Dimensional Analysis of Models and Data Sets: Similarity Solutions and Scaling Analysis

James F. Price

Woods Hole Oceanographic Institution

http://www.whoi.edu/science/PO/people/jprice/

Octobe 4, 2002

Summary. This essay describes a three-step procedure of dimensional analysis that can be applied to all quantitative models and data sets. In rare but important cases the result of dimensional analysis will be a solution; more often the result is an efficient way to display a large or complex data set. The first step of an analysis is to define an appropriate physical model, which is nothing more than a list of the dependent variable and all of the independent variables and parameters that are thought to be significant. The premise of dimensional analysis is that a complete equation made from this list of variables will be independent of the choice of units. This leads to the second step, calculation of a null space basis of the corresponding dimensional matrix (a computer code is made available for this). To each vector of the null space basis there corresponds a nondimensional variable, the number of which is less than the number of dimensional variables. The nondimensional variables are themselves a basis set and in most cases their form is not determined by dimensional analysis alone. The third step is to choose an optimal form for this basis set. One strategy is to nondimensionalize the dependent variable by a physically meaningful 'zero order' solution. When carried to completion this leads naturally to a scaling analysis. The remaining nondimensional variables can then be formed in ways that define the geometry of the problem or that correspond to the ratios of terms in a model equation, e.g., the Reynolds number or Froude number.

Preface. This essay is an introduction to dimensional analysis at about the level appropriate for a first course in fluid dynamics. It is intended to supplement the discussions of dimensional analysis that can be found in most comprehensive fluid dynamics and applied mathematics text books.

A condensed version of this essay, essentially Secs. 2-5 and 8, has been published by J. F. Price, 'Dimensional analysis of models and data sets', American Journal of Physics, **71** (5), 437–447 (2003). If you wish to cite material from those sections, please use that title. If you need to cite material from Secs. 6 or 7 of this manuscript, then cite the present title as an unpublished manuscript available from the web address above.

The method of dimensional analysis described here in Sec. 4 is believed to be somewhat novel, but it is impossible to know all of the literature on a topic that has roots more than a century deep and in many different fields. References are cited wherever they are known, but it is emphasized that the purpose of this essay is to be educational, and not to report research findings.

1 About dimensional analysis

Dimensional analysis is a remarkable tool in so far as it can be applied to any and every quantitative model or data set; recent applications include topics from donuts to dinosaurs[1] and the most fundamental theories of physics[2]. The results of dimensional analysis can be of greater or lesser value. It is most useful, indeed almost indispensible, for problems having no solvable theory. Dimensional analysis can always make a little progress towards a solution, and some of these, the universal spectrum of inertial-range turbulence and the log-layer profile of a turbulent boundary layer, are landmarks in fluid mechanics. More often the result of dimensional analysis is a broad hint at the form of a solution or a more efficient way to display or correlate a large data set. These kinds of results, though seldom complete if taken alone, are nevertheless an essential building block of many investigations.

It is fair to warn that dimensional analysis is not easy to learn, despite that the mathematical steps required are elementary. More than offsetting this is that the premise of dimensional analysis is rather abstract — that complete equations can be written in a form independent of the choice of units — and that the result of a dimensional analysis is usually a family of solutions from which we are free to choose the most useful. These combine to make dimensional analysis seem artful and when done quickly, downright mysterious. We would not want to dispel the mystique of dimensional analysis completely (it persists among those with considerable experience) but this essay does aim to teach a method of dimensional analysis that is rigorous and objective, and, so the hope goes, understandable.

The plan is to demonstrate this method of dimensional analysis on several familiar problems from classical physics; the simple viscous pendulum, Secs. 2-5, diffusion in one dimension, Sec. 6, and projectile motion in variable gravity, Sec. 7. Dimensional analysis has all the makings of a full mathematical analysis, though in a somewhat compressed format. The first and most important step is to define a problem having one dependent variable. The physical model for this problem is nothing more than a list of all of the independent variables and parameters that are thought to be important to determining the outcome of the dependent variable (Sec. 2). That something useful could follow from such a minimal specification is at the heart of what makes dimensional analysis so widely seful and also a bit mysterious. Physical models as first written are likely to be quite general. Anything that can be done to hone the physical model or add physical constraints will make the subsequent analysis much more useful. The second and mathematical step is to find the nondimensional form that the variables must take assuming that they are in a complete equation (Secs. 3 and 4). The usual method[3] of finding the nondimensional variables relies upon the Buckingham Pi theorem and is tedious and arcane. Here the nondimensional variables are computed using a method from linear algebra that is readily automated.[4] The third and final step of a dimensional analysis is to assemble the initial basis set of nondimensional variables into an optimum form (examples in Secs. 5 and 7). This requires some sense of the intent and possible uses for the analysis. When this can be combined with a minimal or zero-order solution of the phenomenon, dimensional analysis develops naturally into scaling analysis (Sec. 7). This essay will emphasize the interpretive aspects of a dimensional analysis — specification of an appropriate physical model and the choice of the basis set — once the purely mathematical second step has been set aside.

2 Models of a simple pendulum

The first problem we take up in some depth is that of a simple pendulum. Consider a pendulum that can be made and observed usefully with very inexpensive tools and materials; a small lead fishing sinker having a mass of a few tens of grams suspended on a thin monofilament line a few meters in length. The motion of this kind of pendulum is only lightly damped by drag with the surrounding air and can be characterized by two distinct time scales – a regular, fast time-scale oscillation having a period, P, and a slow, more-or-less exponential decay with a time-scale, Γ^{-1}. Our specific goal in Secs. 2-5 will be to learn how these time scales and some other variables, e.g., tension in the line, vary with line length, mass of the bob, etc. (If the simple pendulum is too familiar, skip ahead to Sec. 3. If the use of nondimensional variables is also familiar, skip ahead to Sec. 4)

2.1 A physical model

To analyze the motion of this pendulum we begin by listing the variables that are presumed to be relevant to the aspect of the motion that is of interest. To start, consider the fast time-scale, oscillatory motion. The line will be idealized as rigid, so that the bob must swing along a constant radius. The motion of the bob is then defined by the angle of the line to the vertical, $\phi(t)$, and its time derivatives; the angle ϕ is the dependent variable of this physical model and the time, t, is the only independent variable. Several properties of the pendulum would seem to be of importance — the mass of the bob, M, the length of the supporting line, L, and the acceleration of gravity, g. To account for why there is motion at all, the initial angle, ϕ_0, or an initial angular velocity (here assumed to be zero) must also be included. This list of relevant variables constitutes

- A physical model for the oscillatory motion of a simple, inviscid pendulum:

1. the angle of the line, $\phi \doteq$ nond, the dependent variable,
2. time, $t \doteq m^0 l^0 t^1$, the independent variable,
3. mass of the bob, $M \doteq m^1 l^0 t^0$, a parameter,
4. length of the line, $L \doteq m^0 l^1 t^0$, a parameter,
5. acceleration of gravity, $g \doteq m^0 l^1 t^{-2}$, a parameter,
6. the initial angle, $\phi_0 \doteq$ nond, a parameter.

The notation $X \doteq m^a l^b t^c$ indicates the dimensions mass, length and time (or nond if the variable is nondimensional). Parameters are variables that are constant during a particular realization — M, L, g and ϕ_0 in this list — but that vary over some range that defines the family of pendulums and environments that are of interest.

2.2 A mathematical model

Dimensional analysis is most useful in the case that a mathematical model is not known. Mathematical models of the simple pendulum are well-known and we will use them to generate numerical data and to show how dimensional analysis can be applied to a mathematical model. For an inviscid pendulum the rate of change of angular momentum of the bob is due solely to the torque associated with the downward force of gravity acting on the bob,

$$L^2 M \frac{d^2\phi}{dt^2} = -LMg \sin\phi. \tag{1}$$

If we divide by $L^2 M$, the equation of motion becomes

$$\frac{d^2\phi}{dt^2} = -\frac{g}{L} \sin\phi. \tag{2}$$

For experimental purposes it is preferable to start from a state of rest and so the initial conditions at $t = 0$ are taken to be

$$\phi = \phi_0 \quad \text{and} \quad \frac{d\phi}{dt} = 0. \tag{3}$$

It may also be of interest to compute the tension in the line, T, from the radial equation of motion, $dr/dt = 0$, and thus

$$T = gM\cos\phi + LM(\frac{d\phi}{dt})^2. \tag{4}$$

The appropriate solution method to Els. (2) and (3) depends upon the initial angle, ϕ_0. If ϕ_0 is restricted to values less than about 0.1 radian, then $\sin\phi$ in Eq. (2) can be approximated well by ϕ and the resulting linear model has the well-known solution

$$\phi = \phi_0 \cos(t/\sqrt{L/g})\,. \tag{5}$$

In the general case where ϕ_0 may take any value from $-\pi$ to π, Eq. (2) is nonlinear and a solution cannot be given in elementary functions. Numerical integration of the nonlinear model Eqs. (2) and (3) is straightforward, however, and yields (numerical) data (Fig. 1) that we will treat as an intermediary between experiment and theory; we know exactly the physical model, but not the specific parameter dependence.

2.3 Models generally

Model equations are a relation between a dependent variable, the angle ϕ or the tension T, and the independent variables and parameters that make up the physical model. Even if we had no idea of the mathematical model, we could still assert that a complete physical model could be used to define a relation

$$\phi = F(t, g, L, M, \phi_0)\,, \tag{6}$$

where F will be used to indicate an unknown function. If our goal was to solve for the period of the oscillation, then we would evaluate the time at some (arbitrary) repeated value of ϕ to find

$$P = F(g, L, M, \phi_0)\,. \tag{7}$$

For the tension, T, and the maximum tension during an oscillation, $T_{\max}$, we could similarly write

$$T = F(t, g, L, M, \phi_0)\,, \tag{8}$$

and

$$T_{\max} = F(g, L, M, \phi_0)\,. \tag{9}$$

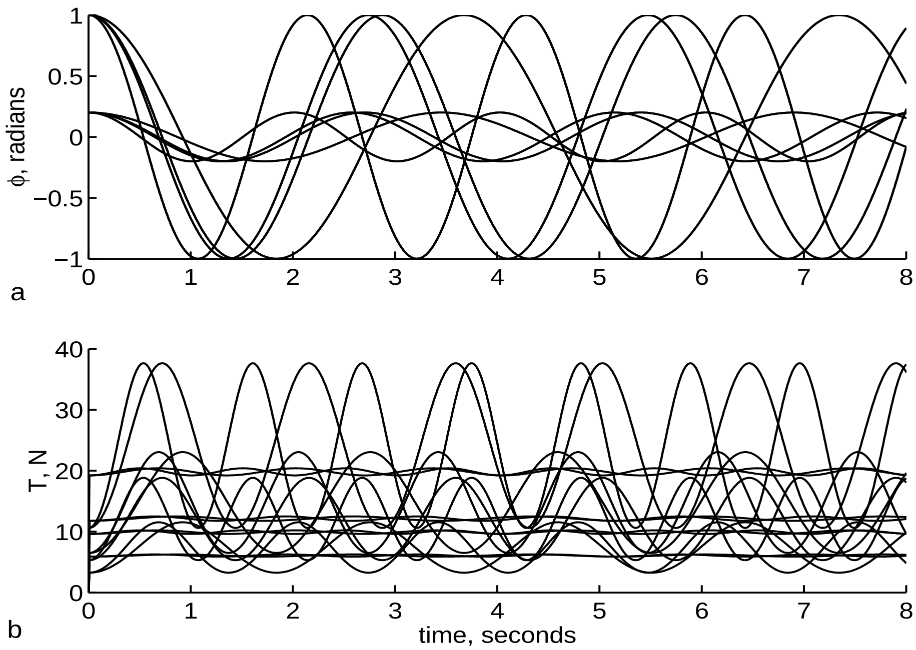

Figure 1: Numerical solutions of the simple, inviscid pendulum for two values each of L $(1, 1.8)$ m, M $(1, 2)$ kg, g $(9.8, 6)$ m^2/s and ϕ_0 $(0.2, 1.0)$ radians or 16 solutions in total. (a) The angle, ϕ. The mathematical model Eqs. (2) and (3) does not depend upon M, and so there are eight distinct solutions here. (b) Tension (Newtons) for the same set of solutions. Here there are 16 distinct solutions, though some are difficult to distinguish. As these data were acquired it was noticed that the maximum tension did not vary with L.

It will often happen that the list of variables for the physical model will include one or more parameters that do not appear in the mathematical model. If we compare Eqs. (6) and (2), the physical model includes the mass, M, while in the mathematical model, the mass appeared as a coefficient in the gravitational force (right side of Eq. (1)) and in the inertial force (left side) and cancels. In this regard, the mathematical model, Eq. (2), is a considerable advance over the physical model, Eq. (6). Note too that the angular velocity, $d\phi/dt$, appears in the mathematical model for the tension, Eq. (4), although not in the physical model Eq. (8). Even if we were aware that the mathematical model of tension depended upon $d\phi/dt$, we should still omit this second dependent variable from Eq. (8) because $d\phi/dt$ must itself depend upon t, g, L, M and ϕ_0 and should not be written into the physical model again.

