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

James F. Price
Dept. of Physical Oceanography, Woods Hole Oceanographic Institution
Woods Hole, Massachusetts, 02543
http:// www.whoi.edu/science/PO/people/jprice   jprice@whoi.edu

Vers. 2024 - A2   October 27, 2024

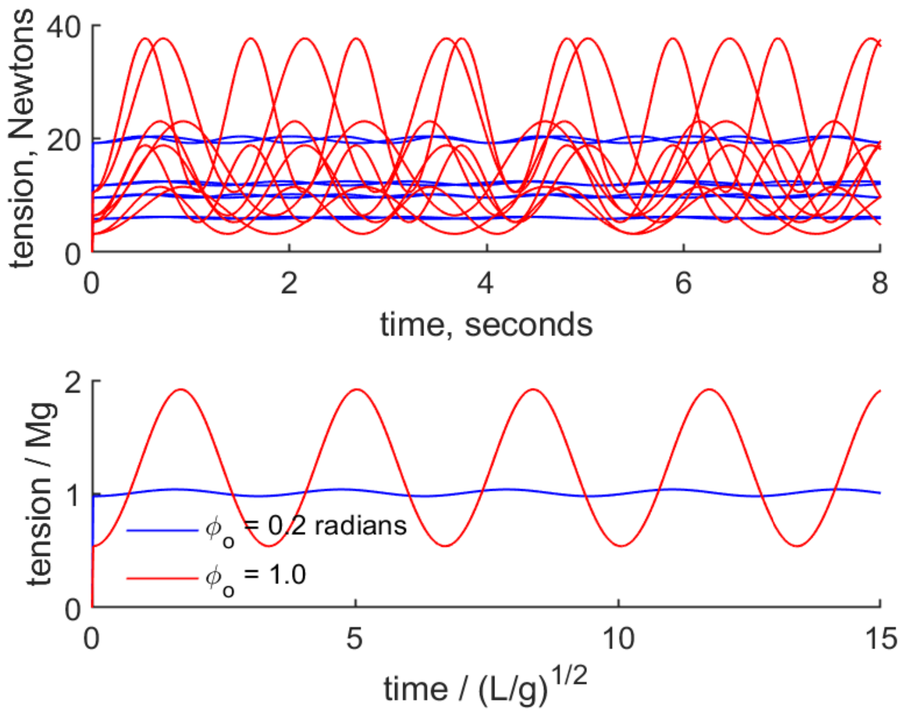


Figure 1: This essay is an introduction to dimensional analysis, one of the most valuable tools of applied mathematics. The essential idea is a change the variables from the usual dimensional variables to nondimensional variables that are formed from scales natural to the phenomenon. An immediate benefit, illustrated here, is that the nondimensional variables will be fewer in number than the dimensional variables. These data are tension in the line of a simple pendulum computed during 16 numerical experiments. In the upper panel, these data are shown in dimensional coordinates, time in seconds and tension in Newtons. In the lower panel, the same data are shown in nondimensional coordinates defined from natural scales of a pendulum: time is scaled by (nondimensionalized by) the reduced period of small amplitude oscillation, $\sqrt{L/g}$, and tension is scaled by the weight of the bob, Mg. With these choices, the 16 distinct dimensional solutions collapse into two solutions that differ only by their initial angular displacement, ϕ_o, large or small amplitude (red or blue lines). These graphs present the same data, but the nondimensional version is far more legible. Better still, it applies to all similar simple pendulums (similar in this case means the same ϕ_o).

Summary. Dimensional analysis may be applied with advantage to virtually every quantitative model and data set. In rare but important cases the result of a dimensional analysis will be a nearly complete solution. More often the result will be a partial solution, an efficient way to display a large data set, or reliable guidance for simplifying model equations.

A method for performing a dimensional analysis proceeds in three steps. The first step is to list 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 and parameters, dubbed a VPlist, is best written in nondimensional variables that are independent of the choice of units, i.e., whether SI or Imperial. The outcome of a dimensional analysis will be most useful when the VPlist is concise, while also including the most important variables that define the phenomenon. This requires a compromise that can best be achieved with reference to observations or numerical solutions.

The second step is calculation of a null space basis of the corresponding dimensional matrix. Each vector of the null space basis corresponds to a nondimensional variable, the number of which is less than the number of dimensional variables. This reduction in the number of variables is a key result of dimensional analysis. The nondimensional variables may themselves be viewed as a basis set, and in most cases their form will not be determined fully by dimensional analysis alone.

The third and artful step is to choose a form of the nondimensional variables that will maximize utility and insight. The properties of a null space basis facilitate this process. The most important choice is the scale used to nondimensionalize the dependent variable. A good strategy is to compute this scale from a physically motivated even if highly simplified, zero order model for the dependent variable. 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 that arises often in models of fluid dynamics.

When applied to data sets, dimensional analysis can help find efficient ways to display a large data set and thus help find correlations. Aerodynamic drag on a moving object is an important example. For a given object, i.e. a sphere or a cylinder, a zero order solution for drag can represent inertial or viscous processes. Either way, the resulting nondimensional drag, often called the drag coefficient, will depend upon a Reynolds number, Re, alone. The choice of the zero order solution (inertial or viscous) can be made to minimize the Re-dependence within a subrange of Re.

When applied to model equations, dimensional analysis can help simplify equations by reducing the number of parameters. If, in addition, the choice of the scales considers magnitude, then the comparative magnitude of terms in an equation will be evidenced by their nondimensional coefficients. This scaling analysis sets the stage for simplifications and further analysis.

