolds stresses of RANS approaches aside from their time dependence.

We remark that in early treatments of SGS modeling each of the contributions listed above was modeled separately; since they correspond to different parts of the energy spectrum one could argue that this should be an advantage. On the other hand, it is shown by Sagaut [73] that for typical simple filters (but not for all filters) the invariance properties of the N.–S. equations are preserved under filtering, but in contrast, the individual parts of the LES decomposition do not all separately preserve all invariances. In particular, neither Leonard nor cross stresses are Galilean invariant, but their sum is; hence, the complete SGS stress is Galilean invariant. This, however, suggests that at least from the standpoint of maintaining N.–S. invariances, it is probably best to model SGS stress as a single entity. This is now the usual practice. Beyond this is the additional fact that unless explicit time integrations are employed for the advective terms, an iteration process is required for construction of Leonard and cross stresses at each time step, although this might be incorporated in the nonlinear iterations of an implicit scheme in any case.

We also note that more elaborate filters can result in loss of invariance of the filtered equations themselves. Specific examples include those with different filter lengths in different directions and those having variable filter lengths in any, or all, direction(s). Such difficulties suggest that alternatives to filtering the governing equations should possibly be sought.

We now observe that the filtered momentum equations still are not yet in a useful form. To obtain such a form we recognize that what is needed for the nonlinear term on the left-hand side of Eq. (3.9) is $\nabla\cdot(\widetilde{\boldsymbol{u}}\widetilde{\boldsymbol{u}})$. But we have seen that this form does not occur in the analysis of the filtered nonlinear term (except in the Leonard stress where it was apparently added artificially). Thus, the only choice for obtaining what is needed is to add and subtract this term from the left-hand side of Eq. (3.9). Once this is done, we replace the filtered nonlinear terms on the left-hand side of Eq. (3.10) with

$$\widetilde{\boldsymbol{u}}\widetilde{\boldsymbol{u}} + \left[\widetilde{(\widetilde{\boldsymbol{u}} + u')\,(\widetilde{\boldsymbol{u}} + u')} - \widetilde{\boldsymbol{u}}\widetilde{\boldsymbol{u}}\right] \equiv \widetilde{\boldsymbol{u}}\widetilde{\boldsymbol{u}} + \boldsymbol{\tau}_{_{SGS}}\,,$$

where $\boldsymbol{\tau}_{_{SGS}}$ is notation for the LES subgrid-scale stress. We now can write Eq. (3.9) as

$$\widetilde{\boldsymbol{u}}_t + \nabla\cdot(\widetilde{\boldsymbol{u}}\widetilde{\boldsymbol{u}}) = -\nabla\widetilde{p} + \nu\Delta\widetilde{\boldsymbol{u}} - \nabla\cdot\boldsymbol{\tau}_{_{SGS}}\,, \tag{3.12}$$

with

$$\boldsymbol{\tau}_{_{SGS,ij}} \equiv L_{ij} + C_{ij} + R_{ij}\,. \tag{3.13}$$

This is the equation for the large- (resolved-) scale part of a LES formulation. We observe that only $\boldsymbol{\tau}_{_{SGS}}$ needs to be modeled, and in that sense Eq. (3.12) takes on the same appearance as would a

time-dependent RANS method. But in viewing (3.12) in this context we emphasize that presence of the time-derivative term in Eq. (3.12) is rigorously correct—in contrast to the RANS case. Namely, the LES decomposition is such as to retain time dependence in both large and small scales because temporal averaging has been replaced with spatial filtering in the construction process. We will also see in the next section that models of $\boldsymbol{\tau}_{SGS}$ have, from the beginning, usually been constructed such that $\boldsymbol{\tau}_{SGS} \to 0$ as $h \to 0$ (or $k_c \to \infty$), where h is a measure of discretization step sizes. Thus, it is clear that Eq. (3.12) converges to the N.–S. equations in this limit for such models, and as we have previously noted, LES $\to$ DNS. We again emphasize that such features are not part of RANS modeling. Indeed, the mathematical nature of the Reynolds decomposition precludes this.

Finally, we note that $-\nabla\cdot\boldsymbol{\tau}_{SGS}$ is typically modeled in a way reminiscent of the Boussinesq hypothesis, so one might argue that all that is actually accomplished with it is stabilization of the numerical algorithms being employed to solve an under-resolved discretization of Eq. (3.12)—*i.e.*, increased (numerical) dissipation. In fact, from a mathematical viewpoint, the only reason for filtering is to eliminate aliasing arising from under resolution imposed by coarse grids of practical discretizations. While it should be clear that filtering $\boldsymbol{U}$, itself, might be more appropriate in this context, presence of the term $-\nabla\cdot\boldsymbol{\tau}_{SGS}$ is analogous to the artificial dissipation schemes widely employed for shock capturing in compressible flow simulations (see, *e.g.*, Hirsch [137]). It is recognition of this that suggests use of ILES-like procedures, but in light of this we expect such methods to perform poorly in situations for which the Boussinesq hypothesis is inadequate, and inparticular, when scale separation does not hold.

3.1.4 Subgrid-scale models for LES

As we have already suggested, there are many different forms of LES models, and we shall not attempt to treat more of than a few of the better known of these in the present lectures. In the first subsection we will introduce the basic Smagorinsky model [130] and discuss some of its advantages and disadvantages. This will motivate considering two other types of models, the dynamic models and mixed models, in subsequent subsections.

The Smagorinsky model

The Smagorinsky model is the oldest of LES SGS models, but because of its simplicity it is still widely used. It is not a particularly good choice for wall-bounded shear flows, but for flows far from solid boundaries it can be quite adequate, especially if the large scale is well resolved with the cut-off wavenumber lying fairly deep within the inertial subrange.

This model is based on the Boussinesq hypothesis employed extensively in RANS models and takes the form

$$\boldsymbol{\tau}_{SGS} = -2\nu_{SGS}\widetilde{\boldsymbol{S}}\,, \tag{3.14}$$

where $\widetilde{\boldsymbol{S}}$ is the usual large-scale strain-rate tensor, and ν_{SGS} is the subgrid-scale eddy viscosity. In Cartesian tensor index notation this is

$$\tau_{SGS,ij} = -2\nu_{SGS}\widetilde{S}_{ij}\,, \qquad i,j = 1,2,3, \quad i \neq j\,.$$

The eddy viscosity is calculated using a formula analogous to the mixing-length formulation of RANS methods:

$$\nu_{SGS} = (C_S\Delta)^2|\widetilde{\boldsymbol{S}}|\,. \tag{3.15}$$

Here, Δ is the *filter width* (proportional to grid spacing), and C_S is the Smagorinsky "constant," typical values of which are $\mathcal{O}(10^{-1})$; these can be estimated theoretically for isotropic turbulence (see [73] and references therein).

We see from this that the term requiring modeling in Eq. (3.12) is now completely specified, but we should also expect from the simplicity of this model that it is not likely to be very accurate except in the context of homogeneous turbulence. (Observe that the unjustifiable tacit assumption "turbulence stress proportional to mean—filtered in this case—strain" is still being used.) Indeed, this is generally the case; as a consequence, early LESs were often performed with resolutions nearly as fine as employed for DNS. In such cases Δ^2 becomes very small, and contributions from the model are rather minimal. It is also easily seen from Eqs. (3.14) and (3.15) that the Smagorinsky model is completely dissipative, and it is due to this that it performs poorly for wall-bounded flows. In particular, it is now known that for such flows as much as one third of the turbulence kinetic energy cascaded to the small scales returns to the large scales without being dissipated. This is termed "backscatter" (recall Def. 1.55), and the basic Smagorinsky model is unable to produce such behavior.

Dynamic models

Subgrid-scale models from the class now termed *dynamic models* were proposed by Germano *et al.* [68] in 1991 as a modification to the basic Smagorinsky model, and constructed so as to permit determination of the Smagorinsky "constant" as a function of both space and time. That is, $C_S = C_S(\boldsymbol{x}, t)$. To accomplish this, additional information is needed, and Germano *et al.* [68] obtain this by means of a second filtering operation. This second filter is usually termed a *test filter* in the present context. In principle, this permits identification of the fluctuating part of the resolved scale, and this is then used (somewhat indirectly in the present case) to obtain an estimate of the unresolved stresses. A scale-similarity hypothesis (see Def. 1.103) is then invoked, leading to a formula for $C_S(\boldsymbol{x}, t)$. Details of constructing dynamic models are provided in [73], and we will follow this reference in the present discussions.

It is worthwhile to first consider the effects of the test filter by considering the location in the energy spectrum of the various filtered quantities. This is depicted in Fig. 3.1. As pointed out in

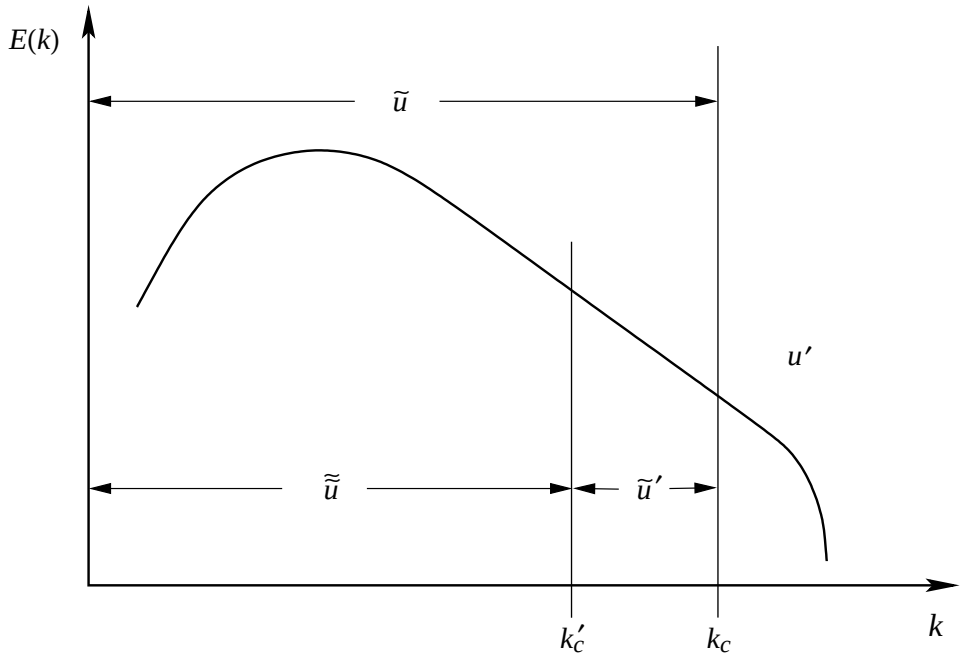

Figure 3.1: Energy spectrum showing cut-off wavenumbers for filtered (k_c) and test-filtered (k_c') quantities.

[73] and elsewhere, it is typical for the filter width of the test filter to be double that of the basic filter, so as indicated in the figure, the cut-off wavenumber corresponding to the test filter is lower than that of the original filter. This permits us to define

$$\widetilde{\boldsymbol{u}}' = \widetilde{\boldsymbol{u}} - \widetilde{\widetilde{\boldsymbol{u}}} \tag{3.16}$$

just as we defined $\boldsymbol{u}' = \boldsymbol{U} - \widetilde{\boldsymbol{u}}$. Then invoking the scale-similarity hypothesis permits us to estimate statistics of $\boldsymbol{u}'$ (which are unknown) in terms of those of $\widetilde{\boldsymbol{u}}'$ (which are known). We remark, however, that while this is what underlies the Germano *et al.* [68] dynamic model, the details are rather different. We outline these here, following Sagaut [73], and return to further direct application of scale similarity in the next subsection.

The starting point for development of the dynamic model is a special case of the *Germano identity* [138], relating the Leonard stress to a similar twice-filtered tensor $\boldsymbol{\mathcal{T}}$ and a second filtering of the usual SGS tensor. In particular, we have

$$L_{ij} = \mathcal{T}_{ij} - \widetilde{\tau}_{SGS,ij}\,, \tag{3.17}$$

where analogous to Eq. (3.11) and following,

$$\begin{aligned}
\tau_{SGS,ij} &= L_{ij} + C_{ij} + R_{ij} = \widetilde{U_iU_j} - \widetilde{u}_i\widetilde{u}_j\,,\\
\mathcal{T}_{ij} &= \widetilde{\widetilde{U_iU_j}} - \widetilde{\widetilde{u}}_i\widetilde{\widetilde{u}}_j\,,\\
L_{ij} &= \widetilde{\widetilde{u}_i\widetilde{u}_j} - \widetilde{\widetilde{u}}_i\widetilde{\widetilde{u}}_j\,.
\end{aligned}$$

Proof of (3.17) is a direct calculation which we leave to the reader, and we omit it here. We note, however, the slightly altered form of L_{ij} from that given earlier. This is a consequence of formally seeking equations for the twice-filtered dependent variables.

We remark that in the context in which many of these quantities were originally introduced the two different filterings necessarily involved identical filters, but here the filters will be different. In particular, as already indicated the second, or test, filter will have filter width usually double that of the original filter. (Hence, the above formulas lead to nontrivial results even for projective filters.)

We next introduce the so-called deviatoric parts of the tensors $\tau_{SGS,ij}$ and $\mathcal{T}_{ij}$, and at the same time assume these can be modeled with analogous formulas employing the same closure constant. Then we can write

$$\begin{aligned}
\tau^d_{SGS,ij} &\equiv \tau_{SGS,ij} - \frac{1}{3}\tau_{SGS,kk}\delta_{ij} = C^d_S\beta_{ij}\,,\\
\mathcal{T}^d_{ij} &\equiv \mathcal{T}_{ij} - \frac{1}{3}\mathcal{T}_{kk}\delta_{ij} = C^d_S\alpha_{ij}\,.
\end{aligned}$$

We recall that $\tau^d_{SGS,ij}$ is simply $\tau_{SGS,ij}$ with zero trace, as the Cartesian tensor operations indicate. It should also be observed that employing the same value of C^d_S for both filter scales is analogous to invoking a scale-similarity hypothesis.

It should be observed that α_{ij} and β_{ij} are the SGS models with their closure constants removed.

$$\beta_{ij} = \Delta^2\widetilde{S}_{ij}|\widetilde{\boldsymbol{S}}|\,,$$

and

$$\alpha_{ij} = \Delta^2_{test}\widetilde{\widetilde{S}}_{ij}|\widetilde{\widetilde{\boldsymbol{S}}}|\,.$$

In the context of dynamic models the Smagorinsky model is most often, but not always, used. The reader should consult [73] to see other alternatives. We also note that with respect to the Smagorinsky model C_S^d is actually $(C_S^d)^2$.

Analogous to the above relations we define the deviatoric part of the Leonard stress tensor and use the above in the corresponding Germano identity:

$$L_{ij}^d \equiv L_{ij} - \frac{1}{3}L_{kk}\delta_{ij} = C_S^d\alpha_{ij} - \widetilde{C_S^d\beta_{ij}}\,. \tag{3.18}$$

Now recall that the whole point of the dynamic modeling process is to arrive at spatially and temporally varying closure constants—C_S^d in the present case. But it is clear from Eq. (3.18) that to proceed further we must employ the approximation

$$\widetilde{C_S^d\beta_{ij}} = C_S^d\widetilde{\beta_{ij}}\,, \tag{3.19}$$

which, as noted in [73], is equivalent to assuming C_S^d is constant over a subdomain corresponding to the test filter width.

One then seeks local values of C_S^d so that the error of approximation in (3.19) is minimized; that is, we define

$$e_{ij} \equiv L_{ij} - \frac{1}{3}L_{kk}\delta_{ij} - C_S^d\alpha_{ij} + C_S^d\widetilde{\beta_{ij}}\,, \qquad i,j = 1,2,3\,, \tag{3.20}$$

the residual of (3.18) when (3.19) is used to replace $\widetilde{C_S^d\beta_{ij}}$. It should be noted that this is a tensor equation, implying that C_S^d, itself, is actually a tensor. In usual treatments of the dynamic modeling approach, however, this is simplified by contraction with the strain-rate tensor $\widetilde{S}_{ij}$ (Germano *et al.* [68]) or with the error itself (Lilly [139]). The former leads to difficulties at flow field locations where the strain rate is very small (or zero), so the latter is recommended. This is equivalent to the least-squares problem of minimizing $|\boldsymbol{e}|^2$ with respect to C_S^d which leads to solving $\partial|\boldsymbol{e}|^2/\partial C_S^d = 0$. It is easily checked that the solution to this problem is

$$C_S^d = \frac{m_{ij}L_{ij}^d}{m_{kl}m_{kl}}\,, \tag{3.21}$$

where $m_{ij} = \alpha_{ij} - \widetilde{\beta}_{ij}$, and we note that Einstein summation is being applied for repeated indices.

As noted in [73], C_S^d given in Eq. (3.21) can exhibit two undesirable properties. First, it is possible for $C_S^d < 0$ to hold, and in the context of the Smagorinsky model this implies a negative SGS eddy viscosity. Since this is employed in the Boussinesq hypothesis, we see that the result is equivalent to a "backward heat equation"—a mathematically ill-posed formulation. But because this typically occurs at only a limited number of points on a computational grid it can be remedied by averaging, typically in directions in which the flow exhibits homogeneity, if such directions exist (see [73] for various options).

Moreover, proponents of dynamic models often view negative viscosities as a desirable feature (ignoring the basic bad mathematics they imply). In particular, the local (usually in both space and time) solution instabilities resulting from ill posedness tend, in at least a qualitative way, to mimic the backscatter that cannot be produced by the highly dissipative Smagorinsky model alone. We should again recall the Galerkin form of the N.–S. equations and observe that no Fourier modes are damped, thus (from a numerical-analytic viewpoint) permitting unresolved high wavenumber effects to alias the low-wavenumber behavior. But, of course, from a pure mathematical perspective, this growth of Fourier coefficients is the hallmark of an ill-posed problem. Hence, it must be emphasized that despite the fact that this approach often produces approximately correct physical outcomes, it does so with completely incorrect mathematics. Thus, its applicability cannot be trusted in general.

The second difficulty with C_S^d as given in Eq. (3.21) is that it can become unbounded, since in principle the denominator can be zero. But Sagaut [73] argues that this is more an implementation problem than a theoretical one because the numerator of this expression also goes to zero at the same time, and at the same rate, as does the denominator.

We conclude this brief treatment of dynamic models by emphasizing that in the end they are still eddy viscosity models, and hence based on the Boussinesq hypothesis which, as we have noted earlier, has no physical basis and moreover represents poor mathematics. But in addition, for the specific case of dynamic models, even more mathematical problems arise which by accident lead to mimicking of true physics. Thus, while dynamic models tend to reproduce physics better than do simple Smagorinsky models—especially for wall-bounded shear flows—the underlying reasons for this are not well founded in either mathematics or physics leaving open the question of whether such models, or any others based on eddy viscosity, could ever be truly predictive.

Mixed models

In general, mixed models consist of some linear combination of a strongly dissipative model, such as the Smagorinsky model, and a scale-similarity (or other deconvolution) model. Numerous such combinations have been studied, and we refer the reader to [73] for an extensive treatment. Here, we will concentrate on only the simplest, the mixed Smagorinsky–Bardina model, to demonstrate the form of this class of SGS models. We will begin with a somewhat more thorough treatment of scale similarity than that introduced earlier, and this will lead to the complete mixed model.

Scale Similarity. As already indicated in the discussion associated with Eq. (3.16), the scale-similarity hypothesis employed for constructing estimates of SGS statistics is based on extracting additional information from the large scale by performing a second filtering. This basic notion is related to what is often termed *deconvolution*, or *defiltering*, and it can be performed in several different ways, as described in [73] and references therein. We will consider one of the simpler versions of this since it leads directly to the Bardina model.

Recall that Eq. (3.16) was constructed by employing two different filter widths; but, in fact, if the filter is not projective a single width is sufficient for the two (or more) different filterings employed in the deconvolution process. This is the approach taken by Bardina *et al.* [140]. In particular, for any variable $\phi(\boldsymbol{x})$, since $\widetilde{\widetilde{\phi}} \neq \widetilde{\phi}$, one can identify the subgrid-scale part of ϕ with the difference between these two filtered quantities (by invoking the scale-similarity hypothesis:

$$\phi'(\boldsymbol{x}) \cong \widetilde{\phi}(\boldsymbol{x}) - \widetilde{\widetilde{\phi}}(\boldsymbol{x}) \,, \tag{3.22}$$

as shown in Fig. 3.1. Furthermore, as indicated in [73], a higher-order approximation can be obtained as

$$\phi'(\boldsymbol{x}) \cong \left(\widetilde{\phi} - \widetilde{\widetilde{\phi}}\right) + \left(\widetilde{\phi} - 2\widetilde{\widetilde{\phi}} + \widetilde{\widetilde{\widetilde{\phi}}}\right) + \cdots \tag{3.23}$$

for filters possessing an invertible kernel. In the Bardina model only (3.22) is used.

From this we can immediately construct approximations to the cross and Reynolds stresses as

$$C_{ij} \cong \left(\widetilde{u}_i - \widetilde{\widetilde{u}}_i\right)\widetilde{\widetilde{u}}_j + \left(\widetilde{u}_j - \widetilde{\widetilde{u}}_j\right)\widetilde{\widetilde{u}}_i \,, \tag{3.24}$$

and

$$R_{ij} \cong \left(\widetilde{u}_i - \widetilde{\widetilde{u}}_i\right)\left(\widetilde{u}_j - \widetilde{\widetilde{u}}_j\right) \,. \tag{3.25}$$

We remark that the approximate equality of Eqs. (3.24) and (3.25) arises from two separate errors. The first is truncation of Eq. (3.23) while the second comes from the following. Recall, *e.g.*, that $C_{ij} \equiv \widetilde{\widetilde{u}_i u'_j} + \widetilde{\widetilde{u}_j u'_i}$. But here we have replaced u'_i with

$$\widetilde{u}_i - \widetilde{\widetilde{u}}_i = \widetilde{u_i - \widetilde{u}_i} \simeq \widetilde{u}'_i$$

so that we have implicitly employed the further approximation

$$\widetilde{u'_i \widetilde{u}_j} \simeq \widetilde{u}'_i \widetilde{\widetilde{u}}_j \, .$$

It can be shown that the error arising from this is $\mathcal{O}(\Delta^2)$, which is relatively large—nearly as large as ν_{SGS}, and hence evidently of the same order as C_{ij} itself.

Then recalling that the Leonard stress is given by

$$L_{ij} = \widetilde{\widetilde{u}_i \widetilde{u}_j} - \widetilde{u}_i \widetilde{\widetilde{u}}_j \, ,$$

and defining the scale-similarity SGS stress as

$$\tau^{(s\text{-}s)}_{SGS,ij} \equiv L_{ij} + C_{ij} + R_{ij} \, , \tag{3.26}$$

we have

$$\tau^{(s\text{-}s)}_{SGS,ij} = \widetilde{\widetilde{u}_i \widetilde{u}_j} - \widetilde{\widetilde{u}}_i \widetilde{\widetilde{u}}_j \, . \tag{3.27}$$

Thus, the SGS stress has been approximated entirely in terms of the filtered resolved-scale quantities—without use of any model!

A Mixed Model. We have earlier alluded to the fact that scale-similarity models such as Eq. (3.27) are not sufficiently dissipative. Indeed, they generally produce far too much backscatter and consequently are generally numerically unstable. This can be remedied by combining them with any of the various highly-dissipative models—artificial dissipation again! Here, we employ the Smagorinsky model treated earlier. The SGS stress tensor for this mixed model can be expressed as

$$\tau^{(m)}_{SGS,ij} - \frac{1}{3}\tau^{(m)}_{SGS,kk}\delta_{ij} = \frac{1}{2}\left(-2\nu_{SGS}\widetilde{S}_{ij} + \tau^{(s\text{-}s)}_{SGS,ij} - \frac{1}{3}\tau^{(s\text{-}s)}_{SGS,kk}\delta_{ij} \right) \, , \tag{3.28}$$

with

$$\nu_{SGS} = (C_S \Delta)^2 |\widetilde{\boldsymbol{S}}| \, ,$$

as already given in Eq. (3.14).

Several comments should be made regarding the mixed model Eq. (3.28). The first is that the simple arithmetic averaging employed here is not always used, and probably is seldom optimal. Indeed, versions of mixed models exist that attempt to estimate the weighting "on the fly," that is, during the simulation. It should also be noted that mixed models have not been as effective as might be hoped, at least in part because the weighting to be employed for the two parts is not easily estimated. But there is a further problem of loss of Galilean invariance (see Speziale [141]) since L_{ij} and C_{ij} are computed separately. On the other hand, it is clear that $\boldsymbol{\tau}^{(m)}_{SGS} \to 0$ as $h \to 0$, since this is true for ν_{SGS}, and it is easily checked for $\boldsymbol{\tau}^{(s\text{-}s)}_{SGS}$. In particular, provided the filter kernel has compact support over the filter width, we see that $\widetilde{u}_i \to U_i$ as the filter width Δ goes to zero. In this limit we then have

$$\tau^{(s\text{-}s)}_{SGS,ij} = \widetilde{\widetilde{u}_i \widetilde{u}_j} - \widetilde{\widetilde{u}}_i \widetilde{\widetilde{u}}_j = U_i U_j - U_i U_j = 0 \, .$$

In closing this section we remark that Speziale [141] has derived conditions under which Bardina's scale-similarity model will preserve Galilean invariance, but unfortunately overall performance of

the model is poorer using constants so prescribed than for others that do not guarantee satisfaction (in fact, violate) this fundamental property of the Navier–Stokes equations. Furthermore, the combination of Boussinesq-based and scale-similarity approaches presents a fundamental contradiction. Namely, any eddy viscosity (Boussinesq-like) method implicitly relies on scale separation since such models incorporate the notion that eddy viscosity "acts like" molecular viscosity—the former plus the latter appear in the (molecular) diffusion terms of the governing equations, and this physical molecular behavior takes place on scales that are not directly influenced by the macroscopic scales of motion. This is shown pictorially in Fig. 3.2, replotted from [73]. In contrast to this, the scale-

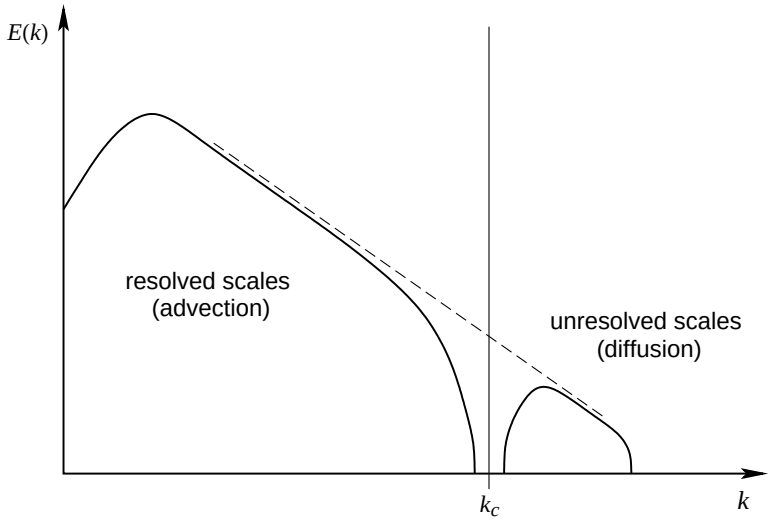

Figure 3.2: Energy spectrum depicting scale similarity.

similarity approaches attempt to bridge this gap near the cut-off wavenumber, and in particular require resolution well into the inertial range to be effective. Thus, even if it were true that mixed models have performed well over a wide range of flows (which is not the case), this fundamental inconsistency would raise serious concerns regarding their ultimate usefulness.

3.1.5 Summary of basic LES methods

In this section we have provided a fairly detailed treatment of the fundamentals of large-eddy simulation. We began by comparing LES with DNS and RANS methods from which we concluded that LES requires far less computation than does DNS ($\mathcal{O}(Re^2)$ total arithmetic instead of the $\mathcal{O}(Re^3)$ required by DNS), and at the same time has a potential for being at least somewhat predictive since in contrast to RANS methods, LES $\rightarrow$ DNS as discretizations are refined. We then presented the basic steps to be followed in constructing typical LES procedures: decomposition of flow variables, filtering of governing equations, construction of SGS models and solution of the resolved-scale, filtered equations. With the exception of the last of these (which is mainly numerical analytic), each was treated in detail in a somewhat general setting with explicit comparisons made with RANS procedures, as appropriate. Finally, we presented outlines of three widely-used classes of LES models: Smagorinsky, dynamic and mixed.