The relations (6)–(9) could be written in one of several forms, for example,

$$\phi/F(t, g, L, M, \phi_0) = 1, \tag{10}$$

or reusing F yet again,

$$F(\phi, t, g, L, M, \phi_0) = 1. \tag{11}$$

What is most important is the assertion that the physical model is complete, meaning that it includes all of the variables required to construct a mathematical model that could in principle yield a unique solution. If we do not know the corresponding mathematical model, then completeness can only be a hypothesis.

While it is essential that the physical model be complete, it is also highly desirable that the physical model be as concise as possible, i.e., that it includes only those variables that have a significant effect upon the dependent variable. The selection of variables for the physical model thus requires considerable judgment.

3 An informal dimensional analysis

3.1 Invariance to a change of units

We take it for granted that every equation must be dimensionally consistent, or homogeneous.[6] But how about the units used to measure length, time, etc.? The premise of dimensional analysis is that the physical relationship expressed by a complete equation is not dependent upon the choice of units, that is, whether SI, British engineering, or any other.[5] Invariance to the choice of units implies a constraint on the form that the dimensional variables can take in a complete equation, and dimensional analysis is a systematic procedure for learning what that form is.

Angles are an interesting and relevant case. An angle is the ratio of two lengths, an arc length and a radius, and is thus inherently nondimensional. (Angles may be specified in units of radians or degrees, among others.) If we compute an angle ϕ by measurements of arc length and radius in units of meters, we will get a certain number. If we then use feet to measure these same lengths, we will get precisely the same number, that is, the same angle. Thus the left side of Eq. (6) is invariant to a change in the units of length. How about the right side of Eq. (6)? For invariance to the choice of units to hold, the length and the acceleration of gravity must appear in the ratio g/L (or any power of the ratio, for example, $\sqrt{L/g}$), and not as g or L separately, because the latter would imply a change of F with a change in the units of length. Thus, we already know something about the invariant form of Eq. (6). Consider the mass, M. A change in the units of mass should also leave F unchanged, and yet it is impossible to see how that could hold since M is the only variable in the physical model having mass dimensions. This informal analysis leads to the conclusion that an equation for ϕ that is invariant to a change of units cannot depend upon the mass of the bob alone. This conclusion is an obvious result of the mathematical model, Eqs. (2) and (3), but can be deduced by dimensional analysis in the absence of the mathematical model. A similar consideration of the units used to measure time indicates that t and g must also appear together in a nondimensional variable, say $t/\sqrt{L/g}$. Again, any power of this variable is possible, but we might as well leave the independent variable t to the first power. The upshot of this reasoning is that the variables that appear in a form of Eq. (6) that is invariant to a change of units can only appear in a nondimensional form, perhaps the simplest (but not the only) form being

$$\phi = F(t/\sqrt{L/g}, \phi_0)\,. \tag{12}$$

The essential result is that in place of a dependence on one independent dimensional variable and four parameters as in the original Eq. (6), we now have a dependence on one nondimensional independent variable, $t/\sqrt{L/g}$, and one nondimensional parameter. When the data of Fig. 1 are plotted using this nondimensional format in Fig. 2, there is a very significant reduction in the volume of data required to display and define the data set, an important benefit of dimensional analysis applied to a presentation of data.

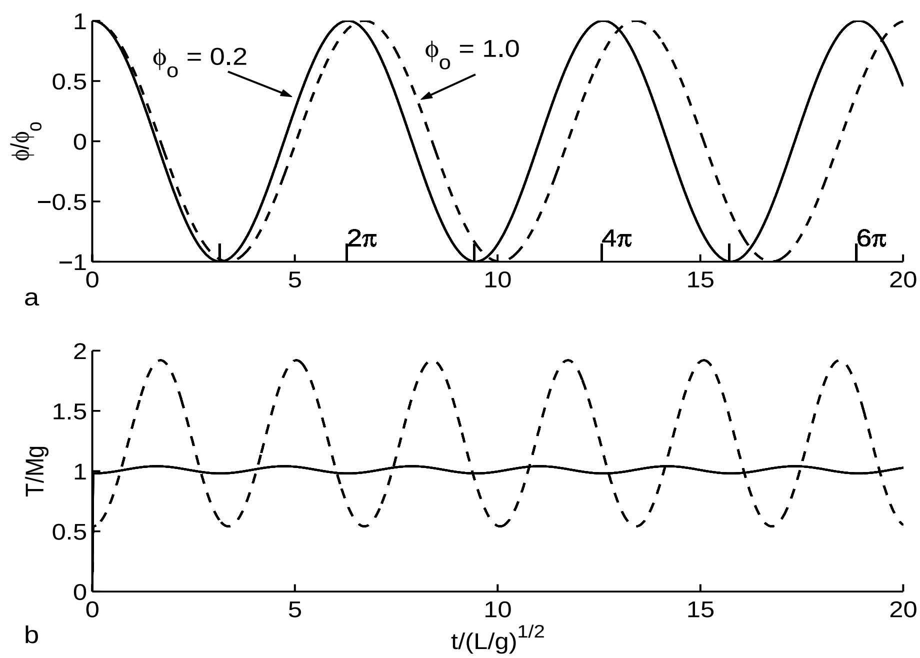

Figure 2: The numerical solutions of Fig. 1 (two values each of L, M, g, and ϕ_0) plotted in a nondimensional format. The time is nondimensionalized by $\sqrt{L/g}$. In (a) the angle ϕ is normalized by the initial angle, ϕ_0. This helps us to compare the period of the two solutions, but obscures the important difference in amplitude. The eight distinct solutions of Fig. 1a collapse to just two curves that correspond to the cases $\phi_0 = 0.2$ (solid curve) and $\phi_0 = 1.0$ (dashed curve). In (b) the tension is nondimensionalized by Mg. The 16 separate curves of Fig. 1b collapse to just two curves that have the ϕ_0 as in (a).

The period of the motion can be written in a way analogous to Eq. (7) as

$$P/\sqrt{L/g} = F(\phi_0). \tag{13}$$

If ϕ_0 is small, say less than about 0.1 radian, the dependence on ϕ_0 is found experimentally to be negligible (Fig. 3a), and $F(\phi_0 \ll 1) = \text{constant}$. The period of a simple pendulum undergoing small amplitude oscillations thus increases in proportion to the square root of the length of the supporting line divided by the local acceleration of gravity, g. This famous result is often attributed to Galileo Galilei, who observed the motion of (inadvertent) pendulums in a Pisa cathedral. The measurement of the period of just one linear pendulum is sufficient to fix the constant, $F(\phi_0 \ll 1) = 2\pi$, for all such pendulums. If ϕ_0 is not small, then from dimensional analysis and Eq. (13) it is evident that the nondimensional period will depend upon the single parameter ϕ_0. The function $F(\phi_0)$, often referred to as a similarity law,[6] might be determined by experiment (assuming that viscous effects are negligible), by the analysis of numerical simulations, or from theory (Fig. 3a).

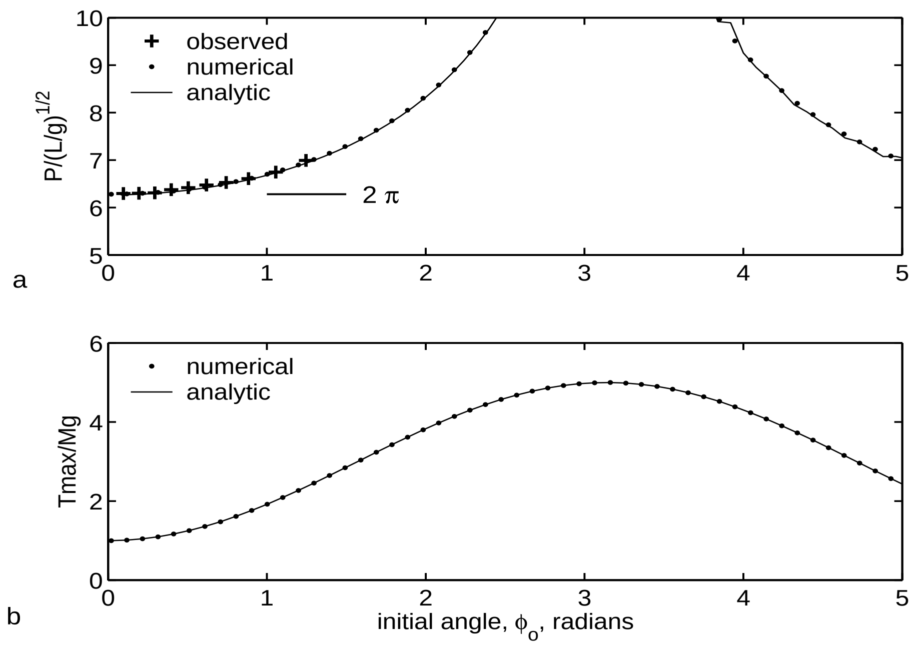

Figure 3: (a) The period of a simple pendulum as diagnosed from a series of numerical solutions (dots), as computed from theory that yields an elliptic integral that is also evaluated numerically (solid line), and as observed (crosses). The observations were acquired by measuring the time required for ten oscillations of a nearly conservative pendulum using an electronic stopwatch (the observed period is accurate to about 0.3%). The flexible line of this pendulum and the initial condition $d\phi/dt = 0$ limit the initial angle to about $-\pi/2 < \phi_0 < \pi/2$. The period goes to infinity as $\phi_0 \rightarrow \pi$ because the initial condition is $\frac{d\phi}{dt}(t = 0) = 0$. From dimensional analysis we expect that this result, $F(\phi_0)$ from the right side of Eq. (13), holds for all simple, inviscid pendulums. (b) The maximum tension (during an oscillation), T_{max}, diagnosed from a series of numerical solutions (dots) and as computed from energy conservation and the radial equation of motion (solid line); we had no means to observe this variable.

3.2 Natural units

A complementary way to come to the same result is to consider the units used to measure time in the mathematical model, Eqs. (2) and (3). There is no compelling reason to use seconds, aside from the practical convenience that clocks are calibrated in this unit. But suppose that our aim was to simplify the mathematical model by choosing a unit of time that is natural to the problem itself. The natural time scale of the pendulum is, of course, the (linear) period, which can be used to define a nondimensional time (omitting the factor 2π),

$$t^* = t/(P/2\pi) = t/\sqrt{L/g}. \tag{14}$$

The variable t^* is a pure number that has the same numerical value regardless of the units used to measure t, g, and L, a hint that there might be something useful here.

Nondimensional time may sound a little esoteric, but amounts to nothing more than counting time in units of the linear period while taking explicit account of the $\sqrt{L/g}$ dependence of the period. If we were to consider only one pendulum, then the whole exercise would amount to dividing the time by a constant. But if we were to consider all possible pendulums (all possible L and g), then there is real merit in this. To see why, let's follow through by rewriting the equation of motion, Eq. (2), using the nondimensional time, t^*. Time derivatives transform as $dt = dt^*\sqrt{(L/g)}$ and so the equation of motion becomes

$$\frac{d^2\phi}{dt^{*2}} = -\sin\phi\,, \tag{15}$$

with the initial conditions as before. The solution will be of the form

$$\phi = F(t^*, \phi_0)\,, \tag{16}$$

which is just like Eq. (12). If the amplitude of the motion is small, then the linear solution of Eq. (15) is just

$$\phi = \phi_0 \cos t^*. \tag{17}$$

The dependence upon L and g has not been omitted, but is rather subsumed into the nondimensional time, t^*, so that Eq. (17) suffices for all L and g.

Recall that the linear pendulum has the solution $\phi = \phi_0 \cos(t/\sqrt{L/g})$, and note that the argument of that cosine function is the nondimensional time — it was there all along! (since the arguments of trigonometric and exponential functions are nondimensional). The difference between Eqs. (5) and (17) is in how you look at them; do you see the dimensional time, t, as the independent variable, or do you see instead the nondimensional time, $t^* = t/\sqrt{L/g}$? The answer will probably depend on the stage of an investigation (and no doubt upon our familiarity with dimensional analysis); experimental data is almost always recorded in dimensional units, and it may be helpful to carry out a numerical integration using dimensional units. But when it comes time to report a mass of data from many experiments or integrations, there is often a great advantage to the use of nondimensional variables defined by dimensional analysis.

3.3 Extra and omitted variables

Dimensional analysis revealed that the period of a simple, inviscid pendulum did not depend upon the mass of the bob, M. This result might suggest that the inclusion of extra or superfluous variables in a physical model will not spoil the result. However, in most cases an extra variable will not be detected in this way and will lead to an extra nondimensional variable. For example, if we had included the bob diameter, D_b, in the physical model of the inviscid pendulum, it would have been carried through to a nondimensional variable, D_b/L. If we had access to an experiment, we would soon find that D_b/L was of no evident importance in determining the period of a nearly conservative pendulum, and would drop it from the final result.