Contents

1 Learning dimensional analysis

This is an introduction to dimensional analysis, one of the most widely useful and rewarding methods of applied mathematics. It may be applied fruitfully to almost any quantitative model or data set on topics ranging from dinosaurs[1] to donuts[2] and the most fundamental theories of physics.[3] Dimensional analysis is most useful, sometimes indispensable, for problems that have no solvable theory. Dimensional analysis can always make a little progress towards a solution, and some of these, e.g., 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 will be a broad hint at the form of a solution or a more efficient way to display a large data set. The pendulum data of Fig. (1) are an example: the nondimensional version of the data is far more legible than is the dimensional version, and it is far more useful in that it applies to all similar, simple pendulums.

author and DALL·E 2

Dimensional analysis can lead to insights that would have been hard to discover in a mass of dimensional data. The maximum tension in the line of a simple pendulum is found to be proportional to the mass of the bob, M (not surprising) and independent of the line length, L, which is not so obvious, Fig. (1), lower. The maximum tension increases significantly with the amplitude of the motion, set by ϕ_o. These results are typical of those coming from dimensional analysis — seldom complete if taken alone, but nevertheless an important step toward improved understanding in many investigations.

1.1 The goal and the plan for this essay

The goal is to help you learn a systematic and partially automated method of dimensional analysis that you will be able to apply to your own research. The key to learning is motivation — why should you devote your time and energy to this task? With that question in mind, the concepts and advantages of dimensional analysis are developed in Sec. 2 by treating a familiar problem, the oscillations of a simple pendulum, using an informal method of dimensional analysis. A more rigorous and partially automated method of dimensional analysis is described in Sec. 3. This method is applied to more advanced applications and topics including a damped, viscous pendulum in Sec. 4, a similarity solution for diffusion in one dimension in Sec. 5, and scaling analysis of

[1] Dinosaurs: J. R. Hutchinson and M. Garcia, 2002, 'Tyrannosaurus was not a fast runner', Nature **415**, 1018–1022.

[2] Donuts: Delaplace, G., K. Loubire, F. Ducept and R. Jeantet, 2015, *Dimensional Analysis of Food Processes,* Elsevier, ISBN : 9781785480409, http://www.iste.co.uk/book.php?id=876.

[3] Physics: Wilczek, F., 1999, 'Getting its from bits,' Nature **397**, 303–306 and https://cds.cern.ch/record/403150/

projectile motion in variable gravity in Sec. 6. Additional applications and homework projects are in Sec. 7.

1.2 A method of dimensional analysis

The method of dimensional analysis espoused here can be envisioned in three steps, previewed below and described in detail in Sec. 3.

1. Definition of a problem. The first step is to define a problem by making a list of the relevant variables. The list starts with one dependent variable, and continues with the independent variables and parameters that are thought to be important for determining that dependent variable. This list of Variables and Parameters is dubbed the VPlist (Sec. 2). That something useful could follow from such a minimal specification is at the heart of what makes dimensional analysis so widely applicable,[4] and also a bit mysterious. The premise of dimensional analysis is that a complete equation is best written in a form that is invariant to the arbitrary choice of units, i.e., whether SI or Imperial, or any other. When that constraint is implemented by a dimensional analysis, the dimensional variables that make up the VPlist will appear in combinations that are nondimensionalized (normalized or divided) by scales that are natural to the problem.

2. Calculating a basis set of nondimensional variables. The usual method of finding the nondimensional variables relies upon the long-running Buckingham Pi theorem to define the number of required nondimensional variables, followed by inspection.[5] That works well for small problems such as the simple pendulum of Sec. 2. For larger problems, e.g., the damped pendulum of Sec. 4, it can be helpful to compute the nondimensional variables from the null space basis set of the dimensional matrix that characterizes the VPlist.[6] This calculation (Sec. 3) leads immediately to a complete, orthogonal set (a basis set) of nondimensional variables.

[4] E. S. Taylor said this regarding dimensional analysis: 'For the amount of time and effort required to understand it and use it, dimensional analysis offers unusually great rewards and it therefore should become a part of the tool kit of every engineer...'. I agree wholeheartedly, and would include physical scientists, data scientists and applied mathematicians among those who will benefit. Taylor, E. S. , 1974, *Dimensional analysis for engineers*, Clarendon Press.

[5] An introduction to dimensional analysis can be found in most comprehensive fluid mechanics textbooks. Recent examples include P. K. Kundu and I. C. Cohen, 2001, *Fluid Mechanics*, Academic Press, and B. R. Munson, D. F. Young, and T. H. Okiishi, 1998, *Fundamentals of Fluid Mechanics*, John Wiley and Sons, 3rd ed. An older but still very useful reference is by H. Rouse, 1946, *Elementary Mechanics of Fluids*, Dover Publications, NY. A particularly good discussion of the relationship between dimensional analysis and other analysis methods is by C. C. Lin and L. A. Segel, 1974, *Mathematics Applied to Deterministic Problems in the Natural Sciences* MacMillan Pub. Also Bender, E. A., 1977, 'An Introduction to Mathematical Modelling', Dover Publications. A very appealing source available online is https://ocw.mit.edu/courses/6-055j-the-art-of-approximation-in-science-and-engineering-spring-2008/

[6] Price, J. F., 2006, 'Dimensional analysis of models and data sets', *Am. J. Phys.*, **71**(5), 437-447.