The conclusion to be drawn from these discussions is simply that there still are many shortcomings of "classical" LES in essentially any form, and much research remains. The book by Sagaut [73] is highly recommended for more details on all of the methods we have considered herein and additional ones we have not discussed; but even beyond this it provides classifications of SGS models that should serve as a useful guide in beginning further research into LES.

3.2 Dynamical Systems and Multi-Scale Methods

In this section we introduce a modern approach to analysis of Navier–Stokes turbulence. Work along the lines to be developed here has been in progress since at least the early to mid 1980s, but results have appeared in the archival literature only somewhat sporadically until fairly recently. As the title of this section suggests, we will consider two main (and in the present discussions, related) topics: dynamical systems and multi-scale methods. Neither of these specifically depends on the other, but their combination presents an interesting and potentially efficient technique for simulating details of turbulent fluid motion, as was already recognized by McDonough *et al.* [184, 185] by 1984. Further studies were continued by McDonough and Bywater [186, 187, 188].

While analysis of dynamical systems, *per se*, is by now very mature, with the initial studies of Poincaré [17] dating to more than a century ago, their use in turbulence models is relatively recent, in a sense beginning with the work of Lorenz [18], but more specifically with that of McDonough and coworkers and Kerstein and coworkers in the 1980s and 90s. On the other hand, formal use of multi-scale techniques is very recent although this is forseen in the rather *ad hoc* procedure of McDonough and Bywater [188] and the inertial manifold techniques (especially, the nonlinear Galerkin procedures) of Temam and students (see, *e.g.*, [189]).

We begin this section with discussion of some basic tools associated with analysis of dynamical systems and expand some of the material presented in the definitions of Chap. 1. We then demonstrate that the N.–S. equations indeed comprise a dynamical system, thus suggesting that use of such tools could prove fruitful in their analysis (as, of course, was already recognized by Ruelle and Takens [8] quite a long time ago). We then view LES in the context of mult-scale methods, and within this framework describe an alternative approach to constructing this method. Finally, we provide a summary/comparison of the various methods that have appeared up to the present time.

3.2.1 Some basic concepts and tools from dynamical systems theory

In this subsection we will expand on several of the somewhat vague definitions presented in Chap. 1 and, in particular, indicate their interconnections and applications. Key among these will be the notions of attractor, bifurcation, phase space—and tools needed to distinguish these, *e.g.*, time series and power spectra—and, ultimately, strange attractor. But we must begin by first indicating what constitutes a dynamical system. Rather loosely, a dynamical system is anything that evolves in time. Clearly, this includes all living systems and essentially everything else in physics, biology, economics, *etc.*, and this is too broad a characterization for our present purposes. On the other hand, the detailed rigorous definition provided by Frisch [80] is far more than we will need. Rather, we view *dynamical systems* as differential (or, in some important cases, algebraic) equations serving as descriptions of some time-evolving phenomenon. For example, as we have already stressed, the N.–S. equations provide a complete formulation for analyzing fluid flow, and we will consider these in the context of dynamical systems in what follows.

Attractor

The concept of attractor (of a dynamical system) can best be viewed by considering time series and phase portraits of a simple, hypothetical dynamical system, say, the system of ODEs

$$\frac{du}{dt} = f(u, v; \boldsymbol{\lambda}) , \qquad u(0) = u_0 , \tag{3.29a}$$

$$\frac{dv}{dt} = g(u, v; \boldsymbol{\lambda}) , \qquad v(0) = v_0 , \tag{3.29b}$$

where $\boldsymbol{\lambda}$ is a (possibly vector) specified parameter. We shall assume this system has a long-time, stationary and bounded solution.

Clearly, if f and g are sufficiently simple, analytical solutions to Eqs. (3.29) can be found; but in any case solutions can be readily computed via numerical methods for any assignment of initial data (u_0, v_0) and parameter(s) $\boldsymbol{\lambda}$. Suppose this has been done, and we have obtained time series $(u(t), v(t))$, at least for a discrete set of times, which exhibit steady behavior after sufficiently long time. Then we can plot these time series in Fig. 3.3 to see the generic behaviors displayed by Eqs. (3.29). It is clear from these time series that for the given initial data and value of $\boldsymbol{\lambda}$, $u(t)$ and $v(t)$ ultimately attain steady values denoted here by (u_s, v_s).

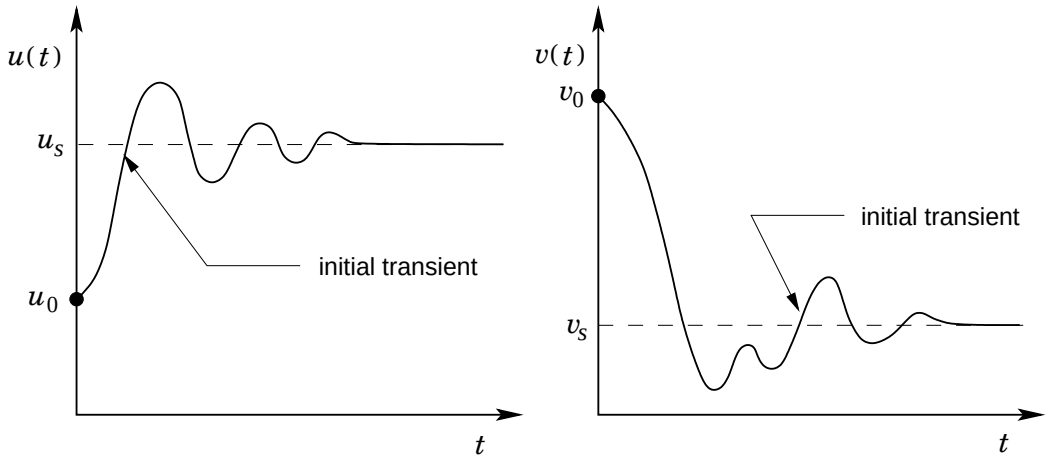

Figure 3.3: Time series of a steady solution to an ODE dynamical system.

Now we plot $u(t)$ *vs.* $v(t)$, or *vice versa*, to obtain a somewhat different view of the evolution of this system. Such a plot is termed a *phase portrait*, and a schematic for our hypothetical system is provided in Fig. 3.4. We see that after a relatively brief transient the solution arrives at a single fixed point, (u_s, v_s), in phase space and remains there. In this steady case this single point is the *attractor* for the dynamical system Eqs. (3.29) associated with the given data, (u_0, v_0) and $\boldsymbol{\lambda}$. We think of this point as "attracting" the initial point (u_0, v_0). Furthermore, it may be (usually is) the

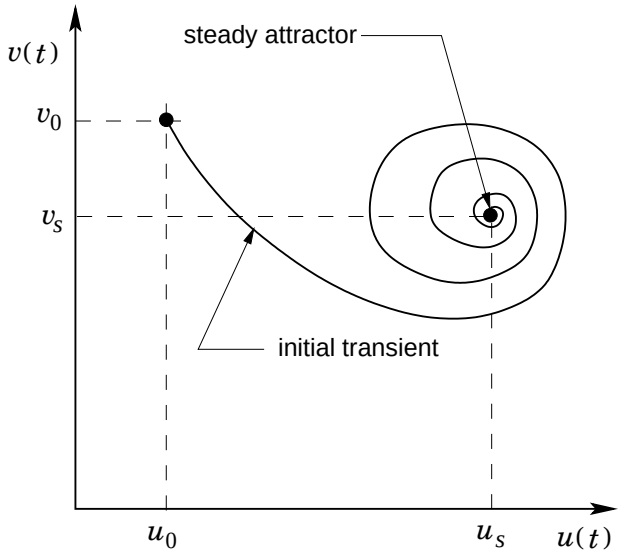

Figure 3.4: Phase portrait of steady attractor.

case that for a given value of $\boldsymbol{\lambda}$ there are many different initial points whose trajectories ultimately end at (u_s, v_s). This set of points is called the *basin of attraction* for the point (u_s, v_s), and the collection of trajectories leading from these points to (u_s, v_s) is termed the *flow* of the dynamical system (3.29).

Bifurcation

Next, consider changing the value of $\boldsymbol{\lambda}$ in Eq. (3.29). One would expect that if the mathematical initial-value problem for these equations is well posed, then a small change in $\boldsymbol{\lambda}$ should result in only a small change in $(u(t), v(t))$ for fixed initial data (u_0, v_0). Indeed, this can usually be guaranteed for linear systems, but if (f, g) are nonlinear functions of (u, v) additional, more interesting, things can occur. In particular, it is possible that at certain *critical values* small changes in $\boldsymbol{\lambda}$ will result in large, even qualitative, changes in the behavior of (u, v) both during the initial transient, and especially in the nature of the stationary state. Such qualitative changes are associated with a *bifurcation* of the dynamical system, and correspond to singularities of the Jacobian matrix of the right-hand side functions (f, g) of Eqs. (3.29). This leads to a new "branch" of solutions coming into existance at the *bifurcation point* associated with a *critical value* of $\boldsymbol{\lambda}$. (It is worthwhile to observe that the linear algebraic eigenvalue problem exhibits analogous behavior: the solution vector X is identically zero except when the matrix minus an eigenvalue times the identity is singular.)

A widely-studied (both theoretically and experimentally) and easily-understood example arises in thermal convection in the form of the *Rayleigh–Bénard problem*. Physically, this consists of a layer of fluid heated from below and/or cooled from above as shown schematically in part (a) of Fig. 3.5. The temperature difference, $\Delta T = T_{H,1} - T_{C,1} > 0$, is such that the Rayleigh number,

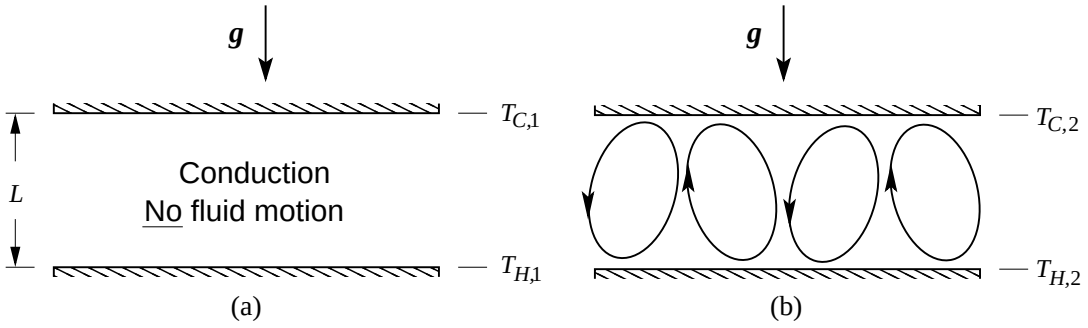

Figure 3.5: Bifurcation (transition) to convection in Rayleigh–Bénard problem; (a) conduction, and (b) convection.

defined as $Ra \equiv g\beta\Delta T L^3/\kappa\nu$, is lower than the critical value Ra_{cr} for transition to fluid motion, and heat is transferred across the fluid layer by conduction only. At the first critical value of Ra a bifurcation occurs, and fluid motion in the form of steady counter-rotating cells, often called "rolls," as depicted in part (b) of the figure, begins. (In the definition of Ra, g is gravitational acceleration; β is thermal volumetric expansion coefficient; L is thickness of the fluid layer, and κ and ν are, respectively, thermal diffusivity and kinematic viscosity.)

This transition is a physical example of a bifurcation, and after this occurs heat transport in the fluid layer becomes predominantly convective. Hence, it is of a qualitatively different nature from the conductive heat transfer prior to the bifurcation. In particular, the dimensionless heat transfer quantity known as Nusselt number, Nu (representing the ratio of total heat transport to conductive heat transport), suddenly begins to increase due to onset of convection. This is indicated in Fig. 3.6, a *bifurcation diagram*, which indicates the states that a dynamical system might attain,

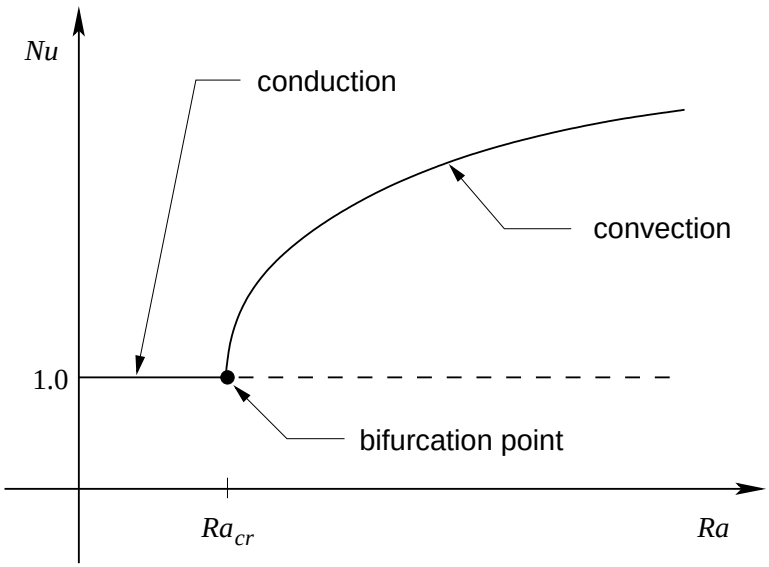

Figure 3.6: Qualitative bifurcation diagram for Rayleigh–Bénard convection.

and locations in terms of bifurcation parameter values where the system changes from one state to another.

There are several important points to be noted regarding Fig. 3.6. First, the branch corresponding to unit Nusselt number representing conduction only, is shown extended as a dashed line beyond the bifurcation point. This is intended to imply that the conduction solution is still a physical and mathematical solution even beyond the critical Rayleigh number; that is, this solution still satisfies the equations of motion and required boundary conditions. On the other hand, a stability analysis (see, *e.g.*, Chandrasekhar [190]) shows that for $Ra \geq Ra_{cr}$ this solution becomes unstable to small perturbations and thus will no longer be observed in laboratory experiments (or in machine computations). It is replaced by the now stable convective solution indicated in Fig. 3.6 by the curve of increasing Nu.

Second, it is also important to note that the bifurcation shown in Fig. 3.6 is only the first of several that occur in a sequence ultimately ending in turbulent convection as Ra is increased in the physical system. Such bifurcation sequences have been widely studied in laboratory experiments (see,*e.g.*, Gollub and Benson [58], Cioni *et al.* [191], Libchaber [192]), and results confirm the Ruelle–Takens [8] theory in a general way. In particular, there appear to be only a finite (and relatively small) number of bifurcations prior to the onset of turbulent fluid motion, in sharp contrast to the Landau–Hopf view mentioned earlier ([28],[29]) which predicts an infinite sequence.

Finally, it should be emphasized that Fig. 3.6 is not the only form taken by bifurcation diagrams. We will later display others that will be more quantitative and detailed. But the present one very clearly represents the basic ideas that need to be understood, and this simplicity is an advantage at this point.

It is now worthwhile to return again to a view of time series and phase portraits and examine their character as further bifurcations occur. Figure 3.7 displays two different time series that we might associate with the hypothetical dynamical system consisting of the ODEs of Eqs. (3.29). The first of these, part (a), corresponds to a periodic solution that essentially always (but there are some exceptions) occurs at the first bifurcation from the steady solution presented earlier. Part (b) of the figure shows a time series corresponding to behavior after the first subharmonic bifurcation (which is often—but not always—the second bifurcation), often termed a "period-doubling" bifurcation

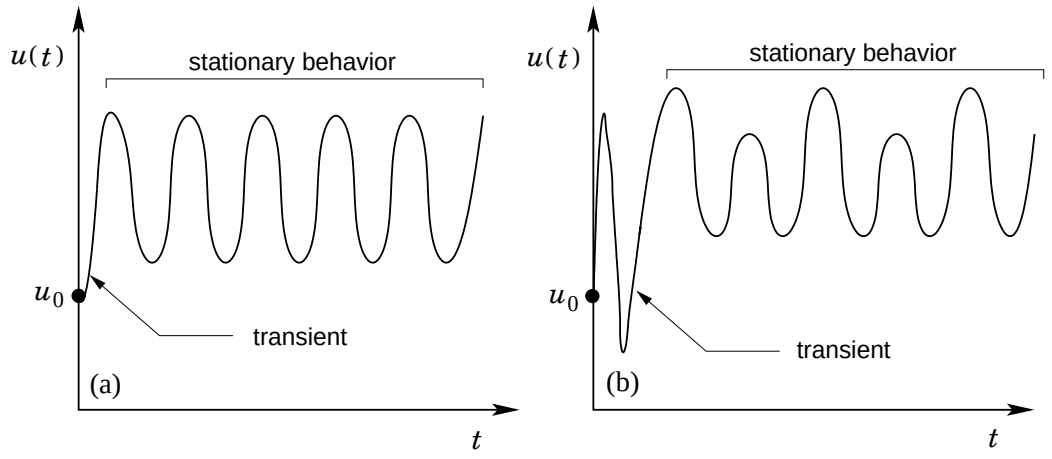

Figure 3.7: Time series of (a) periodic and (b) subharmonic solutions to an ODE dynamical system.

since, as is clear from the figure, the period of the time series in part (b) is twice that of part (a).

We are now to a point in our discussions where use of phase portraits is again useful; as we have seen earlier in the nearly trivial case of Fig. 3.4, these provide a graphical representation of the "topology" of the attractor. Figure 3.8 provides a comparison of these representations for the attractors associated with the respective time series of parts (a) and (b) of Fig. 3.7.

These figures depict several important points associated with periodic and subharmonic behaviors. The first is that, especially for the latter, the initial transient may be rather long; but with regard to this we also note that the length of an initial transient generally depends on how far the bifurcation parameter value is from its next critical value—in general, the closer it is, the longer the initial transient. (The reader may wish to consider why this might be expected, at least qualitatively, by recalling that the original behavior prior to a bifurcation is still a solution to the equations of motion after the bifurcation. An important issue is stability of solutions.)

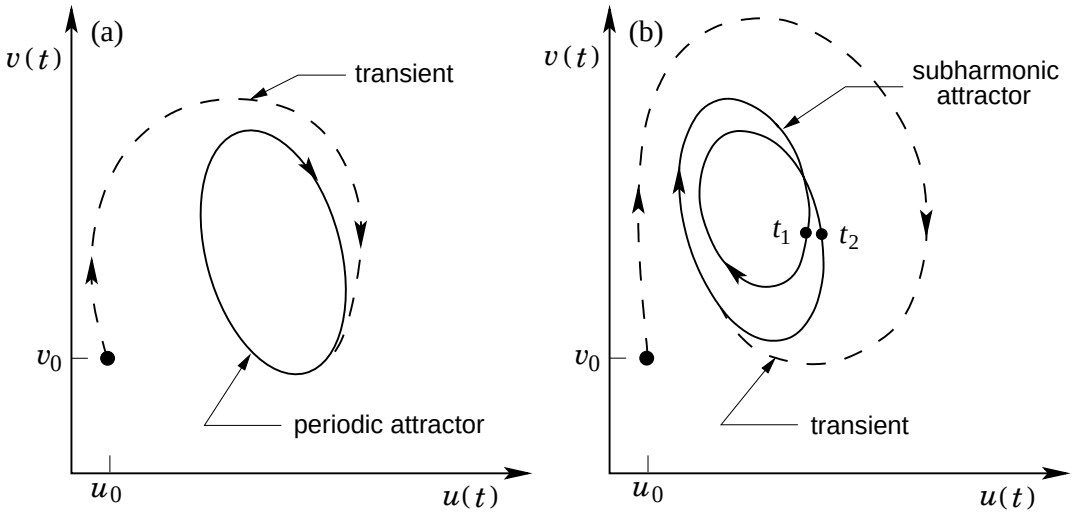

Figure 3.8: Comparison of phase portraits of (a) periodic, and (b) subharmonic attractors.

The second point is associated with part (a) of Fig. 3.8; it is that the structure of the attractor (its topology) takes the form often termed a "limit cycle." From the figure we see that once the

trajectory has completed the initial transient and has arrived at the attractor, it continues to "cycle" through the points comprising the attractor in a completely persistent and regular manner $\forall\ t \to \infty$, as is already suggested by the time series plots of Fig. 3.7.

In Fig. 3.8(b), representing a subharmonic attractor, we see a somewhat more complicated structure replacing the simple limit cycle. We have labeled two points to provide some insight into what has happened to the dynamical system at the subharmonic bifurcation. Namely, at time t_1 for $\lambda < \lambda_{cr}$ the trajectory would already have returned to the location marked as t_2 taking on values (u_2, v_2) in phase space; hence, the periodic orbit would have been completed, and the trajectory would begin a new traversal of this same limit cycle. But in contrast, following bifurcation to a subharmonic regime the trajectory arrives at the phase-space point $(u_1, v_1) \neq (u_2, v_2)$ at time t_1. The orbit is not complete, and the trajectory must be continued until it ultimately arrives at (u_2, v_2). We remark that proof that this should ever occur, and moreover that it occurs in such a way that $t_2 = 2t_1$, is not trivial and is beyond the intended scope of these lectures.

We next note that as λ is increased beyond λ_{cr} it is common for second, third, and so on subharmonic bifurcations to occur. This is termed a *Feigenbaum bifurcation sequence* [56], and it can be proven that as this sequence progresses the distance between successive bifurcation points decreases (*i.e.*, $\lambda_{cr,n+1} - \lambda_{cr,n} < \lambda_{cr,n} - \lambda_{cr,n-1}$, *etc.*) resulting in a "cascade of bifurcations," each of which results in more complicated phase-space trajectories and associated time series, ultimately leading to chaotic behavior. While this formally represents an infinite sequence of bifurcations, from any practical standpoint they begin to occur so close together (usually by $n \simeq 8$) that at least from a computational perspective it is not possible to distinguish the behavior from that of a chaotic attractor.

Power Spectra

We now briefly discuss another important diagnostic tool, power spectral analysis, for study of dynamical systems. The result of applying this to a data set (usually a time series, but sometimes a spatial distribution) is the *power spectral density* (PSD), a Fourier-space representation of energy (in the L^2 sense) of a signal as a function of frequency (or wavenumber). This can be used to identify attractors on a qualitative basis; and while in some cases it does not provide a conclusive characterization, it is often quite useful.

Figure 3.9 presents three PSDs associated with a hypothetical subharmonic bifurcation sequence such as the one discussed above. Comparison of parts (a) and (b) indicates the effect of the first

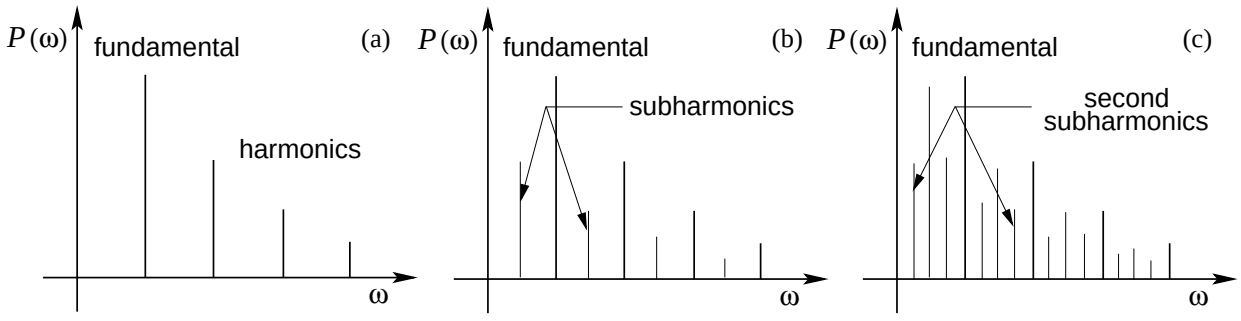

Figure 3.9: Subharmonic bifurcation sequence (a) periodic, (b) subharmonic and (c) second subharmonic.

subharmonic bifurcation (usually, as noted above, the second bifurcation, overall) on the frequency content of a fairly general periodic signal (part (a)) containing a fundamental and three harmonics. It is clear that the new frequencies (shown as lighter lines in parts (b) and (c)), and termed *subharmonics*) have been generated at the midpoints between successive harmonics. The first of

these is one half the frequency of the fundamental, corresponding to "frequency halving" that is equivalent to period doubling displayed in the time series and phase portraits of Figs. 3.7 and 3.8, respectively. Comparisons of parts (b) and (c) of Fig. 3.9 shows a repeat of the process seen in parts (a) and (b). Of particular note is the fact that power, $P(\omega)$, of the half frequencies is generally quite low for λ just above λ_{cr}; but as λ is further increased, half-frequency powers increase; and as they approach the power(s) of the original frequencies (fundamental and harmonics), a subsequent subharmonic bifurcation typically occurs, as indicated in parts (b) and (c). This suggests that the frequency content of the spectrum will be further and further filled as bifurcations continue, resulting in a (nearly) continuous, broad-band (hence, noisy) power spectrum which in the context of fluid flow would be associated with turbulence.

Strange Attractor

In the preceding subsection we introduced the notion of attractor for trajectories of a dynamical system because without this, consideration of a "strange" attractor would be nearly impossible. In the present section we will describe, in a non-rigorous way, what strange attractors are, how they are constructed, how they can be characterized (recognized?) in data, and finally, what they might have to do with fluid turbulence.

The simplest (non-rigorous) description of a *strange attractor* is "a manifold constructed as the Cartesian product of a smooth manifold and a Cantor set." For the nonmathematician this contains many unfamiliar terms (but with some defined in Chap. 1), and we will elaborate on these here because an at least heuristic understanding of them is important. It is sufficient to view a *manifold* as being like a "surface," but of dimension possibly greater than two. The simplest manifolds with which we are familiar are the two- and three-dimensional Euclidean spaces $\mathbb{E}^d$, $d = 2, 3$, in which we have always lived, and which are denoted $\mathbb{R}^2$ and $\mathbb{R}^3$ in the mathematics literature. (Of course, these are not surfaces, but common surfaces, *e.g.*, the surface of a sphere or torus, are often embedded in them.)

The next term to consider is *Cartesian product*. The easiest way to understand this notion is via example: $\mathbb{E}^2 = \mathbb{R}^1 \times \mathbb{R}^1$. This implies that the points, say (x, y) of a two-dimensional Euclidean space are obtained by selecting their coordinates, independently, from two separate copies of the real line $\mathbb{R}^1$. That is, we have $(x, y) = \{x \in \mathbb{R}^1, y \in \mathbb{R}^1\}$. The notation $\mathbb{R}^1 \times \mathbb{R}^1$ is often simplified to $\mathbb{R}^2$.

Now, a smooth manifold requires rather precise definition for rigorous purposes, but we will forego this here and simply observe that tori (donut-shaped objects) are typically smooth manifolds. We will see, moreover, that these are precisely the shapes needed to build a strange attractor, at least of the type envisioned in [8] as being associated with turbulent solutions to the N.–S. equations.

This leaves us needing to consider the Cantor set as the final piece of information leading to an understanding of the structure of a strange attractor. The Cantor set, and others like it, have an important place in pure mathematical analysis because they provide examples of sets containing an uncountable infinity of points, and yet possessing zero measure (think length). But in addition, they have non-zero fractal dimension (which will be defined below). It is clear that these are rather peculiar objects. Their construction, however, is quite simple as we demonstrate in Fig. 3.10.

This figure shows the unit interval as the starting point, but other intervals could easily be employed. At step 1 we discard the points of the middle third of the interval, retaining the point set $\{[0, \frac{1}{3}] \cup [\frac{2}{3}, 1]\}$. Clearly, this is an uncountable set, and its length is 2/3. Now we repeat the process of removing the middle third from each of the two remaining non-empty segments, thus obtaining the set $\{[0, \frac{1}{9}] \cup [\frac{2}{9}, \frac{1}{3}] \cup [\frac{2}{3}, \frac{7}{9}] \cup [\frac{8}{9}, 1]\}$. Obviously, this set is still uncountable, and its length is now $(2/3)^2 = 4/9$. It is easily checked by induction that at the n^{th} stage of the construction the length of the union of the remaining subintervals must be $(2/3)^n$. Hence, as $n \to \infty$, the length

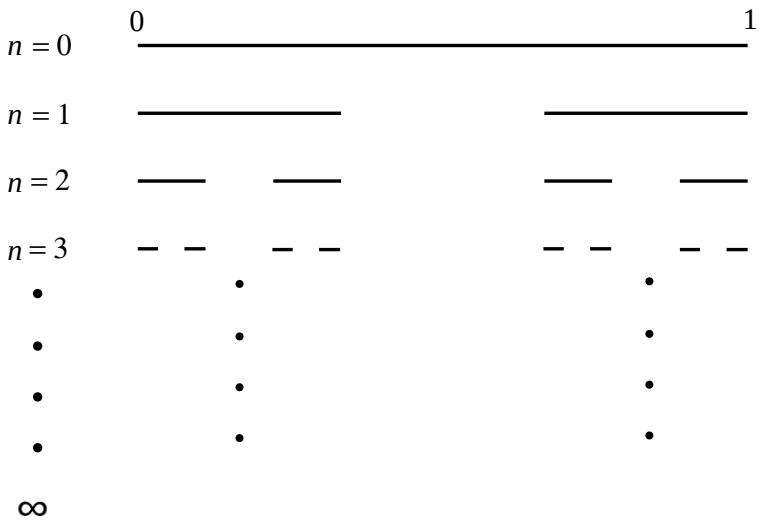

Figure 3.10: Construction of the "1/3" Cantor set.