It may occur to ask whether the omission of a relevant variable would be detected. The answer is yes, rarely, if the omission makes it impossible to nondimensionalize the dependent variable. For example, if we analyzed tension under the assumption that the mass would be irrelevant as it was for the period, then it would not be possible to find a nondimensional tension. That would be a clear signal that something important had been omitted from the physical model. However, if the dependent variable can be nondimensionalized with the variables that are included, then the purely formal procedure of dimensional analysis is not able to identify an incomplete model.

4 A basis set of nondimensional variables

Once a preliminary physical model has been defined, the second and mathematical step of a dimensional analysis is to find a complete set of nondimensional variables for that model. With a little experience and for small problems such as the simple, inviscid pendulum, this can be done by inspection. For larger problems it may be helpful to use the following technique[4] that relies upon the matrix methods of linear algebra. Elements of linear algebra are commonly used in dimensional analysis,[3,5] and an exhaustive exposition of matrix methods can be found in Ref. 8. Brückner and colleagues[9] show how matrix methods can be applied to very large problems. The following development differs from most others in that it does not rely on the Buckingham Pi theorem, although it comes to the same result, and utilizes the null space basis to find a basis set of nondimensional variables.[10]

4.1 The mathematical problem

What can we infer about a function given only that it is invariant to a change of units? An arbitrary change of units for the dimensional variable X_i can be written as

$$X'_i = \alpha_1^{D_{1i}} \alpha_2^{D_{2i}} \dots \alpha_M^{D_{Ji}} X_i \,, \tag{18}$$

where α_1 is the scale change associated with mass, α_2 the scale change associated with length, α_3 is for time and so on up to J fundamental units.[12] For pendulum problems and for mechanics generally, $J = 3$ (mass, length and time), which is assumed to simplify later expressions. The doubly indexed object, D_{ji}, is the dimensionality of the ith dimensional variable with respect to the jth fundamental unit, and when written as a matrix is called the dimensional matrix, $\mathbf{D}$. We have already listed the elements of $\mathbf{D}$ as part of the physical model. Assuming a physical model with I dimensional variables, invariance generally (though for $J = 3$) may be written as

$$F(X_1, X_2, \ldots, X_I) = F(\alpha_1^{D_{11}}\alpha_2^{D_{21}}\alpha_3^{D_{31}}X_1,\ \alpha_1^{D_{12}}\alpha_2^{D_{22}}\alpha_3^{D_{32}}X_2\ \ldots\ \alpha_1^{D_{1I}}\alpha_2^{D_{2I}}\alpha_3^{D_{3I}}\ X_I) \tag{19}$$

for all α (all possible changes of units). Thus for α_j, for example, we can write that

$$\frac{\partial F}{\partial \alpha_j} = \frac{\partial F}{\partial X_1}\frac{\partial X_1}{\partial \alpha_j} + \frac{\partial F}{\partial X_2}\frac{\partial X_2}{\partial \alpha_j} + \ldots\ \frac{\partial F}{\partial X_I}\frac{\partial X_I}{\partial \alpha_j} = 0\,. \tag{20}$$

If we multiply Eq. (20) by α_j/F (assuming F to be non-zero as in Eq. (10)), and use

$$D_{ji} = \frac{\alpha_j}{X_i}\frac{\partial X_i}{\partial \alpha_j}\,, \tag{21}$$

which follows from Eq. (18), we obtain $J = 3$ equations, one for each α:

$$D_{11}\frac{X_1}{F}\frac{\partial F}{\partial X_1} + D_{12}\frac{X_2}{F}\frac{\partial F}{\partial X_2} + \ldots\ D_{1I}\frac{X_I}{F}\frac{\partial F}{\partial X_I} = 0, \tag{22}$$

$$D_{21}\frac{X_1}{F}\frac{\partial F}{\partial X_1} + D_{22}\frac{X_2}{F}\frac{\partial F}{\partial X_2} + \ldots\ D_{2I}\frac{X_I}{F}\frac{\partial F}{\partial X_I} = 0, \tag{23}$$

$$D_{31}\frac{X_1}{F}\frac{\partial F}{\partial X_1} + D_{32}\frac{X_2}{F}\frac{\partial F}{\partial X_2} + \ldots\ D_{3I}\frac{X_I}{F}\frac{\partial F}{\partial X_I} = 0. \tag{24}$$

This set of equations is best written and solved in matrix form

$$D_{ji}S_i = 0\,, \tag{25}$$

where $\mathbf{D}$ is a known $J \times I$ matrix, and $\mathbf{S}$ is an unknown $I \times 1$ vector of the (logarithmic) derivatives of F with respect to the dimensional variables that we seek to find;

$$S_i = \frac{\partial \log F}{\partial \log X_i}. \tag{26}$$

We will discuss a solution method in the following, but anticipate here that there will usually be several solution vectors denoted by $\mathbf{S_k}$, with $k = 1 \ldots\ K$ (a bold subscript denotes which vector, not an element of the vector as in Eq. (26)). For example, let's say that there are $I = 4$ dimensional variables and $K = 2$ solution vectors (written in row form) that happened to be $\mathbf{S_1} = [\beta_1\ 0\ \beta_2\ 0]$ and $\mathbf{S_2} = [0\ \beta_3\ 0\ 0]$, where the β are usually small rational numbers. The first solution vector indicates that

$$\frac{X_1}{F}\frac{\partial F}{\partial X_1} = \beta_1, \quad \frac{X_2}{F}\frac{\partial F}{\partial X_2} = 0, \quad \frac{X_3}{F}\frac{\partial F}{\partial X_3} = \beta_2, \quad \frac{X_4}{F}\frac{\partial F}{\partial X_4} = 0\,. \tag{27}$$

A solution for F is thus

$$F = X_1^{\beta_1} X_3^{\beta_2}\,, \tag{28}$$

where it is useful to term the right hand side a "Pi-variable," that is,

$$\Pi_1 = X_1^{\beta_1} X_3^{\beta_2}\,, \tag{29}$$

with the subscript on $\Pi_{()}$ referring to the subscript on the solution vector $\mathbf{S}_{()}$. Any multiple of Π_1 is a solution to Eq. (27), as is any power of Π_1, as is any sum of any power; evidently any function having the argument Π_1 is a solution to Eq. (27). Another solution can be found from the second solution vector $\mathbf{S_2}$ and is some function of $\Pi_2 = X_2^{\beta_3}$. In effect, we have integrated a partial differential equation but without supplying boundary or initial data; thus we learn something about the argument of an otherwise arbitrary function. We find that the dimensional variables can appear in F only in certain combinations that correspond one-to-one with the solution vectors $\mathbf{S_k}$,

$$\Pi_k = X_1^{S_{1k}} X_2^{S_{2k}} \ldots X_I^{S_{Ik}} = \mathbf{X}^{\mathbf{S}_k}, \tag{30}$$

where $\mathbf{X} = [X_1 \; X_2 \; \ldots \; X_I]$ is a vector of the dimensional variables in the order they were entered into the dimensional matrix, $\mathbf{D}$. As anticipated in Sec. 2, these Pi-variables are nondimensional. The relationship among the Pi-variables can be written as

$$\Pi_1 = F(\Pi_2, \Pi_3 \; \ldots \; \Pi_K) \tag{31}$$

with no loss of generality. In the uncommon case that $K = 1$ and there is only one nondimensional variable, the function F must be a constant whose value cannot be determined from dimensional analysis alone. The period of the linear pendulum is an example, and in that case $F = 2\pi$ could be determined by experiment or theory (Fig. 3a). Neither can dimensional analysis determine anything further about the form of F in the much more common case that $K > 1$.

4.2 The null space

Eq. (25) is under determined in the usual case that there are more unknown exponents than there are equations. There will thus be many possible solution vectors that collectively make up the null space of the matrix $\mathbf{D}$. To represent the null space we seek a basis set from which any solution vector can be constructed. The computation of a null space basis is readily automated[4] and so we will not delve into the solution method (see Ref. 11). It is essential, however, to know the following two properties of null space bases:

P1) *The number of solution vectors, K, is the same for all basis sets and is given by the number of dimensional variables in the physical model minus the rank of the dimensional matrix, $K = I - R$.* K is also the number of nondimensional variables and in that respect all basis sets are equally efficient. One particular basis set may be more useful than others, and so it is often necessary to transform from one basis to another. A transformation is easily accomplished because of the following.

P2) *The basis set vectors are orthogonal and span the null space.* Any vector that is a solution of the homogeneous system Eq. (25) can be found as a linear combination of the vectors in any basis set. For example, if $\mathbf{S_1}$ and $\mathbf{S_2}$ are a null space basis, then their linear combination, say $\mathbf{S_3} = a_1\mathbf{S_1} + a_2\mathbf{S_2}$, with a_1 and a_2 any real number, is in the null space and is thus a solution. The corresponding nondimensional variable is $\Pi_3 = \Pi_1^{a_1}\Pi_2^{a_2}$. If $\mathbf{S_3}$,$\mathbf{S_1}$ and Π_3,Π_1 are preferred over, say, $\mathbf{S_2}$,Π_2, then a revised basis set can be taken as $\mathbf{S_1}$,Π_1, and $\mathbf{S_3}$,Π_3 while omitting $\mathbf{S_2}$,Π_2. The revised basis set has the same number of vectors as the initial basis set and it too spans the null space. An initial basis set of nondimensional variables can thus be transformed to another basis set simply by multiplying or dividing the Πs in any order (example are in Secs. 5 and 7).

4.3 A basis set for the simple, inviscid pendulum

An application to the fast time-scale oscillation of the simple, inviscid pendulum may help clarify the use of the null space basis. The dimensional matrix $\mathbf{D}$ can be read directly from the physical model:

$$\mathbf{D} = \begin{array}{c} \\ m \\ l \\ t \end{array} \begin{array}{c} \begin{array}{cccccc} \phi & t & M & L & g & \phi_0 \end{array} \\ \begin{bmatrix} 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 & -2 & 0 \end{bmatrix} \end{array}, \tag{32}$$

where the first row is the dimensionality for mass, the second row is the dimensionality of length, and the third row is for time. The dependent variable ϕ is represented by the first column, 0 0 0, all zeros because angles are nondimensional; the time t by the second column, 0 0 1; the mass M by the third column, 1 0 0, etc. The order of listing the dimensional variables is important only insofar as the algorithm seeks to make the first few dimensional variables appear in the nondimensional variables with exponents of 1. The calculation of a null space basis yields three vectors here concatenated into a matrix whose columns are the solution vectors, $\mathbf{S} = [\mathbf{S_1}; \mathbf{S_2}; \mathbf{S_3}]$,

$$\mathbf{S} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & -1/2 & 0 \\ 0 & 1/2 & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{33}$$

and the corresponding basis set of nondimensional variables has three elements:

$$\Pi_1 \; = \; \mathbf{X}^{\mathbf{S_1}} = \phi^1 \; t^0 \; M^0 \; L^0 \; g^0 \; \phi_0^0 = \phi \tag{34}$$

$$\Pi_2 = \mathbf{X}^{\mathbf{S_2}} = \phi^0\ t^1\ M^0\ L^{-1/2}\ g^{1/2}\ \phi_0^0 = t/\sqrt{L/g} \tag{35}$$

$$\Pi_3 = \mathbf{X}^{\mathbf{S_3}} = \phi^0\ t^0\ M^0\ L^0\ g^0\ \phi_0^1 = \phi_0. \tag{36}$$

The functional relationship among these may be written as $\Pi_1 = F(\Pi_2, \Pi_3)$, or

$$\phi = F(t/\sqrt{L/g}, \phi_0)\,, \tag{37}$$

and in analogy with Eqs. (6) and (7)

$$P/\sqrt{L/g} = F(\phi_0), \tag{38}$$

which is beginning to look familiar. Notice that mass M has an exponent of zero in all of the solution vectors, consistent with the informal analysis of Sec. 3 that showed that there was no way to construct a nondimensional variable from a single parameter having dimensions of mass. Also note that the angles ϕ and ϕ_0 sailed into the null space untouched because they were already nondimensional.