3. Constructing a preferred basis set. More important than the calculation *per se* is that the properties of a null space basis provide very useful guidance for the third, and artful step of an analysis — the reordering, as necessary, of the initial basis set of nondimensional variables into a form that maximizes utility and insight. This is emphasized in the examples of Secs. 4, 6 and 7. The most important choice is the scale[7] used to nondimensionalize the dependent variable. A good choice for this scale is likely to come from a highly simplified, zero order model of the dependent variable. From that perspective, dimensional analysis develops into scaling analysis discussed in Sec. 6.

1.3 About this essay

This essay has been written for students who are at the level of a first or second course in classical physics or applied mathematics, e.g., fluid dynamics or partial differential equations. The intent is to provide a resource that is suitable for self-study. A first edition was published online in 2006. The present 2024 edition aims for improved clarity and offers a wider range of homework projects.

The author was supported by the Seward Johnson Chair in Physical Oceanography, MIT/WHOI Joint Program in Oceanography, and by the U.S. Office of Naval Research. My thanks to Sidney Batchelder of WHOI for the Python scripts linked in Sec. 3.4, and to Jack Whitehead of WHOI for his thoughtful comments on a draft manuscript.

[7] The word 'scale' is used here as a noun and a verb, and alongside 'normalize' and 'nondimensionalize'. The present usage is standard, but merits a brief explanation. When used as a noun, scale refers to a quantity used to measure a variable via normalization, i.e., division. For example, a standard meter is a length scale that may be divided into some variable length, L. The result is nondimensional, a length/length, but typically reported in the units of the normalization, 'L = xx meters'. In place of a standard length, a length scale may also be formed from variables that are intrinsic to a problem, and then the scale is said to be a 'natural scale'. This is an important aspect of dimensional analysis. In that case, the result of a normalization is again nondimensional, but the units are generally not attached directly and so must be explained elsewhere. When 'scale' is used as a verb or an adjective, as in 'scaling analysis', it usually means to perform a normalization in a way that considers also the magnitude of the scale.

2 An informal dimensional analysis of a simple pendulum

The first problem taken up in some depth involves the oscillatory motion of a tabletop-size, simple pendulum, Fig. (2).[8] A significant appeal of a simple pendulum is that it can be made and observed with very inexpensive materials. The bob position is defined by the angular displacement away from vertical, ϕ. The amplitude of the displacement, Φ, may be readily estimated from the measured cord, c, at the endpoint of the arc as $\Phi = \sin^{-1}(c/L)$, and the period of an oscillation easily measured with a stopwatch, uncertain to about 0.3%. These data will be referred to as 'observed' in later Fig. (4). What was not measured was the time-dependent displacement, $\phi(t)$, and the tension in the line. These observations are primitive compared to those made in some other studies[9] but they nevertheless serve an imporrant purpose, *viz.,* to help ensure that the analysis methods discussed here — theoretical, numerical, and dimensional analysis — are relevant to a real, physical system.

The motion of such a pendulum is only lightly damped by drag with the surrounding air and can be characterized by two distinct time scales – a very regular, fast time-scale oscillation having a period, P, of a few seconds, and a slow, more-or-less exponential decay of the amplitude having a time-scale, $\Gamma^{-1} \gg P$, of tens of minutes. This Sec. 2 will study the fast, oscillatory motion, idealized as if energy conserving. The aim is to show how dimensional analysis can help us understand how the oscillatory motion depends upon the properties of the pendulum, e.g., the length of the line, mass of the bob, etc. Sec. 4 will take up the more complex dynamics associated with damping.

2.1 Mathematical models of a simple, inviscid pendulum

Dimensional analysis is especially valuable when dealing with problems that do not admit a known or solvable mathematical model. That is not the case for a simple pendulum, for which mathematical models are well-known and very useful for generating data sets including the tension in the line, Fig. (1), and angular displacement, Fig. (3).

[8] A simple pendulum is a common starting point for discussion of dimensional analysis including the classic text by P. W. Bridgman, 1937, *Dimensional Analysis*, 2nd ed., Yale Univ. Press, New Haven, CT, which will always be an excellent introduction to this topic. More advanced is Sedov, L. I., 1959, *Similarity and Dimensional Methods in Mechanics*, Academic Press, NY.

[9] Mathevet, R. N. Lamrani, P Ferrand, J.P. Castro, P. Marchou and C. M. Fabre, 2022, 'Quantitative analysis of a smartphone pendulum beyond linear approximation: A lockdown practical homework', Am. J. Phys., 90, 344 - 350. https://doi.org/10.1119/10.0010073

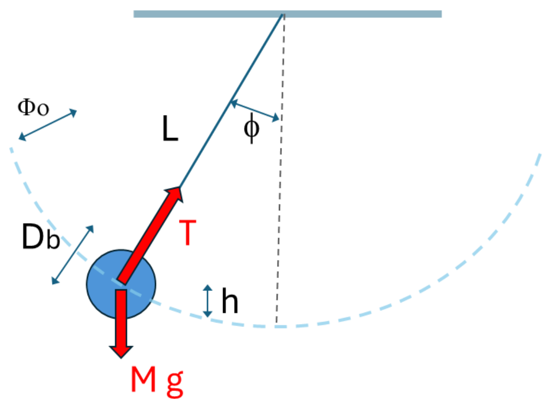


Figure 2: A simple (one bob) pendulum. In the example of the real pendulum observed here, the line is a thin monofilament of length $L = 1$ m. In models, the line is idealized as a rigid, massless rod. The bob is a small, spherical lead fishing weight having a mass of 20 grams. The distance of the bob above the lowest point of the arc is h. Not indicated here are the diameter of the line, D_l, and the density and viscosity of the air through which the pendulum swings.