(measure) of the Cantor set goes to zero. But at the same time, the point set remains uncountable as $n \to \infty$ due to the basic topology of the real line (and thus, the unit interval).

In summary, we have generated a point set consisting of an uncountable infinity of points, but this collection of points has zero length. In the context of the original unit interval it is clear that what remains is mainly gaps containing no points, as is clear from Fig. 3.10, but at the same time there are 2^n small concentrations of points each having length $\frac{1}{2^n}\left(\frac{2}{3}\right)^n = \left(\frac{1}{3}\right)^n$.

An interesting question is, "What might be the dimension of this point set?" The original unit interval has unity dimension, but it also has unit length. We have seen that what remains of length in the Cantor set is zero; moreover, we know that the dimension of a single point is zero (which is true also of its length). It turns out that to determine the dimension of the Cantor set we need a more general notion of dimension than the simple "geometric" one we have implicitly been applying in the preceding discussion. There are many such general definitions, and the following one attributed to Hausdorff [84] will be sufficient for our purposes.

Definition 3.1 *Let N_ϵ denote the number of sets of size ϵ needed to "cover" the "object" whose dimension D is being sought. Then*

$$D \equiv \lim_{\epsilon \to 0} \frac{\ln N_\epsilon}{\ln 1/\epsilon}\,, \tag{3.30}$$

if the limit exists.

If, for example, we consider the unit interval and choose to cover this with subintervals of size ϵ, it is clear that the number of these needed is $N_\epsilon = 1/\epsilon$. Hence, the limit in Eq. (3.30) exists, and we see that $D = 1$, the same as the usual "geometric" dimension. We leave as an exercise to the reader the demonstration that squares and cubes have Hausdorff dimension equal to their respective geometric dimensions.

For the Cantor set we have already shown that there are 2^n small concentrations of points, each of size 3^{-n}, at the n^{th} stage of construction. Thus, if we take $\epsilon = 3^{-n}$, then $N_\epsilon = 2^n$ will provide a sufficient covering, and $\epsilon \to 0$ as $n \to \infty$; it follows that $D = \ln 2/\ln 3$ is the dimension of the Cantor set. Clearly, this is a non-integer value, implying that the Cantor set is a *fractal*; *i.e.*, it has a non-integer (fractional) dimension.

We are now prepared to attempt "construction" of a strange attractor. There are several different ways to go about this. Here, we will rely on simple heuristics with relatively easy geometric interpretations that utilize preceding discussions. We remark that this will not be rigorous, and probably cannot be made entirely rigorous. Readers seeking more precise descriptions are encouraged to consult treatment of the Smale "horse-shoe map," found *e.g.*, in sections of Alligood *et al.* [85].

From a purely geometric viewpoint (see, *e.g.*, Lanford [86]) a strange attractor can be constructed by causing the flow of a differential system to repeatedly undergo the following sequence of transformations as its trajectories traverse the interior of a bounding manifold (for example, a torus): *i*) simultaneous expansion (stretching), and compression (contraction), *ii*) rotation, *iii*) folding, and *iv*) scaling to force the flow to stay withing the manifold. Schematics of these are presented in Fig. 3.11. This figure displays a 2-torus in part (a) with some representative Poincaré sections

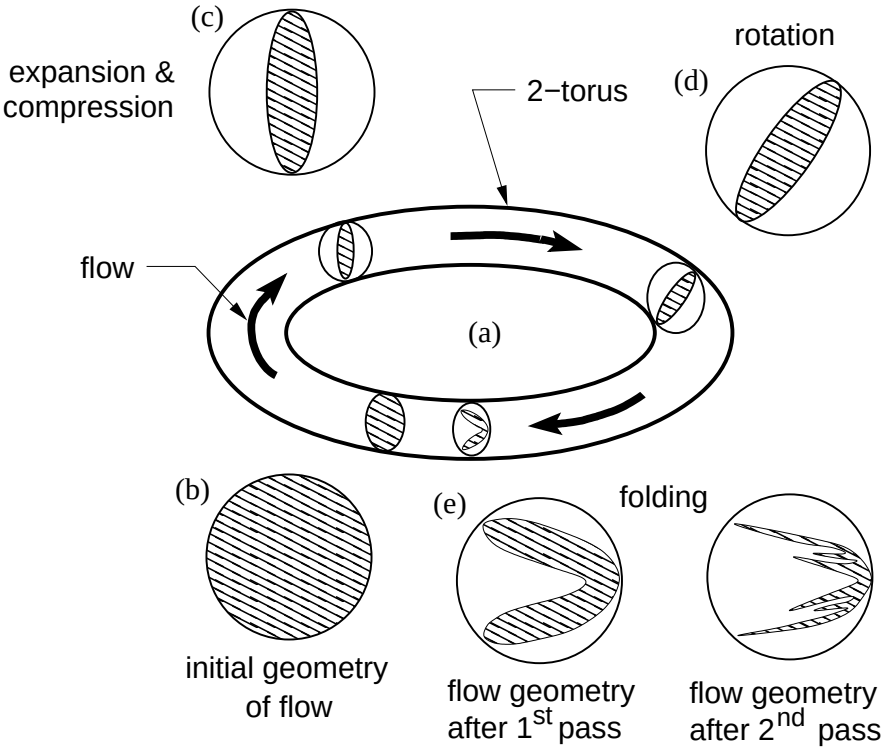

Figure 3.11: Geometric representation of construction of a strange attractor.

corresponding to the above-noted sequence of transformations indicated in parts (b) through (e). The key observation is that by the time the flow has completed one orbit around the 2-torus its cross-sectional area has been significantly reduced—analogous to removal of the middle third of the unit interval during the first step of constructing a Cantor set. Then, the process is repeated (just as occurs in Cantor set constructions). Part (e) of the figure shows results from passes 1 and 2 through the torus, while part (b) shows the area of the hypothetical initial flow. The Cantor set features are obvious from these comparisons, and it is not hard to imagine that as the number of orbits of the flow approaches infinity, what remains will be a structure with Cantor set properties—but, in general, in higher dimensions. Moreover, the figure exhibits the geometric characterization of a strange attractor mentioned earlier, namely, a Cartesian product of a smooth manifold (the 2-torus) and a Cantor set. In particular, one can view each Poincaré section as marking a coordinate of the smooth manifold, with individual Cantor set points within this section providing the other entries

of the Cartesian product.

There are a couple important items to consider with regard to this construction. First, for continuous dynamical systems these heuristics suggest that dimension three (in phase space) is the smallest system that can produce a strange attractor. Moreover, the fractal dimension of any such attractor is greater than two. We remark that these are provable statements (see, *e.g.*, [77]).

The next point to consider is what is required in a dynamical system to permit it to carry out the required transformations. There are two key properties. The first is nonlinearity. If we recall the Galerkin representation of the N.–S. equations of Chap. 1 where we noted that the nonlinear terms are capable of generating modes that were not present in prescribed initial data, we see that there is an intrinsic capacity to generate complicated phase-space behaviors associated with stretching and rotation. The second property required to create a strange attractor is dissipation. Especially the expansive behaviors of the nonlinear terms must be controlled in order to produce an attractor; and as was emphasized in our Galerkin procedure analyses of the N.–S. equations, this is accomplished with dissipation. In particular, in the phase-space context we can also view dissipation as producing contraction of the flow, and scaling in general, to allow it to remain within a bounding mainfold. But in addition, it is reasonable to consider the combination of nonlinearity and dissipation as contributing to the folding process. In summary, we observe that in order for a continuous dynamical system to be able to generate a strange attractor, it must be of at least dimension three in phase space; it must be nonlinear, and it must be dissipative.

We remark that by way of contrast, for a discrete dynamical system (algebraic map), three dimensionality is not required (but, of course, is permitted) for existence of a strange attractor. Indeed, simple 1-D quadratic maps such as the *logistic map*

$$x^{(m+1)} = \beta x^{(m)} \left(1 - x^{(m)}\right) , \qquad 0 < \beta \leq 4 , \tag{3.31}$$

studied by May [193] exhibit the chaotic temporal behavior of a strange attractor. Moreover, a 2-D discrete map due to Hénon exhibits another property often associated with strange attractors, namely *scale similarity*. This can be seen in the Cantor set of Fig. 3.10; in particular, one sees that if construction were started at the first step rather than at the zeroth one, the results would be the same up to a rescaling of lengths, as hinted earlier; clearly, this is true $\forall\ n < \infty$. The Hénon map is generated from the algebraic system

$$x^{(m+1)} = 1 - a x^{(m)\,2} + y^{(m)} , \tag{3.32a}$$

$$y^{(m+1)} = b x^{(m)} , \tag{3.32b}$$

usually computed using $a = 1.4$ and $b = 0.3$. This is presented by Hénon [194] as a Poincaré map of the Lorenz equations [18], although this is in the framework of an argument and construction that is not especially rigorous. Figure 3.12, a phase portrait for this discrete dynamical system (DDS) (3.32), displays the scale similarity property.

We observe that as we continue to focus on finer and finer scales we see repetition of the basic structure. This is evidence of the Cantor set-like construction, now for a concrete example. In particular, part (a) of Fig. 3.12 displays the usual phase portrait of the Hénon attractor generated from Eqs. (3.32) starting with initial data $x_0 = y_0 = 0$. This figure at first appears rather unremarkable, consisting of a few seemingly smooth curves. Moreover, a first zoom-in in a neighborhood of $[-0.1, 0.1]\times[0.2, 0.3]$, part (b) of the figure, seems to confirm this. But a further finer-scale region interior to $[-0.002, 0.002]\times[0.268, 0.272]$ shown in part (c) indicates significant structure. Here we see a sequence of bands of points arranged successively (from bottom to top) in groups consisting of a single band, two bands, and then three bands. (We remark that these bands are clearly not continuous curves; they contain many gaps just as do the later steps of the Cantor set.)

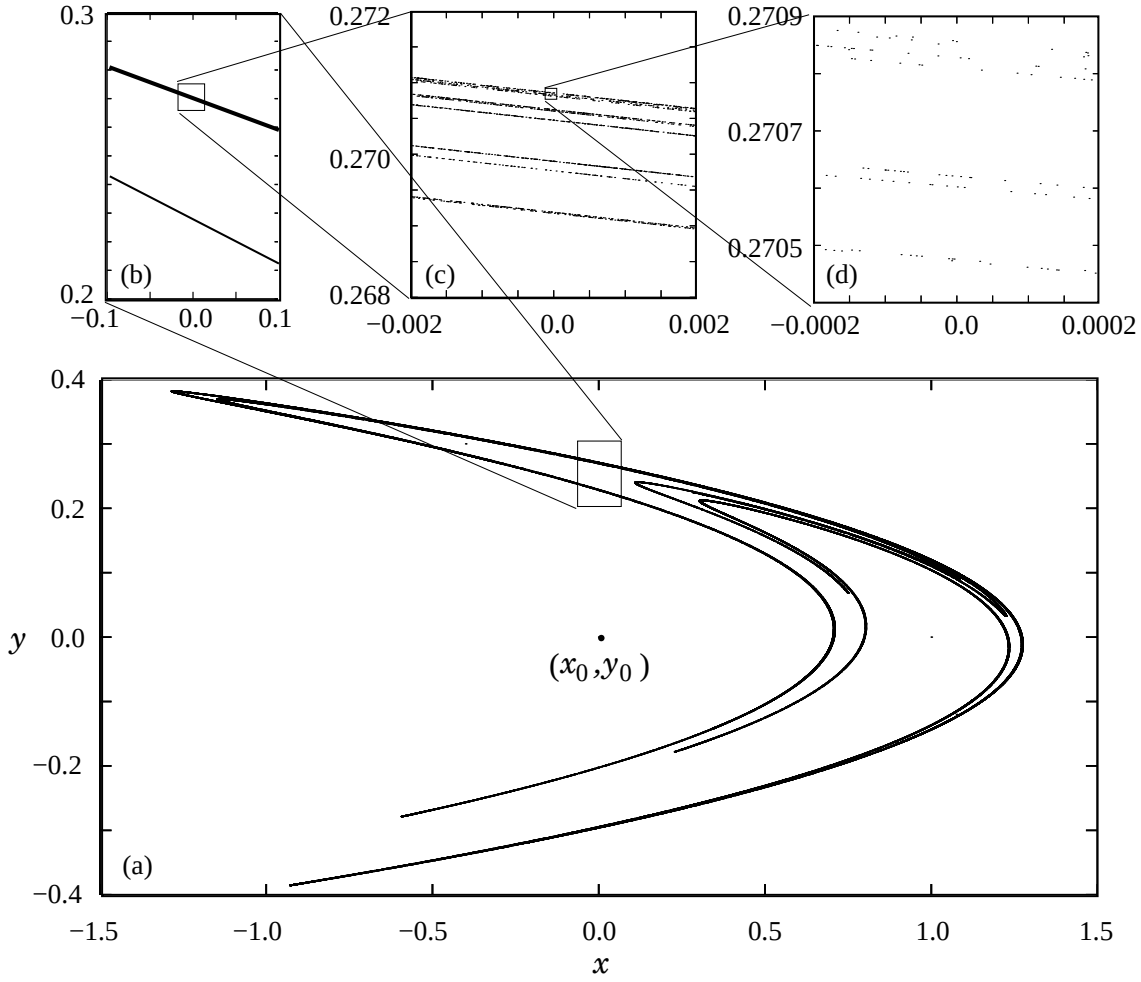

Figure 3.12: Demonstration of scale similarity in Hénon map.

Beyond this is the structure within these bands. We have chosen to further zoom in on the upper-most band of the three-band structure of part (c). Results of this are presented in Fig. 3.12(d) and show that the three sets of bands exhibited by part (c) are repeated on finer scales, again analogous to what occurs in Cantor set constructions. It is clear that with the number of computed points (5×10^6) in the trajectory we will not be able to observe further fine-scale structure in detail, but consistency of relative spacing of bands on the two different scales displayed here suggests that such structure will persist on ever finer scales.

Finally, in Fig. 3.13(a) we display what is often (in fact, usually) considered the defining characteristic of a strange attractor: sensitivity to initial conditions (SIC). This plot shows the same subinterval of time for time series of a DDS known as the "poor man's Navier–Stokes equation" (PMNS) derived and studied by McDonough and Huang [81, 82]. For the results shown here a 2-D PMNS equation was used, so two pieces of initial data are needed to start the time evolution. Calculations leading to the time series plotted in red (dotted line in greyscale) were computed with only a 0.001% change in one component of the initial data (and no change in the other). It is clear that while the overall structure of the two time series is very similar (hence, it is reasonable to expect they are on the same attractor—and thus belong to the same basin of attraction), they are very different in detail. In particular, if one chooses a point in time where the two series attain the same value, it is impossible to predict details of subsequent evolution of one of the trajectories from

knowledge of the behavior of the other.

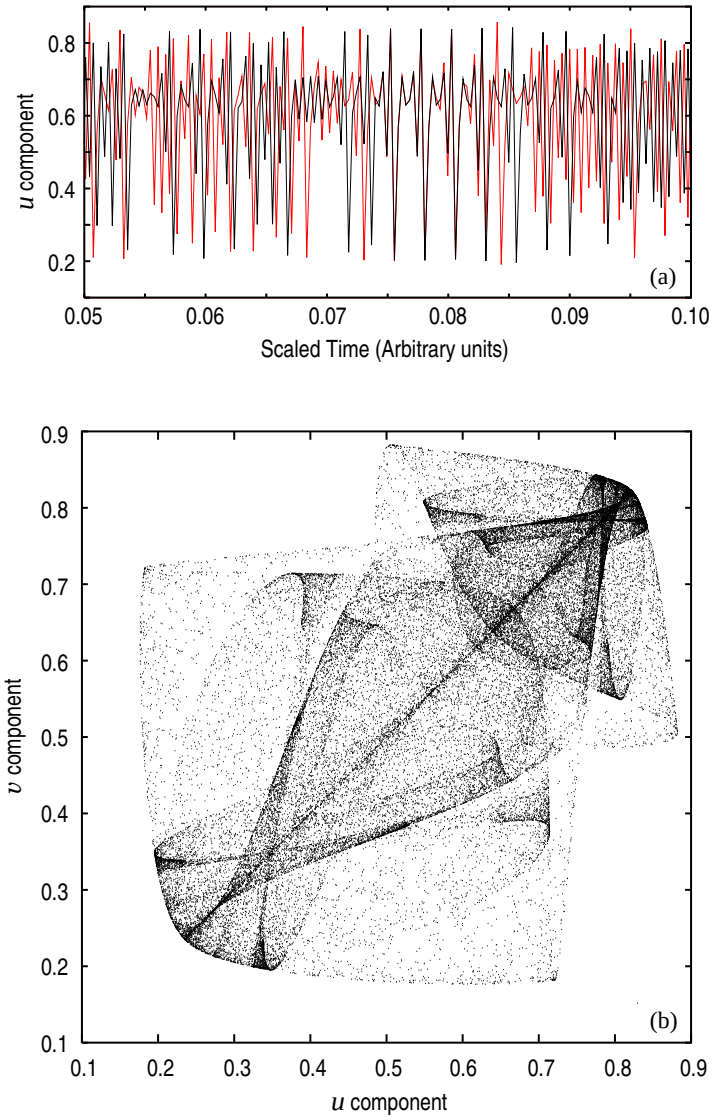

Figure 3.13: (a) Sensitivity to initial conditions in the PMNS equation DDS, and (b) corresponding phase portrait.

Part (b) of Fig. 3.13 presents the phase portrait displaying the topology of this attractor. It is clearly far more complicated than that of the Hénon attractor which is also associated with a 2-D DDS, but it is also important to point out that the PMNS equation exhibits multiple strange atractors as well as many non-chaotic attractors, as is true of the N.–S. equations, themselves. Further discussion of this can be found in [82] and in a later subsection devoted to alternative forms of LES.

3.2.2 The Navier–Stokes equations as a dynamical system

With a basic understanding of strange attractors in hand, we now are prepared to consider how they might be linked to turbulence. This involves recognition that the N.–S. equations are, in fact, a nonlinear dynamical system, as we previously noted in Chap. 1. Here, we begin by recalling that this link was first proposed in the seminal paper by Ruelle and Takens [8] where it was shown

that phase-space trajectories corresponding to a flow of the N.–S. equations (viewed as a dynamical system) become attracted to a topological object having properties we have just discussed. They termed this a "strange attractor," and proposed that its temporally chaotic behavior should be associated with physical turbulence. It is interesting to also note that Lorenz [18] had earlier computed similar trajectories for a severely truncated Galerkin approximation of the N.–S. and thermal energy equations (also a dynamical system) and used sensitivity to initial conditions of solutions to these equations to conclude that long-range weather prediction would never be possible. We should observe, however, that while the equations due to Lorenz do exhibit a strange attractor in their solutions for certain ranges of parameter values, the bifurcation sequence leading to such states is different from that studied in [8].

But beyond the fact, already used in [8], that the N.–S. equations comprise a dynamical system is the experimental observation that typically only short bifurcation sequences occur before flow behaviors are chaotic and sensitive to initial conditions, and thus associated with physical turbulence. In particular, experiments with fluid flow (and other physical phenomena) have identified at least three distinct short bifurcation sequences leading to physical chaos, and they have also been observed and studied analytically and computationally. These are the following (characterized by qualitative system description of the various states):

$$\begin{aligned}
&\text{steady} \longrightarrow \text{periodic} \longrightarrow \text{subharmonic} \longrightarrow \text{chaotic}\,,\\
&\text{steady} \longrightarrow \text{periodic} \longrightarrow \text{quasiperiodic} \longrightarrow \text{chaotic}\,,\\
&\text{steady} \longrightarrow \text{periodic} \longrightarrow \text{intermittent} \longrightarrow \text{chaotic}\,.
\end{aligned}$$

The first of these was studied by Feigenbaum [56] in the context of the simple algebraic map, Eq. (3.31), first studied by May [193], as noted earlier. But numerical solutions to the N.–S. equations also exhibit such a sequence in the context of boundary layer transition, as reported by Pulliam and Vastano [195]. The second sequence is that originally proposed by Ruelle and Takens [8], and which, as noted earlier, has been observed (with some modification) in numerous laboratory experiments (*e.g.*, those of Gollub and Benson [58]) and simulations. The final sequence given here was first studied by Pomeau and Manneville [57] (see also [77]). We remark that in these works, and in further studies cited therein, the authors identified three distinct types of intermittency, all of which have been observed in physical experiments but not yet conclusively in N.–S. equation simulations (possibly because they tend to occur at values of Re well above those currently accessible via DNS). Finally, we observe that various combinations of these sequences have been seen in both experiments and computations, but in all cases the route to chaos consists of relatively few bifurcations.

Thus, both computations with the N.–S. equations and laboratory investigations of physical fluid flows generally support the Ruelle and Takens view and contradict predictions of the Landau–Hopf theory which proposed that turbulence was the result of an infinite sequence of quasiperiodic bifurcations, each of which produced an additional incommensurate frequency. While any such flow could be very complicated, as would be its power spectrum, it would lack SIC and the fractal dimension seen in modern laboratory experiments and DNS of the N.–S. equations.

While the preceding discussions appear to support the chaotic dynamics/strange attractor view of turbulence, there are two main arguments often given against this interpretation. The first of these is the claim that dynamical systems produce only temporal chaos while it is well known that turbulence exhibits both temporal and spatial chaotic behavior. Here, we propose that this argument is completely fallacious, arising from a lack of understanding of the N.–S. equations and the (mathematical) nature of their solution structure. In particular, it ignores the fact that Fourier representations with temporally chaotic coefficients will automatically produce spatial chaos as a consequence of different linear combinations of these coefficients (via different values taken on by

basis functions) at each spatial location and effects arising from SIC associated with spatially-distributed initial data. This is clearly observed in DNS, and even in well-resolved LES. Those adherring to claim of no connection between dynamical systems and turbulence must forego use of (DNS of) the N.–S. equations for their argument to self consistent—and very few researchers in the modern era would be willing to do this (fortunately) after all the success experienced with simuation via these equations.

The second criticism of the "strange attractor theory of turbulence" is more serious—especially from a practical, computational viewpoint. It is that, to date, the theory of dynamical systems (aside from use of DNS) has failed to produce any computational tools for practical simulations of turbulence. Although there have been scattered exceptions to this (for example, the shell models described in Bohr *et al.* [83]), they have for the most part represented rather anecdotal results, as opposed to well-founded, general formulations capable of wide application. We note that the one-dimensional turbulence (ODT) model of Kerstein and coworkers (see, *e.g.*, Echekki *et al.*)[71]), an outgrowth of Kerstein's linear-eddy models (LEMs), is to some extent proving to provide fairly wide applicability. In addition, we will in the next subsection provide details of a new modeling approach that appears to hold significant promise, primarily because of its firm foundations in theory of the N.–S. equations.

3.2.3 Multi-scale methods and alternative approaches to LES

In this subsection we will treat two main, related, topics: *i*) the basics of multi-scale methods, and *ii*) application of these techniques to an alternative formulation for LES. The first of these is a current topic of accelerating research in numerical analysis, and complete journals are now devoted to this area. The second is not new; but in the past, work has been sporadic and not necessarily well founded on a mathematical basis. Here, we will attempt to remedy this by providing a different approach to constructing LES procedures in considerable detail and justifying this with formal mathematics—but in the absence of rigorous proofs. As part of this, we will supply a pseudo-language algorithm presenting an outline of the steps needed to produce a computational LES code based on this approach.

Basics of Multi-Scale Methods

The goal of multi-scale formalisms, in general, is to provide computationally-efficient techniques for solving problems possessing a wide range of space and/or time scales. The need to do this has been recognized from the beginning of turbulence simulation, but in recent years it has been identified in many other areas (*e.g.*, laminar flow in porous media and solidification of crystals in liquid melts) giving rise to attempts to construct a general theory applicable across a broad spectrum of problem types (physical situations) and at the same time able to be analyzed by formal mathematical techniques. In the present section we provide a qualitative discussion of such methods.

To begin we should make clear precisely why computing over a wide range of scales is a "problem" in order to motivate approaches now being employed in constructing multi-scale methods. As we have already discussed, specifically in the context of turbulent fluid flow, if a complete solution is to be computed it is necessary to resolve all physically-relevant scales—DNS in this context. Moreover, as we have also already emphasized, the range of scales can be very large. For example, in considering prediction of lift or drag for an aircraft, the structure of turbulent ambient air contains scales ranging from at least tens of kilometers down to millimeters, and the flow containing these scales interacts with the surface of the aircraft on scales of tens of meters down to smaller than millimeters. Hence, there are seven to eight orders of magnitude in the range of spatial scales needed to accurately predict lift and drag on an aircraft flying through a cumulous cloud. This is presently not a tractable calculation, and various approximations must be made—which, of course, motivates

the need for computational procedures much more efficient than DNS but at the same time able to capture much of the physics currently only accessible via DNS.

But this represents only the beginnings of the difficulties. If, for example, heat transfer and/or chemistry are important parts of the physics then, especially in the latter case, far smaller spatial scales are needed; and relevant time scales can span a range of minutes down to elementary reaction times of 10^{-8} seconds, or less. Furthermore, it is now the case that different physics, requiring diverse mathematical representations, may occur in different ranges of scales. For example, at nanoscales the continuum hypothesis is no longer valid, and the N.–S. equations must be replaced with the Boltzmann equation.

It is clear in all such cases that a brute-force use of direct simulation will probably not be feasible for many years—some would argue, never—without radical changes in computing hardware architectures, and yet analysis of such problems is becoming increasingly important. The multi-scale formalisms now being developed represent attempts to address this issue.

There are several key ideas that are combined to produce multi-scale algorithms, in general. But not all such algorithms contain all of these, at least not explicitly. The first is to identify the important scales needed to sufficiently represent the problem and attempt to obtain equations representing the physics on each of these. In some cases, as noted by E and Engquist [131, 196], these equations may not be known; then one of the goals of a multi-scale formalism is to bypass the need for such unknown equations. This will not be the case for our specific purposes herein, and we will have little more to say regarding this aspect of multi-scale procedures.

After identifying the scales (and their appropriate mathematical representations), it is then necessary to construct efficient solution procedures on (for) each of these. As we will see, it is not usually efficient to employ the same type of algorithm on all scales; moreover, it is typical to resort to more approximation as inherent arithmetic complexity increases. In particular, models employed on the smallest scales where high resolution is most difficult are often the most approximate. But we remark that this can often be reasonably-well justified with both physical and mathematical arguments.

From an algorithmic standpoint, the final aspect to consider in a multi-scale formalism is combining results from the various represented scales to obtain a complete solution. In [196] it is pointed out that this is a highly nontrivial mathematical problem, and much research regarding this is in progress. Furthermore, we remark that there are numerous ways (some of them rather obvious) in which this can be done, but it must be emphasized, as done in [196], that essentially any procedure brings with it questions associated with stability and consistency of the overall (numerical) procedure. Thus, the final aspect of any multi-scale method is its mathematical analysis; it is important to be able to prove that solutions obtained from a multi-scale algorithm converge to those that would be obtained from a highly-accurate (fully-resolved) approach if this is appropriate. Of course, as already hinted, this is not always the case. We will not herein devote much space to such analyses despite their importance, as they are the subjects of current research, and instead we refer the reader to the cited literature.