Tension can be analyzed in the same manner; the dimensional matrix is

$$\mathbf{D} = \begin{matrix} \\ m \\ l \\ t \end{matrix} \begin{matrix} \begin{matrix} T & t & M & L & g & \phi_0 \end{matrix} \\ \begin{bmatrix} 1 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 1 & 0 \\ -2 & 1 & 0 & 0 & -2 & 0 \end{bmatrix} \end{matrix} \tag{39}$$

and the null space basis vectors (again in matrix form) are

$$\mathbf{S} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & -1/2 & 0 \\ -1 & 1/2 & 0 \\ 0 & 0 & 1 \end{bmatrix}. \tag{40}$$

A basis set of nondimensional variables is thus

$$\Pi_1 = T/Mg \tag{41}$$

$$\Pi_2 = t/\sqrt{L/g} \tag{42}$$

$$\Pi_3 = \phi_0. \tag{43}$$

The functional relationship for the tension and the maximum tension can be written as

$$T/Mg = F(t/\sqrt{L/g}, \phi_0), \tag{44}$$

and

$$T_{\max}/Mg = F(\phi_0). \tag{45}$$

Notice that the mass, M, has been retained in the nondimensional tension. That the mass must appear is evident when one considers that the tension in the line will equal the weight of the bob, $T = Mg$, in the absence of motion (the tension of a moving pendulum will exceed this value due to centrifugal acceleration). Notice too that the length, L, has been eliminated from the maximum tension. A little thought will reveal that a length cannot be made nondimensional with T, g, and M in any combination, and thus dimensional analysis reveals that the maximum tension of a simple, inviscid pendulum started from rest must be independent of L. This conclusion was suggested by inspection of a few numerical solutions (see Fig. 1b) and now dimensional analysis assures us that it holds rigorously.

4.4 A remark on epistemology

There is no doubt that dimensional analysis has just added something significant to what we had observed in numerical integrations, i.e., that maximum tension is independent of L. It is interesting to reflect briefly upon whether dimensional analysis has provided a satisfactory *explanation* of this observation. This will be a matter of degree and opinion; my opinion is mixed. The present explanation is rigorous in that dimensional analysis has deduced a very clear statement of the observation at hand from a general principle (invariance to the choice of units) and a set of specific conditions (the physical model). Rigorous or not, this explanation from dimensional analysis seems oddly shallow and unsatisfying. There is no connection made to a larger physical principle, and not the slightest hint of limits, i.e., whether maximum tension would depend upon L if the physical model included a small viscosity, for example. In this instance and frequently, we will have to look beyond dimensional analysis when we seek explanations having sufficient depth to confer a useful understanding.[13]

5 The viscous pendulum

Now consider the decay rate defined by

$$\Gamma = \frac{1}{\Phi}\frac{d\Phi}{dt}, \tag{46}$$

where Φ is the amplitude of the motion. We begin with observations of the amplitude Φ made by measuring visually the cord length at intervals of 30 sec to 2 min (the crosses of Fig. 4a). For this purpose it was advantageous to use a longer pendulum, $L = 3.70\,\mathrm{m}$, to minimize the noise associated with the rather coarse least count on the measured cord, about 10^{-3} m. This pendulum was supported on a needle bearing (a fishhook on a hard metal surface) to minimize interactions with the pivot, and the line was smooth monofilament having a diameter $D_l = 0.40 \times 10^{-3}$ m. The bob was a nearly spherical, more or less smooth lead fishing sinker with a diameter $D_b = 0.0211\,\mathrm{m}$ and mass $M = 0.055\,\mathrm{kg}$. The observed amplitude history, $\Phi(t)$, was quite repeatable and can be roughly characterized as an exponential decay with a time scale of about 10 min.

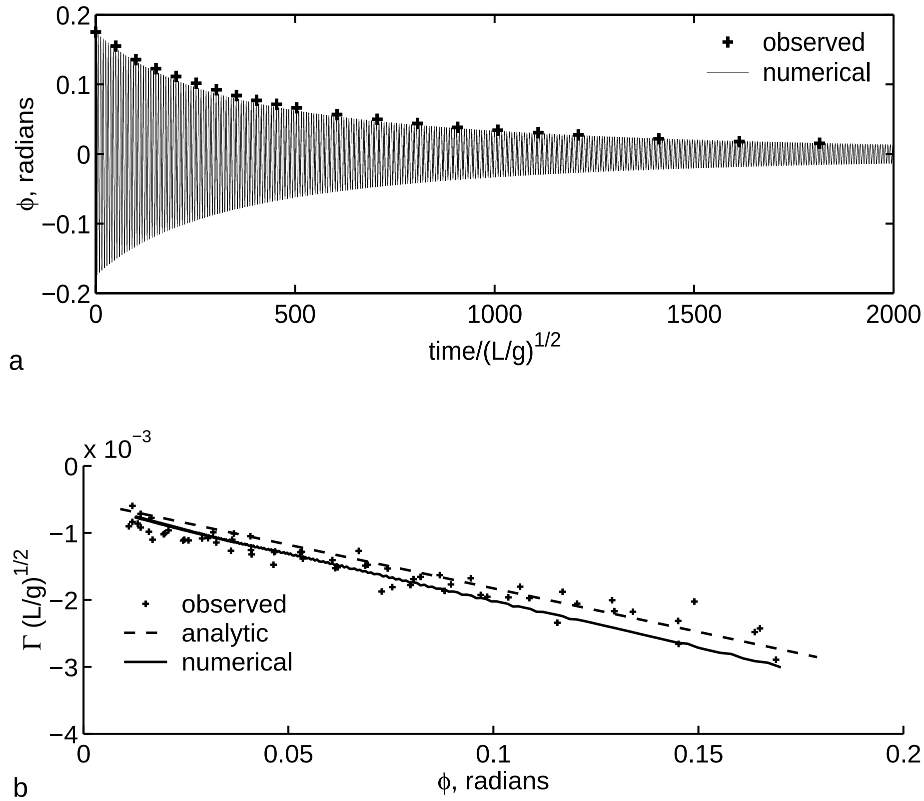

Figure 4: (a) Observations (crosses) and a numerical solution (the thin solid line) of the motion of a simple, viscous pendulum. The crosses are observations[18] of the amplitude at intervals of 30 seconds to 2 minutes. (b) The decay rate computed directly from the observations (crosses, from three repetitions of the experiment), as estimated from an approximate analytic solution Eq. (61) (dashed line) and as diagnosed from the numerical model solution (solid line). Drag that is linear in the angular velocity produces a constant decay rate (simple exponential decay in time), and drag that is quadratic in the angular velocity produces a decay rate that increases linearly with the amplitude, ϕ.

5.1 A physical model of the viscous pendulum

We presume that hydrodynamic drag with the surrounding air is the primary damping process,[14] and that the diameter of the bob, D_b, and of the line, D_l, are now relevant, as are the density and kinematic viscosity of air, ρ and ν. When we amend the inviscid model of Sec. 2 to include these variables we have

- A physical model for the decay rate of a simple, viscous pendulum:

1. the decay rate, $\Gamma \doteq m^0 l^0 t^{-1}$, the dependent variable;
2. mass of the bob, $M \doteq m^1 l^0 t^0$, a parameter,
3. length of the line, $L \doteq m^0 l^1 t^0$, a parameter,
4. acceleration of gravity, $g \doteq m^0 l^1 t^{-2}$, a parameter,
5. the amplitude of the motion, $\Phi \doteq$ nond, a parameter,
6. diameter of the line, $D_l \doteq m^0 l^1 t^0$, a parameter,
7. diameter of the sphere, $D_b \doteq m^0 l^1 t^0$, a parameter,
8. density of air, $\rho \doteq m^1 l^{-3} t^0$, a parameter (1.2 kg m^{-3}, nominal),
9. kinematic viscosity of air, $\nu \doteq m^0 l^2 t^{-1}$, a parameter ($1.5 \times 10^{-5}$ m^2 s^{-1}, nominal).

(For the purpose of defining the amplitude of the motion we might have used ϕ_o in place of Φ.) Dimensional analysis, from here on omitting all of the intermediate steps, indicates six nondimensional variables;

$$\Gamma\sqrt{L/g} = F(\Phi, D_b/L, D_l/L, \rho D_b^3/M, g^{1/2} L^{3/2}/\nu). \tag{47}$$

The first five nondimensional variables have an obvious interpretation, but the last one involving the viscosity, ν, does not. In any event, we are not ready to make use of such a comprehensive model. We may still be thinking of the nearly conservative pendulum of Sec. 2, but the nine-variable physical model includes all possible pendulums and fluid mediums. Before we can expect a useful result from dimensional analysis we will have to identify the most relevant parameters for the kind of nearly conservative pendulum that we have in mind.

5.2 Drag on a moving sphere

A piecewise approach is tried next. Consider in isolation the hydrodynamic drag on a smooth sphere (the bob) due to a steady motion through an infinite viscous fluid (air) that is otherwise at rest.

- A physical model for drag on a sphere moving through viscous fluid:

1. drag (a force), $H \doteq m^1 l^1 t^{-2}$, the dependent variable,
2. speed of the sphere, $U \doteq m^0 l^1 t^{-1}$, a parameter,
3. diameter of the sphere, $D_b \doteq m^0 l^1 t^0$, a parameter,
4. density of air, $\rho \doteq m^1 l^{-3} t^0$, a parameter,
5. kinematic viscosity of air, $\nu \doteq m^0 l^2 t^{-1}$, a parameter.

Despite the highly idealized configuration of this problem, it is very difficult to compute the drag in the common case that the flow around the sphere is turbulent. However, dimensional analysis combined with laboratory measurement leads to a useful result. The initial basis set of nondimensional variables for this physical model comes out to be

$$\Pi_1 = \frac{H}{\rho D_b^2 U^2} \quad \text{and} \quad \Pi_2 = \frac{\nu}{U D_b}, \tag{48}$$

where we recognize that Π_2 is the inverse of an important nondimensional variable called the Reynolds number,

$$\mathrm{Re} = \frac{U D_b}{\nu}, \tag{49}$$

which we prefer. We know from P2 of the null space (Sec. 3) that Π_1 in Eq. (48) is not uniquely determined by dimensional analysis, and that a general basis set can be written as

$${}_n\Pi_1 = \frac{H}{\rho D_b^2 U^2}(\frac{U D_b}{\nu})^n \quad \text{and} \quad \Pi_2 = \frac{U D_b}{\nu}, \tag{50}$$

where n is any real constant. In this we are assuming that H and $\Pi_2 = \mathrm{Re}$ should both remain to the first power. The functional relation between these nondimensional variables could be written as ${}_n\Pi_1 = F(Re)$, where F depends on n. We will consider next how to choose the value of n that gives the best or most useful form.[15] Regardless of the form finally chosen, an essential result is that the nondimensional drag, ${}_n\Pi_1$, is expected to be a function of the Reynolds number alone (more on Re in the following). Laboratory measurements can thus be used to define $F(\mathrm{Re})$ that should hold for all steadily moving spheres, Fig. 5a, just the way that the function $F(\phi_0)$ (Sec. 3.2) sufficed to define the period for all inviscid, simple pendulums. More recent text books[3], like this article, show only the curve that runs through the middle of a tight cloud of data points that have accumulated from many laboratory experiments, see for example, Rouse[3]. What is most important, but not evident from this kind of presentation, is that drag coefficients inferred from experiments made using a very wide range of spheres and cylinders moving at widely differing speeds and through many different viscous fluids (Newtonian fluids) do indeed collapse to a well-defined function of Reynolds number alone, just as dimensional analysis had indicated.

This is a result, characteristic of dimensional analysis generally, that is at once profound and trivial. One might say trivial because, after all, dimensional analysis told us that the drag coefficient must depend upon Re alone. From this perspective, an effective collapse of the data verifies that carefully controlled laboratory conditions can indeed approximate the idealized physical model. It is profound in that dimensional analysis has shown the way to a useful result (Fig. 5), where there would otherwise have been be an unwieldy mass of highly specific data (as in going from Fig. 1 to Fig. 2).

5.2.1 Zero order solution

The crucial (and in this case the only) choice is that of the dependent nondimensional variable, Π_1. One strategy is to form Π_1 so that it reflects a physically meaningful "zero order" solution for the dependent variable. This strategy is equivalent to scaling a dimensional variable in a model equation so that the corresponding nondimensional variable has a maximum size of about one (see Ref. 3, Lin and Segel, Chapter 6).

A zero order solution requires some sense of the physics of the problem. Visual observations of the flow around a sphere provide hints that drag can arise from two distinct processes. If the sphere is moving very slowly so that the wake behind the sphere is nearly undisturbed, then the drag will be mainly viscous, that is, due to the shear of the flow around the sphere and directly proportional to the viscosity of the fluid. The shear can be estimated by U/D_b, and the viscous stress by $\rho\nu U/D_b$. If this viscous stress acts over an area proportional to D_b^2, then the zero order solution for viscous drag on the sphere would be $H \propto \rho\nu D_b U$. This is the basis set $n = 1$ of Eq. (50). If we expected that this was the dominant drag-producing process, then it would be appropriate to nondimensionalize the drag as

$$\frac{H}{\rho\nu D_b U} = F(\mathrm{Re}) = C_v(\mathrm{Re})\,, \tag{51}$$

because the Re-dependence of C_v, the so-called viscous drag coefficient, would then be minimized.