The mathematical models considered here in Sec. 2 are appropriate to an undamped, energy-conserving simple pendulum. The basis is conservation of angular momentum: the angular momentum of the swinging bob, $M\,V_\phi L = M\,L^2\,d\phi/dt$, is presumed to change only by virtue of the torque associated with the downward force of gravity acting on the bob,[10]

$$ML^2\frac{d^2\phi}{dt^2} = -MLg\,\sin\phi, \tag{1}$$

and after minor rearrangement,

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

This will be referred to as the equation of motion. For comparison with the available observations, it is preferable to release the bob from a state of rest at a known displacement, ϕ_o, and so the initial conditions are

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

Energy in this system is the sum of gravitational potential energy and kinetic energy

$$E = E_p + E_k. \tag{4}$$

The potential energy referenced to the lowest point in the arc, $\phi = 0$, is

$$E_p = Mgh = MgL(1-\cos\phi), \tag{5}$$

[10] https://en.wikipedia.org/wiki/Pendulum and follow the link to mechanics.

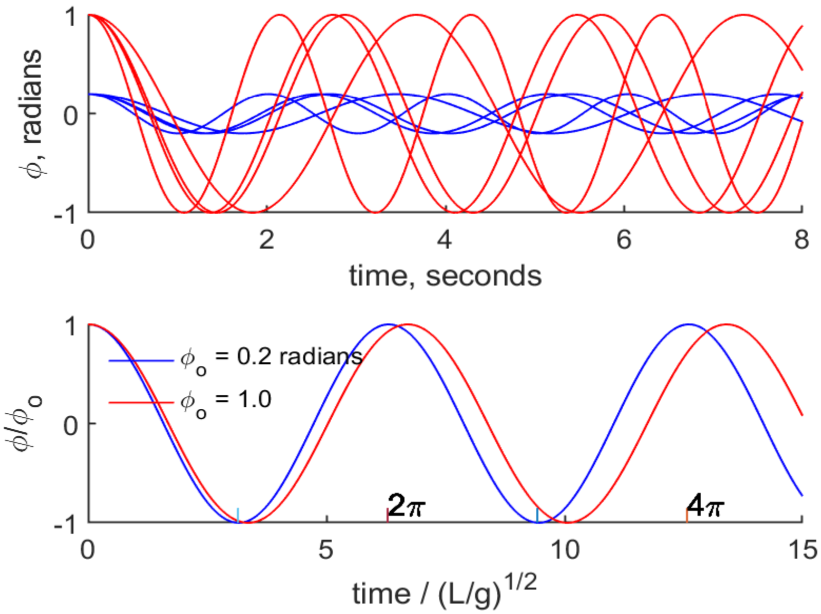


Figure 3: Numerical solutions for the angular displacement $\phi(t)$ of a simple pendulum. These are the same 16 solutions used to calculate the tension in the line of Fig. (1). **(upper)** Displacement plotted in dimensional units. The mathematical model of ϕ, Eqs. (2) and (3), does not depend upon M, and so there are eight distinct solutions here. **(lower)** The displacement data plotted in nondimensional units that will be discussed in Sec. 2.3. The eight distinct solutions shown above collapse to just two solutions that vary by the initial displacement, ϕ_o, a considerable simplification.

and the kinetic energy is

$$E_k = \frac{1}{2}M(L\frac{d\phi}{dt})^2. \tag{6}$$

When the bob is released from rest it has zero kinetic energy, and the total energy is then the initial potential energy,

$$E_o = MgL(1 - \cos\phi_o) = \text{constant} = E_p + E_k, \tag{7}$$

which is conserved (ignoring damping), and so from Eqs. (4) - (7),

$$MgL(1 - \cos\phi) + \frac{1}{2}M(L\frac{d\phi}{dt})^2 = MgL(1 - \cos\phi_o) \tag{8}$$

or

$$\frac{1}{2}(L\frac{d\phi}{dt})^2 = gL(\cos\phi - \cos\phi_o), \tag{9}$$

a useful relation for $\frac{d\phi}{dt}(\phi)$.

Tension in the line, T, appears in the radial equation of motion,

$$M\frac{d^2L}{dt^2} = -T + M\,g\,\cos\phi + ML(\frac{d\phi}{dt})^2.$$

The line is presumed to be inextensible, $dL/dt = 0$, and so the tension may be diagnosed from a known $\phi(t)$ and

$$T = M\,g\cos\phi + ML(\frac{d\phi}{dt})^2. \tag{10}$$

2.1.1 Analytic solutions

The solution method for Eqs. (2) and (3) depends upon the initial angle, ϕ_0. If ϕ_0 is restricted to small amplitude, values less than about 0.1 radian, then $\sin\phi$ in Eq. (2) can be approximated well enough for many purposes by ϕ. The resulting linear model is that of a simple harmonic oscillator having the well-known solution

$$\phi(t) = \phi_0\cos(\frac{t}{\sqrt{L/g}}), \tag{11}$$

and hence the period of small amplitude oscillation is

$$P = 2\pi\sqrt{L/g}. \tag{12}$$

It is convenient to refer to the reduced period, $P/2\pi$, which is $\sqrt{L/g}$.