Probably the oldest examples of multi-scale approaches are the homogenization and renormalization group (RNG) techniques. Aspects of these methods have motivated some of the formal analysis of modern multi-scale procedures. Moreover, as we will consider in more detail below, at least SGS models for LES, but possibly even RANS models, might be viewed as forms of homogenization—and some have been derived via RNG (see, *e.g.*, [93]). The basic idea in any such approach is to alter physical transport properties to enable taking into account physics on scales that have not been explicitly resolved by the discretization employed for the numerical solution process. That is, *homogenization* provides a modeling technique that accounts for the effects of small-scale physics without actually representing that physics in detail, as is also true of RNG models. Recall that this is precisely what takes place in both RANS and LES SGS models. At the same time it

must be recognized that there are formal mathematical requirements and techniques associated with homogenization—see Marchenko and Khuslov [197] (and RNG approaches), and these are not generally followed during construction of RANS and SGS models.

In the present discussions we will not specifically emphasize homogenization aspects of multi-scale methods but instead provide some basic heuristics. We begin with Fig. 3.14 which depicts a portion of a coarse-resolution grid on which regions of high resolution have been indicated. This

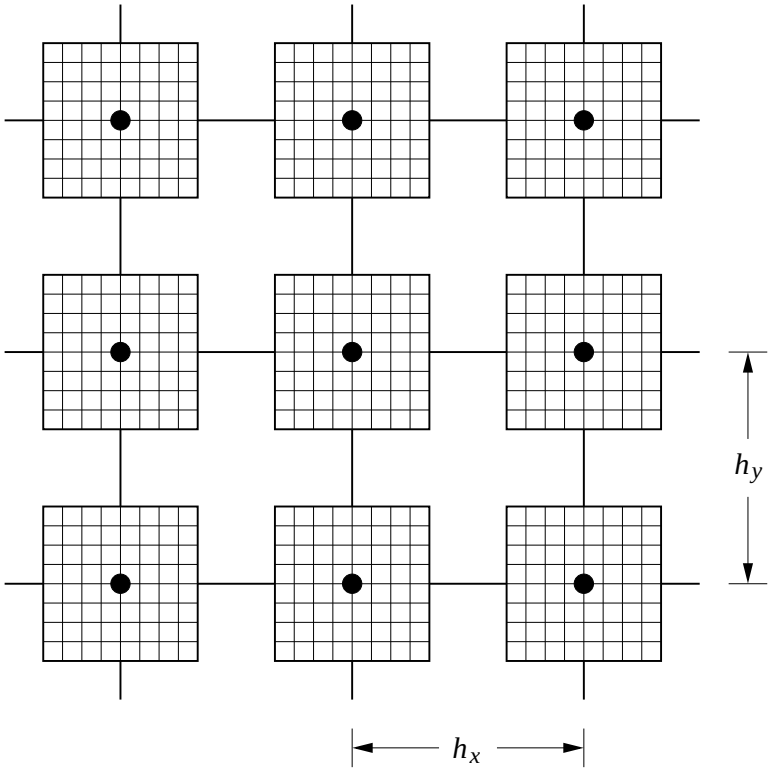

Figure 3.14: Example multi-scale gridding.

type of discrete structuring would be appropriate in situations for which the coarse-grid resolution is globally inadequate, but in which the indicated local resolution could not be supported globally. In principle, multi-scale formalisms provide techniques by means of which relatively inexpensive local fine-grid calculations can be combined with the also relatively inexpensive global coarse-grid results.

In particular, in the context of the grid displayed in Fig. 3.14, the high-resolution portion, as shown, covers only about one third of the complete domain area, so even if the fine-grid calculations were completely coupled, total arithmetic would be no more than about one-nineth that required on a complete high-resolution grid (in 2D—and relatively less in 3D). But these subdomains are not directly coupled. Hence, total arithmetic is further reduced, and, beyond this, the decoupling of subdomains provides an opportunity for parallelization to further improve computational efficiency.

It is clear that if the coarse- and fine-grid results can be combined correctly, it should be possible to improve the coarse-grid solution. The basic question is "How should this be done?" Here, we will discuss some possible alternatives. One can envision at least two main classes of approaches to this problem. First, one could perform local calculations on the small, highly-resolved subdomains and compare computed results at the embedded large-scale grid points to estimate required modifications to large-scale transport property values. This might be viewed as a form of homogenization. A second possibility is more closely related to function decompositions employed in LES (and in

RANS) and involves using fine-grid local calculations to directly enhance coarse-grid results. In either approach, calculations on the subdomains must be performed, and it is mainly the manner in which these are used that distinguishes the two techniques. We remark that mathematical analysis of these approaches is, for the most part, still lacking.

Thus, we begin with a description of these calculations. Because, as they are presented in Fig. 3.14, the small-scale subdomains are completely disjoint, and as a consequence the small-scale solution in one subdomain cannot "know" anything about the corresponding solutions in any neighboring domain. In physical situations dominated by diffusion one might argue that if the large-scale time step is less than the diffusion time between coarse-grid points, then this lack of communication is not a serious concern—at least if effects of this are somehow included in the coarse-grid calculations, *e.g.*, via alteration of transport properties (as in RNG or homogenization methods). But if the physics of the problem under consideration is not diffusion dominated, calculations based on completely separated high-resolution small-scale domains will probably be inaccurate, and possibly completely incorrect. Hence, a general approach should attempt to provide some communication between neighboring small-scale subdomains.

We can envision at least two specific situations, each arising from a different representation of the overall solution associated with the classes of techniques noted above. First, assume no decomposition of dependent variables has been performed. In this case the purpose of the small-scale, high-resolution subdomains is to directly achieve high resolution in these local regions—in a sense, a special case of adaptive mesh refinement (AMR, see, *e.g.*, Berger and Oliger [199]). In such cases we might interpolate the coarse-grid solution at a given time step onto the fine grids and solve on each of these using either "frozen" interpolated boundary conditions or periodicity conditions. It should be clear that gaining improved accuracy from such an approach is not entirely trivial. If improved temporal resolution is also required, the fine-grid calculations would be performed with smaller time steps than those used on the coarse grid. Furthermore, either the same, or different, governing equations might be employed on the two (or more) different scales.

Finally, when integrations on these high-resolution subdomains have been advanced to the current coarse-grid time, we must transfer the results to the coarse grid. Again, there are various alternatives for accomplishing this. The simplest is to directly use the result computed at the coarse-resolution grid point contained within each small-scale region as the advanced-time value. Observe that this is basically a predictor-corrector approach: predict on the coarse grid and correct on the fine grid. This, along with notions from multi-grid methods for solving linear systems, provides a starting point for analyses of this form of multi-scale technique. An alternative is to average spatially (and temporally, if appropriate) over each subdomain and apply this averaged result at each coarse-grid point (an heuristic form of renormalization).

If a dependent-variable decomposition has been employed (as in LES), then before interpolation of coarse-grid results to the fine-grid subdomains, these large-scale computations must be high-pass filtered to obtain data for the small-scale calculations. This is a form of deconvolution that has been widely used in LES constructions in recent years (see, *e.g.*, Adams and Stolz [198] and references therein). Options for treating problem formulation and boundary conditions on the high-resolution subdomains are the same as for the previous approach. But it is important to observe that details of the mathematical formulation must now be different because the fine-grid solutions now represent only a portion of the overall solution. Clearly, this must also be considered in processing computed fine-grid results prior to their combination with coarse-grid calculations to produce a complete solution. In particular, one must view the governing equations on each scale in light of an appropriate projection of some over-riding formulation, and for nonlinear problems this leads to terms from all scales in the equations representing each scale; see Hylin and McDonough [136] for a treatment of this.

Example: 1-D Burgers' equation To demonstrate construction of multi-scale methods we begin with a simple Burgers' equation problem introduced many years ago in [184]–[188]. This is given by

$$U_t + \left(U^2\right)_x - \nu U_{xx} = F(x,t)\,, \quad (x,t) \in (0,1)\times(0,t_f]\,, \tag{3.33}$$

with initial data

$$U(x,0) = U_0(x)\,, \qquad x \in [0,1]\,,$$

and Dirichlet boundary conditions

$$U(0,t) = U(1,t) = 0\,.$$

We might suppose that the forcing function $F(x,t)$ in Eq. (3.33) has been constructed such that it cannot be represented on a coarse grid, and hence, this will also be true of $U(x,t)$ if ν is sufficiently small. Of course, for a 1-D problem it is now nearly always the case that full resolution can be obtained, but even for 2-D problems this may not be true.

The grid to be considered is displayed in Fig. 3.15, a 1-D analogue of Fig. 3.14. For purposes of

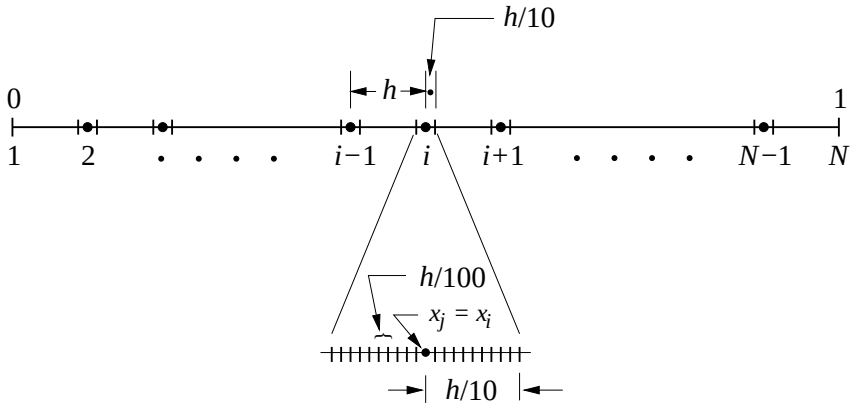

Figure 3.15: Example multi-scale gridding.

demonstration we have employed a large-scale grid of uniform spacing h with small-scale subdomains of size $h/5$ and uniform grid spacing $h/100$ centered on each large-scale grid point.

With a multi-scale grid in place, we can now consider how to solve the differential equation (3.33). We have earlier described some alternatives for doing this, and here we will employ a technique in which the dependent variables are decomposed, *a la* LES, but for which equations of the same form are solved on the (in this case) two different scales. This will constitute what would be called a *homogeneous multi-scale method* in the terminology employed in [131, 196]. The solution approach will be reminiscent of the *additive turbulent decompositions* (ATD) introduced in [184]–[188], and further studied by Yang and McDonough in 2-D [200] and for the compressible N.–S. equations in 1-D by McDonough and Wang [201]. Although there will be some differences in details (in part because there are many possible alternatives in implementing such procedures), we begin with the usual LES decomposition, expressed as

$$U(x,t) = \tilde{u}(x,t) + u^*(x,t)\,, \tag{3.34}$$

as in the cited works. We note that $U_0(x)$ should be similarly decomposed, for example via low-pass filtering, to produce $\tilde{u}_0(x)$ and $u_0^*(x)$, with the latter computed as

$$u_0^*(x,t) = U_0(x) - \tilde{u}(x,t)\,.$$

We now substitute (3.34) into (3.33) to obtain

$$(\tilde{u}+u^*)_t + \left((\tilde{u}+u^*)^2\right)_x - \nu(\tilde{u}+u^*)_{xx} = \widetilde{F}(x,t) + F^*(x,t)\,, \tag{3.35}$$

where we have also spatially low-pass filtered the forcing function F.

To make use of the proposed grid structure of Fig. 3.15 we must decompose Eq. (3.35) into separate equations to be employed on each of the coarse and fine grids. There are many alternatives for doing this (amongst them temporal averaging of RANS methods and spatial filtering of LES), but the ATD approach of the references cited above is particularly straightforward and possesses strong mathematical underpinnings, so we will use this here. Application of this to (3.35) results in

$$\tilde{u}_t + \left(\tilde{u}^2\right)_x + (2-\beta)(\tilde{u}u^*)_x - \nu\tilde{u}_{xx} = \widetilde{F}(x,t)\,, \tag{3.36a}$$

$$u^*_t + \left(u^{*\,2}\right)_x + \beta(\tilde{u}u^*)_x - \nu u^*_{xx} = F^*(x,t)\,. \tag{3.36b}$$

We remark that the *splitting parameter* β is not specified, *a priori*, but analysis by Brown *et al.* [202] of ATD applied to the 2-D N.–S. equations indicates that $\beta = 1$ is optimal for convergence rate of a Galerkin approximation in the context of the present Burgers' equation problem. Hence, in this case the large- and small-scale equations are identical in form.

It is also worthwhile to note that the decomposition of equations contained in Eqs. (3.36) is motivated by operator-splitting techniques such as described by Yanenko [203]. In this framework, we see that the sum of (3.36a) and (3.36b) is precisely the original undecomposed equation (3.33) when the LES decomposition of dependent variables, (3.34) is employed—despite the nonlinear term of Burgers' equation. Moreover, for $\beta \neq 0$ and $\beta \neq 2$, the large- and small-scale parts are coupled through terms analogous to the cross stresses of LES SGS models, but here these terms are computed directly, without modeling, via the small-scale equations of the multi-scale formalism.

At this point we have produced sufficient structure to permit construction of a coarse-grained solution algorithm associated with the multi-scale formalism.

Algorithm 3.1 *Suppose we have computed n time steps for a discretization of Eqs.* (3.36) *on a grid structure such as shown in Fig. 3.15. To advance the computed grid function $\{u_i\}_{i=1}^N$ to time level $n+1$, perform the following calculations.*

1. *Solve Eq.* (3.36a) *on the coarse grid of spacing h using $\widetilde{F}(x_i,t^{n+1}) = F(x_i,t^{n+1})$ evaluated at the coarse-grid points, and $u^*(x_i,t^{n+1}) = u^*(x_i,t^n)$.*

2. *Solve Eq.* (3.36b) *on each of the fine-grid subdomains of Fig. 3.15 employing interpolation of $\tilde{u}(x_i,t^{n+1})$ to the x_j locations required on the fine grid, and with $F^*(x_j,t^{n+1})$ directly evaluated using the given function $F(x,t)$ in conjunction with high-pass filtering. Note that if large-scale time steps are too large to resolve the temporal behavior of F, multiple small-scale time steps will be needed.*

3. *Calculate the complete solution at time level $n+1$ at the coarse-grid locations only using Eq.* (3.34) *and the results from steps 1 and 2:*

$$U(x_i,t^{n+1}) = \tilde{u}(x_i,t^{n+1}) + u^*(x_i = x_j, t^{n+1})\,.$$

We should provide several remarks regarding this procedure. First, setting $\widetilde{F}(x_i,t^{n+1}) = F(x_i,t^{n+1})$ in the first step may result in aliasing. Hence, formal filtering of $\tilde{u}(x_i,t^{n+1})$ may be necessary during each time step. (In fact, it may be desirable to filter F, itself, as suggested earlier.) Indeed, if $\tilde{u}$ is badly aliased, the interpolations required to begin step 2 will be very inaccurate. Second, in

step 3 of the algorithm we have employed the simplest possible approximation, namely, that u^* is taken directly from the small-scale solution at the grid point that coincides with the large-scale point at the large-scale time. There are numerous other alternatives. A fairly simple one would be to employ a spatially-averaged value of u^* with averaging over the whole subdomain, or over some subset near the large-scale point of application, and do this either at the final time of the small-scale time stepping, or use a time-averaged value. A somewhat more sophisticated approach might involve calculating the energy on the small-scale domain and then using a value of u^* in step 3 based on this. We note that little is presently known regarding relative merits of these various alternatives. Finally, we comment that if the small-scale subdomain is constructed to have a size equal to the large-scale interval on which it is centered, then the above algorithm is, in reality, simply a highly-parallelable form of DNS as studied in various of the cited references of McDonough and coworkers. In these works, however, a different numerical procedure (Galerkin) is employed for the small-scale calculations than for the large-scale ones (finite difference).

Synthetic-Velocity LES as a Multi-Scale Procedure

In this section we describe a relatively new form of LES and show that it can be categorized as a *heterogeneous multi-scale method* in the terminolgy of [131, 196]. We first observe that the term "synthetic velocity" refers to the fact that dependent variables, rather than their statistics, are explicitly modeled on the sub-grid scales. The form of these models varies widely, as has been alluded to above, and herein we will consider details of only one specific model, a modern version of the chaotic-map SGS model first introduced by McDonough *et al.* [178] with details provided in [136] and a summary in [73].

We begin by observing that there are three main differences between synthetic-velocity LES and the typical versions that have been considered up to the present time. These are:

i) solutions are filtered, rather than the equations of motion;

ii) SGS physical variables are modeled instead of modeling their statistics;

iii) subgrid-scale results are directly added to those from the resolved scale(s), as they should be, *a la* the LES decomposition.

There are numerous advantages arising from such an approach. By filtering solutions we immediately avoid having to model SGS stresses (statistics) and can focus attention on modeling physical variables. Furthermore, filtering solutions is far easier than filtering equations in the context of generalized coordinates; in particular, commutation errors arising from commuting differentiation and the filtering operator when it is applied to the differential equations are formally avoided. But we note that even when solutions are filtered some additional errors arise, and these can be significant in generalized coordinates. On the other hand, they are more easily analyzed in this form than in the equation-filtering approaches; moreover, this constitutes a direct numerical application of mollification, as done theoretically in analytical studies of PDEs.

By modeling physical variables we are able to directly produce interactions of flow physics with other phenomena (*e.g.*, interactions of turbulence with heat transfer and/or reaction chemistry) on sub-grid scales. It is essentially impossible to achieve this when employing techniques based on statistical correlations (*i.e.*, using SGS stresses and scalar fluxes) as in typical LES procedures. On the other hand, direct modeling of SGS physical variables is potentially very difficult; and use of a poor model is probably worse than applying no model at all, at least in the context of the methods employed herein. At the same time, there are many possible approaches to this SGS modeling problem as will be apparent in the sequel, and one should expect to be able to find one that is both relatively accurate and computationally efficient.

Finally, direct application of a physical-variable SGS model when constructing terms of an LES decomposition also brings with it both advantages and disadvantages. On the positive side is the potential for combining resolved-scale and modeled results in a way that does not formally demand scale separation and, in general, provides direct interaction between large-and small-scale physics. But it is this step, more so than the others, that requires a multi-scale mathematical formalism. In particular, as already noted, a poor model—and/or its incorrect implementation—could lead to very inaccurate, possibly even inconsistent, results (as well as destabilization of numerical methods employed), so rigorous mathematical analyses are crucial. At the same time, however, it is precisely this aspect of multi-scale methods that is least developed.

From the first of the ideas given above for constructing a synthetic-velocity LES procedure it is clear that the equations to be solved are simply the usual unfiltered N.–S. equations, which we repeat here for convenience:

$$\boldsymbol{U}_t + \nabla\cdot\big(\boldsymbol{U}^2\big) = -\nabla P + \nu\Delta\boldsymbol{U}\,, \qquad \boldsymbol{x}\in\Omega\subset\mathbb{R}^d\,,\ \ d=2,3 \tag{3.37a}$$

$$\nabla\cdot\boldsymbol{U} = 0\,, \tag{3.37b}$$

on a prescribed time interval $(0,t_f]$ with appropriate boundary and initial conditions. We remark that these are precisely the equations employed in ILES (see, *e.g.*, Fureby and Grinstein [133]), as mentioned earlier, where under resolution is controlled with a combination of monotone discretization of $\nabla\cdot\big(\boldsymbol{U}^2\big)$ and filtering of $\boldsymbol{U}$. Thus, the approach being presented here bears some resemblence to ILES in the latter regard.

But unlike ILES (and like ATD, treated in the preceding section), we retain the usual LES decomposition of dependent variables in the form given for solution to the Burgers' equation problem earlier in Eq. (3.34): *viz.*,

$$\boldsymbol{U}(\boldsymbol{x},t) = \widetilde{\boldsymbol{u}}(\boldsymbol{x},t) + \boldsymbol{u}^*(\boldsymbol{x},t)\,. \tag{3.38}$$

This is formally substituted into Eqs. (3.37), as done in ATD; but instead of splitting the resulting equation(s) into large- and small-scale parts, we construct models to permit direct evaluation of $\boldsymbol{u}^*(\boldsymbol{x},t)$ via efficient algebraic expressions that hold locally in space and time. We remark that the actual numerical procedure must be somewhat more involved, and in particular, it must include a filtering process. At each time step this is applied during calculation of $\widetilde{\boldsymbol{u}}$. Then $\boldsymbol{u}^*$ is computed via a procedure to be described below, and appeal to the LES decomposition (3.38) is made to obtain $\boldsymbol{U}$. These basic steps comprise one time step of the discrete solution operator.

The large-scale calculations are carried out with a typical projection method, and filtering is employed prior to the projection step (since this step involves only linear operators) to guarantee that the filtered velocity field is divergence free (see, *e.g.*, McDonough [204]). It is the small-scale calculations that require considerable description, and we begin this here.

Small-Scale Calculations, General. In particular, we will develop models for constructing the small-scale velocity field $\boldsymbol{u}^*$ (and, in principle, any scalar quantities associated with problem physics—except for the small-scale pressure which, for incompressible flow, will be computed in the usual way via projection in order to satisfy Eq. (3.37b) on sub-grid scales). A basic hypothesis regarding this construction is that any small-scale flow variable can be expressed as

$$q_i^* = A_i M_i\,, \qquad i=1,\ldots\,,N_v\,, \tag{3.39}$$

where q_i^* is the i^{th} one of N_v small-scale dependent variables; A_i is its amplitude (employed locally in space, and during the current large-scale time step); and M_i is a chaotic map that can exhibit bifurcations leading to a strange attractor, thus producing small-scale turbulent temporal fluctuations, also locally in space and time. Equation (3.39) is a modification, first introduced by McDonough [205], of the form of the SGS model proposed in [178].

This form can be motivated as follows. First, we know the N.–S. equations admit solutions in the form of Fourier series (recall discussions in Chap. 1) as suggested earlier in Eq. (3.5), and obviously their remainders (the small-scale q_i^*s) can also be expressed in this way:

$$\begin{aligned} q_i(\boldsymbol{x},t) &= \sum_{|\boldsymbol{k}|}^{\infty} a_{\boldsymbol{k},i}(t)\varphi_{\boldsymbol{k}}(\boldsymbol{x}) \\ &= \sum_{|\boldsymbol{k}|\le \boldsymbol{k}_c} a_{\boldsymbol{k},i}(t)\varphi_{\boldsymbol{k}}(\boldsymbol{x}) + \sum_{|\boldsymbol{k}|=\boldsymbol{k}_c+1}^{\infty} a_{\boldsymbol{k},i}(t)\varphi_{\boldsymbol{k}}(\boldsymbol{x}) \\ &= \tilde{q}_i(\boldsymbol{x},t) + q_i^*(\boldsymbol{x},t)\,. \end{aligned}$$

Now if we consider this representation restricted to a small subdomain and recognize that, in general, the number of wavevectors required for adequate representation falls significantly as the extent of the spatial domain is reduced, we see that if h_s is sufficiently small and $\boldsymbol{k}_c$ is sufficiently large, we should be able to represent q_i^* with at most a few well-chosen wavevectors. (Here, h_s is a characteristic size of a small-scale domain such as shown in Fig. 3.15, and $\boldsymbol{k}_c$ is the chosen cut-off wavevector, which depends on large-scale grid spacing.)

Indeed, we currently choose a single wavevector $\boldsymbol{k}$ so that the above collapses to

$$q_i^*(\boldsymbol{x},t) = a_{\boldsymbol{k},i}(t)\varphi_{\boldsymbol{k}}(\boldsymbol{x})\,.$$

But this is only formal, and in practice our models include additional wavenumbers, as will be apparent as we proceed. This representation, however, is required to hold only on a small subdomain within which evaluation of $\varphi_{\boldsymbol{k}}(\boldsymbol{x})$ produces an $\mathcal{O}(1)$ constant (which we do not need to explicitly determine). Then we have

$$q_i^*(\boldsymbol{x},t) \cong C a_{\boldsymbol{k},i}(t)\,.$$

If we determine the magnitude of $a_{\boldsymbol{k},i}$, we can write this in the form (3.39) where we now have

$$A_i = C|a_{\boldsymbol{k},i}|\,, \tag{3.40}$$

and values of $M_i = M_i(t)$ are restricted to $[-1,1]$. In what follows, we will provide details of calculating both the A_is and M_is.

Amplitude Factors. Construction of the A_is at each grid point of the large-scale solution begins by recalling the Parseval identity,

$$\|q_i^*\|^2 = |A_i|^2 = \sum_{|\boldsymbol{k}|=\boldsymbol{k}_c+1}^{\infty} |a_{\boldsymbol{k},i}|^2 \simeq |a_{\boldsymbol{K},i}|^2\,, \tag{3.41}$$

and recognizing that this corresponds to energy—in fact, an energy increment in the sense of Kolmogorov structure functions—due to restriction to high-wavevector content and the fact that $|M_i| \le 1$. Here, $\boldsymbol{K}$,i represents a wavevector, usually fairly deep within the inertial subrange, that must be determined at each discrete large-scale grid point (and for each time step) for each i^{th} dependent variable.

Now observe that $|a_{\boldsymbol{K},i}|^2 \sim E_{\boldsymbol{K},i}$ is energy of q_i^* in the L^2 sense, and $E_{\boldsymbol{K},i}$ has a direct relationship to second-order Kolmogorov structure functions of q_i^*. These can be computed (approximately) from the high-pass filtered $\tilde{q}_i$s. In particular, similar to scale-similarity methods discussed earlier, we can compute

$$q_i^{**} = \tilde{q}_i - \tilde{\tilde{q}}_i\,,$$

and construct the second-order structure functions

$$S_2(q_i^{**}, r) \simeq \langle (q_i^{**}(\boldsymbol{x} + \boldsymbol{r}) - q_i^{**}(\boldsymbol{x}))^2 \rangle \,, \qquad i = 1, \ldots, N_v \,, \tag{3.42}$$

where the average (denoted by $\langle \cdot \rangle$) is taken over all discrete points of the large-scale finite-difference grid at a distance r from the current point $\boldsymbol{x}_{i,j,k}$ in a 3-D cube containing 27 grid points with $\boldsymbol{x}_{i,j,k}$ at the center. Note that with uniform gridding in all three directions, as depicted (but not-to-scale) in Fig. 3.16, there will be only four possible values of $r = r_m$, $m = 1, 2, 3, 4$, with only six, four,

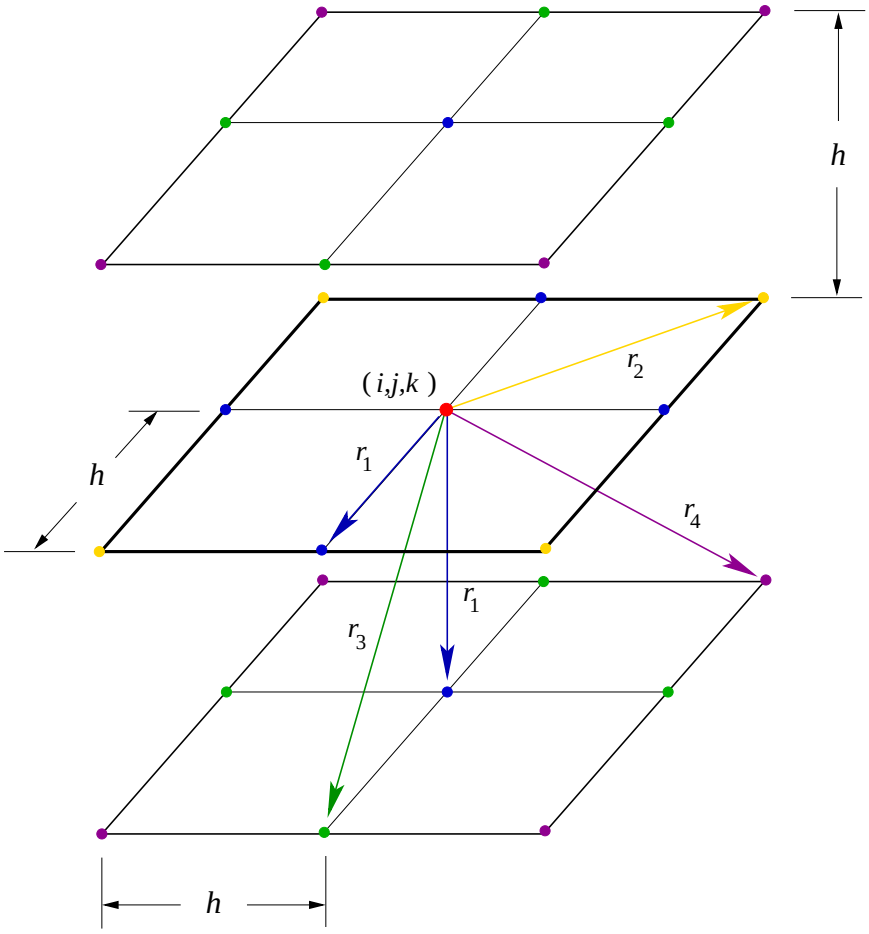

Figure 3.16: Grid point locations used for second-order structure function averaging in the 3-D uniform-grid case.

eight and eight samples, respectively, for these. As a consequence, statistics (the averages) are sometimes not as representative as would be desired, and it should be clear that this problem is exacerbated on non-uniform grids. There are several ways in which this can be remedied—the main one being construction of averages over larger numbers of grid cells, but we will not pursue this implementation issue here.