Even if the fluid were nearly inviscid, there would still be drag because fluid must be accelerated as it is displaced by the moving sphere. If the displaced fluid is carried along in a highly disturbed wake, as is more or less observed behind a rapidly moving sphere (we will clarify what is meant by rapidly), then the drag would be roughly proportional to the density of the fluid times the speed squared (a momentum flux) multiplied by the frontal area, $A = \pi D_b^2/4$. Thus the inertial drag would be estimated as $H \propto \rho A U^2$. If we expected that this inertial drag process was dominant, then the initial basis set corresponding to $n = 0$ would be appropriate:

$$\frac{H}{\rho A U^2} = F(\mathrm{Re}) = C_i(\mathrm{Re})\,, \tag{52}$$

and the Re-dependence of the inertial drag coefficient C_i would show the departures from inertial drag due to viscous effects. Either form of the drag coefficient effectively conveys the laboratory data and in that regard there is nothing to choose between them.

5.2.2 The other nondimensional variables: remarks on the Reynolds number

Once the dependent nondimensional variable, Π_1, has been selected, the remaining nondimensional variables can be formed in ways that most clearly define the geometry of the problem, that reflect a balance of terms in a governing equation, or that follow conventions in the field. This is necessarily vague because the possibilities are limitless, however the task is often easier than might be expected. In the example of drag on a moving sphere, there is only one remaining nondimensional variable, the Reynolds number or its inverse. There are many other such ratios, often termed nondimensional "numbers," that succinctly characterize the balances among terms in mathematical models and thus are the natural terminology of theoretical mechanics. Like any language, these nondimensional numbers are conventions that have to be learned. Here a few that are likely to be encountered in geophysical fluid mechanics:

- Reynolds number, $Re = \frac{UL}{\nu}$, where U is the speed of a current, L is the spatial scale over which U changes by roughly 100%, and ν is the kinematic viscosity of the fluid. The Reynolds number is the ratio of advective to viscous terms in the Navier-Stokes momentum balance and arises very frequently in fluid mechanics (more on the Reynolds number below). There are many different definitions of the Reynolds numbers, differing most often by the choice of the length scale. For drag on spheres it is conventional to use the diameter, for example.

- Rossby number, $R_o = \frac{U}{Lf}$, where U and L are as above and f is the Coriolis parameter, proportional to the earth's rotation rate (units are t^{-1}). The Rossby number is the ratio of advective to Coriolis terms in a momentum equation for fluid observed in a rotating reference frame (the earth).

- Strouhal number, $S = \frac{\omega L}{U}$, where ω is the frequency or the inverse of the time scale, $\frac{\partial}{\partial t}$, of the current, U. The Strouhal number is the ratio of the local rate of change of U to the advective rate of change. The inverse of this ratio is often called the temporal Rossby number in geophysical fluid dynamics.

- Froude number, $U/\sqrt{gH}$, where H is the thickness of a fluid layer, or some other length scale. The Froude number (or its square) is the ratio of advection of momentum to the pressure gradient and arises in problems in which the acceleration of gravity is important, for example, in the wave drag on ships.

- Ekman number, $E = \frac{K}{f}$, where K is the drag coefficient that appears in a linear (Rayleigh) drag law. The Ekman number is a measure of the ratio of frictional to Coriolis terms in a momentum equation. There are as many definitions of the Ekman number as there are parameterizations of frictional terms.

Recall that for the purpose of modeling drag, a slowly moving sphere is one that has a nearly undisturbed wake. Observational evidence shows that this kind of flow occurs when Re ≤ 1, regardless of speed *per se*; dimensional analysis tells us as much in that the drag coefficient depends only upon Re. The small Re range is that of a very small bug swimming through water, for example. Notice that in this very small Re range the viscous drag coefficient C_v is O(1) and roughly constant (Fig. 5a). For creatures and objects anywhere near our size, Reynolds numbers of $O(10^5)$ and greater are the norm, and inertial drag (often termed form drag) is generally more important for runners and bicyclers than is viscous drag. Notice that for moderately large values of Re, $10^3 \leq \mathrm{Re} \leq 10^5$, the inertial drag coefficient C_i is O(1) and very roughly constant within subranges.[17] We can anticipate the our pendulum will have Reynolds numbers in an intermediate range in which both viscous and inertial drag are liable to be important.

5.3 A numerical simulation

To model the decay process we will include hydrodynamic drag on the line and bob in the angular momentum balance (1). Drag will be estimated by means of the steady drag laws discussed above, and so it is implicitly assumed that the instantaneous speed of the bob or line gives the same drag as would a steady motion of the same speed. Whether this is appropriate remains to be seen.

The main task is to account for the Re-dependence of the drag coefficients Because the line is quite thin, the Reynolds numbers of the line are rather small, $\mathrm{Re}_l = UD_l/\nu \leq 20$, where $U = rd\phi/dt$, r is the distance from the pivot and an *a priori* estimate of $d\phi/dt$ is $\phi_o/\sqrt{L/g}$. In that small Re range the viscous drag coefficient on a cylinder can be approximated well by $C_v = 3/2 + \mathrm{Re}/3$ (the heavy dotted line of Fig. 5b). The drag per unit length of the line, $\delta = dr$, can then be computed by the drag law corresponding to Eq. (51) as $H = \pi\rho\nu C_v U dr$, and the (dimensional) torque due to drag over the length of the line is then

$$\tau_l = \int_0^L rH dr = \rho\left(\frac{\pi}{2}\nu L^3 + \frac{1}{12}D_l L^4 \mid \frac{d\phi}{dt} \mid\right)\frac{d\phi}{dt}. \tag{53}$$

The absolute value operator insures that the drag force always opposes the motion. The bob has a much larger diameter and thus a much larger Reynolds number; $\mathrm{Re}_b = L\frac{d\phi}{dt}D_b/\nu$ is in the range $\mathrm{Re}_b \leq 1000$ where no simple formula for a drag coefficient is highly accurate. Thus we will allow an arbitrary $C_i(\mathrm{Re}_b)$ and compute the drag-induced torque on the bob as

$$\tau_b = \frac{\pi\rho}{8}C_i(Re_b)D_b^2 L^3 \mid \frac{d\phi}{dt} \mid \frac{d\phi}{dt} \tag{54}$$

where Re_b and C_i are evaluated at each time step of the numerical integration using the data of Fig. 5a. The amended angular momentum balance (in dimensional variables),

$$\frac{d^2\phi}{dt^2} = -\frac{g}{L}\sin(\phi) - \frac{\tau_l + \tau_b}{L^2 M}, \tag{55}$$

together with Eqs. (53) and (54) and the data of Fig 4.1 plus the initial condition Eq. (3) make a complete if rather cumbersome model that can be integrated numerically.

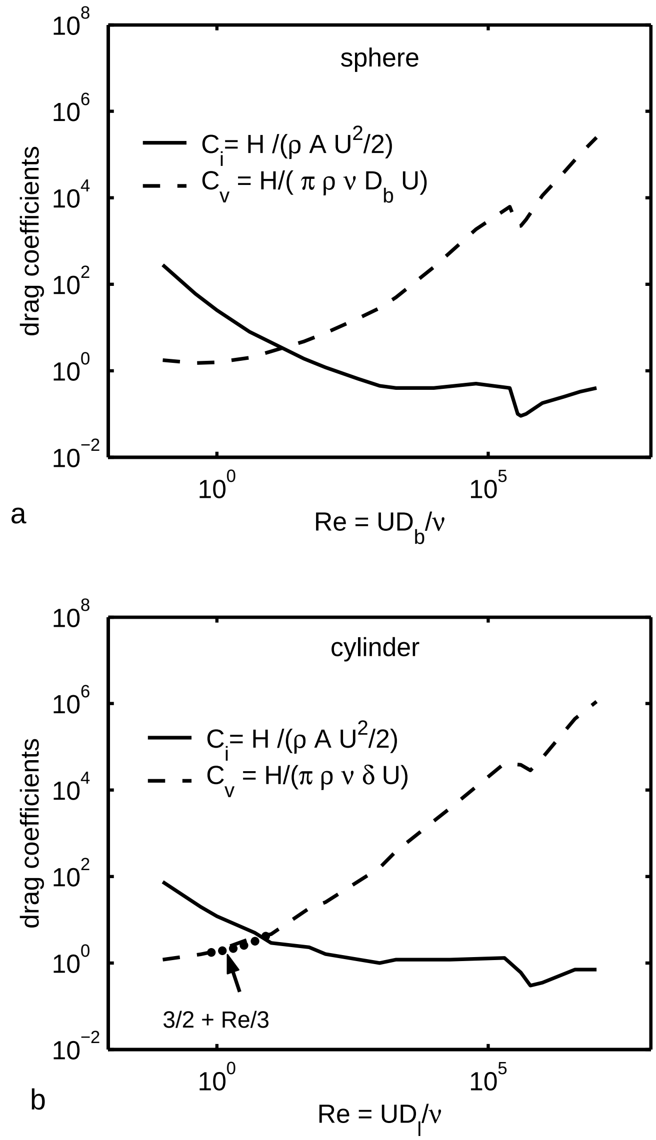

Figure 5: Drag coefficients of a sphere (a) and a cylinder (b) moving at a steady speed U through viscous fluid. Two forms of drag coefficient are shown here, the viscous drag coefficient is denoted by C_v (the dashed line), and the inertial drag coefficient denoted by C_i (the solid line, usually denoted C_d, and by far the most commonly encountered form). Note that C_v is O(1) if Re is very small, and that C_i is O(1) if the Reynolds number is very large. The inertial drag coefficients were read from Munson et al., Fig. 7.7 and Rouse, Figs. 125-126.[3,16]

With drag terms included, the period of the oscillation is nearly unchanged, but of course the amplitude slowly decays (Fig. 4a). The decay simulated by numerical solution looks plausible when compared with the observations, suggesting that the steady drag laws have the gist of it (a more critical appraisal is given below).

5.4 An approximate model of the decay rate

Numerical solutions are not revealing of parameter dependence, but given two modest approximations we can go on to deduce a model of the viscous pendulum that has transparent solutions. First, the angle ϕ is small enough in the case shown in Fig. 4a that $\sin\phi$ of Eq. (55) can be approximated well as ϕ. Second, the drag overall is due mostly, $\approx 85\%$, to the line, and so it should be acceptable to make the strong approximation that the inertial drag coefficient for the bob is a constant, $C_i = 0.7$, an average for the Re_b range of the bob in the present case. With these approximations we obtain a solvable model for the simple, viscous pendulum (now in nondimensional variables)

$$\frac{d^2\phi}{dt^{*2}} = -\phi - a\frac{d\phi}{dt^*} - b \left| \frac{d\phi}{dt^*} \right| \frac{d\phi}{dt^*}, \tag{56}$$

where the coefficient in the linear drag term is

$$a = \frac{\pi}{2}\frac{\rho\nu L^{3/2}}{Mg^{1/2}} \tag{57}$$

and the coefficient in the quadratic term is

$$b = \frac{\rho}{8M}(0.7D_b^2 L + \frac{2\pi}{3}D_l L^2). \tag{58}$$

Approximate solutions for small damping are given in Ref. 14; linear drag causes the amplitude to decay at a rate (nondimensional)

$$\frac{1}{\Phi}\frac{d\Phi}{dt^*} = -\frac{a}{2} \tag{59}$$

and the quadratic term causes decay at a rate

$$\frac{1}{\Phi}\frac{d\Phi}{dt^*} = -\frac{8b}{6\pi}\Phi\,, \tag{60}$$

where again Φ is the slowly varying amplitude. For small damping, these can be added together and evaluated to give an approximate decay rate,

$$\Gamma\sqrt{L/g} = \frac{1}{\Phi}\frac{d\Phi}{dt^*} \approx -5.2 \times 10^{-4} - 1.6 \times 10^{-2}\Phi \tag{61}$$

shown as the dashed line of Fig. 4b. This approximate model shows very clearly how the decay rate is expected to vary with the parameters that characterize the pendulum and the fluid medium (and in fact it gives an excellent account even for quite strong damping). All of the pieces of this model were present in our first attempt at dimensional analysis of the viscous pendulum, Eq. (47), though we had no way to recognize them at the time.

The decay rate can be estimated from the observations and from the numerical solution by a direct (no smoothing required) first differencing (Fig. 4b). A comparison of decay rates makes a much more sensitive test of the drag formulation than does the amplitude itself (cf. Fig. 4a) and reveals that the decay is not simple exponential as it first appears. In fact, there is a significant dependence of the decay rate upon amplitude, which in the approximate model follows from the quadratic drag term, Eq. (58). This could be interpreted to show that hydrodynamic drag on the pendulum is due mostly to inertial drag rather than the purely viscous drag of the very small Re range (not unexpected).