In the general case that ϕ_0 may take any value from $-\pi$ to π, a solution for the period cannot be written in elementary functions. However, the period can be calculated using the first integral provided by energy conservation (9) to find $d\phi/dt$ as a function of $\phi(t)$. Then, separate the variables ϕ and t, and integrate over one quarter of an oscillation, ϕ from ϕ_o to 0, and t from 0 to $P/4$, to arrive at an elliptic integral called 'theory' in Fig. (4),

$$P = 4\sqrt{\frac{L}{2g}}\int_0^{\phi_0}\frac{d\phi}{\sqrt{\cos\phi - \cos\phi_o}}. \tag{13}$$

2.1.2 Numerical solutions.

The equations of motion (2) and (3) are straightforward ODEs that may be solved by elementary finite difference methods for any value of ϕ_o. An ensemble of 16 numerical solutions was generated for two values each of line length, $L = (1, 1.8)$, mass of the bob, $M = (1, 2)$ kg, the acceleration of gravity, $g = (9.8, 6)$ m^2 s^{-1}, and the initial angle, $\phi_0 = (0.2, 1.0)$ radians, Figs. (1) and (3).

These numerical solutions occupy an important middle ground between the observations made on a real pendulum and the solutions of a solvable theory. The mathematical model that underlies the numerical solutions is known, and the numerical solutions for this problem can be very precise. However, all that we get from a numerical solution is the answer, the $\phi(t)$, for one specific case.

Developing insight for the parameter dependence and for the dynamics of a system will usually require an ensemble of solutions and further (dimensional?) analysis.

2.2 A VPlist of the most important Variables and Parameters

The starting point of a dimensional analysis is the assertion that there exists a solution for the angular displacement, $\phi(t)$,

$$\phi \; = \; F(t,\, g,\, M,\, L,\, \phi_o), \tag{14}$$

where F is for now an unknown function (and F will be used repeatedly in that role). A solution is a relation between the dependent variable, ϕ, and the independent variables, here just the time, t, and the pendulum properties, or parameters, that we expect will define the displacement.[11] Several properties of a simple pendulum would seem to be important and so are listed as parameters in the argument of F — the acceleration of gravity, g, the mass of the bob, M, and the length of the supporting line, L. To account for why there is motion at all, the initial displacement, ϕ_0, is also necessary. Parameters are variables that are constant during a particular realization — g, M, L and ϕ_0 in this list — but that vary over some range that defines the family of pendulums and environments that are of interest.

The Variables and Parameters that enter (14) will be compiled into a list, dubbed the VPlist, that defines a problem. A VPlist[12] includes the physical dimensions of each variable; the acceleration of gravity, g, has physical dimensions length time^{-2}, that is written here as a row vector whose components are the powers on $[\,\text{mass}\;\; \text{length}\;\; \text{time}\,]$, and hence $g \;\doteq\; [\,0\;1\;-2\,]$. Following (14), the VPlist for $\phi(t)$ is

(15)

- A preliminary VPlist for the oscillatory motion of a simple pendulum:

1. the angular displacement, $\phi \;\doteq\; [\,0\;\;0\;\;0\,]$, the one and only dependent variable,
2. time, $t \;\doteq\; [\,0\;\;0\;\;1\,]$, an independent variable,
3. acceleration of gravity, $g \;\doteq\; [\,0\;\;1\;\;\text{-}2\,]$, a parameter,
4. mass of the bob, $M \;\doteq\; [\,1\;\;0\;\;0\,]$, a parameter,
5. length of the line, $L \;\doteq\; [\,0\;\;1\;\;0\,]$, a parameter,
6. the initial angle, $\phi_0 \;\doteq\; [\,0\;\;0\;\;0\,]$, a parameter.

---

[11] The (putative) solution (14) is written as an explicit function for ϕ. This has been done to keep some focus on the dependent variable, but is not essential in what follows; we could just as well presume $F(\phi,\, t,\, g,\, M,\, L,\, \phi_o) \;=\; 0$.

[12] Why this made-up word? In most of the literature, including my previous work, this list of variables is referred to as a 'physical model'. A list of variables is some sort of model, but to call it a 'physical' model began to seem like an overuse of the term physical.

A VPlist of this kind is the entire specification required for the dimensional analysis of a simple pendulum. This can be turned around: in effect, this VPlist defines what is meant by 'simple pendulum'. Notice that there isn't much here: the meaningful, quantitative part of a VPlist is the collection of physical dimension vectors, in this case six vectors having three components, the powers of [mass length time]. Of the 18 components, 13 are zero, and the rest are small integers. The variable names that we associate with these vectors carry an important meaning for us, but play no part in the analysis.

What is most important is the assertion that the VPlist is complete, meaning that it includes all of the variables required to construct a mathematical model that could in principle yield a solution (14). If the corresponding mathematical model is not known, then an assertion of completeness can only be a hypothesis. It is highly desirable that the VPlist 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 a VPlist thus requires considerable judgment and will often entail trial and error that is best guided by reference to experimental or numerical data.

The VPlist (15) makes no mention of the viscosity or density of air, as if the pendulum was swinging in a vacuum. The same is true of the mathematical models of Sec. 2.1, and thus the numerical integrations that are made with those equations. We can be confident that this airless, energy conserving VPlist is consistent with those models and solutions. Whether these omissions are acceptable as a model of the real, observed pendulum of Fig. (2) is a different matter that will be decided by comparison to observations.