Next, we recall the power-law form of structure function scaling in the framework of the Kolmogorov K41 theory, which was given earlier (up to notation) in Eq. (2.92) as

$$S_2(r) = C \left(\langle \varepsilon \rangle r \right)^{2/3}$$

for homogeneous isotropic turbulence at high Re. But, in general, we cannot expect any of these conditions to hold at an arbitrary location in an arbitrary flow field. (Moreover, this result is

specifically for longitudinal structure functions although it generally holds for transverse ones as well; but Eq. (3.42) is neither of these.) On the other hand, at least locally with respect to values of the independent variables, most quantities in physics can be well represented by a power law. Thus, we replace the above with

$$S_{2,i}(r) = C\left(\langle\varepsilon\rangle r\right)^{\beta} \tag{3.43}$$

locally in space and time, and for each individual dependent variable, $i = 1, \ldots, N_v$. Now observe that for each value of r we can compute $S_{2,i}(r)$ if we know $\langle\varepsilon\rangle$, and this can be obtained directly from its definition, Eq. (1.53), applied to the high-pass filtered $\tilde{q}_i$ (a form of deconvolution [198]) and averaged over the 27 grid points centered on the (i, j, k) point indicated in Fig. 3.16. Then the general exponent β and the constant C in Eq. (3.43) can be found using a least-squares minimization of the form find β and C such that

$$\sum_{m=1}^{4}\left[S_{2,i}(q_i^{**}, r_m) - C\left(\langle\varepsilon^{**}\rangle r_m\right)^{\beta}\right]^2$$

is a minimum, where the values of the $S_{2,i}$ are directly calculated from the definition, Eq. (3.42), and are averaged over the number of entries for each r_m. Clearly, this minimization problem can be expressed in terms of simple closed-form analytical formulas due to use of the power-law form. Hence, even though this computation must be done at every resolved-scale grid point, at every time step for all dependent variables, there is only quite minimal arithmetic needed. Moreover, calculations at each grid point are independent of all others, so the process is easily parallelized.

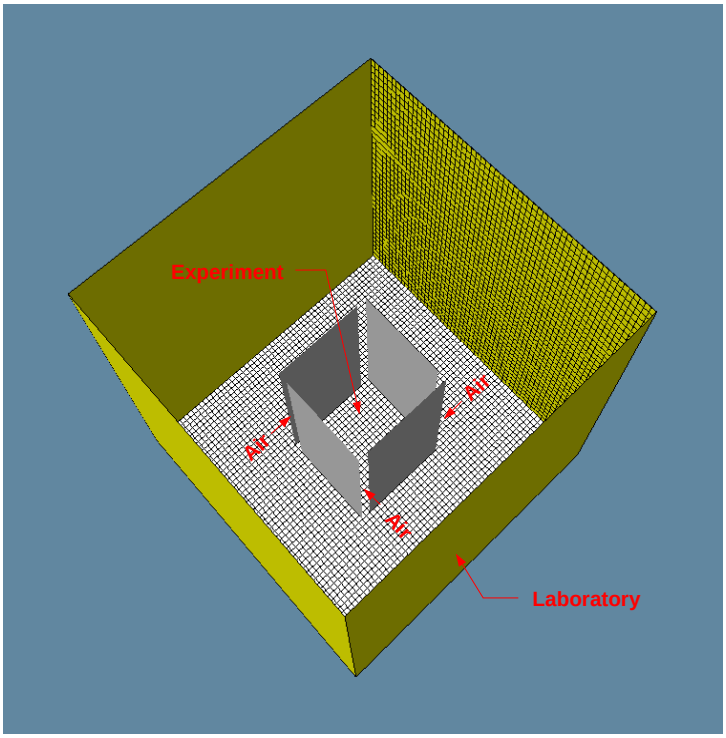

Figure 3.17: Physical model employed for synthetic-velocity LES of swirling, buoyant plume in an enclosure.

It is worthwhile to study details of the values of β and C that arise from this formulation during synthetic-velocity LES calculations that employ all of the above noted aspects. McDonough and Yang [72] have reported one such calculation which we summarize here. The physical problem, the domain of which is sketched in Fig. 3.17, consisted of a turbulent, swirling buoyant plume confined

to a square enclosure 0.9m wide by 1.8m tall with 0.1m slits in each corner off set to establish the swirl direction—all located in the center of a closed (including ceiling) room (laboratory).

The resolved-scale calculations were quite coarse: $61\times 41\times 61$ uniformly-spaced grid points in a physical domain that is 3m in each direction. Spatial planes of the solution and other variable values (such as β and C) were recorded at various times and locations during a simulation that ultimately reached a nearly stationary state. Figures 3.18(a,b) display such data for β and C, respectively, in a horizontal xz plane 1m above the floor of the enclosure.

It should be clear that values of β differ from the theoretical value of 2/3 for homogeneous, isotropic turbulence—as they should. In fact, most of the flow outside the central square region of the experimental apparatus is nearly stagnant, and is certainly not turbulent, except within the first grid cell adjacent to the laboratory wall where transition to a turbulent boundary layer is occurring due to down flow of cooled gases emanating from a "wall" jet spreading across the ceiling of the laboratory from the center of the plume.

It is clear that values of β cover a range from what would be expected for integral-scale behavior ($\beta = 0.0$) to values corresponding to the beginning of the dissipation scales ($\beta = 3.0$). Moreover, values near 2/3 occur only within the experimental apparatus containing the buoyant plume. Similar observations can be made for the constant C. In particular, $C \simeq 0$ holds in most of the flow field outside the experimental region, thus showing that the model automatically produces zero amplitude small-scale results when there is no turbulence.

The final step in computing values of the amplitude factors involves relating $S_2(r)$ to $E(k)$ $(= E_{K,i})$, as is done for the Kolmogorov 5/3 law, for each dependent variable. In particular, since $r \sim 1/k$, we see that the general power law (3.43) leads to

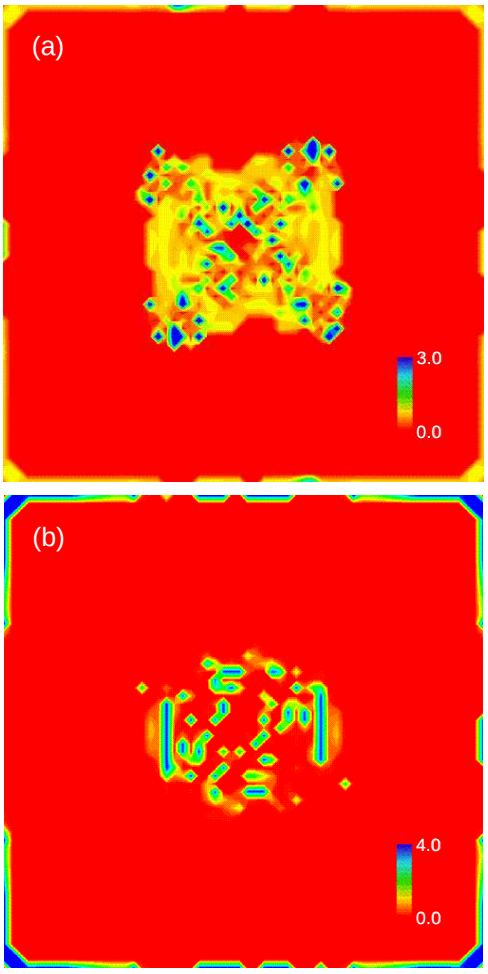

Figure 3.18: Kolmogorov exponent and constant in plane 1m above floor; (a) exponent, β, and (b) constant, C.

$$E(k) = C\langle\varepsilon\rangle^{\beta}k^{-(\beta+1)}\,. \tag{3.44}$$

Then from (3.41) we have

$$A_i = [E_i(k_i)]^{1/2}\,.$$

We observe that once K has been determined, we can replace the above with

$$A_i = \left[\sum_{k=K}^{K+N} E_i(k_i)\right]^{1/2}\,, \tag{3.45}$$

where N is prescribed (and, consistent with earlier discussion, relatively small—typically, no more than 10).

Currently, we calculate K locally in both space and time by finding the Taylor microscale length in each resolved-scale grid cell using Eq. (1.56):

$$\lambda = \left[\frac{\nu \langle |\boldsymbol{u}'|^2 \rangle}{\varepsilon} \right]^{1/2} .$$

Then for each separate velocity component we calculate

$$\lambda_i \simeq \left[\frac{\nu \langle u_i^{**\,2} \rangle}{\langle \varepsilon \rangle} \right]^{1/2} , \qquad i = 1, 2, 3 , \tag{3.46}$$

where u_i^{**} is the high-pass filtered i^{th} velocity component. Then the wavevector components are $1/\lambda_i$, resulting in

$$K = \left[\sum_{i=1}^{3} \left(\frac{1}{\lambda_i} \right)^2 \right]^{1/2} . \tag{3.47}$$

Observe that working with individual velocity components and calculating component λ_is removes any need for homogeneity and/or isotropy assumptions. Our only assumption has been that the local in space and time physics (represented as a "general" second-order structure function) can be reasonably well approximated with a power law.

Temporal Fluctuations. Our next task is to find formulas for the M_is. We remark that in the earliest synthetic-velocity models produced by McDonough and students, these were calculated from linear combinations of logistic maps, as analyzed in detail in [136] and in Hylin [206]. This approach was motivated by an epithet of Frisch [80], mentioned earlier, "poor man's Navier–Stokes equation," applied to a quadratic map that can be transformed to a logistic map, and also by the success, demonstrated by Mukerji *et al.* [44], in fitting physical chaotic measurements to linear combinations of logistic maps. It was, however, later discovered by the present author in analyzing results containing heat transfer in addition to fluid flow, that this approach could not be guaranteed to work. Beyond that is the fact that it is rather *ad hoc* and difficult to justify, beyond the fact that it seemed to work in some specific situations.

This motivated McDonough and Huang [81] to seek a replacement for the logistic map as the source of small-scale temporal fluctuations, especially in problems where additional physics (combustion chemistry and heat transfer in the cited reference) must be modeled. The approach taken was to begin with a Galerkin representation of the N.–S. equations such as Eqs. (1.42) and (1.44) (initially in 2D) and any others needed for complete specification of problem physics. This was simplified to a Leray-projected form not including pressure gradient terms, and the result was written for a single wavevector for each momentum equation (and other equations, as appropriate). In the case of the 3-D N.–S. equations, the Fourier representation of, *e.g.*, the x-momentum equation is

$$\sum_{\boldsymbol{\ell}} \dot{a}_{\boldsymbol{\ell}} \varphi_{\boldsymbol{\ell}} + \sum_{\boldsymbol{\ell},m} (\ell_1 + m_1) a_{\boldsymbol{\ell}} a_{\boldsymbol{m}} \varphi_{\boldsymbol{\ell}} \varphi_{\boldsymbol{m}} + \sum_{\boldsymbol{\ell},m} (\ell_2 + m_2) a_{\boldsymbol{\ell}} a_{\boldsymbol{m}} \varphi_{\boldsymbol{\ell}} \varphi_{\boldsymbol{m}} + \sum_{\boldsymbol{\ell},m} (\ell_3 + m_3) a_{\boldsymbol{\ell}} a_{\boldsymbol{m}} \varphi_{\boldsymbol{\ell}} \varphi_{\boldsymbol{m}} = -\frac{1}{Re} \left(\sum_{\boldsymbol{\ell}} |\boldsymbol{\ell}|^2 a_{\boldsymbol{\ell}} \varphi_{\boldsymbol{\ell}} \right) ,$$

and forming the Galerkin inner product with $\varphi_{\boldsymbol{k}}$ yields

$$\dot{a}_{\boldsymbol{k}} + \sum_{\boldsymbol{\ell},\boldsymbol{m}} A^{(1)}_{\boldsymbol{k}\boldsymbol{\ell}\boldsymbol{m}} a_{\boldsymbol{\ell}} a_{\boldsymbol{m}} + \sum_{\boldsymbol{\ell},\boldsymbol{m}} B^{(1)}_{\boldsymbol{k}\boldsymbol{\ell}\boldsymbol{m}} b_{\boldsymbol{\ell}} a_{\boldsymbol{m}} + \sum_{\boldsymbol{\ell},\boldsymbol{m}} C^{(1)}_{\boldsymbol{k}\boldsymbol{\ell}\boldsymbol{m}} c_{\boldsymbol{\ell}} a_{\boldsymbol{m}} = -\frac{C|\boldsymbol{k}|^2}{Re} a_{\boldsymbol{k}} ,$$

where, for example,

$$A^{(1)}_{\boldsymbol{k\ell m}} \equiv (\ell_1 + m_1) \int_\Omega \varphi_{\boldsymbol{k}} \varphi_{\boldsymbol{\ell}} \varphi_{\boldsymbol{m}} \, d\boldsymbol{x} \, .$$

Then restricting the above to a single wavenumber and writing the result for all three momentum equations leads to

$$\begin{aligned}
\dot{a} + A^{(1)} a^2 + B^{(1)} ab + C^{(1)} ac &= \frac{|\boldsymbol{k}|^2}{Re} a \, , \\
\dot{b} + A^{(2)} ab + B^{(2)} b^2 + C^{(2)} bc &= \frac{|\boldsymbol{k}|^2}{Re} b \, , \\
\dot{c} + A^{(3)} ab + B^{(3)} bc + C^{(3)} c^2 &= \frac{|\boldsymbol{k}|^2}{Re} c \, .
\end{aligned}$$

These equations are numerically integrated using simple Euler methods (usually forward, but sometimes backward for passive scalars), and the result is subjected to two transformations, one due to McDonough and Huang [81, 82] to introduce the logistic map, and a second due to May [193] to obtain the usual logistic-map scaling of the bifurcation parameters. The result is a true "poor-man's Navier–Stokes equation:"

$$a^{(n+1)} = \beta_1 a^{(n)} \left(1 - a^{(n)}\right) + \gamma_{12} a^{(n)} b^{(n)} + \gamma_{13} a^{(n)} c^{(n)} \, , \tag{3.48a}$$

$$b^{(n+1)} = \beta_2 b^{(n)} \left(1 - b^{(n)}\right) + \gamma_{21} a^{(n)} b^{(n)} + \gamma_{23} b^{(n)} c^{(n)} \, , \tag{3.48b}$$

$$c^{(n+1)} = \beta_3 c^{(n)} \left(1 - c^{(n)}\right) + \gamma_{31} a^{(n)} c^{(n)} + \gamma_{32} b^{(n)} c^{(n)} \, , \tag{3.48c}$$

where the β_is and γ_{ij}s, $i, j = 1, 2, 3$, $i \neq j$, are bifurcation parameters.

We remark that one should be concerned about the large number of bifurcation parameters and how they might be evaluated—is this the turbulence closure problem in just another form? But one readily sees from details of the derivation of Eqs. (3.48) given in 2D in [81, 82] that all of these parameters can be directly related to physical variables which, in turn, can be calculated from the high-pass filtered large-scale solution as has already been done in constructing the A_is. In particular, the transformations of [81, 82] and [193] show that

$$\beta_i = 4 \left(1 - \frac{\tau |\boldsymbol{k}|^2}{Re_i} \right) \, , \qquad i = 1, 2, 3 \, , \tag{3.49}$$

where in the context of SGS models the Re_i are small-scale component Reynolds numbers defined as

$$Re_i \equiv \frac{h^2 |\partial u_i^{**} / \partial x_i|}{\nu} \, ,$$

with no Einstein summation and h being the discretization step size for the large-scale numerical approximation. Also, for example,

$$\gamma_{21} \equiv \tau A^{(2)} = \tau A^{(2)}_{\boldsymbol{k\ell m}} = \tau A^{(2)}_{\boldsymbol{k\ell m}} \equiv (\ell_2 + m_1) \int_\Omega \varphi_{\boldsymbol{k}} \varphi_{\boldsymbol{\ell}} \varphi_{\boldsymbol{m}} \, d\boldsymbol{x} \, . \tag{3.50}$$

In Eqs. (3.49) and (3.50) τ is the small-scale time scale, obtained in dimensional form as the reciprocal strain rate norm of the high-pass filtered resolved-scale velocity. McDonough *et al.* [207] show that it is possible to estimate the γ_{ij}s by $\tau \partial u_j^{**} / \partial x_i$, and there are other possible efficient approximations. Thus, up to implementational details, we can consider the β_is and γ_{ij}s known at each large-scale grid point, and for each instant of time—independent of any "adjustable constants."

It is interesting at this point to view some of the values taken on by these automatically-computed bifurcation parameters. Results shown here are again taken from [72]. Figures 3.19(a,b) display spatial distributions at a single instant of time for the parameters β_2 and γ_{21} of the vertical momentum equation (3.48b) in this case (which here contains a typical natural convection body-force term with bifurcation parameter corresponding to Ra, see [72]). We observe that with the scalings employed, we would expect the β_is to lie in the interval $[0, 4]$; although it is possible for values to lie outside this range, it is seldom that they lead to stable stationary behavior of the corresponding attractor of the DDS. The γ_{ij}s in 3-D do not have theoretically-known restrictions, but numerical experiments (in 2-D) show that they are typically (but not always) in the interval $(-1, 1)$.

In general, we expect relatively large (near $\beta = 4$) values of the β_i to imply chaotic behavior of the PMNS equation, but this depends on values of all other bifurcation parameters, as shown, *e.g.*, in [207]. In particular, it is possible for chaotic dynamics to be exhibited with (at least one) β_i as low as unity in some cases. Moreover, as will be clear from a later regime map, for any particular value of γ_{ij}, essentially any type of behavior can occur as the β_i are varied.

Figure 3.19(a) indicates relatively large values of β_2 only within the experimental apparatus where the swirling buoyant plume contributes to turbulence. It is also interesting to note that, while there is an apparent approximate geometric symmetry of the β_2 distribution, this is indeed only approximate. Effects of turbulent fluctuations from the preceding time step are captured in the high-pass filtered $\boldsymbol{u}^{**}$ leading to loss of detailed symmetry, just as should be expected in actual fluid physics. Moreover, as should be expected, details of these distributions change form plane to plane, and, of course, they vary in time. In Fig. 3.19(b) we see generally similar, but more pronounced, asymmetries. In addition, it is clear that all values of γ_{21} are relatively small (as should be expected in a buoyant plume), and generally larger values appear in the central apparatus and in the boundary layers on the laboratory walls. In general, these features are all consistent with the physics expected for this flow field.

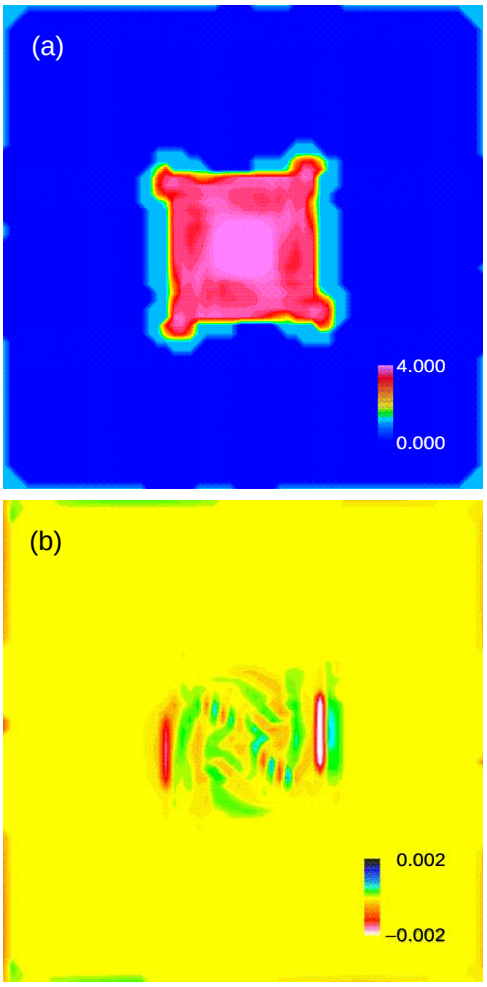

Figure 3.19: Bifurcation parameters in plane 0.9m above floor; (a) β_2, and (b) γ_{21}.

Further Features of PMNS Equation. It is useful to here consider some further properties exhibited by the discrete dynamical system we term the poor man's Navier–Stokes equation. This will demonstrate the degree of realism that is possible when using this factor in our proposed SGS model. We begin with a direct comparison of time series produced by the 2-D PMNS with thermal energy equations, and corresponding buoyancy term in the y-momentum equation, with data from the Gollub and Benson [58] Rayleigh–Bénard convection experiments. These measured time series shown in Figs. 3.20 were of the horizontal velocity component and present a bifurcation sequence consisting of periodic, quasiperiodic, phase-locked, noisy quasiperiodic and turbulent regimes in parts (a) through (e) as the Rayleigh number, Ra, is increased. Parts (f) through (j) present the corresponding PMNS equation time series generated by changing only the buoyancy-related bifur-

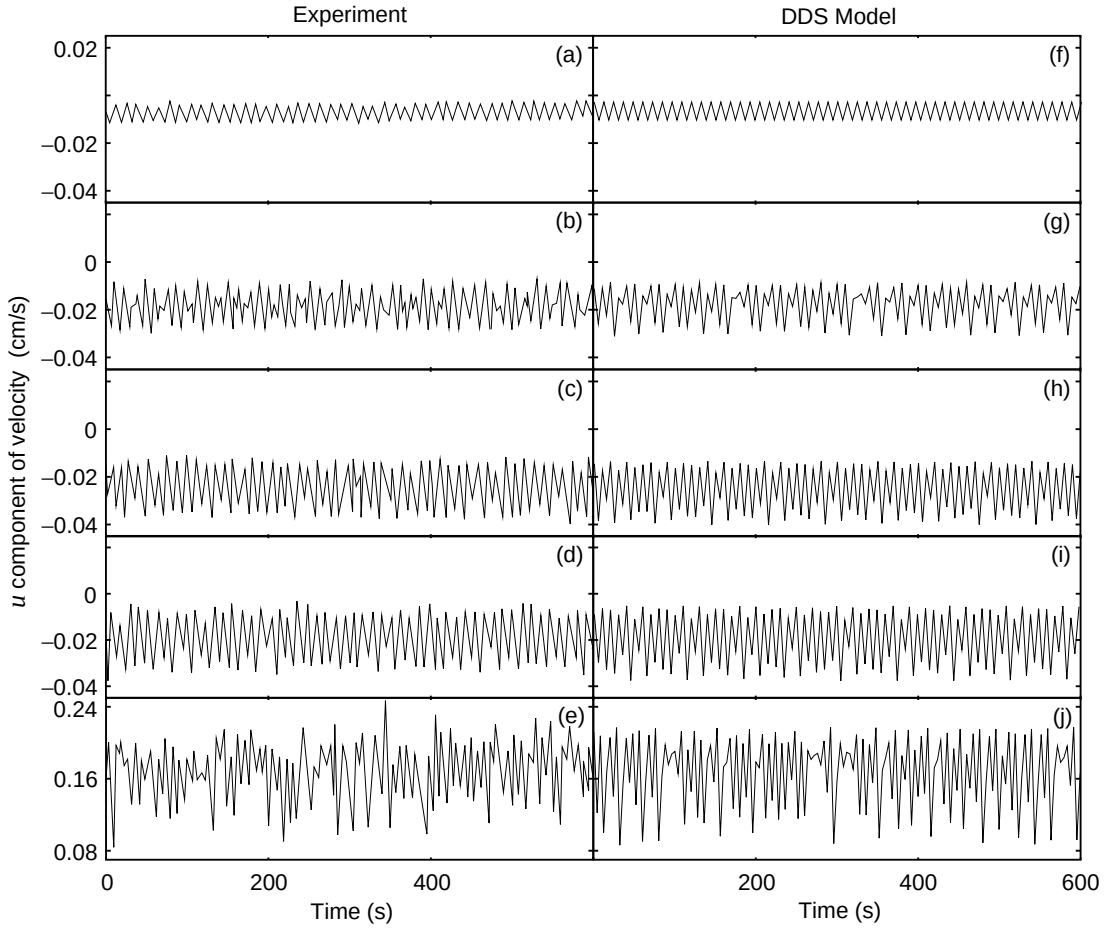

Figure 3.20: Left side is sequence of experimental time series from [58]; right side presents corresponding computational time series from the PMNS equation DDS.

cation coefficient in the y-momentum equation except in part (j) where some strain rate parameters were also altered to better account for the enhanced turbulence. The significance of this result is twofold: first, it shows that physically-realistic time series can be produced by the PMNS equation (even in the presence of a scalar quantity), and second, the set of time series demonstrates that this easily-evaluated discrete dynamical system has produced a bifurcation sequence completely analogous to the physical one by systematically varying the corresponding bifurcation parameter. We note that the PMNS results have not been scaled in any way. Their amplitudes grow with increasing bifurcation parameter values just as occurs physically.

We also remark that, as pointed out in [58], the location and details of the phase-locked regimes within the bifurcation sequence were not completely reproducible in repetitions of the experiments, and the authors offered no explanation for this. Simulating the experiments in detail to study this interesting effect via DNS would have been a completely insurmountable task in the late 1970s when they were performed, and even today, enormous computational resources would be needed for a thorough study. But it turns out that examination of easily-generated bifurcation maps of the PMNS equation (plus thermal energy in this case) leads to an explanation. In particular, the boundaries in β_T-α_T space between quasiperiodic and phase-locked behaviors are quite complicated, as can be seen from Fig. 3.21 which displays a bifurcation diagram for each of the two velocity components and temperature in parts (a) through (c). Here β_T and α_T are the bifurcation parameters corresponding to diffusive terms in the thermal energy equation (the Péclet number, Pe) and to the buoyancy

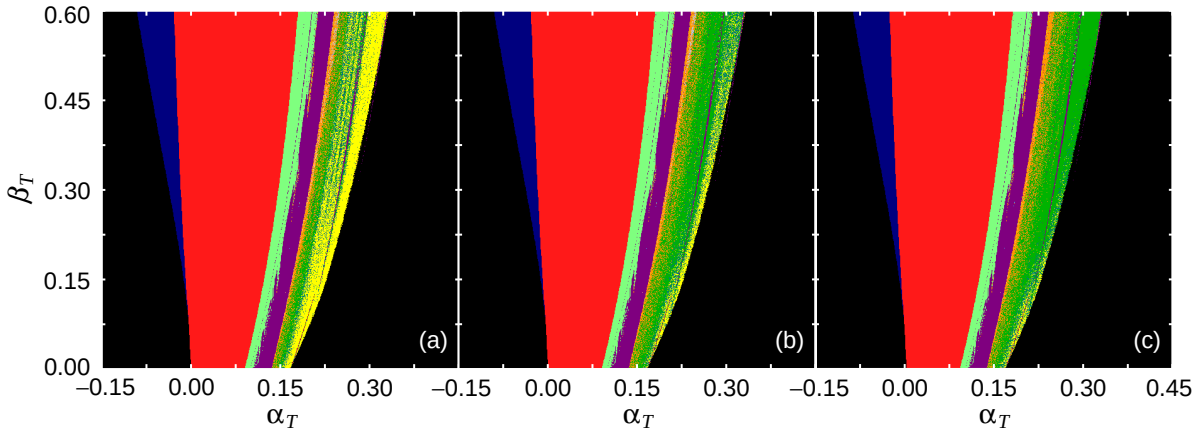

Figure 3.21: PMNS plus thermal energy equation β_T *vs.* α_T bifurcation diagrams; (a) horizontal velocity, (b) vertical velocity, and (c) temperature.

term of the vertical momentum equation (Rayleigh number, Ra), respectively. This implies that a slight change in either β_T or α_T can lead to a qualitatively different regime, and such changes are analogous to slight changes that are inevitable in experimental conditions in repetitions of experiments. The detailed color table is provided in a later figure. Here, it is sufficient to recognize that purple contours correspond to phase lock, and light and darker green represent quasiperiodic and noisy quasiperiodic behaviors, respectively.