While the modeled decay rate is fairly accurate, there is at least a hint that the appropriate drag law for this pendulum overall (that is, the entire system, including the supporting structure) has a somewhat greater linear drag than is found in the models, and slightly less quadratic drag. This behavior is found consistently over a range of conditions, but further study of drag phenomena is outside the scope of this paper.[18]

6 A similarity solution for diffusion in one dimension

Sometimes the use of dimensional analysis can lead to solutions that might otherwise have been missed. A good example is afforded by the study of Stokes First Problem, in which a fluid column is driven from rest by an imposed surface speed, V_o (standard text book fare done well by e.g., Kundu and Cohen[3]). A key physical assumption is that the momentum supplied at the upper boundary is assumed to diffuse downward into the fluid at a rate set by a kinematic viscosity, ν, that is presumed to be a given constant (and specifically not dependent upon the flow itself). In that case the governing equation for the current, U, is the elementary one-dimensional diffusion equation,

$$\frac{\partial U}{\partial t} = \nu \frac{\partial^2 U}{\partial z^2}. \tag{62}$$

The initial condition is presumed to be a state of rest,

$$U(z, t = 0) = 0 \tag{63}$$

and the boundary conditions are that the fluid sticks to an upper surface that is moving at speed Vo, and to a lower surface at $z = -L$ that is at rest,

$$U(z = 0, t \geq 0) = Vo, \quad \text{and} \quad U(z = -L, t) = 0. \tag{64}$$

It is easy to generate solutions to this linear problem; Fourier transform leads to an infinite trigonometric series that can be summed to very high accuracy, and numerical solution is almost deceptively easy. What role can dimensional anlysis have in this case? There is yet another avenue, known as a similarity solution, that may yield more insight than we are likely to derive from an infinite series or from numerical data. To arrive at this solution we will begin with a dimensional analysis of the elementary diffusion model. As usual, we will start by making a straightforward list of the important variables.

- A physical model of one-dimensional, elementary diffusion:

1. current speed, $U \doteq l^1\ t^{-1}$, the dependent variable
2. time, $t \doteq t^1$, an independent variable
3. depth, $z \doteq l^1$, a second independent variable
4. surface boundary value, $Vo \doteq l^1\ t^{-1}$, a parameter
5. the kinematic viscosity, $\nu \doteq l^2\ t^{-1}$ a parameter
6. the depth of the fluid column, $L \doteq l^1$, a parameter.

Given this physical model, the nondimensional functional relationship might then read

$$U/Vo = F(zVo/\nu,\ tVo^2/\nu,\ z/L). \tag{65}$$

6.1 Honing the physical model

Solutions to Eqs. (62 - 64) show that the current at a given depth and time is directly proportional to the boundary value, Vo, as might have been inferred also from inspection of the mathematical model. This important property has not been built into the physical model nor is it reflected in the initial basis set of nondimensional variables Eq (65). Instead, this physical model covers a much more general problem in which Vo appears in the nondimensional variables combined with z or t as if Vo effected the diffusion process. On physical grounds this can be expected to happen when the diffusion process results from turbulence (rather than molecular diffusion) generated by the boundary forcing. Whether the flow is turbulent or laminar depends upon the distance from the boundary, the current speed at that point, and the fluid viscosity, i.e., a Reynolds number! If we insist that the diffusion process be represented by a constant viscosity (this is what is meant by the elementary diffusion model) then we are implicitly limiting the analysis to small Reynolds number flows, or to something like heat diffusion in a solid. To assert this important physical property in the physical model we will make a small but highly significant change — we will replace the dependent variable by U/Vo and then remove Vo from the list of parameters. The basis set for this revised physical model is then

$$U/Vo = F(z/\sqrt{t\nu}, z/L), \tag{66}$$

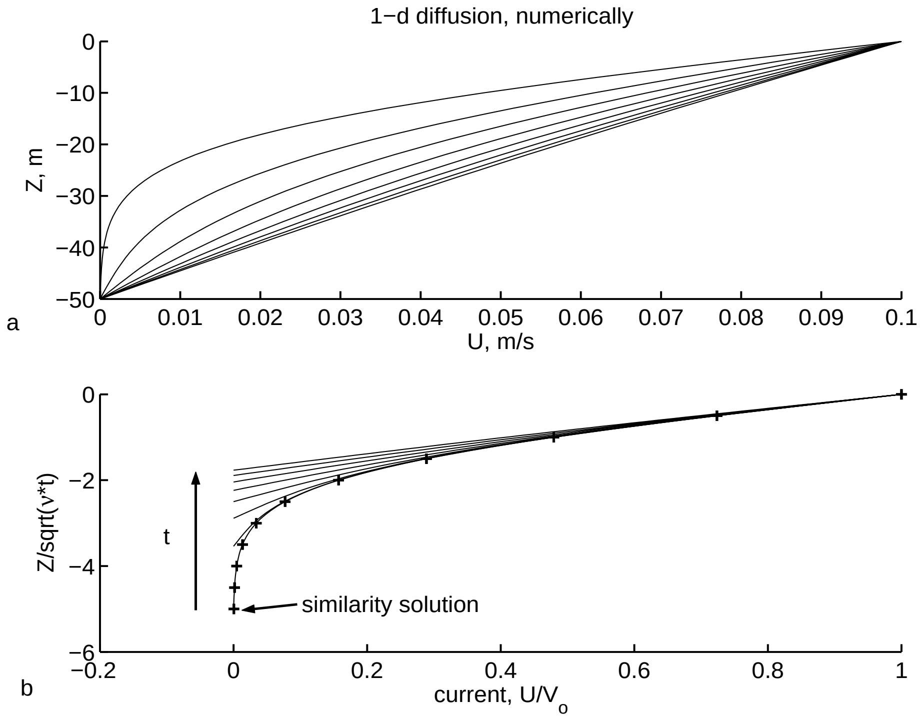

Figure 6: (a) Solutions of $U(z,t)$ from an elementary diffusion (numerical) model plotted at intervals of 10^4 s. The dimensional values are roughly those of an upper ocean ($\nu = 10^{-2}\mathrm{m}^2\mathrm{s}^{-1}$). (b) The numerical solutions nondimensionalized and plotted along with the similarity solution (the small crosses). At small time, the numerical solutions lie exactly on top of the similarity solution. At longer times, the effect of the lower no-slip boundary condition becomes appreciable and the numerical solutions depart more and more from the similarity form.

which notice has one fewer variables than before. In the event that L is very large compared to the depth that diffusion has reached (we will elaborate on this point below), then the nondimensional current depends upon only the single independent variable

$$U/Vo = F(\eta), \tag{67}$$

where

$$\eta = z/\sqrt{t\nu}, \tag{68}$$

and not upon z and t separately as the first set of nondimensional variables indicated. The current profiles at various times thus have a similar shape, being more less stretched out depending upon $\sqrt{t}$, Fig. 6. The variable η is said to be a 'similarity' variable and the function $F(\eta)$ a similarity function, as noted in connection with the drag coefficients of Sec. 5.

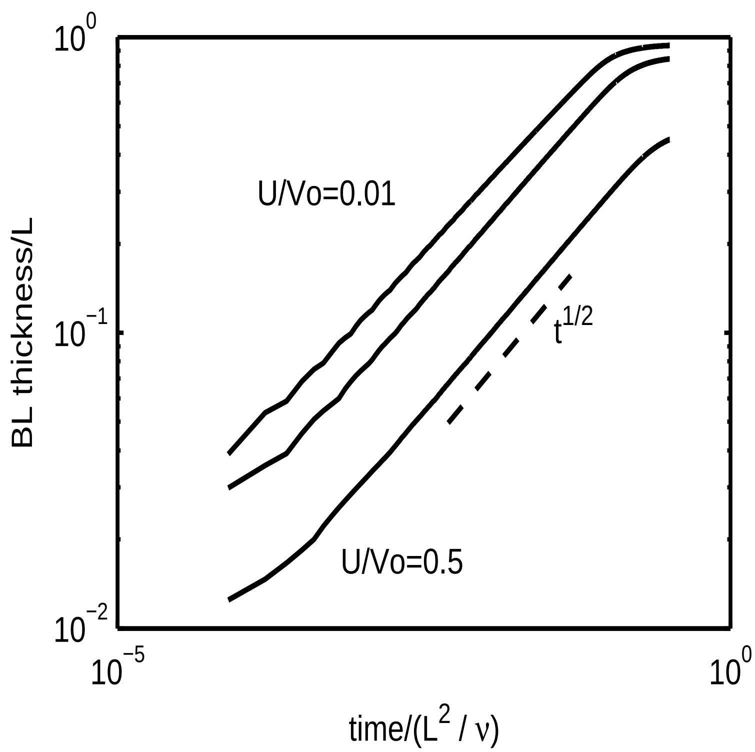

Figure 7: The boundary layer thickness (the depth at which the current is 0.5, 0.1 or 0.01 of the surface value) from the numerical solution of Fig. 6. The time step and grid interval are shown as dt and dz. Note that there is an intermediate nondimensional time from about 10^{-3} to 10^{-1} during which the boundary layer thickness grows like time$^{1/2}$.

6.2 A similarity solution

The analysis above suggests that the governing equation (a partial differential equation in z and t) might be transformable into an ordinary differential equation in the single independent variable η. To see if this is true we will substitute (67) into the governing equation; the partial time derivative becomes

$$\frac{\partial U}{\partial t} = Vo\frac{\partial F}{\partial \eta}\frac{\partial \eta}{\partial t} = -VoF'\frac{\eta}{2t}, \tag{69}$$

where $F' = \frac{dF}{d\eta}$, and the second derivative with respect to z becomes

$$\frac{\partial^2 U}{\partial z^2} = VoF''\frac{1}{t\nu}. \tag{70}$$

Substitution into the governing equation and noting that z and t appear only in the combination (68) shows that the second order partial differential equation is indeed transformed into the second order oridnary differential equation

$$F'' + \frac{1}{2}\eta F' = 0. \tag{71}$$

Substitution into the upper and lower boundary conditions gives

$$F(\eta = 0) = 1, \quad \text{and} \quad F(\eta = -\infty) = 0. \tag{72}$$

So far as the current is concerned, small time and large z are now equivalent, and the initial condition is identical to the lower boundary condition. These two boundary conditions are then sufficient to define the solution,

$$U/Vo = 1 - erf(\eta/2). \tag{73}$$

The error function

$$erf(x) = \frac{2}{\sqrt{\pi}}\int_0^x exp(-y^2)dy \tag{74}$$

is tabulated and this similarity solution can be considered exact and closed. This similarity solution is valuable in at least two ways. Because it is exact, it can serve as a precise test of numerical or numerically evaluated solutions whose accuracy might be hard to evaluate *a priori*. Secondly, because of its simplicity we can see qualitative features in a similarity solution that might have been missed in a mass of numerical data. In particular, note that a given value of η, say $\eta = C$, and thus a given value of $U/Vo = F(C)$, moves downward over time as

$$z = C\sqrt{\nu t}. \tag{75}$$

Thus the thickness of the layer directly effected by the boundary condition, i.e., the boundary layer, grows like the square root of time Fig. 7). This is characteristic of

1-dimensional elementary diffusion and random processes alike. This result can be turned around; the time needed to diffuse a distance L is roughly L^2/ν, which forms a natural time scale for this one-dimensional diffusion problem.[19]

This elegant and precise solution has built-in limitations, of course. In practice this similarity solution will hold only for an intermediate time; the time has to be long enough that the growing boundary layer does not retain the detailed imprint of the startup, which could never be a pure step function as is assumed, or, in a numerical model, the boundary layer must be thick enough that the structure is not highly dependent upon the vertical resolution, Fig. 6). The time has to be short enough that diffusion has not caused an appreciable current near the lower, no-slip boundary at $z = -L$. Once the current becomes appreciable near the lower boundary, we can expect that similarity will no longer hold accurately, and the more general form Eq. (66) will be relevant from then on.[20]

7 Scaling analysis

Dimensional analysis can be taken as the staring point of another general and important procedure called scaling analysis, that we consider here, briefly. The specific goal of a scaling analysis is often to identify small terms in a model equation so that an approximate solution can be sought. A scaling analysis can be done without going through the three step procedure of a dimensional analysis touted here. However, one way to think about the method of a scaling analysis is that it seeks a nondimensional basis set that reflects a meaningful 'zero order' solution of just the sort that we have already had occasion to discuss in Sec. 5.2. A dimensional analysis can be performed with no thought given to scaling *per se*, and we have done that, too. However, the results of a dimensional analysis will often be much more useful if aspects of a scaling analysis are considered either from the outset, or at the third (interpretive) part of a dimensional analysis, the main point that we hope to make here.

7.1 A nonlinear projectile problem

To illustrate the purpose and the elementary methods of a scaling analysis we follow the nonlinear projectile problem that was developed in superb detail by Lin and Segel[3]. The problem is to calculate the motion of a projectile that is launched upwards with speed V_o from the surface of a planet having a radius R and mass M. The only external force on the projectile is presumed to be the acceleration of gravity. To make the problem interesting, the variation of the gravitational acceleration with height above the surface, z, is considered. The equation of motion for the projectile would then be

$$m\frac{d^2z}{dt^2} = -mMG/(R+z)^2, \qquad (76)$$

where G is the universal gravitational constant. Suitable initial conditions are

$$z(t = 0) = 0 \quad \text{and} \quad \frac{dz}{dt}(t = 0) = V_o. \tag{77}$$

The acceleration of gravity on the planet's surface can be defined as $g = MG/R^2$, a parameter, and the equation of motion rewritten

$$\frac{d^2 z}{dt^2} = -gR^2/(R + z)^2. \tag{78}$$

As the height of the projectile becomes comparable to R, the local gravitational acceleration thus decreases. For a sufficiently large z the projectile could escape the gravitational tug of the planet altogether and continue into deep space. Clearly this is going to have something to do with the parameters g, R and V_o. To see just how, we will analyze a nondimensional form of the model equations (77) and (78), beginning with

- A physical model of projectile motion in variable gravity:

1. height above the planet surface, $z \doteq l^1$, the dependent variable,
2. time, $t \doteq t^1$, an independent variable,
3. the acceleration of gravity on the planet surface, $g \doteq l^1 \ t^{-2}$, a parameter,
4. radius of the planet, $R \doteq l^1$, a parameter,
5. initial (vertical) speed, $V_o \doteq l^1 \ t^{-1}$ a parameter.