2.3 The premise of dimensional analysis — invariance to a change of units

It is taken for granted that every valid equation must be dimensionally consistent, or homogeneous: if the left side of an equation defines a length, then the right side had better be a length as well, or otherwise there has been an error.[13] This is the most basic application of dimensional analysis and is of considerable utility when checking complex expressions.

[13] A thoughtful, in-depth discussion of measurement and dimensional analysis is by A. A. Sonin, 2001, *The physical basis of dimensional analysis*, MIT. https://www2.whoi.edu/staff/jprice/wp-content/uploads/sites/199/2024/04/AASonin.pdf and Sonin, A. A., 2004, 'A generalization of the Π-theorem and dimensional analysis'. Proc. Nat. Acad. Sci., 101, No. 23, 8525 - 8526.

Now consider the units used to measure mass, length and time. Shall we measure a length by comparison to the standard meter, or by comparison to the standard foot?[14] If this seems to you to be a rather empty question, then you will be sympathetic to

The premise of dimensional analysis: The relationship expressed by a complete equation is best written in nondimensional variables that are invariant to the arbitrary choice of units.

The key word here is 'best', and an important goal for this section is to show what that means.

Dimensional analysis is a systematic procedure for learning the form that the variables can take in an equation that satisfies units invariance. Angles are relevant here, and a good example. An angle is the ratio of two lengths, an arc length divided by the radius, and is thus inherently nondimensional. If the arc length and the radius are measured in units of meters, we will get a certain number. If units of feet are used to measure the same lengths, we will get precisely the same number for their ratio, that is, the same angle. Thus the left side of Eq. (14) is invariant to a change in the units of length. Consider the right hand side, which has two variables with fundamental dimension length, L and g. For the same invariance to the choice of units to hold also on the right hand side, the length and the acceleration of gravity must appear in the ratio L/g, or any power of the ratio, for example, $(L/g)^{1/2}$ or $(L/g)^2$ or $(L/g)^{-1}$. Otherwise a change in the units of length would almost certainly change the argument and hence the value of the (unknown) function F on the right hand side of Eq. (14). From these simple considerations we already know something useful about the units-invariant form of Eq. (14).

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 VPlist having dimensions of mass. This informal dimensional analysis leads to the conclusion that a solution for ϕ that is invariant to a change of units cannot depend upon the mass of the bob only. Either there must be another parameter having a dimension mass that was omitted from the VPlist, or, the mass should be excluded from the list of relevant parameters. This is an immediate result also in the mathematical models (Eqs. (2) and (3)) where it follows from the equivalence of gravitational mass and inertial mass. The same result can be deduced by dimensional analysis in the absence of a mathematical model (provided that there is just one mass in the VPlist, and not distinct inertial and gravitational masses).

A similar consideration of the units used to measure time indicates that t, L, and g must also appear together in a nondimensional variable, say $t/\sqrt{L/g}$, or just as well $t^2/(L/g)$, since any power of this ratio is equally suitable. Absent some reason otherwise, we may as well leave the independent variable t to the first power and so choose the former, $t/\sqrt{L/g}$.

[14] The standard meter was first defined in terms of Earth's circumference, which was exceedingly cumbersome and not precise. The meter is now defined in terms of universal constants, the speed of light divided by the transition frequency between two hyperfine ground states of cesium. The standard foot was once defined by an average of actual human feet, convenient but not precise. It is now defined as a fixed fraction of a meter.

The upshot of this informal dimensional analysis is that the variables and parameters that appear in a units-invariant form of Eq. (14) will appear in combinations that are nondimensional, i.e., nondimensional variables. The simplest, but not the only possible version is

$$\text{the nondimensional form,} \qquad \phi \; = \; F(\frac{t}{\sqrt{L/g}}, \phi_0). \tag{16}$$

This may be compared to the starting point, Eq. (14), repeated here,

$$\text{the dimensional form,} \qquad \phi \; = \; F(t, g, M, L, \phi_o). \tag{17}$$

2.3.1 How many nondimensional variables?

The nondimensional version (16) represents significant progress. In place of a dependence upon one independent dimensional variable and four parameters as in (17), the nondimensional version (16) has a dependence upon one nondimensional independent variable, $t/\sqrt{L/g}$, and just one nondimensional parameter, ϕ_o. This is the kind of result that can be expected generally, and is important enough to state as a corollary to the premise,

> *Corollary 1: The number of nondimensional variables will be fewer than the number of corresponding dimensional variables.*

The significance of this becomes evident when the data from an ensemble of numerical solutions is plotted using a nondimensional format in Figs. (1) and (3). The 16 (or 8) distinct dimensional solutions collapse into two nondimensional solutions that depend upon a nondimensional time and just one parameter, the initial displacement, ϕ_o, consistent with (16).

It is important to be able to anticipate the number of nondimensional variables, K, that is required to have a complete nondimensional form like (16). A useful rule of thumb: K is generally equal to the number of variables in the VPlist minus the number of fundamental dimensions in those variables. In the case of (15), there are six dimensional variables having three fundamental dimensions, mass, length and time, and so we can expect $K = 3$ nondimensional variables: ϕ, $t/\sqrt{L/g}$, and ϕ_o. Be cautioned, though, that while this informal estimate of K is correct in most practical problems, it is not fool proof. A better, rigorous calculation of K will be coming in Sec. 3.2.