It is also interesting to observe from Figs. 3.21 that the flow is steady (dark blue), corresponding to stable stratification, when $\alpha_T \leq 0$ if β_T is relatively low. But as β_T (Pe) increases, time-dependent periodic behavior (red) can occur since there is insufficient dissipation in the thermal energy equation to completely damp the turbulent oscillations of the momentum equations. In moving from left to right in each individual part of the figure (increasing Ra), one can see evidence of a Ruelle and Takens bifurcation sequence, but with the added features of phase-locked and noisy quasiperiodic state (just as reported by Gollub and Benson [58]). At the far right of the stable portion of each of these one sees yellow contours which we hypothesize might be related to so-called "strong" turbulence in natural convection (see discussion of Cioni *et al.* [191] data below). Finally, with regard to this figure, we see that the individual bifurcation diagrams are all quite similar, but with horizontal velocity showing a significantly larger region of strong turbulence. This detail seems to contradict the Takens theorem [75], suggesting that careful experiments should be undertaken to determine its (range of?) validity. Associated with this is the somewhat counter-intuitive result that vertical velocities appear to be possibly less chaotic than horizontal ones. But one must recall that these diagrams are local in space and time, and moreover include effects from six other physically-based bifurcation parameters associated with local strain rates, scalar fluxes and Re, all of which are held fixed during calculations leading to Fig. 3.21. Thus, we contend that this aspect of the figure is to be expected and probably is of little intrinsic significance.

Our second result shows that the PMNS equation also shows favorable comparisons with DNS. Figure 3.22 displays DNS time series, again for a free convection problem, due to Paolucci [208]. The important aspect of this comparison is the ability of the PMNS equation to reproduce the time series "structures" that appear in direct numerical simulations, and even more, the sequencing of these structures within a given time series. We have not in this case attempted to scale the PMNS computed results, so magnitudes of results do not match those of the DNS. We also note that both the DNS and the PMNS calculation were two dimensional in this case. In both cases (DNS and PMNS) we see strong correlation, in a qualitative sense, among the three solution components, as

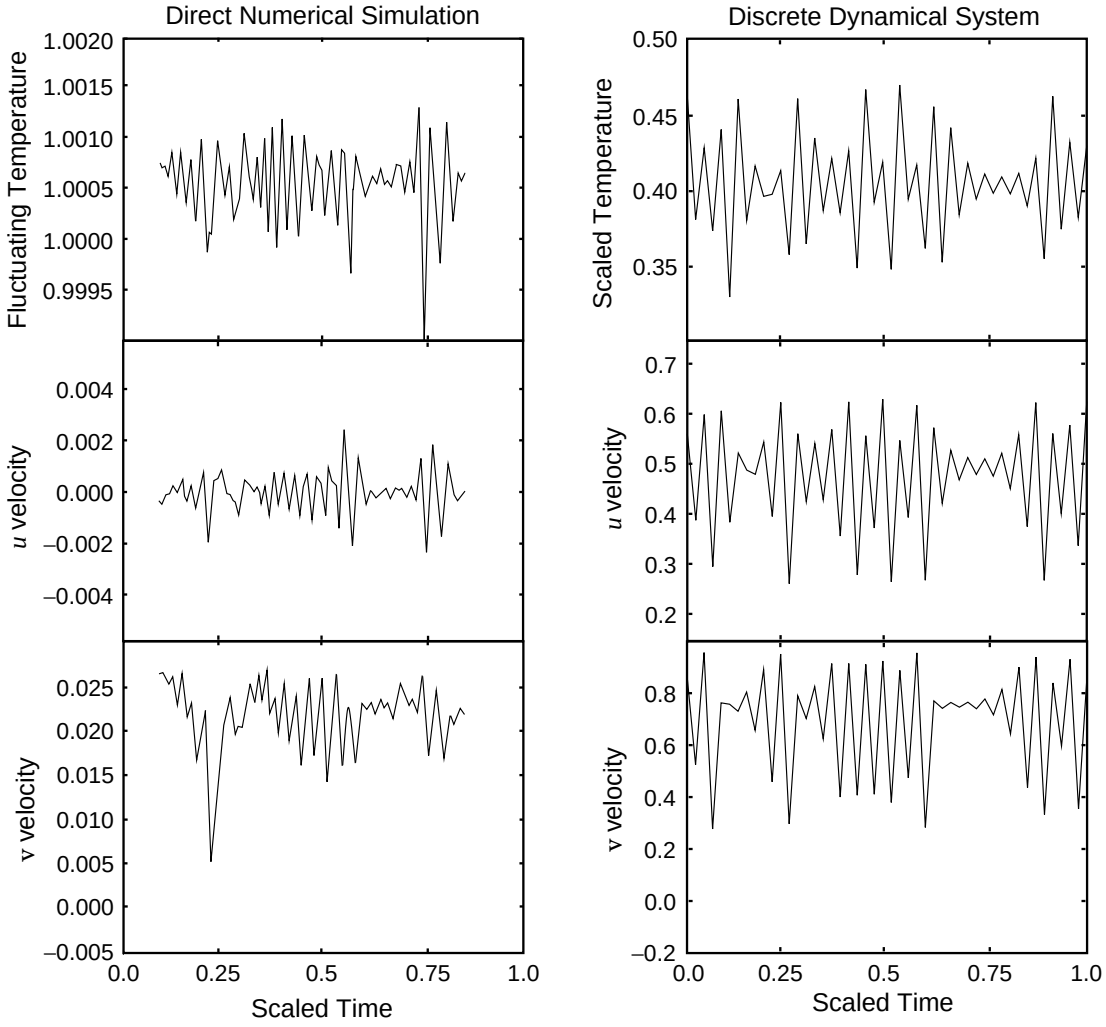

Figure 3.22: Comparison of PMNS plus thermal convection DDS with DNS.

should be expected for natural convection. Moreover, we remark that this is significantly degraded, physically, in forced convection where coupling between momentum and energy equations is in only one direction; *i.e.*, the energy equation is coupled to the momentum equations, but not *vice versa*. This too has been demonstrated for the PMNS plus thermal energy equations by McDonough and Joyce [209].

As a final result associated with thermal convection we demonstrate that dimensionless heat transfer in the form of a Nusselt number, Nu, can be matched between the (2-D) PMNS plus thermal energy DDS and experimental results of Cioni *et al.* [191]. This is displayed in Fig. 3.23. These experiments we performed at quite low Prandtl number, Pr, and sufficiently high Ra to permit investigation of the transition from "soft" to "hard" turbulence in natural convection. It is interesting to note that not only was the PMNS DDS able to match the slopes of the Nu *vs.* Ra experimental correlations, but that it also detected the change in slope at the correct Ra corresponding to the transition to hard turbulence.

In many respects, the foregoing results may seem rather anecdotal. The PMNS equation is extremely simple, and it may appear rather surprising that it is able to reproduce so much physics. Moreover, we remark that what has been displayed here is only a small fraction of the total results available in this regard. (In particular, we have chosen not to display results associated with

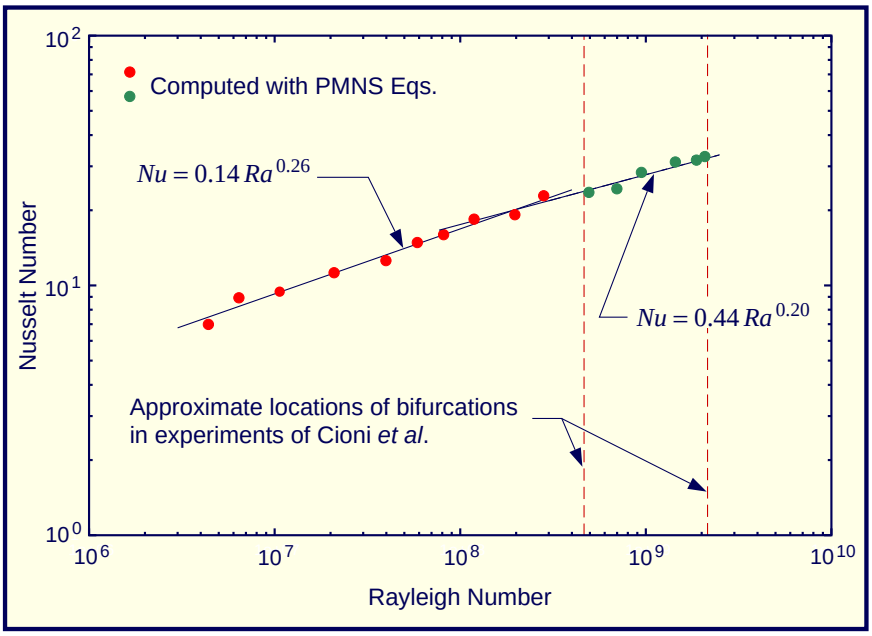

Figure 3.23: Comparison of PMNS plus thermal convection equations with high-Ra data of Cioni *et al.* [191].

chemical kinetics, due to McDonough and Huang [81] and McDonough and Zhang [210, 211], which are equally accurate.) But what we now present is a comparison of the PMNS equation results with Kolmogorov K41 theory discussed earlier in Chap. 2 of these notes). Unlike most of the preceding comparisons, this will involve computations with the 3-D PMNS DDS given in Eqs. (3.48). But recall from Chap. 2 that the K41 theory is appropriate mainly for reasonable approximations of very-high Re, homogeneous, isotropic turbulence. This implies that all β_is in (3.48) should be equal, and similarly, this should also be true of the γ_{ij}s. At the same time, we expect values of the βs should be near to 4. Figure 3.24 presents a "regime map" similar to those presented in McDonough and Huang [82], and elsewhere, displaying the regimes that can be accessed by this DDS. As in the cited reference, and as needed for present purposes, all βs and γs are set equal, respectively.

The figure shows the 14 possible states that can be attained by stationary attractors of the PMNS equation. Of these, the first is steady and of no importance to turbulence modeling, while the 14^{th} is divergent and of no value for anything in particular, other than to indicate that the PMNS equation, like the N.–S. equations from which it is derived, has unbounded solutions in some cases. This bifurcation diagram is composed of several million points, each corresponding to a (β, γ) pair with Eqs. (3.48) iterated for 50,000 iterations with these fixed values to guarantee existence of stationarity beyond the final 10,000 iterations from which the last 8192 points were selected for processing in a radix-2 fast Fourier transform to determine the type of behavior. We see from the table included in the figure that these states include (beyond steady and divergent) periodic, periodic with a different fundamental, subharmonic, phase locked, quasiperiodic, and then so-called "noisy" states corresponding to each of these, thus culminating in a completely broad-band frequency spectrum which generally can be associated with stochastic ("turbulent") behavior of the time series. Examples of these can be found for the 2-D PMNS equation in [82], and our recent investigations have shown that no new states exist in 3D, despite a greater tendency for anisotropy in the latter case.

Figure 3.24 also displays a white dot near the upper right-hand corner in the high-β and relatively high-γ regions corresponding to broad-band behavior of the time series. It is at this point that

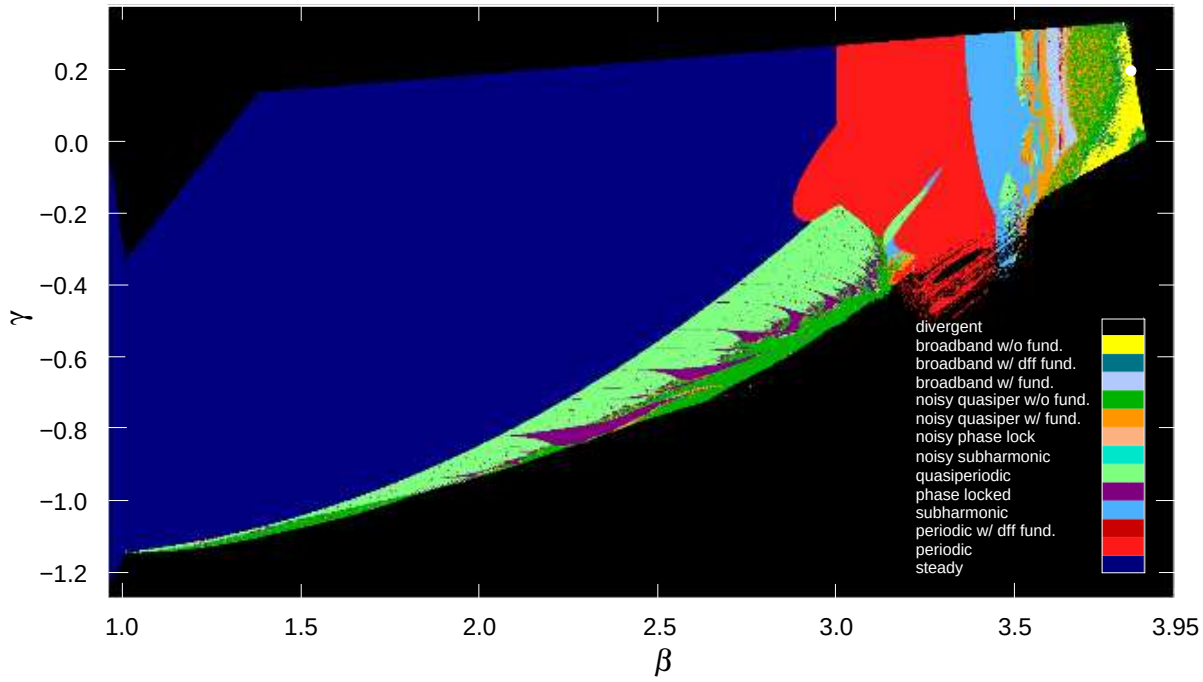

Figure 3.24: Regime map displaying possible types of time series from the PMNS equation.

all calculations to be subsequently reported were performed. These involve specific aspects of Kolmogorov's K41 theory; in particular, we will provide results corresponding to scaling of the 2^{nd}-, 3^{rd}-, 4^{th}- and 6^{th}-order structure functions, and the 1-D energy *vs.* wavenumber spectrum.

We begin with the 3^{rd}-order structure function because, as we noted in Chap. 2, this is one of the most fundamental of all Navier–Stokes equation results; it is exact, with no adjustable constants, so any model that is expected to produce accurate turbulence results must reasonably well duplicate this. This is the so-called 4/5 law which states that 3^{rd}-order structure functions of longitudinal velocity increments should scale linearly with their displacement distances.

In Fig. 3.25 we present results derived from PMNS time series produced with a set of bifurcation parameters that can be associated with high-Re homogeneous isotropic turbulence, as indicated above. We should recall, from Eq. (3.43), that structure function correlations include the dis-

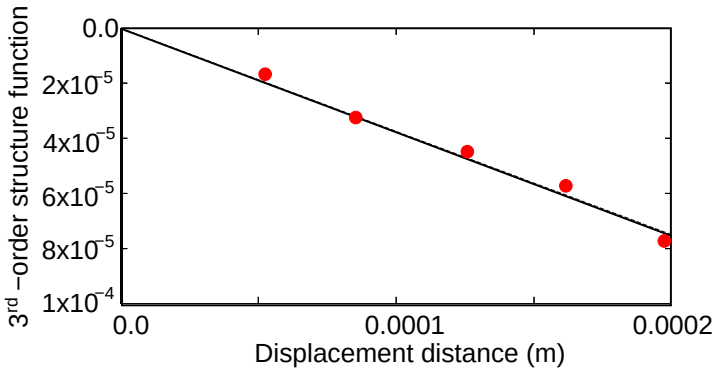

Figure 3.25: Third-order structure function of PMNS time series for homogeneous, isotropic bifurcation parameters.

placement distance times the average turbulence kinetic energy dissipation rate raised to a power. Theoretically, as we discussed in Chap. 2, this power is given by $3/p$, where p is the order of the structure function—hence, the linear dependence in the 3^{rd}-order case. If one employs the Taylor frozen-flow hypothesis it is possible to convert temporal data to approximate spatial data; this has been employed throughout the present calculations to obtain displacements. On the other hand, this approach is generally not sufficiently accurate to permit numerical approximations of spatial derivatives needed to calculate dissipation rates (recall Eq. (1.53)). Thus, in the pressent case we first plotted normalized values of S_3 *vs.* displacement r and determined $\langle\varepsilon\rangle$ so that the slope of the straight line would be $-4/5$. This value of $\langle\varepsilon\rangle$ was then used in constructing scalings for all other orders of structure functions. In Fig. 3.25 the solid line is the theoretical one corresponding to a slope of $-4/5$, and the points were obtained from the PMNS results. It is clear that these computed values are, indeed, linear with displacement distance. Furthermore, it will be clear from subsequent results that the value of $\langle\varepsilon\rangle$ needed to attain the $-4/5$ slope is consistent with scalings for other orders of structure function.

Figure 3.26 presents these additional results. Here, the solid lines represent least-squares curve

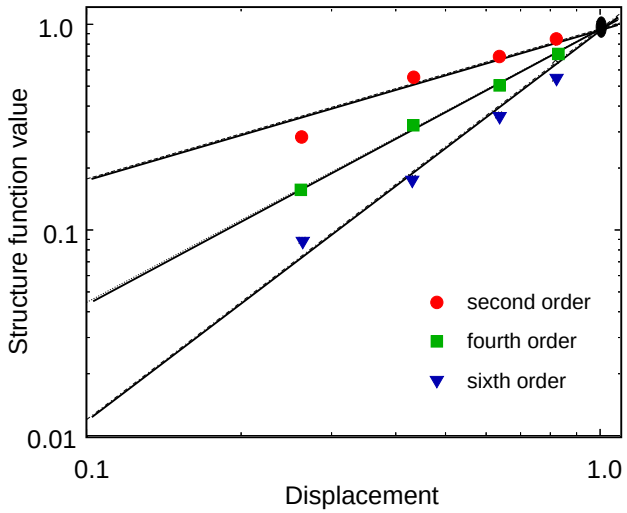

Figure 3.26: Second-, fourth- and sixth-order structure functions from PMNS time series for homogeneous, isotropic bifurcation parameters.

fits of the plotted PMNS points. (Correlation coefficients ranged from 0.89 for the second-order case to 0.98 for fourth order.) We remark that the correlations are not in perfect agree with K41 theory, but they are generally quite good. What is of further interest is that the observed deviations from theory are very similar (same magnitude and direction) to those seen in experimental data for which it is argued that intermittency resulting in multi-fractal scalings (which differ somewhat from K41 theory—see [80]) is the source of discrepancies. It is known (see [82]) that bifurcation parameters in the range of the values corresponding to the point in Fig. 3.24 lead to intermittent time series, so it appears that the PMNS equation actually possesses solutions leading to the true, physical multi-fractal exponents for structure functions.

With the success seen in reproduction of structure function scaling, one must wonder whether the 5/3 law of K41 theory can also be produced by PMNS solutions. It turns out that this 1-D energy spectrum is far easier to obtain via the Taylor hypothesis than are the structure functions because the dissipation rate is no longer needed. Moreover, the formula corresponding to the Taylor hypothesis is simple and produces a dispersion relation which permits exact transformation between temporal and spatial data. (This has been widely used by experimentalists.) Figure 3.27 displays

a PMNS 1-D energy spectrum on which the $k^{-5/3}$ slope is indicated. We observe that the integral

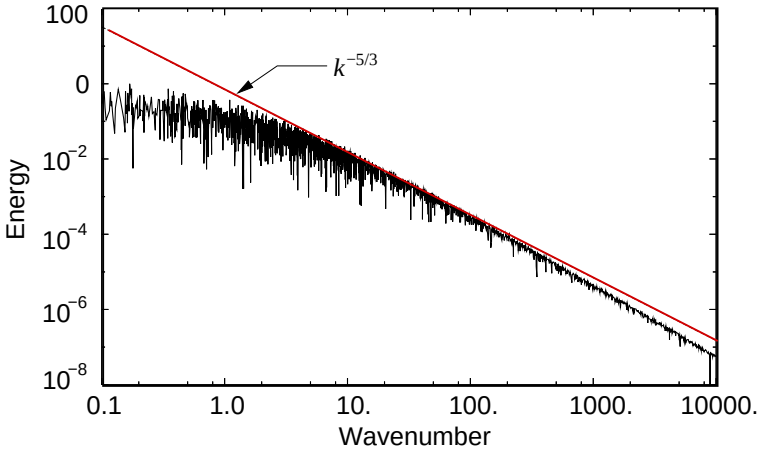

Figure 3.27: PMNS equation 1-D energy spectrum.

scale is apparent, with approximately two decades of inertial subrange behavior, before some decay toward dissipation scales is seen. In particular, the PMNS equation is carrying a considerable amount of the physics associated with the full system of N.–S. equations. But, of course, from a modeling standpoint this is only local. At this point one might wonder how a model constructed from at most a few wavenumbers could produce results over a wavenumber range such as indicated in the figure. The "mechanism" for this is the same as occurs in analogous treatments of experimental time series. Namely, the longer the time series, the larger the possible range of frequencies that will be produced by an FFT of the data—just and immediate consequence of the Nyquist criterion for signal processing. But this, in turn, implies a large range of wavenumbers, *a la* the dispersion relation that corresponds to the Taylor hypothesis.

PMNS Equation Viewed as Pseudodifferential Operator of Navier–Stokes Equations. The preceding results appear to be extremely favorable, and one must ask how such a relatively simple system of algebraic maps could possibly reproduce so much of the behavior of a theoretically very complicated system of partial differential equations. We will provide an answer to this question in the present subsection.

The goal of this section is demonstration that Eqs. (3.48) comprise a mathematical object that is closely related to a time-dependent, nonlinear pseudodifferential operator corresponding to the N.–S. equations. To do this we start with a brief review of the simplest notions associated with pseudodifferential operators; we then use these to show that the PMNS equation arises from an analogous structure, and then discuss the significance of this result.

Definition 3.2 *Let $\Omega \subseteq \mathbb{R}^n$ and $\boldsymbol{\xi} \in \mathbb{R}^n$. Then for $\boldsymbol{u} \in C_0^\infty(\Omega)$ a simple (vector) pseudodifferential operator on Ω can be expressed as (see, e.g., Treves [212])*

$$\boldsymbol{\mathcal{P}}\boldsymbol{u}(\boldsymbol{x}) = \frac{1}{(2\pi)^n} \int e^{i\boldsymbol{x}\cdot\boldsymbol{\xi}} \boldsymbol{p}(\boldsymbol{x},\boldsymbol{\xi}) \widehat{\boldsymbol{u}}(\boldsymbol{\xi}) d\boldsymbol{\xi}\,, \tag{3.51}$$

where $\widehat{\boldsymbol{u}}$ is the Fourier transform of $\boldsymbol{u}$:

$$\widehat{\boldsymbol{u}}(\boldsymbol{\xi}) \equiv \int \boldsymbol{u}(\boldsymbol{x}) e^{-i\boldsymbol{x}\cdot\boldsymbol{\xi}} d\boldsymbol{x}\,.$$

In (3.51) $\boldsymbol{p}(\boldsymbol{x},\boldsymbol{\xi})$ *is a smooth (vector) function on $\Omega \times \mathbb{R}^n$ and is called the symbol of the pseudodifferential operator $\boldsymbol{\mathcal{P}}$.*

We observe that (3.51) arises by applying the Fourier transform to the differential operator, and then using the inverse transform. In this way we formally construct a pseudodifferential operator for the Navier–Stokes equations in a form found in Leslie [30] and elsewhere in the turbulence literature, but not generally identified as a pseudodifferential operator.

We now develop a relationship of the Galerkin form of N.–S. equations to a pseudodifferential operator. We begin by expressing the Fourier representations of N.–S. solutions, originally given in Chap. 1 in 2D as Eqs. (1.39), in the more concise 3-D form

$$\boldsymbol{U}(\boldsymbol{x},t) = \sum_{|\boldsymbol{k}|\geq 0} \widehat{\boldsymbol{U}}_{\boldsymbol{k}}(t)\varphi(\boldsymbol{k},\boldsymbol{x})\,, \tag{3.52}$$

where, corresponding to our earlier notations, $\widehat{\boldsymbol{U}}_{\boldsymbol{k}} \equiv (a_{\boldsymbol{k}}, b_{\boldsymbol{k}}, c_{\boldsymbol{k}})^T$, and by definition

$$\widehat{\boldsymbol{U}}_{\boldsymbol{k}} \equiv \langle \boldsymbol{U}, \varphi_{\boldsymbol{k}}\rangle = \int_\Omega \boldsymbol{U}(\boldsymbol{x},t)\varphi(\boldsymbol{k},\boldsymbol{x})d\boldsymbol{x}\,.$$

At this point we observe that substitution of (3.52) into the N.–S. equations leads to a form analogous to (3.51) except that the integral must be replaced with summation, and $\boldsymbol{\xi} \to \boldsymbol{k}$ with the range of $\boldsymbol{k}$ being only countable. In particular, corresponding to equations given above during construction of Eqs. (3.48), we have

$$\sum_{\boldsymbol{\ell}} \left(\widehat{\boldsymbol{U}}_{\boldsymbol{\ell}}\right)_t \varphi(\boldsymbol{\ell},\boldsymbol{x}) + \sum_{\boldsymbol{\ell},\boldsymbol{m}} (\boldsymbol{\ell}_1 + \boldsymbol{m}_1)\widehat{\boldsymbol{U}}_{\boldsymbol{\ell}}\widehat{\boldsymbol{U}}_{\boldsymbol{m}}\varphi(\boldsymbol{\ell},\boldsymbol{x})\varphi(\boldsymbol{m},\boldsymbol{x}) + \sum_{\boldsymbol{\ell},\boldsymbol{m}} (\boldsymbol{\ell}_2 + \boldsymbol{m}_2)\widehat{\boldsymbol{U}}_{\boldsymbol{\ell}}\widehat{\boldsymbol{U}}_{\boldsymbol{m}}\varphi(\boldsymbol{\ell},\boldsymbol{x})\varphi(\boldsymbol{m},\boldsymbol{x})$$
$$+ \sum_{\boldsymbol{\ell},\boldsymbol{m}} (\boldsymbol{\ell}_3 + \boldsymbol{m}_3)\widehat{\boldsymbol{U}}_{\boldsymbol{\ell}}\widehat{\boldsymbol{U}}_{\boldsymbol{m}}\varphi(\boldsymbol{\ell},\boldsymbol{x})\varphi(\boldsymbol{m},\boldsymbol{x}) = -\frac{1}{Re}\sum_{\boldsymbol{\ell}} |\boldsymbol{\ell}|^2 \widehat{\boldsymbol{U}}_{\boldsymbol{\ell}}\varphi(\boldsymbol{\ell},\boldsymbol{x})\,, \tag{3.53}$$

where it should be observed that the notation $\widehat{\boldsymbol{U}}_{\boldsymbol{\ell}}\widehat{\boldsymbol{U}}_{\boldsymbol{m}}$ is somewhat ambiguous, although completely analogous to that in Eq. (3.37a). As can be inferred from our preceding more detailed Galerkin construction we have

$$\widehat{\boldsymbol{U}}_{\boldsymbol{\ell}}\widehat{\boldsymbol{U}}_{\boldsymbol{m}} = \begin{bmatrix} a_{\boldsymbol{\ell}}a_{\boldsymbol{m}} & b_{\boldsymbol{\ell}}a_{\boldsymbol{m}} & c_{\boldsymbol{\ell}}a_{\boldsymbol{m}} \\ a_{\boldsymbol{\ell}}b_{\boldsymbol{m}} & b_{\boldsymbol{\ell}}b_{\boldsymbol{m}} & c_{\boldsymbol{\ell}}b_{\boldsymbol{m}} \\ a_{\boldsymbol{\ell}}c_{\boldsymbol{m}} & b_{\boldsymbol{\ell}}c_{\boldsymbol{m}} & c_{\boldsymbol{\ell}}c_{\boldsymbol{m}} \end{bmatrix}\,,$$

with the first row corresponding to terms of the first component of the above vector equation with entries from the columns inserted in the terms of the individual spatial operators, and similarly for the second and third components.

Observe that up to scaling, existence of nonlinearities and time dependence, if we were to replace summations with integrals in (3.53) we would obtain a form analogous to (3.51). Also note that if we had first linearized the N.–S. equations, instead of using (3.53) we would have obtained a form containing the vector polynomial

$$\boldsymbol{p}(\boldsymbol{x},\boldsymbol{k}) = \sum_{|\alpha|\leq 2} \boldsymbol{p}_\alpha(\boldsymbol{x})\boldsymbol{k}^\alpha\,,$$

the vector symbol of the pseudodifferential operator arising from the linearized N.–S. equations. Thus, it seems reasonable to view (3.53) as a nonlinear, time-dependent pseudodifferential operator for the Navier–Stokes equations with a symbol constructed from the various wavenumber factors.