For this arbitrary ordering of the physical model, the initial basis set of nondimensional variables comes out to be

$$\Pi_1 = \frac{z}{R}, \quad \Pi_2 = \frac{t}{V_o/R} \quad \text{and} \quad \Pi_3 = \frac{V_o^2}{gR}, \tag{79}$$

and the relation between these three nondimensional variables could be written

$$\frac{z}{R} = F(\frac{t}{V_o/R}, \frac{V_o^2}{gR}). \tag{80}$$

Since there are three nondimensional variables in the initial basis set, there is a somewhat greater range of possible basis sets for this problem than we have encountered before. There are some constraints; for example, we may as well leave the dependent and independent variables z and t to the first power (as they already are). Other possible basis sets can be then be formed only if we choose to multiply the initial Π_1 and Π_2 and by some n-th power of Π_3, or,

$$\Pi_1 \Pi_3^n, \quad \Pi_2 \Pi_3^n \quad \text{and} \quad \Pi_3. \tag{81}$$

As we will see shortly, the basis set $n = -1$ is especially useful.

The maximum height that the projectile will reach, say Z, is of special interest, and from the above we can see that the nondimensional form corresponding to this initial basis set will be

$$Z/R = F(\epsilon), \tag{82}$$

where $\epsilon = \frac{V_o^2}{gR}$ is the only nondimensional parameter in the problem, i.e., g, R and V_o will appear only in this combination. Thus, in the usual, partial way of dimensional analysis, we have already learned something useful about this problem. The parameter ϵ is loosely a measure of the magnitude of the initial velocity compared to gravity; more precisely it is (proportional to) the ratio of the kinetic energy of the projectile to the depth of the potential energy well of the planet. For larger ϵ we would expect a larger maximum height, and at some threshold value the projectile will escape altogether (V_o is then said to be the escape velocity, approx. 12 km s^{-1} for Earth, ignoring air drag).

Numerical solutions of Eqs. (78 and 77) nondimensionalized by this initial basis set look entirely reasonable (Figs. 8 and 9). As we should have by now expected, Z/R is a well-defined function of the nondimensional parameter ϵ. This is a mathematical certainty since these 'data' are solutions of a numerical model that is exactly consistent with the physical model. (This can be contrasted with the drag coefficient data of Fig. 5 which came from observations of physical systems. The collapse of those data to a single function of the Reynolds number is a result of considerable *physical* significance.) This initial basis set thus serves one of the main goals of dimensional analysis - to make a compact and useful presentation of what would otherwise be an unwieldy mass of data (imagine plotting the maximum height as a function of the three relevant dimensional variables).

7.2 Small parameter → small term?

An aspect of nondimensionalization that goes beyond our previous discussions is the question of how or whether a nondimensional model equation can be used to develop an approximate solution. The basis of approximation considered here is that one or more of the most difficult terms of an equation might be dropped to yield a solvable problem. Once an approximate solution is at hand, then the equation can be iterated to arrive at successively better solutions that take account of the term dropped on the first pass (an excellent discussion of iteration procedures is Ch. 7 of Lin and Segel[3]).

Our initial basis set amounts to nondimensionalizing the projectile height by the radius of the planet, R, and nondimensionalizing time by the time interval required to move the distance R at a rate V_o. For the purpose of writing the model equation in nondimensional form let $z' = z/R$, and $t' = \frac{t}{R/V_o}$, and using the change of variable procedure, one quickly

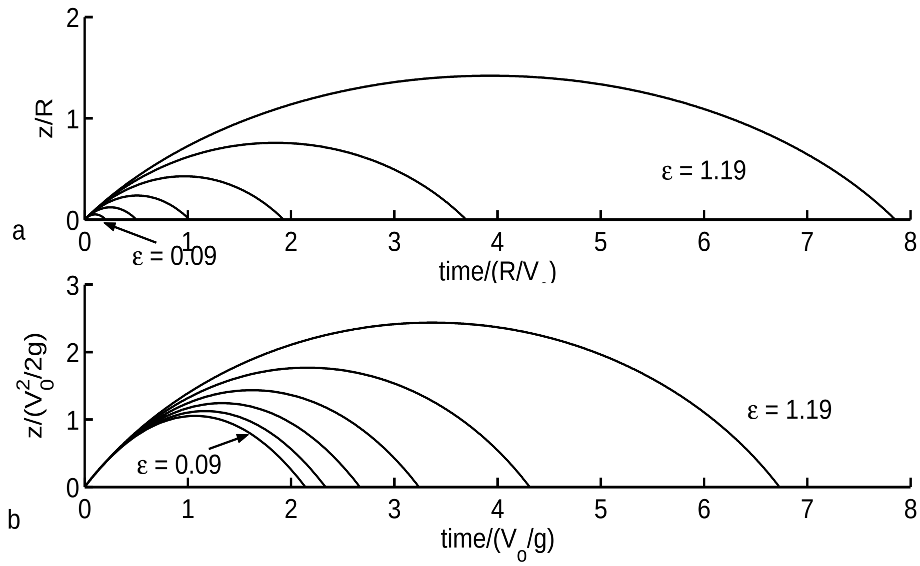

Figure 8: Projectile height computed numerically for several ϵ varying from 0.09 to 1.19. The solutions have been nondimensionalized in two ways; in (a) using the initial basis set, and in (b) using a second basis set built around a zero order solution. Notice that for small values of ϵ the set of curves in (b) appear to collapse toward one curve, while the set of curves in (a) do not.

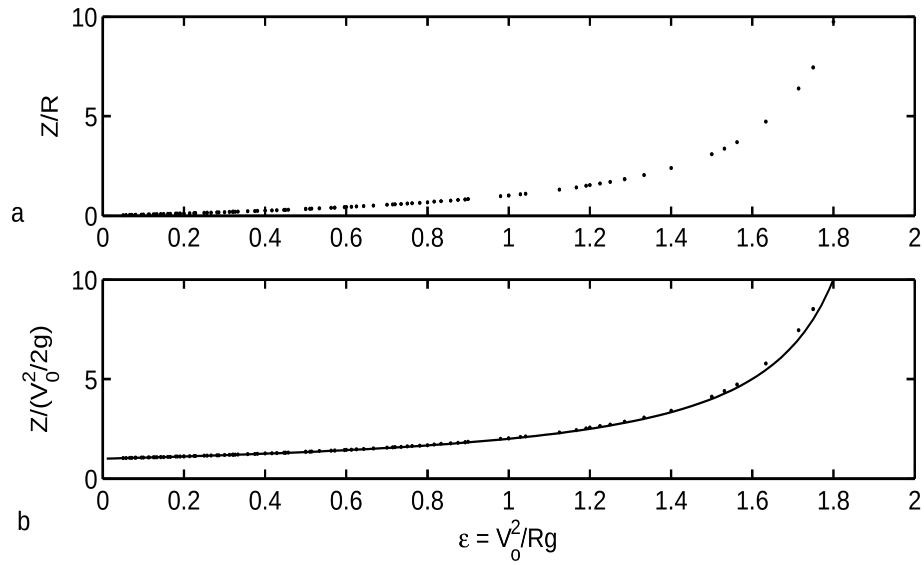

Figure 9: The maximum height of a projectile launched upwards at various speeds V_o from planets with various R and g. In (a), the maximum height has been nondimensionalized by the initial basis set. In (b), the basis set is one deduced from a scaling analysis. Either of these indicate a clear-cut dependence upon the single nondimensional parameter $\epsilon = V_o^2/gR$. Notice that as ϵ approaches 2 the nondimensional height goes to infinity as the projectile escapes the tug of the planet's gravity. The solid line in (b) is an approximate solution in which the effective (or average) gravitational acceleration is $g(1 - \epsilon/2)$ (developed in Sec. 7.4).

finds

$$\epsilon \frac{d^2 z'}{dt'^2} = -1/(1+z')^2 \approx -(1-2z'). \tag{83}$$

This nondimensional model equation contains the single parameter, ϵ. For earth-like values of R and g, and for $V_o = 2000$ m s^{-1} or less, say, ϵ is a small parameter, roughly 10^{-2}.

In this problem, the difficult (nonlinear) term is the z-dependent gravitational term of Eq. (78). It is plausible that the z-dependence could be ignored if $z \ll R$, and by extension, if $\epsilon \ll 1$. Thus, we should be able to solve the problem in the limit that $\epsilon \rightarrow 0$ and then go on to find a better solution by iteration. On that basis we might well guess that a first approximation to Eq. (83) might be had by simply dropping the term multiplied by the small parameter ϵ, which happens to be the acceleration term. However, the solution to the reduced equation would be $z' = \infty$, which is contrary to our assumption of small z', and is nonsensical, generally. Either the idea that we could find a useful approximation starting from small ϵ is wrong, or, we erred in dropping the acceleration term. In fact, it was the latter step that failed; there was no reason to conclude that the acceleration term could be dropped simply because it is multiplied by the small factor ϵ because at this point we have no idea how big the nondimensional acceleration $\frac{d^2 z'}{dt'^2}$ is compared with the terms kept, i.e., compared with 1. Neither did we consider that $\frac{d^2 z'}{dt'^2}$ might vary with ϵ. It turns out that $\frac{d^2 z'}{dt'^2}$ varies inversely with ϵ and is thus quite large for small ϵ. As a result, if we start by dropping the acceleration term we will not be able to proceed toward an improved solution. It is important to understand that the nondimensional Eq. (83) is not at fault here. The terms of (83) still have the ratio one to another of the dimensional Eq. (78) since all that has been done is to divide by parameters. It is only the inferences that might be drawn from Eq. (83) that could be at fault.

The important lesson to be learned is this — if we intend to make inferences about the size of terms that are multiplied by nondimensional parameters (that we can estimate *a priori*), then we have to take care that the nondimensional variables, in this case z' and $\frac{d^2 z'}{dt'^2}$, are themselves O(1) in the limit that the small parameter, ϵ, goes to zero. This additional requirement on the choice of nondimensional variables is often termed 'scaling', implying a thoughtful choice of the scales.[21]

7.3 Scaling the dependent variable

To insure that a nondimensional dependent variable is O(1) we have to nondimensionalize the dimensional dependent variable using a 'zero order' solution for the dependent variable. That is, the scale that we use to nondimensionalize the dependent variable must be roughly the maximum size that the (dimensional) variable will take on. Moreover, this scale must vary with parameters as does the dependent variable. In the initial nondimensional basis set we used R as the length scale by which to nondimensionalize z. Though R is certainly an intrinsic length scale for the problem, it has no direct relation to the size of z and more to the point, it obviously does not have the small ϵ parameter dependence of z. R thus fails to make a useful scale for z in the way that we now require.[21]

In this problem we can form a zero order solution for the dimensional z by ignoring the height dependence of the gravitational acceleration and so find

$$z_0 = V_o t - gt^2/2. \tag{84}$$

From this we can estimate that the maximum height, and thus an appropriate length scale for the height, is $Z_0 = V_o^2/2g$. An appropriate time scale can be estimated as the time it takes the acceleration g to produce (or erase) the initial velocity V_o, and is just $t_0 = V_o/g$. This second form of nondimensionalization is a second basis set, derivable from Eq. (81) when $n = -1$, i.e.,

$$\Pi_1 = \frac{z}{V_o^2/2g}, \quad \Pi_2 = \frac{t}{V_o/g} \quad \text{and} \quad \Pi_3 = \frac{V_o^2}{gR}. \tag{85}$$

The relation between these three nondimensional variables can be written

$$\frac{z}{V_o^2/2g} = F(\frac{t}{V_o/g}, \frac{V_o^2}{gR}). \tag{86}$$

The maximum height corresponding to this basis set will be

$$\frac{Z}{V_o^2/2g} = F(\epsilon), \tag{87}$$

where $\epsilon = \frac{V_o^2}{gR}$ as before (and notice that ϵ is proportional to the ratio of the zero order maximum height to the radius of the planet). The result is, again, a clearly defined functional dependence of maximum height upon the single parameter ϵ (Fig. 9, lower). This second $F(\epsilon)$ happens to look something like the initial form, though with one significant difference. At small values of ϵ the new F goes to a constant, 1, where the initial F decreased to zero as $\epsilon \to 0$. Thus the new scaling is consistent with an underlying zero order (vanishing ϵ) solution, while the initial basis set was evidently not. This new basis set of nondimensional variables seems awfully obvious now that we have it in front of us, but then the first basis set seemed obvious, too.