2.3.2 Nondimensional variables formed with the natural scales of a VPlist

Nondimensional variables have a second inherent advantage over their dimensional cousins when, as here, they are formed with scales that are natural (or intrinsic) to the problem at hand. For example, the natural time scale of a simple pendulum is, as you might guess, the reduced period of small amplitude oscillation, which has been used as the scale to form a nondimensional time,

$$t^* = \frac{t}{P/2\pi} = \frac{t}{\sqrt{L/g}}. \tag{18}$$

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 in this.

Nondimensional time may sound a bit esoteric, but amounts to nothing more than measuring time in units of the reduced, linear period while taking explicit account of the $\sqrt{L/g}$ dependence of the period. If the analysis was to consider one pendulum only, then the whole exercise would amount to dividing the time by a constant. But if the scope of the investigation is all possible pendulums, i.e., all possible L and g, then there is considerable merit to measuring time in these natural units, including that the linear (small amplitude) period of the oscillation is 2π for all simple pendulums, Sec. 2.4.1 and Fig. (4).

This second and in some ways deeper benefit of nondimensional variables may be stated as

> *Corollary 2: Nondimensional variables that are formed with natural scales will often have a physical significance that the dimensional variables, scaled in the arbitrary units of kilograms, meters and seconds, do not.*

The advantage extends to model equations: a time derivative transforms as $dt = dt^*\sqrt{(L/g)}$, and the equation of motion, Eq. (2), becomes

$$\frac{d^2\phi}{dt^{*2}} = -\sin\phi\,, \tag{19}$$

with the initial conditions (ICs) as before, Eq. (3). The solution will be of the form

$$\phi = F(t^*, \phi_0)\,, \tag{20}$$

which is just like Eq. (16). If the amplitude of the motion is small in the sense that $\sin\phi \approx \phi$ in Eq. (19), then the solution to the linearized Eq. (19) with (3) is

$$\phi = \phi_0 \cos t^*. \tag{21}$$

The dependence upon L and g has not been omitted, but rather has been subsumed into the nondimensional time, t^*, so that Eq. (21) suffices for all L and g. This new equation (19) has the small advantage that it is simpler notationally, and it may be easier to recognize a canonical form. However, a switch to nondimensional units will preclude the prospect of catching algebraic errors that could be detected easily in a dimensional equation as a dimensional *in*homogeneity.

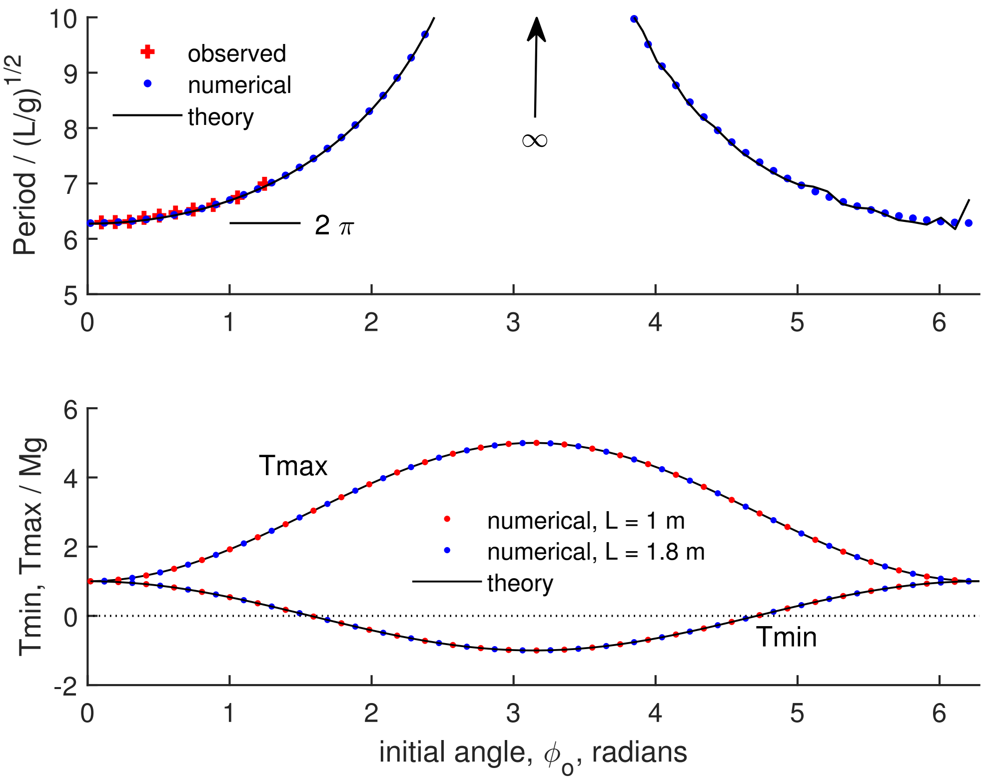


Figure 4: Observations and solutions for the period and the tension in the line of a simple pendulum as a function of the initial angle, ϕ_o. The flexible line of the real, observed pendulum and the initial condition $d\phi/dt = 0$ limit the initial angle to the range $\phi_0 < \pi/2$. In the mathematical models of Sec. 2.1, the line was idealized as a rigid, massless rod, and so all initial angles were possible, including $\phi_0 \rightarrow \pi$. **(upper)** The period in nondimensional coordinates. The data are from three sources: observations (red crosses), the period diagnosed from an ensemble of numerical solutions (blue dots), and as computed from theory, Eq. (13) (the black line). From dimensional analysis it is expected that the similarity function $F(\phi_0)$ evident here will hold for all simple pendulums that are consistent with the VPlist (15). **(lower)** The nondimensional minimum and maximum tension in the line. Data are from an ensemble of numerical solutions (the alternating red and blue dots are length $L = 1$ m or $L = 1.8$ m, and are indistinguishable), and from theory (the black line). Negative values of tension indicate that the line (massless rod) is in compression. There are no observations of tension.