Now recall that in the Galerkin procedure we form inner products of (3.53) with elements of the basis set $\{\varphi_{\boldsymbol{k}}(\boldsymbol{x})\} = \{\varphi(\boldsymbol{k},\boldsymbol{x})\}_{|\boldsymbol{k}|\geq 0}$ and employ assumed orthonormality to simplify the result

to the form

$$\frac{\partial \widehat{\boldsymbol{U}}_{\boldsymbol{k}}}{\partial t} + \frac{|\boldsymbol{k}|^2}{Re}\widehat{\boldsymbol{U}}_{\boldsymbol{k}} + \sum_{\boldsymbol{\ell},\boldsymbol{m}} \boldsymbol{A}_{\boldsymbol{k\ell m}}\widehat{\boldsymbol{U}}_{\boldsymbol{\ell}}\widehat{\boldsymbol{U}}_{\boldsymbol{m}} + \sum_{\boldsymbol{\ell},\boldsymbol{m}} \boldsymbol{B}_{\boldsymbol{k\ell m}}\widehat{\boldsymbol{U}}_{\boldsymbol{\ell}}\widehat{\boldsymbol{U}}_{\boldsymbol{m}} + \sum_{\boldsymbol{\ell},\boldsymbol{m}} \boldsymbol{C}_{\boldsymbol{k\ell m}}\widehat{\boldsymbol{U}}_{\boldsymbol{\ell}}\widehat{\boldsymbol{U}}_{\boldsymbol{m}} = 0\,,$$
$$\boldsymbol{k} = 0, 1, \ldots, \infty\,.$$

Although this is no longer in the form (3.51) because the $\varphi(\boldsymbol{k}, \boldsymbol{x})$ have been integrated (to form $\boldsymbol{A}_{\boldsymbol{k\ell m}}$, *etc.*), it does represent a direct simplification of (3.53)—one that is essential if a solution consisting of the $\widehat{\boldsymbol{U}}_{\boldsymbol{k}}$ is to be obtained. Moreover, the final form of the PMNS equation (3.48) follows immediately from this via discretization in time, as shown in the preceding subsection. Hence, it appears reasonable to view the PMNS equation as being closely related to a simple discretization of a microlocal (because only a single wavevector is retained) pseudodifferential operator for the N.–S. equations.

The importance of the relationship between the PMNS equation and pseudodifferential operators of the N.–S. equations cannot be overemphasized. For simple linear PDEs pseudodifferential operators provide an algebra leading to exact solutions in some cases, and at least simplified analysis in more complicated situations. In any case, estimating the integral operator on the right-hand side of (3.51) is typically far more tractable than attempting to estimate the original differential operators, especially in the context of microlocal analyses employed herein and elsewhere. In the nonlinear, time-dependent case being treated here, this is true locally (in physical and/or wavenumber space) via microlocal analysis, and it also leads to an efficient computational tool. As we have demonstrated in the preceding subsection, and in various of the cited references, we are already employing the PMNS equation both as a part of SGS models for LES and for "curve fitting" chaotic (presumably turbulent) experimental flow data. Identification of the PMNS equation with a pseudodifferential operator of the N.–S. equations provides an explanation for the previous (seemingly somewhat fortuitous) success of the curve fits (see [45]) and suggests its use should be valuable in constructing high-fidelity SGS models for LES, as has been initiated in [72]. Moreover, in this latter application it provides a means to introduce mathematical rigor (via microlocal analysis) into study of SGS models.

Application of Multi-Scale Formalism to Produce Complete Solution. During each resolved-scale time step Eqs. (3.48) are iterated a number of times corresponding to formal integration of the small-scale equations across the large-scale time step. This number is set by using a small-scale time scale obtained as $\tau \equiv 1/\|\boldsymbol{S}^{**}\|$, as noted above. The result of this is evaluation of a discrete dynamical system directly related to the N.–S. equations, but restricted to a single wavevector (per dependent variable), and associated with the M_is of Eq. (3.39) as

$$M_1 = a\,, \qquad M_2 = b\,, \qquad M_3 = c\,. \tag{3.54}$$

But these "velocity components" require rescaling; recall that the M_is are restricted to $[-1, 1]$ to make sense of use of the amplitude factors, A_is. In addition, these velocity components are not divergence free (just as is the case for velocity components computed from essentially any form of (3.37) on any scale). Thus, after calculating $q_i^*(\boldsymbol{x}, t)$, $i = 1, \ldots, N_v$, of Eq. (3.40) at all resolved-scale grid points, we must project the portion of this solution vector corresponding to the velocity field onto a divergence-free subspace by first solving a (pseudo-) pressure Poisson equation,

$$\Delta p^* = \frac{\nabla\cdot\boldsymbol{u}^*}{\Delta t}\,, \tag{3.55}$$

and then projecting the velocity field in the usual way via

$$u^* = u^* - \Delta t\, p_x^*\,, \qquad v^* = v^* - \Delta t\, p_y^*\,, \qquad w^* = w^* - \Delta t\, p_z^*\,. \tag{3.56}$$

These are the fluctuating (small-scale) velocity components on the large-scale grid at time level $n+1$. Furthermore, as is well known, p^* provides a reasonable approximation to the small-scale fluctuating pressure if Re is large (despite the fact—or, maybe because of it—that p^* is actually a velocity potential). See McDonough [204] for more detailed discussions of projection methods applied to the incompressible N.–S. equations.

At this point both large- and small-scale solution compoments are available at time level $n+1$, and we are now ready to consider precisely how to use the SGS quantities in the context of a formal multi-scale procedure. We again remind the reader that LES is intrinsically a multi-scale approach due to the basic decomposition of dependent variables:

$$q_i(\boldsymbol{x},t) = \tilde{q}_i(\boldsymbol{x},t) + q_i^*(\boldsymbol{x},t)\,, \qquad i = 1,2,\ldots,N_v\,. \tag{3.57}$$

Indeed, as should be clear from the preceding treatment, both $\tilde{q}_i$ and q_i^* have been constructed such that each serves as at least a consistent approximation to the corresponding Hilbert (sub)space quantities. Up to now this has been implicit for $\tilde{q}_i$, and we here provide a basic description of this part of the method.

We earlier noted that the synthetic-velocity approach presented here begins with a formulation not too different from ATD; in fact, the large-scale equations are those of ATD. Thus, we formally solve

$$(\widetilde{\boldsymbol{U}} + \boldsymbol{U}^*)_t + \nabla\cdot\Big((\widetilde{\boldsymbol{U}} + \boldsymbol{U}^*)^2\Big) = -\nabla(\widetilde{P} + P^*) + \nu\Delta(\widetilde{\boldsymbol{U}} + \boldsymbol{U}^*)\,, \tag{3.58a}$$

$$\nabla\cdot(\widetilde{\boldsymbol{U}} + \boldsymbol{U}^*) = 0\,, \tag{3.58b}$$

for $\widetilde{\boldsymbol{U}}$ with $\boldsymbol{U}^*$ computed as given in the preceding analyses. The same treatment is also applied to any additional quantities related to the physics of a particular problem. Formal numerical-analytic consistency (but not necessarily accuracy) of this approach is obvious if a consistent discretization is applied to (3.58) because $\boldsymbol{U}^*$, and $P^* \to 0$ as $\Delta t,\, h_i \to 0$.

But it is well known from a numerical-analytic standpoint that consistency does not imply stability (nor *vice versa*, of course), and it is often observed that small-scale perturbations arising in multi-scale algorithms lead to destabilization of the underlying numerical method. Beyond this is yet a separate and subtle issue that occurs for heterogeneous multi-scale algorithms, in general (see [196]), of which the present synthetic-velocity LES is an example. Namely, because the q_i^*s have been produced with a completely separate algorithm from that leading to the $\tilde{q}_i$s, but at the same time the q_is are required to satisfy the same discrete equations as do the $\tilde{q}_i$s, it follows that time-derivative information is needed for the q_i^*s. In particular, if one considers the discrete solution operator

$$\tilde{q}_i^{n+1} = \tilde{q}_i^n + \int_{t^n}^{t^{n+1}} F_i(\boldsymbol{q})\,dt \tag{3.59}$$

for the large-scale part of the solution, and takes this to be the same as the complete solution operator

$$q_i^{n+1} = q_i^n + \int_{t^n}^{t^{n+1}} F_i(\boldsymbol{q})\,dt\,,$$

then since from (3.57)

$$q_i^{n+1} = \tilde{q}_i^{n+1} + q_i^{*\,n+1}\,,$$

we see that

$$d\boldsymbol{q} = d\,\widetilde{\boldsymbol{q}} + d\boldsymbol{q}^*$$

must hold. Then it follows that to obtain consistency for the complete solution operator we must have the following temporal difference:

$$\boldsymbol{q}^{n+1} - \boldsymbol{q}^{n} = \widetilde{\boldsymbol{q}}^{\,n+1} - \widetilde{\boldsymbol{q}}^{\,n} + \boldsymbol{q}^{*\,n+1} - \boldsymbol{q}^{*\,n} \,.$$

But the typical construction is to use (3.59) to calculate $\widetilde{\boldsymbol{q}}^{n+1}$ and then simply add the separately computed $\boldsymbol{q}^{*\,n+1}$, as suggested by the LES decomposition. It is clear, however, from the foregoing solution operator analyses that instead we must employ

$$\boldsymbol{q}^{n+1} = \widetilde{\boldsymbol{q}}^{\,n+1} + \boldsymbol{q}^{*\,n+1} - \boldsymbol{q}^{*\,n} \tag{3.60}$$

to obtain consistent time evolution. We remark that in early implementations where this was not done, numerical instabilities could develop fairly rapidly; and (3.60) was first discovered during numerical experiments undertaken to remedy the stability problem.

We close this subsection by providing an example of time series resulting from the overall multi-scale synthetic-velocity algorithm. Figure 3.28 provides representative results from the center of the buoyant swirling plume studied in [72] and from which earlier results have been presented in Figs. 3.17–3.19. The smooth solid curve in each figure is the large-scale solution corresponding to what would be obtained with a highly under resolved ILES procedure. The fluctuating curves are those of the complete solution (large scale plus small scale) as given by the LES decomposition Eq. (3.57). All quantities are dimensional, with velocity components having dimensions m/sec. We have not displayed the second horizontal velocity component because it is similar (but, or course, not identical) to the one shown. It is clear that at this particular location in the plume the vertical component of velocity is far greater than the horizontal one, as would be expected in much of a buoyant plume.

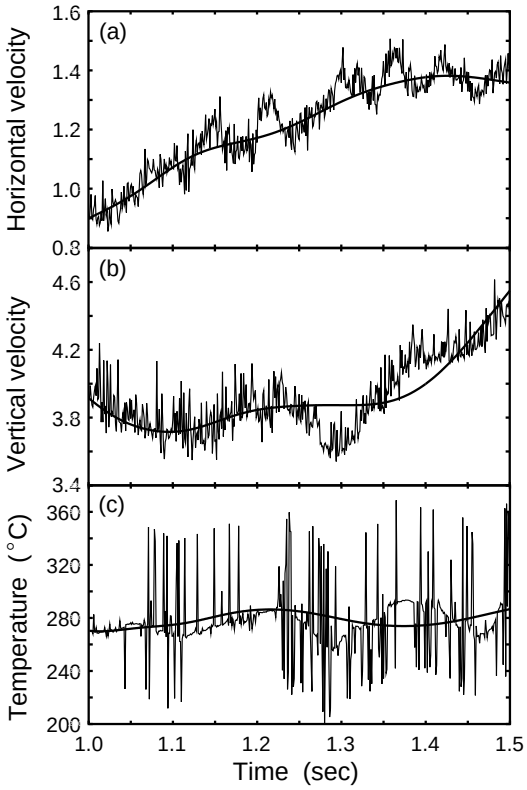

Figure 3.28: Velocity components and temperature in center of swirling buoyant plume; (a) u-component of horizontal velocity, (b) vertical velocity, and (c) temperature.

Furthermore, the temperature fluctuations are significantly more intermittent than are those of the velocity fluctuations. This, too, is a physically-expected outcome. In particular, it has long been observed that passive scalars in a turbulent flow exhibit more intermittency than does the flow field (see, *e.g.*, [111]). This is due to several effects, but from a straightforward mathematical viewpoint it occurs because equations governing transport of scalars are "less nonlinear" than the N.–S. equations, thus allowing dissipation terms to be more effective in damping oscillations.

Finally, we note that the large-scale part of the solution displayed in Fig. 3.28 does not perfectly follow the mean of the complete solution. For horizontal velocity it is fairly close to this mean, but rather significant departures can be seen for vertical velocity and temperature, implying that even in an averaged sense, omission of the $\boldsymbol{q}^*$ terms of the LES decomposition can lead to noticeable errors.

Pseudo-Language Algorithm for Multi-Scale Synthetic-Velocity LES. We will end this treatment of the multi-scale, dynamical-systems-based approach to LES by providing a coarse-grained pseudo-language algorithm from which, in principle, computer code can be written. We note, however, that considerable additional effort will be needed to produce a complete operational program. The present algorithm represents only an outline of the overall process.

Algorithm 3.2 *(Multi-Scale, Synthetic-Velocity LES). Suppose n resolved-scale time steps have been completed. To calculate $n+1$ time-level results, perform the following steps.*

1. *Solve Eq.* (3.58a) *for* $\widetilde{\boldsymbol{U}}^{n+1}$ *by any valid CFD method (see, e.g., McDonough [204] for various possibilities) using values of* $\boldsymbol{U}^*$ *from the preceding time step. Note that the* $\widetilde{\boldsymbol{U}}^{n+1}$ *velocity field must be divergence free before proceeding to the next step.*

2. *Construct the SGS contribution* $\boldsymbol{U}^*$*. Note that this can be done with either the new time level* $\widetilde{\boldsymbol{U}}^{n+1}$ *velocity field, or with the previous time-level result* $\widetilde{\boldsymbol{U}}^{n}$*. (Use of the latter permits extensive parallelization of the overall algorithm.)*

   a. *High-pass filter* $\widetilde{\boldsymbol{U}}$ *to obtain* $\boldsymbol{U}^{**}$*.*

   b. *Determine turbulent time scale at each large-scale grid point using* $\tau = 1/\|\boldsymbol{S}^{**}\|$*.*

   c. *Calculate Taylor microscale length at each grid point via Eq.* (3.46)*.*

   d. *For every grid point at which* $\tau > \Delta t$ *and* $\lambda_i > h_i \ \forall\ i = 1,2,3$*, set all* $q_i^* = 0$*, and proceed to next grid point. Here,* Δt *is the large-scale time step, and the* h_i *are the large-scale spatial discretization step sizes.*

   e. *Compute amplitude factors,* A_i*, for each solution component using Eqs.* (3.43)–(3.47)*.*

   f. *Evaluate bifurcation parameters of PMNS equation from Eq.* (3.49) *and (usually, a surrogate involving a difference approximation of appropriate high-pass filtered velocity components for) Eq.* (3.50)*.*

   g. *Set* $N_{max} = \Delta t/\tau + 1$*, and iterate the PMNS equation for* N_{max} *iterations. Note:* N_{max} *will not, in general, be the same at all grid points.*

   h. *Obtain provisional* $q_i^*\,,\ \ i = 1,\ldots,N_v$*, from Eq.* (3.39)*.*

   i. *Project entire* q_i^* *velocity field onto divergence-free subspace via Eqs.* (3.55) *and* (3.56)*.*

3. *Employ Eq.* (3.60) *to obtain the complete* $n+1$ *time-level solution.*

There are several points to be made regarding this algorithm. First, an implementation is in existence and was used to produce Figs. 3.18, 3.19 and 3.28, as already indicated. In fact, various other turbulence problems have also been solved with this code; but it is still under development, and little has been published in the archival literature. Second, as noted in step 1 of the algorithm, the time level $\widetilde{\boldsymbol{U}}^{n+1}$ velocity field must be divergence free before proceeding to the rest of the algorithm (among other things, making solution algorithms of the form of SIMPLE not very appropriate). This is the only part of the algorithm devoted to mass conservation of large-scale results. Moreover, if $\widetilde{\boldsymbol{U}}^{n+1}$ used in step 2 is not divergence free, turbulence in the small scales will have been generated with an incorrect large-scale velocity field.

The third point specifically concerns step 2. We have indicated that either time level n or $n+1$ large-scale velocity fields can be used in this step. The reason for this flexibility is that we can view the small-scale calculations in two, nearly equivalent, ways. If we look on this as simply an explicit time integration continuing from the preceding time step, then use of $\widetilde{\boldsymbol{U}}^{n}$ is appropriate. On the other hand, we can view this overall multi-scale algorithm as a predictor-corrector method in which

we make a rough prediction (not including high-wavenumber content) via large-scale calculations at the beginning of a new time step, and then correct this with small-scale, high-wavenumber results. In this context $\widetilde{\boldsymbol{U}}^{n+1}$ would be preferred. It should be clear, however, that detailed results from these two approaches cannot be identical; but we should expect their statistics to be the same (under the assumption that the two different solutions lie on the same N.–S. attractor).

We further note that the nature of step 2.d is such as to detect under resolution in both space and time, and to generate small-scale solutions if this occurs in either (or both). Conversely, if Δt and the h_is are sufficiently small (compared with respective inertial range scales), then the calculations correspond to DNS, and no SGS models are needed. In addition, it should be clear from step 2.g that independent of whether $\boldsymbol{U}^{**}$ is based on $\widetilde{\boldsymbol{U}}^{n}$ or $\widetilde{\boldsymbol{U}}^{n+1}$, we formally integrate the PMNS equation over a large-scale time step. Indeed, solutions to this equation are stored from the preceding time step and used as initial data at the new time step. Thus, the choice of $\widetilde{\boldsymbol{U}}^{n}$ or $\widetilde{\boldsymbol{U}}^{n+1}$ directly effects amplitude factors in the small-scale solution representation and bifurcation parameters in the PMNS equation, but it does not effect the small-scale integration process.

Finally, we note that step 2.i is the arithmetically expensive part of the SGS model because it requires solution of a Poisson equation to impose mass conservation on $\boldsymbol{U}^{*}$. At the same time, this introduces global effects into otherwise completely local small-scale solutions (which is desirable). But, as is well know, Poisson equation solves are the expensive part of any incompressible flow simulation, so we have essentially doubled the total arithmetic over that required for a laminar flow calculation (on the same grid). On the other hand, total arithmetic is nearly the same for both large and small scales, leading to good load balancing in a parallel computing environment. Hence, in principle, even if only two processors are available, the present algorithm can produce a turbulent solution in essentially the same wall-clock time as needed for a laminar solution.

3.2.4 Summary of dynamical systems/multi-scale methods

In this section we first provided an elementary discussion of general aspects of dynamical systems, including such notions as attractor, bifurcation, fractal, and strange attractor; and we described several of the more inportant diagnostic tools used in analyzing these objects, including use of power spectra, phase portraits, *etc.* We then provided a brief, heuristic overview of multi-scale ideas and then used these in combination with dynamical systems to develop a synthetic-velocity version of large-eddy simulation. The keys to this approach include filtering solutions rather than equations, directly modeling SGS physical quantities instead of their statistics, and finally, explicitly combining resolved-scale and SGS quantities to produce the entire LES (Hilbert space) representation of flow variables.

The portions of this analysis that are quite new are first the (formally) single-mode Fourier representations of SGS quantities as a product of an amplitude and a time-dependent oscillation, and second the details by means of which each of these is calculated. Amplitudes are produced as a generalization of Kolmogorov's K41 theory to power laws for which exponents must be determined based on resolved-scale results—a form of deconvolution which removes the homogeneity and isotropy assumptions of the K41 theory. Temporal fluctuations are computed from the poor man's Navier–Stokes equation, a discrete dynamical system derived from the N.–S. equations via a Galerkin procedure, and which represents a discretization of microlocalization of a pseudodifferential operator of the N.–S. equations. We demonstrated that the PMNS equation is able to replicate both experimental and DNS results, and that it furthermore reproduces the Kolmogorov K41 scalings—in fact, display possible multi-fractal effects more like K62 theory. We concluded that this type of SGS model shows very good promise for future development, and a pseudo-language algorithm was provided for its implementation.

3.3 Summary

This chapter dealt with modern turbulence simulation techniques of the general nature of large-eddy simulation. We presented the usual methods of LES in some detail, beginning with the basic steps of construction—the LES decomposition, filtering of governing equations, formulating SGS models, and neglecting small-scale results—including treatment of Smagorinsky, dynamic and mixed SGS models. Then much more analysis was provided for a very modern synthetic-velocity approach to LES which eliminates some of the shortcomings of the more classical approaches. The key aspects of this new approach appeal to rigorous mathematics of PDEs and include use of mollification of large-scale solution components to control aliasing caused by the (deliberately) under-resolved representations, modeling of SGS dependent variables via extension of Kolmogorov K41 theory and direct application of dynamical systems in the form of the poor man's Navier–Stokes equation, and introduction of multi-scale methods to construct the complete LES decomposition. The result of this combination of techniques is one of the first turbulence modeling procedures to explicitly employ theories from dynamical systems which, in the context of the Navier–Stokes equations—because these are a dynamical system—correctly establish the mathematical nature of the intriguing and elusive physical phenomenon known as *turbulence*.

References

[1] G. T. Chapman and M. Tobak. Observations, Theoretical Ideas, and Modeling of Turbulent Flows — Past, Present and Future, in *Theoretical Approaches to Turbulence*, Dwoyer *et al.* (eds), Springer-Verlag, New York, pp. 19–49, 1985.

[2] P. Constantin and C. Foias. *Navier–Stokes Equations*, University of Chicago Press, Chicago, 1988.

[3] A. Tsinober. *An Informal Introduction to Turbulence*, Kluwer Academic Publishers, Dordrecht, 2001.

[4] L. F. Richardson. *Weather Prediction by Numerical Process*, Cambridge University Press, 1922.

[5] T. von Kármán. Some remarks on the statistical theory of turbulence, *Proc.* 5^{th} *Int. Congr. Appl. Mech.*, Cambridge, MA, 347, 1938.

[6] J. O. Hinze. *Turbulence*, McGraw-Hill, New York, 1959.

[7] H. Tennekes and J. L. Lumley. *A First Course in Turbulence*. MIT Press, 1972.

[8] D. Ruelle and F. Takens. On the nature of turbulence. *Comm. Math. Phys.* **20**, 167–192, 1971.

[9] S. A. Orszag and G. S. Patterson. Numerical simulation of turbulence: statistical models and turbulence, *Lecture Notes in Physics* **12**, 127–147, Springer-Verlag, Berlin, 1972.

[10] J. W. Deardorff. A numerical study of three-dimensional turbulent channel flow at large Reynolds numbers, *J. Fluid Mech.* **41**, 453, 1970.

[11] B. E. Launder and D. B. Spalding. *Mathematical Models of Turbulence*, Academic Press, 1972.

[12] R. S. Rogallo and P. Moin. Numerical simulation of turbulent flows, *Annu. Rev. Fluid Mech.* **16**, 99–137, 1984.

[13] J. Kim, P. Moin and R. Moser. Turbulence statistics in fully developed channel flow at low Reynolds number, *J. Fluid Mech.* **177**, 133–166, 1987.

[14] M. Lesieur. *Turbulence in Fluids*, Second Revised Edition, Kluwer Academic Press, Dordrecht, 1990.

[15] J. Boussinesq. Essai sur la théorie des eaux courantes, *Mém. prés. par div. savant à l'Acad. Sci.* **23**, 1–680, 1877.

[16] O. Reynolds. On the dynamical theory of turbulent incompressible viscous fluids and the determination of the criterion, *Phil. Trans. R. Soc. London A* **186**, 123–161, 1894.

[17] H. Poincaré. *Les Methodes Nouvelles de la Mécanique Celeste*, Vol. 1–3, Gauthier Villars, Paris, 1899.

[18] E. N. Lorenz. Deterministic nonperiodic flow. *J. Atmos. Sci.* **20**, 130–141, 1963.

[19] L. Prandtl. Bericht über Untersuchungen zur ausgebildeten Turbulenz, *Zs. agnew. Math. Mech.* **5**, 136–139, 1925.

[20] G. I. Taylor. Statistical theory of turbulence, *Proc. Roy. Soc. London A* **151**, 421–478, 1935.

[21] T. von Kármán. On the statistical theory of turbulence, *Proc. Nat. Acad. Sci., Wash.* **23**, 98, 1937.

[22] T. von Kármán and L. Howarth. On the statistical theory of isotropic turbulence, *Proc. Roy. Soc.* A, **164**, 192, 1938.

[23] N. Weiner. The use of statistical theory in the study of turbulence, *Proc.* 5^{th} *Int. Congr. Appl. Mech.*, Cambridge, MA, 356, 1938.

[24] J. Leray. Etude de diverses équations intégrals non linéaires et de quelques problèmes que pose l'hydrodynamique. *J. Math. Pures Appl.* **12**, 1–82, 1933.

[25] J. Leray. Essai sur les mouvemnts d' un liquide visqueaux que limitent des parois. *J. Math. Pures Appl.* **13**, 331–418, 1934.

[26] A. N. Kolmogorov. The local structure of turbulence in incompressible viscous fluid for very large Reynolds number, *Dokl. Acad. Nauk. SSSR* **30**, 9–13, 1941; On degeneration (decay) of isotropic turbulence in an incompressible viscous liquid, *Dokl. Acad. Nauk. SSSR* **31**, 538–540, 1941; Dissipation of energy in locally isotropic turbulence, *Dokl. Acad. Nauk. SSSR* **32**, 16–18, 1941.

[27] A. N. Kolmogorov. A refinement of previous hypotheses concerning the local structure of turbulence in a viscous incompressible fluid at high Reynolds number, *J. Fluid Mech.* **13**, 82–85, 1962.

[28] E. Hopf. A mathematical example displaying the features of turbulence, *Commun. Pure Appl. Math.* **1**, 303, 1948.

[29] L. D. Landau and E. M. Lifshitz. *Fluid Mechanics* (translated by J. B. Sykes and W. H. Reid from original Russian published in 1944), Pergamon Press, Oxford, 1959.

[30] D. C. Leslie. *Developments in the Theory of Turbulence*, Oxford University Press, Oxford, 1973.

[31] G. K. Batchelor. Energy decay and self-preserving correlation functions in isotropic turbulence, *Quart. Appl. Math.* **6**, 97, 1948.

[32] J. M. Burgers. A mathematical model illustrating the theory of turbulence, *Adv. Appl. Mech.* **1**, 171, 1948.

[33] S. Corrsin. An experimental verification of local isotropy, *J. Aero. Sci.* **16**, 757, 1949.

[34] W. Heisenberg. On the theory of statistical and isotropic turbulence, *Proc. Roy. Soc.* A, **195**, 402, 1948.

[35] T. von Kármán. Progress in the statistical theory of turbulence, *Proc. Nat. Acad. Sci., Wash.* **34**, 530, 1948.

[36] A. M. Obukhov. The local structure of atmospheric turbulence, *Doklady. Akad. Nauk. SSSR* **67**, 643, 1949.

[37] A. A. Townsend. The measurement of double and triple correlation derivatives in isotropic turbulence, *Proc. Camb. Phil. Soc.* **43**, 560, 1947.

[38] A. M. Yaglom. Homogeneous and isotropic turbulence in a viscous compressible fluid, *Izv. Akad. Nauk. SSSR* Ser. Geogr. i Geofiz., **12**, 501, 1948.

[39] G. K. Batchelor. *The Theory of Homogeneous Turbulence*, Cambridge University Press, Cambridge, 1953.

[40] A. A. Townsend. *The Structure of Turbulent Shear Flow*, Cambridge University Press, Cambridge, 1956.

[41] H. W. Emmons. The laminar-turbulent transition in a boundary layer. Part I, *J. Aero. Sci.* **18**, 490–498, 1951.

[42] H. L. Dryden. Recent Advences in the Mechanics of Boundary Layer Flow, in *Adv. Appl. Mech.* **1**, Academic Pressm, New York, 1–40, 1948.

[43] J. M. McDonough, S. Mukerji and S. Chung. *J. Comput. Appl. Math.* 1998.

[44] S. Mukerji, J. M. McDonough, M. P. Mengüç, S. Manickavasagam and S. Chung. *Int. J. Heat Mass Trasfer* 1998.

[45] T. Yang, J. M. McDonough and J. D. Jacob. Two-dimensional "poor man's Navier–Stokes equation" model of turbulent flows, *AIAA J.* **41**, 1690–1696, 2003.

[46] O. Ladyzhenskaya. *The Mathematical Theory of Viscous Incompressible Flow*, revised English edition (translated from the Russian by Richard A. Silverman). Gordon & Breach, New York, 1963.