7.4 Approximate and iterated solutions

We now have length and time scales that are appropriate specifically to the vertical motion of a free projectile, as opposed to just any obvious length and time scales that may happen to be in the physical model. Using this new scheme, the nondimensional velocity of the projectile is

$$\frac{dz*}{dt*} = \frac{dz/(V_o^2/2g)}{dt/V_o g} = \frac{2}{V_o}\frac{dz}{dt}, \tag{88}$$

and the acceleration is

$$\frac{d^2z^*}{dt^{*2}} = \frac{dz/(V_o^2/2g)}{(dt/V_o g)^2} = \frac{1}{g}\frac{d^2z}{dt^2}. \tag{89}$$

Thus, if the dimensional acceleration is equal to g (and it is for small ϵ) then this nondimensional acceleration will be O(1) for all values of the parameters that are in the small ϵ range. If we rewrite the equation of motion using this new scheme the result is

$$\frac{d^2z^*}{dt^{*2}} = -1/(1 + \frac{\epsilon z^*}{2})^2 \approx -(1 - \epsilon z^*). \tag{90}$$

Now when we drop the term multiplied by ϵ we recover a sensible first approximation to the projectile problem; in fact, we recover our 'zero order' model! It seems that we have gone in a circle, but with Eq. (90) we know where to go next. If we use the first 'zero order' solution as an estimate of z^*, we can then solve for a new maximum height given the reduced gravity (compared to the value at the planet surface, $z = 0$). After several iterations of this sort the solution for, say, the maximum height approaches very near the solution found in the full problem. Inspection of (90) indicates that the effect of the height-dependence of gravity is as if the gravity were reduced by the fraction $\epsilon/2$ (recall that z^* is O(1), and so the average value of z^* is 1/2). The first order (first iteration) solution for maximum height computed in this way looks quite good, even for rather large values of ϵ, Fig. 9. It commonly happens that the first iteration gives valuable insight into the parameter dependence of a phenomenon in a way that numerical model solutions, no matter how extensive and precise they may be, can not.

8 Concluding remarks

The claim was made in Sec. 1 that dimensional analysis was occasionally quite powerful. With some experience we can see that dimensional analysis is most useful in cases wherein mathematical model is either not known or cannot be solved usefully. Dimensional analysis can always make a little progress toward a solution merely by showing the form that variables must take in an equation that is invariant to a change of units. That, in a nutshell, is what dimensional analysis does. In the case that there are only two or three nondimensional variables in a problem, dimensional analysis can be an immensely powerful tool leading almost directly to a solution (the inviscid pendulum) or an efficient way to correlate a large data set (drag on a moving sphere). If the problem has many variables (the viscous pendulum), then dimensional analysis alone will probably not suffice to reveal all that we may need to know, and further analysis will be required.

We have emphasized that an equation written in nondimensional variables, for example Eq. (17), is more efficient than its dimensional counterpart, Eq. (5). There is something to keep in mind, however. An equation written in nondimensional variables must be accompanied by a definition of the nondimensional variables. Better yet is an explanation of just why a particular definition was used, and what its advantages and limitations may be. The thoughtful use of dimensional analysis is a hallmark of insightful analysis, while the cavalier use of nondimensional variables can obscure what might otherwise have been a valuable message.

The mathematical steps that produce a nondimensional basis set are certain and quick (indeed, automatic), and the physical model is a finite list of variables. The ease with which a dimensional analysis can be done may engender confidence that the procedure is without risk of error. When dimensional analysis is applied to a mathematical (or numerical) model, this may well be true. But when dimensional analysis is meant to describe a real, physical system, that is not the case. Though the mathematical analysis is certain, it remains that the definition of an appropriate physical model is seldom as straightforward as the examples here might suggest. The absolute requirement that the physical model be complete is always at odds with the practical need to keep the physical model concise. The success of a dimensional analysis depends upon finding a satisfactory compromise; this requires judgment that comes with experience and from continual reference to relevant observations and numerical integrations.

Acknowledgements: The author was supported primarily by the Seward Johnson Chair in Physical Oceanography, administered by the Massachusetts Institute of Technology — Woods Hole Oceanographic Institution Joint Program in Oceanography. Additional support during the period of manuscript preparation was provided by the U.S. Office of Naval Research. Thanks to James Girton, Eric Kunze and Andy Solow for their comments on a draft manuscript, and to Jim Hansen and Harvey Gould for comments and suggestions that have helped to clarify this manuscript.

References

[1] E. Thurairajasingam, E. Shayan, and S. Masood, "Modelling of a continuous food pressing process by dimensional analysis," Computer and Industrial Engineering, in press. J. R. Hutchinson and M. Garcia, "Tyrannosaurus was not a fast runner," Nature **415**, 1018–1022 (2002).

[2] F. Wilczek, "Getting its from bits," Nature **397**, 303–306 (1999).

[3] This essay builds upon the introduction to dimensional analysis that can be found in most comprehensive fluid mechanics textbooks. Recent examples include P. K. Kundu and I. C. Cohen, *Fluid Mechanics* (Academic Press, 2001), B. R. Munson, D. F. Young, and T. H. Okiishi, *Fundamentals of Fluid Mechanics* (John Wiley and Sons, NY, 1998), 3rd ed., D. C. Wilcox, *Basic Fluid Mechanics* (DCW Industries, La Canada, CA, 2000), and F. M. White, *Fluid Mechanics* (McGraw-Hill, NY, 1994), 3rd ed. An older but very useful reference is by H. Rouse, *Elementary Mechanics of Fluids* (Dover Publications,

NY, 1946). A particularly good discussion of the relationship between dimensional analysis and other analysis methods is by C. C. Lin and L. A. Segel, *Mathematics Applied to Deterministic Problems in the Natural Sciences* (MacMillan Pub., 1974).

[4] An algorithm for computing nondimensional variables has been implemented in Matlab and can be downloaded from the author's web page, `<http://www.whoi.edu/science/PO/people/jprice/misc/Danalysis.m>` or from the Matlab File Central archive (the file name is Danalysis.m).

[5] The simple pendulum is the starting point for most discussion of dimensional analysis including the classic text by P. W. Bridgman, *Dimensional Analysis* (Yale Univ. Press, New Haven, CT, 1937), 2nd ed., which is an excellent introduction to the topic, and the more advanced treatment by L. I. Sedov, *Similarity and Dimensional Methods in Mechanics* (Academic Press, NY, 1959). Still more advanced is G. I. Barenblatt, *Scaling, Self-Similarity and Intermediate Asymptotics* Cambridge Univ. Press, Cambridge, 1996).

[6] An excellent discussion of physical measurement and much else that is relevant to dimensional analysis is given by A. A. Sonin, "The physical basis of dimensional analysis," 2001. This unpublished manuscript is available from `<http://me.mit.edu/people/sonin/html>`.

[7] What would be the result if the acceleration of gravity, g, was omitted, that is, what phenomenon would that entail? What if g were omitted, but an initial angular velocity $\frac{d\phi}{dt}$ was included? What if in place of g we used the acceleration due to the earth's rotation, $\frac{\Omega^2}{R}$? ($\Omega \doteq m^0 l^0 t^{-1}$ is the rotation rate of the earth and R is the distance normal to the rotation axis.)

[8] T. Szirtes, *Applied Dimensional Analysis and Modeling* (McGraw-Hill Publishing, 1997).

[9] S. Brückner and the University of Stuttgart Pi-Group, `<http://www.pigroup.de/>`, is an excellent resource for advanced applications of dimensional analysis.

[10] The calculation of a null space basis is, in effect, what all computational methods accomplish, see E. A. Bender, *An Introduction to Mathematical Modelling* (Dover Publications, 1977). Here we point this out, and then delegate the calculation to the computer.

[11] G. Strang, *Introduction to Linear Algebra* (Wellesley-Cambridge Press, Wellesley, MA, 1998).

[12] For example, suppose that X_5 is a speed in British engineering units, feet/second, and we wish to compute X_5' in SI units, meters/second. This variable has dimensionality, $D_{15} = 0$ (X_5 does not have units of mass), $D_{25} = 1$ for length, and $D_{35} = -1$ for time. The appropriate scale change factors are $\alpha_1 = 0.435$ (pounds to kilograms for nominal g), $\alpha_2 = 0.3048$ (feet to meters), and $\alpha_3 = 1$ (seconds to seconds). Thus $X_5' = 0.3048 X_5$.

[13] Scientific explanation can take many forms, and may not even be the aim of all investigations. An interesting, concise discussion of explanation is by Karl Popper, 'The aim of science', Ch. 12 of *Popper Selections* Ed. by D. Miller. Princeton Univ. Press (1985)).

Notice that the maximum tension in Fig. 3b for any ϕ_0 is exactly 5 and occurs at $\phi_0 = \pi$. Can you explain this using energy conservation and the radial equation of motion?

[14] Detailed treatment of damping processes are by P. T. Squire, "Pendulum damping," Am. J. Phys. **54**, 984–991 (1986) and R. A. Nelson, and M. G. Olsson, "The pendulum: Rich physics from a simple system," Am. J. Phys. **54**, 112–121 (1985).

[15] One criterion is to follow conventions of the field. In this case Π_1 is a drag coefficient, usually defined as $C_d = H/\frac{1}{2}\rho A U^2$, where A is the frontal area of the object. For the purpose of this essay we will consider other possible forms for Π_1.

[16] An open question of considerable practical importance is whether the steady drag laws are robust in the sense of giving useful estimates in practical problems, say our pendulum, in which the idealized conditions are inevitably violated. Other data sets have been developed to define the effects of idealized surface roughness, for example, but our pendulum has a long list of violations, time-dependence, a nearby solid boundary (the floor), slight surface roughness ... all present at once, so that we are on our own. About all that can be said is that it is important to understand the full set of assumptions under which a similarity law has been defined, and to be skeptical of applications outside of those bounds.

[17] Even at very large Re it does not follow that viscosity is entirely irrelevant. Significant changes in the drag coefficient occur at around Re $\approx 2 \times 10^5$ due to changes in the viscous boundary layer and the width of the wake behind a moving sphere. This is the Re range of a well-hit golf ball or tennis ball, and is part of the reason that aerodynamic drag on these objects has a surprising sensitivity to surface roughness or spin. For much more detail on these phenomenon see S. Vogel, *Life in Moving Fluids* (Princeton Univ. Press, 1994) and P. Timmerman, and J. P van der Weele, "On the rise and fall of a ball with linear and quadratic drag," Am. J. Phys. **67**, 538–546 (1999).

[18] Can you calculate a Reynolds number for the bob and the line from the original six nondimensional variables of Eq. (47)? Which nondimensional variable is present in Eq. (47) but not in Eqs. (57) and (58)? How or why was it omitted? Under what conditions (what parameter range) would you expect to see a significant effect of the time-dependent motion? How could you test (in principle and in practice) that the steady drag formulations really are appropriate for modeling the damping of a simple pendulum? You might, for example, consider that the fluid medium was water in place of air (the approximate density and kinematic viscosity of water are $\rho = 1.0 \times 10^3$ kg m^{-3} and $\nu = 1.8 \times 10^{-6}$ m^2 s^{-1} at a temperature = 0°C, and $\rho = 1.0 \times 10^3$ kg m^{-3} and $\nu = 0.7 \times 10^{-6}$ m^2 s^{-1} at a temperature = 40°C). Can you think of a name more apt than 'viscous' pendulum?

[19] We should perhaps qualify this. When we say that diffusion reaches a certain distance we mean that an appreciable (or given) fraction of the boundary value amplitude will be found at that distance from the boundary, say. The continuous diffusion equation has the property that any given point is effected by the entire domain instantaneously. However, if the point is far away from the boundary in the sense that η is large, then the boundary effect will be correspondingly small, though never literally zero.

[20] One way to think about this problem in the large is that the governing equation (62) determines the structure of the solution for short and intermediate times, but the boundary conditions, and specifically the lower no-slip boundary condition, wins out in the end (at steady state). This serves to show how crucially important boundary conditions are for fluid mechanics problems generally, since fluid mechanics problems usually require solution over an entire domain defined by some kind of boundary. Moreover, fluid motion is often forced or initiated by or through the boundary.

Can you infer the steady solution for Stokes First Problem? Suppose that the lower boundary was one of free-slip (which in practice means no gradient normal to the surface)? Suppose that the upper boundary condition was an imposed stress (rather than a speed), what then? (Imposed stress implies that the gradient normal to the boundary is imposed.)

[21] This new use of the word 'scale' is confusing but conventional terminology. Before when we said that R was a length 'scale' we meant only that it had the right dimensions to serve as a unit measure of height. We imply much more when use the phrase 'scaling'; that the unit measure is more or less the same size as the maximum value of the dimensional variable.