[47] S. Smale. Diffeomorphisms with many periodic points, in *Differential and Combinatorial Topology*, Princeton University Press, Princeton, 63–80, 1965.

[48] V. I. Arnol'd. Instability of dynamical systems with several degrees of freedom, *Sov. Math. Dokl.* **5**, 581–585, 1964.

[49] R. H. Kraichnan. Irreversible statistical mechanics of incompressible hydromagnetic turbulence, *Phys. Rev.* **109**, 1047, 1958.

[50] R. H. Kraichnan. The structure of isotropic turbulence at very high Reynolds number, *J. Fluid Mech.* **5**, 497, 1959.

[51] G. Comte-Bellot and S. Corrsin. The use of contraction to improve the isotropy of grid-generated turbulence, *J. Fluid Mech.* **25**, 657–682, 1966.

[52] H. J. Tucker and A. J. Reynolds. The distortion of turbulence by irrotational plain strain, *J. Fluid Mech.* **32**, 657–673, 1968.

[53] I. Wygnanski and H. E. Fiedler. *J. Fluid Mech.* **38**, 577, 1969.

[54] C. H. Gibson. *Phys. Fluids* **11**, 2305, 1968.

[55] J. L. Lumley and G. R. Newman. The return to isotropy of homogeneous turbulence, *J. Fluid Mech.* **82**, 161–178, 1977.

[56] M. J. Feigenbaum Quantitative universality for a class of nonlinear transformations, *J. Stat. Phys.* **19**, 25–52, 1978.

[57] Y. Pomeau and P. Manneville. Intermittent transition to turbulence in dissipative dynamical systems, *Commun. Math. Phys.* bf 74, 189–197, 1980.

[58] J. P. Gollub and S. V. Benson. Many routes to turbulenct convection, *J. Fluid Mech.* **100**, 449–470, 1980.

[59] R. F. Blackwelder and L. S. G. Kovasznay. *Phys. Fluids* **15**, 1545, 1972.

[60] R. A. Antonia, A. Prabhu and S. E. Stephenson. *J. Fluid Mech.* **72**, 455, 1975.

[61] W. C. Reynolds and A. K. M. F. Hussain. *J. Fluid Mech.* **54**, 263, 1972.

[62] D. H. Wood and P. Bradshaw. *J. Fluid Mech.* **122**, 57, 1982.

[63] B. E. Launder, G. J. Reece and W. Rodi. Progress in the development of a Reynolds stress turbulence closure, *J. Fluid Mech.* **68**, 537–566, 1975.

[64] J. H. Ferziger. Large Eddy Simulation: Its Role in Turbulence Research, in *Theoretical Approaches to Turbulence*, Dwoyer *et al.* (eds), Springer-Verlag, New York, pp. 51–72, 1985.

[65] W. C. Reynolds. The Potential and Limitations of Direct and Large Eddy Simulation, in *Wither Turbulence? Turbulence at the Crossroads*, J. L. Lumley (ed.), Springer-Verlag, Berlin, pp. 313–343, 1990.

[66] M. Lesieur and O. Métais. New trends in large-eddy simulations of turbulence, *Annu. Rev. Fluid Mech.* **28**, 45–82, 1996.

[67] C. Meneveau and J. Katz. Scale-invariance and turbulence models for large-eddy simulation, *Annu. Rev. Fluid Mech.* **32**, 1–32, 2000.

[68] M. Germano, U. Piomelli, P. Moin and W. H. Cabot. A dynamic subgrid-scale eddy viscosity model, *Phys. Fluids A* **3**, 1760–1765, 1991.

[69] U. Piomelli. High Reynolds number calculations using the dynamic subgrid scale stress model, *Phys. Fluids A* **5**, 1484–1490, 1993.

[70] J. A. Domaradzki and E. M. Saiki. A subgrid-scale model based on the estimation of unresolved scales of turbulence, *Phys. Fluids* **9**, 1–17, 1997.

[71] T. Echekki, A. R. Kerstein and T. D. Dreeben. *Combustion and Flame* **125** 1083–1105, 2001.

[72] J. M. McDonough and T. Yang. A new SGS LES model applied to an internally-heated, swirling buoyant plume, presented at *Fall 2003 Meeting of Western States Section Combustion Symposium*, Los Angeles, Oct. 19–21, 2003.

[73] P. Sagaut. *Large Eddy Simulation for Incompressible Flows, an Introduction*, Springer, Berlin, 2001.

[74] G. B. Schubauer and H. K. Skramstad. Laminar Boundary-Layer Oscillations and Transition on a Flat Plate, *NACA Rep.* **909**, 1948.

[75] F. Takens. Detecting strange attractors in turbulence, in *Dynamical Systems and Turbulence*, Rand and Young (eds.), Springer Lecture Notes in Mathematics, Vol. 898, Springer-Verlag, NY, pp. 366–381, 1980.

[76] A. Leonard. Energy cascade in large-eddy simulations of turbulent fluid flows, *Adv. Geophys.*, C18AD, 237, 1974.

[77] P. Bergé, Y. Pomeau and C. Vidal. *Order within Chaos, Towards a deterministic approach to turbulence*, John Wiley & Sons, NY, 1984.

[78] A. Wolf, J. B. Swift, H. L. Swinney and J. A. Vastana. Determining Lyapunov exponents from a time series, *Physica* **16D**, 285–317, 1985.

[79] S. Wiggins. *Global Bifurcations and Chaos*, Springer-Verlag, NY, 1988.

[80] U. Frisch. *Turbulence, the Legacy of A. N. Kolmogorov*, Cambridge University Press, Cambridge, 1995.

[81] J. M. McDonough and M. T. Huang. "A low-dimensional model of turbulence–chemical kinetics interactions," Paper ISSM3-E8 in *Proceedings of Third Int. Symp. on Scale Modeling*, Nagoya, Japan, Sept. 10–13, 2000.

[82] J. M. McDonough and M. T. Huang. A "poor man's Navier-Stokes equation": derivation and numerical experiments—the 2-D case, *Int. J. Numer. Meth. Fluids* **44**, 545–578, 2004.

[83] T. Bohr, M. H. Jensen, G. Paladin and A. Vulpiani. *Dynamical Systems Approach to Turbulence*, Cambridge University Press, Cambridge, 1998.

[84] F. Hausdorff. Dimension und äusseres mass, *Math. Ann.* **79**, 159–179, 1919.

[85] K. T. Alligood, T. D. Sauer and J. A. Yorke. *CHAOS, An Introduction to Dynamical Systems*, Springer, New York, 1997.

[86] O. E. Lanford. Strange attractors and turbulence, in *Hydrodynamics Instabilities and the Transition to Turbulence*, Swinney and Gollub (Eds.), Springer-Verlag, Berlin, pp 7–26, 1981.

[87] I. Stakgold. *Boundary Value Problems of Mathematical Physics I, II.* SIAM, Philadelphia, 2000. (Originally published by Macmillan Co., New York, 1967)

[88] R. H. Kraichnan. On Kolmogorov's inertial-range theories, *J. Fluid Mech.* **62**, 305–330, 1974.

[89] C. Foias, O. Manley, R. Rosa and R. Temam. *Navier–Stokes Equations and Turbulence*, Cambridge University Press, Cambridge, 2001.

[90] D. C. Wilcox. *Basic Fluid Mechanics*, DCW Industries, Inc., La Cañada, CA, 1997.

[91] P. A. Libby. *Introduction to Turbulence*, Taylor & Francis, Washington, DC, 1996.

[92] W. D. McComb. *The Physics of Fluid Turbulence*, Oxford University Press, Oxford, 1990.

[93] V. Yahot and S. A. Orszag. Renormalization group analysis of turbulence. I. Basic theory, *J. Sci. Comput.* **1**, 3, 1986.

[94] C. J. Freitas. Perspective: selected benchmarks from commercial CFD codes, *J. Fluids Engineering* **117**, 208–218, 1995.

[95] P. W. Berg and J. L. McGregor. *Elementary Partial Differential Equations*, Holden-Day, San Francisco, 1966.

[96] S. Pope. Pdf methods for turbulent reactive flows, *Prog. Energy Combust. Sci.* **11**, 119–192, 1985.

[97] R. L. Bishop and S. I. Goldberg. *Tensor Analysis on Manifolds*, The Macmillan Co., 1968.

[98] D. C. Wilcox. *Turbulence Modeling for CFD*, DCW Industries, Inc., La Cañada, CA, 1993.

[99] C. G. Speziale. Turbulence modeling in non-inertial frames of reference, *Theor. Comput. Fluid Dyn.* **1**, 3–19, 1989.

[100] C. Canuto, M. Y. Hussaini, A. Quarteroni and T. A. Tang. *Spectral Methods in Fluid Dynamics*, Springer-Verlag, New York, 1988.

[101] R. J. Garde. *Turbulent Flow*, John Wiley & Sons, New York, 1994.

[102] K. S. Choi and J. L. Lumley. Return to isotropy of homgeneous turbulence revisited, in *Turbulence and Chaotic Phenomena in Fluids*, Tatsumi (ed.), North-Holland, NY, pp. 267–272, 1984.

[103] C. G. Speziale. Analytical methods for the development of Reynolds-stress closures in turbulence, *Annu. Rev. Fluid Mech.* **23**, 107–157, 1991.

[104] J. G. Brasseur and P. K. Yeung. Large and small-scale coupling in homogeneous turbulence: analysis of the Navier–Stokes equations in the asymptotic limit, presented at *Eighth Symposium on Turbulent Shear Flows*, Technical University of Munich, September 9–11, 1991.

[105] X. Shen and Z. Warhaft. The anisotropy of small-scale structure in high Reynolds number ($Re_\lambda = 1000$) turbulent shear flow, *Phys. Fluids* **12**, 2976–2989, 2000.

[106] A. N. Kolmogorov. Equations of turbulent motion of an incompressible turbulent fluid, *Izv. Akad. Nauk. SSSR* Ser. Phys. VI, No. 1–2, 1942.

[107] J. Rotta. Statistische Theorie Nichthomogener Turbulenz, *Zeitsch für Physik* **129** and **131**, 1951.

[108] W. P. Jones and B. E. Launder. The prediction of laminarization with a two-equation model, *Int. J. Heat Mass Transfer* **15**, 301–314, 1972.

[109] B. E. Launder, G. J. Reece and W. Rodi. Progress in the development of Reynolds-stress turbulence closure, *J. Fluid Mech.* **68**, 537–566, 1975.

[110] P. Bradshaw. *Turbulence*, Second Edition, Springer-Verlag, Berlin, 1978.

[111] Z. Warhaft. Passive scalars in turbulent flows, *Annu. Rev. Fluid Mech.* **32**, 203–240, 2000.

[112] F. A. Williams. *Combustion Theory* Second Edition, The Benjamin/Cummings Pub. Co., Inc., Menlo Park, CA, 1985.

[113] J. M. McDonough. On Intrinsic Errors in Turbulence Models Based on Reynolds-Averaged Navier-Stokes Equations, *Int. J. Fluid Mech. Res.* **22**, 27–55, 1995.

[114] H. L. Royden. *Real Analysis, 2nd ed.* Macmillan, New York, 1971.

[115] O. Métais and M. Lesieur. Spectral large-eddy simulation of isotropic and stably stratified turbulence, *J. Fluid Mech.* **256**, 157–194, 1992.

[116] W. G. Vincenti and C. H. Kruger. *Introduction to Physical Gas Dynamics*, John Wiley & Sons, Inc., New York, 1965.

[117] J. Chen, J. Katz, and C. Meneveau The Implication of Mismatch between Stress and Strain-Rate in Turbulence Subjected to Rapid Straining and Destraining on Dynamic LES Models, *J. Fluids Eng.* bf 127, 840–850, 200).

[118] P. Y. Chou. On the velocity correlations and the solution of the equations of turbulent fluctuation, *Quart. Appl. Math.* **3**, 3, 1945.

[119] B. E. Launder and B. I. Sharma. Application of the energy dissipation model of turbulence to the calculation of flow near a spinning disc, *Letters in Heat and Mass Transfer* **1**, 131–138, 1974.

[120] J. O. Hirschfelder, C. F. Curtiss and R. B. Bird. *Molecular Theory of Gases and Liquids*, John Wiley & Sons, New York, 1954.

[121] E. R. Van Driest. On turbulent flow near a wall, *J. Aeronaut. Sci.* **23**, 1007, 1956.

[122] T. Cebeci and A. M. O. Smith. Analysis of turbulent boundary layers, *Ser. in Appl. Math. & Mech.*, Vol. XV, Academic Press, 1974.

[123] B. S. Baldwin and H. Lomax. Thin-layer approximation and algebraic model for separated turbulent flows, AIAA Paper 78-257, 1978.

[124] P. R. Spalart and S. R. Allmaras. A one-equation turbulence model for aerodynamic flows, AIAA Paper 92-439, 1992.

[125] B. Mohammadi and O. Pironneau. *Analysis of the K-Epsilon TURBULENCE MODEL*, John Wiley & Sons, Chichester, 1994.

[126] P. A. Durbin and B. A. Pettersson Reif. *Statistical Theory and Modeling for Turbulent Flows*, John Wiley & Sons, Chichester, 2001.

[127] D. J. Tritton. *Physical Fluid Dynamics*, Van Nostrand Reinhold, New York, 1977.

[128] J. R. Chasnov. Simulation of the Kolmogorov inertial subrange using an improved subgrid model, *Phys. Fluids A* **3**, 188–200, 1991.

[129] A. M. Obukhov. On the distribution of energy in the spectrum of turbulent flow, *Dokl. Akad. Nauk SSSR* **32**, 22–24, 1941; Spectral energy distribution in a turbulent flow, *Izv. Akad. Nauk SSSR Ser. Geogr. Geofiz.* **5** 453–466, 1941.

[130] J. Smagorinsky. General circulation experiments with the primitive equations. I: The basic experiment, *Month. Weath. Rev.* **91**, 99–165, 1963.

[131] W. E and B. Engquist. Multiscale modeling and computation, *Notices AMS* **50**, 1062–1070, 2003.

[132] J. Bardina, J. H. Ferziger and W. C. Reynolds. Improved subgrid scale models for large eddy simulation, AIAA Paper 80-1357.

[133] C. Fureby and F. F. Grinstein. Large eddy simulation of high-Reynolds-number free and wall-bounded flows, *J. Comput. Phys.* **181**, 68–97, 2002.

[134] J. P. Boris, F. F. Grinstein, E. S. Oran and R. L. Kolbe. New insights into large-eddy simulation, *Fluid Dyn. Res.* **10**, 199–228, 1992.

[135] A. R. Kerstein. *J. Fluid Mech.* **216**, 411, 1990.

[136] E. C. Hylin and J. M. McDonough. Chaotic small-scale velocity fields as prospective models for unresolved turbulence in as additive decomposition of the Navier–Stokes equations, *Int. J. Fluid Mech. Res.* **26**, 539–567, 1999.

[137] C. Hirsch. *Numerical Computation of Internal and External Flows, Volume 2: Computational methods for inviscid and viscous flows*, John Wiley & Sons, Chichester, 1990.

[138] M. Germano. Turbulence: the filtering approach, *J. Fluid Mech.* **238**, 325–336, 1992.

[139] D. K. Lilly. A proposed modification of the Germano subgrid-scale closure method, *Phys. Fluids A* **4**, 633–635, 1992.

[140] J. Bardina, J. H. Ferziger and W. C. Reynolds. Improved turbulence modelsbased on large eddy simulation of homogeneous incompressible turbulent flows, Report TF-19, Thermosciences Division, Dept. Mechanical Engineering, Stanford University, 1983.

[141] C. G. Speziale. Galilean invariance of subgrid-scale stress models in the large-eddy simulation of turbulence, *J. Fluid Mech.* **156**, 55–62, 1985.

[142] O. Ladyzhenskaya. A dynamical system generated by the Navier–Stokes equations. *J. Soviet Math.* **3**, 458–479, 1973.

[143] O. Ladyzhenskaya. *Attractors for Semigroups and Evolution Equations*, Cambridge University Press, Cambridge, 1991.

[144] R. Temam. *Navier–Stokes Equations: Theory and Numerical Analysis*, North-Holland Pub. Co., Amsterdam, 1979. (new edition published by Amer. Math. Soc., Providence, RI, 2001)

[145] R. Temam. *Navier–Stokes Equations and Nonlinear Functional Analysis*, Soc. Indust. Appl. Math., Philadelphia, 1983. (2^{nd} edition published by SIAM, 1995)

[146] R. Temam. *Infinite Dimensional Dynamical Systems in Mechanics and Physics*, Springer-Verlag, New York, 1988.

[147] C. R. Doering and J. D. Gibbon. *Applied Analysis of the Navier–Stokes Equations*, Cambridge University Press, Cambridge, 1995.

[148] K. E. Gustafson *Introduction to Partial Differential Equations and Hilbert Space Methods*, John Wiley & Sons, New York, 1980.

[149] L. Schwartz. *Théorie des Distributions I, II.* Hermann, Paris, 1950, 1951.

[150] A. R. Mitchell and D. F. Griffiths. *The Finite Difference Method in Partial Differential Equations*, John Wiley & Sons, Inc., Chichester, 1980.

[151] F. Treves. *Basic Linear Partial Differential Equations*, Academic Press, New York, 1975.

[152] J.-P. Aubin. *Applied Functional Analysis*, John Wiley & Sons, New York, 1979.

[153] J. A. Goldstein. *Semigroups of Linear Operators and Applications*, Oxford University Press, New York, 1985.

[154] F. G. Shuman. Numerical method in weather prediction: smoothing and filtering, *Mon. Weath. Rev.* **85**, 357–361, 1957.

[155] P. R. Garabedian. *Partial Differential Equations*, John Wiley & Sons, New York, 1964.

[156] P. M. Gresho and R. L. Sani. On pressure boundary conditions for the incompressible Navier–Stokes equations, *Int. J. Numer. Methods Fluids* **7**, 1111–1145, 1987.

[157] P. M. Gresho. Incompressible fluid dynamics: some fundamental formulation issues, *Annu. Rev. Fluid Mech.* **23**, 413–453, 1991.

[158] H.-O. Kreiss and J. Lorenz. *Initial-Boundary Value Problems and the Navier–Stokes Equations*, Academic Press, Boston, 1989.

[159] T. B. Gatski, C. E. Grosch and M. E. Rose. A numerical study of the two-dimensional Navier–Stokes equations in vorticity–velocity variables, *J. Comput. Phys.* **48**, 1–22, 1982.

[160] F. H. Harlow and J. E. Welch. Numerical calculation of time-dependent viscous incompressible flow of fluid with free surface, *Phys. Fluids* **8**, 2182–2189, 1965.

[161] C. W. Hirt, A. A. Amsden and J. L. Cook. An arbitrary Lagrangean-Eulerian computing method for all flow speeds, *J. Comput. Phys.* **14**, 227, 1774.

[162] C. M. Rhie and W. L. Chow. Numerical study of the turbulent flow past an airfoil with trailing edge separation, *AIAA J.* **11**, 1525–1532, 1983.

[163] S. V. Patankar. *Numerical Heat Transfer and Fluid Flow*, McGraw-Hill Book Co., New York, 1980.

[164] Y. Zang, R. L. Street and J. R. Koseff. A non-staggered grid, fractional step method for time-dependent incompressible Navier–Stokes equations in curvilinear coordinates, *J. Comput. Phys.* **114**, 18–33, 1994.

[165] J. M. McDonough. *Lectures in Basic Computational Numerical Analysis*, available in PDF format at the following URL: http://www.engr.uky.edu/∼egr537

[166] J. Douglas, Jr. and J. E. Gunn. A general formulation of alternating direction methods, part I. parabolic and hyperbolic problems, *Numer. Math.* bf 6 428–453, 1964.

[167] J. M. McDonough. *Lectures in Computational Numerical Analysis of Partial Differential Equations*, available in PDF format at the following URL: http://www.engr.uky.edu/∼me690

[168] M. D. Gunzburger. *Finite Element Methods for Viscous Incompressible Flows*, Academic Press, Boston, 1989.

[169] J. Kim and P. Moin. Application of a fractional step method to incompressible Navier–Stokes equations, *J. Comput. Phys.* **59**, 308–323, 1985.

[170] D. Tafti. Alternate formulations for the pressure equation Laplacian on a collocated grid for solving the unsteady incompressible Navier–Stokes equations, *J. Comput. Phys.* **116**, 143–153, 1995.

[171] C. A. J. Fletcher. *Computational Techniques for Fluid Dynamics, Volume II*, Springer-Verlag, Berlin, 1988.

[172] R. E. Mickens. *Difference Equations*, Van Nostrand Rheinhold Co., New York, 1987.

[173] G. de Vahl Davis and G. D. Mallinson. An evaluation of upwind and central difference approximations by a study of recirculating flow, *Computers and Fluids* **4**, 29–43, 1976.

[174] D. B. Spalding. A novel finite-difference formulation for differential expressions involving both first and second derivatives, *Int. J. Num. Methods Eng.* **4**, 551, 1972.

[175] W. Shyy. A study of finite difference approximations to steady-state convection-dominated flow problems, *J. Comput. Phys.* **57**, 415–438, 1985.

[176] T. Yang and J. M. McDonough. Solution filtering technique for solving Burgers' equation, accepted for special issue of *Discrete and Continuous Dynamical Systems*, 2003.

[177] J. M. McDonough and T. Yang. Solution filtering technique for solving turbulence problems, to be submitted to *J. Comput. Phys.*, 2003.

[178] J. M. McDonough, Y. Yang and E. C. Hylin. Modeling Time-dependent turbulent flow over a backward-facing step via additive turbulent decomposition and chaotic algebraic maps, presented at *First Asian Computational Fluid Dynamics Conference*, Hong Kong, Jan. 16–19, 1995.

[179] J. M. McDonough, V. E. Garzon and D. Schulte. Effect of film-cooling hole location on turbulator heat transfer enhancement in turbine blade internal air-cooling circuits," presented at *ASME TURBO EXPO 99*, Indianapolis, IN, June 7–10, 1999.

[180] J. M. McDonough, T. Yang and M. Sheetz. Parallelization of a modern CFD incompressible turbulent flow code, to be presented at *Parallel CFD 2003*, Moscow, May 13–15, 2003.

[181] W. F. Ames. *Numerical Methods for Partial Differential Equations* Second Edition, Academic Press, New York, 1977.

[182] R. Shapiro. Smoothing, filtering and boundary effects, *Rev. Geophys. and Space Phys.* **8**, 359–387, 1970.

[183] A. Majda, J. McDonough, and S. Osher. The Fourier method for nonsmooth initial data, *Math. Comput.* **32**, 1041–1081, 1978.

[184] J. M. McDonough, R. J. Bywater and J. C. Buell. An Investigation of Strange Attractor Theory and Small-Scale Turbulence, AIAA Paper # AIAA-84-1674, presented at *AIAA* 17^{th} *Fluid Dynamics, Plasma Dynamics, and Lasers Conference*, Snowmass, CO, June 25–27, 1984.

[185] J. M. McDonough, J. C. Buell and R. J. Bywater. A Comparison of Routes to a Strange Attractor in One-Dimensional Local Models of Turbulent Free and Forced Convection, ASME Paper # 84-WA/HT-16, presented at *ASME Winter Annual Meeting*, New Orleans, LA, Dec. 9–14, 1984.

[186] J. M. McDonough and R. J. Bywater. Effects of local large-scale parameters on the small-scale chaotic solutions to Burgers' equation, AIAA Paper 85-1653, presented at *AIAA* 18^{th} *Fluid Dynamics, Plasmadynamics and Lasers Conference*, Cincinnati, OH, July 16–18, 1985.

[187] J. M. McDonough and R. J. Bywater. Large-scale effects on local small-scale chaotic solutions to Burgers' equation, *AIAA J.* **24**, 1924–1930, 1986.

[188] J. M. McDonough and R. J. Bywater. Turbulent solutions from an unaveraged, additive decomposition of Burgers' equation, presented at *ASME Fluids Engineering Conference*, San Diego, CA, July 9–12, 1989 and contained in the proceeding volume *Forum on Turbulent Flows—1989*, Bower and Morris (Eds.), American Society of Mechanical Engineers, New York, 1989.

[189] M. Marion and R. Temam. Nonlinear Galerkin Methods, *SIAM J. Numer. Anal.* **26**, 1139–1157, 1989.

[190] S. Chandrasekhar. *Hydrodynamic and Hydromagnetic Stability*, Dover Publications, Inc., New York, 1981.

[191] S. Cioni, S. Ciliberto and J. Sommeria. Strongly turbulent Rayleigh–Bénard convection in mercury: comparison with results at moderate Prandtl number, *J. Fluid Mech.* **335**, 111–141, 1997.

[192] A. Libchaber. From Chaos to Turbulence in Benard Convection, *Proc. Roy. Soc. London, Series A* **413**, No. 1844, Dynamical Chaos, 63–69, 1987.

[193] R. M. May. Simple mathematical models with very complicated dynamics, *Nature* **261**, 149–157, 1976.

[194] M. Hénon. A two-dimensional mapping with a strange attractor, *Commun. Math. Phys.* **50**, 69–77, 1976.

[195] T. H. Pulliam and J. A. Vastano. *J. Comput. Phys.* **105**, 133–149, 1993.

[196] W. E and B. Engquist. The heterogeous multiscale methods, *Comm. Math. Sci.* **1**, 87–132, 2003.

[197] V. A. Marchenko and E. Ya. Khruslov. *Homogenization of Partial Differential Equations*, Springer, 2006.

[198] N. A. Adams and S. Stolz. Deconvolution methods in subgrid-scale approximations in LES, in *Modern Simulation Strategies for Turbulent Flow*, B. J. Geurts, (Ed.), R. T. Edwards, Inc., pp 21–44, 2001.

[199] M. J. Berger and J. Oliger. *J. Comp. Phys.* **53**, 484–512, 1984.

[200] Y. Yang and J. M. McDonough. Bifurcation studies of Navier-Stokes equations via additive turbulent decomposition, presented at *ASME Winter Annual Meeting*, Anaheim, CA, Nov. 8–13, 1992.

[201] J. M. McDonough and D. Wang. Parallelization of algorithms based on additive decomposition of the Navier-Stokes equations, presented at *Parallel CFD '94*, Kyoto, Japan, May 16–19, 1994.

[202] R. Brown, P. Perry and Z. Shen. The Additive Turbulent Decomposition for the Two-Dimensional Incompressible Navier-Stokes Equations: Convergence Theorems and Error Estimates, *SIAM Appl. Math.* **59**, 139–155, 1999.

[203] N. N. Yanenko. *The Method of Fractional Steps*, Springer-Verlag, Berlin, 1971.

[204] J. M. McDonough. *Lectures in Computational Fluid Dyamics of Incompressible Flow: Mathematics, Physics and Algorithms*, available at http://www.engr.uky.edu/∼acfd, 2007.

[205] J. M. McDonough. A 'synthetic scalar' subgrid-scale model for large-eddy simulation of turbulent combustion, *Proc. 2002 Spring Tech. Mtg. Central States Sec., Combust. Inst.*, Knoxville, TN, April 7–9, 2002.

[206] E. C. Hylin. Doctoral Dissertation in Mechanical Engineering, University of Kentucky, Lexington, 1997.

[207] J. M. McDonough, S. A. Bible and J. Scoville, Response to Strain Rate in a Discrete Dynamical System Model of the High-Wavenumber Navier–Stokes Equation, *Journal of Turbulence* **JOT4**, Dec. 2003.

[208] S. Paolucci. Direct numerical simulation of two-dimensional turbulent natural convection in an enclosed cavity, *J. Fluid Mech.* **215**, 229–262, 1990.

[209] J. M. McDonough and D. Joyce. A discrete dynamical system subgrid-scale model of turbulent convection, *AIAA Thermophysics and Heat Transfer Meeting and Exhibit*, St. Louis, MO, June 24–27, 2002.

[210] J. M. McDonough and Sha Zhang. LES subgrid-scale models of turbulence-chemical kinetics interactions based on discrete dynamical systems, *AIAA Fluid Dynamics Conference and Exhibit*, St. Louis, MO, June 24–27, 2002.

[211] J. M. McDonough and Sha Zhang. Discrete dynamical systems models of turbulence-chemical kinetics interactions, *Intersociety Energy Conversion Engineering Conference*, Washington, DC, July 28–31, 2002.

[212] F. Treves. *Introduction to Pseudodifferential and Fourier Integral Operators, Vol. 1 Pseudodifferential Operators*, Plenum Press, New York, 1980.