Recall that the linear pendulum has the solution $\phi = \phi_0 \cos(t/\sqrt{L/g})$, Eq. (12), and note that the argument of the cosine is the same nondimensional time — it was there all along.[15] Evidently the difference between Eqs. (11) and (21) 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 both the stage of an investigation and upon your familiarity with dimensional analysis. Experimental data will generally be recorded in dimensional units (or the native units of the instruments), and it may also be preferable to carry out a numerical integration in dimensional units, if the magnitude of intermediate results is more familiar in those units. But when it comes time to report and interpret a collection of data from an ensemble of experiments, there can be a significant advantage to the use of nondimensional variables, evidenced in Figs. (1) and (3).[16]

2.4 Other dependent variables of a simple pendulum

2.4.1 Period of the oscillation

The most characteristic property of a pendulum is its period of oscillation, P. To observe the period, we simply measure the time interval between two extremes of the displacement. What can dimensional analysis tell us? The period does not vary with the dependent variable time, t, which should therefore be omitted from

- A VPlist for the period of oscillation of a simple pendulum: (22)

1. the period, $P \doteq [0 \ \ 0 \ \ 1]$, the dependent variable,
2. acceleration of gravity, $g \doteq [0 \ \ 1 \ \ \text{-}2]$, a parameter,
3. length of the line, $L \doteq [0 \ \ 1 \ \ 0]$, a parameter,
4. the initial angle, $\phi_0 \doteq [0 \ \ 0 \ \ 0]$, a parameter.

Compared with Eq. (14), this VPlist also omits the mass of the bob, which we learned just above. Following the same reasoning that led to (16), and expecting that there will be $K = 4 - 2 = 2$ nondimensional variables,

$$\frac{P}{\sqrt{L/g}} = F(\phi_0). \tag{23}$$

For a given ϕ_o, the (dimensional) period of a simple pendulum is expected to increase in proportion to the square root of the length of the supporting line and is independent of the mass. These are among the truly fundamental results of mechanics, deduced from observations and experiments by Galileo Galilei, Fig. (5), near the beginning of the Scientific Revolution, ca., 1600.

[15] The arguments of trigonometric, exponential and logarithmic functions are always nondimensional.

[16] A far more profound example of natural units comes in Sec. 7.1 which considers Planck units that are thought to be universal — literally the same throughout the universe. However, for all of their considerable appeal, Planck units are remarkably awkward size-wise for the day-to-day phenomena encountered here on Earth.

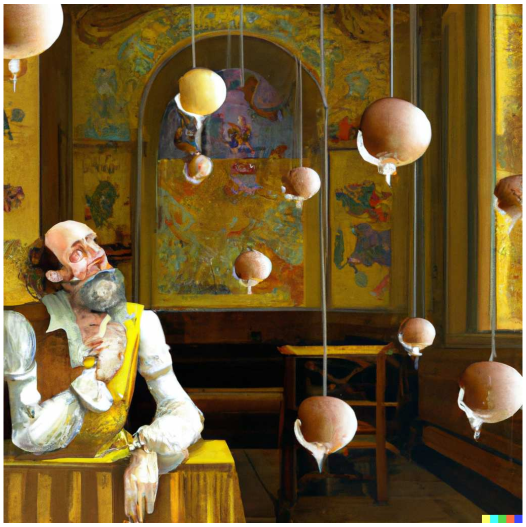


Figure 5: Galileo Galilei began a lifelong study of kinematics as a medical student observing chandeliers swinging in the breeze of a Pisa cathedral, ca. 1580. Clocks were then very primitive and Galileo timed their oscillations by counting his pulse. He reached several important conclusions: 1) every simple pendulum has a characteristic period of oscillation that is proportional to the square root of the length of the supporting line, 2) the period is independent of the mass of the bob, and, 3) the period of the oscillation is independent of the amplitude of the oscillation (now known to hold only for small amplitude motion.) These properties of a simple pendulum must have been noticed by many other cathedral-goers, but it was Galileo who quantified and reported the isochronicity property 3) of simple pendulums, a key concept that led to the later development of practical, accurate clocks.

Dimensional analysis has taken us part of the way to a solution in the form of Eq. (23), but not to the finish; the function $F(\phi_o)$ remains undefined and is not accessible to the methods of dimensional analysis alone. A function of the sort $F(\phi_o)$ will be referred to as a 'similarity function', in the sense that two simple pendulums having the same ϕ_o will have the same *non*dimensional period; the dimensional period will depend upon L and g as $P \propto \sqrt{L/g}$. The similarity function $F(\phi_o)$ might be determined by observations, the red crosses of Fig. (4), upper. A key contribution of dimensional analysis is that we can expect that the $F(\phi_o)$ inferred from observations made on our little table-top size pendulum will apply for all simple pendulums that are described by the VPlist (25). In this case we also have numerical and theoretical solutions that are consistent with the observations, and which are able to access the full range of ϕ_o.

The similarity function $F(\phi_o)$ of a simple pendulum has an important limit as the amplitude approaches zero; the similarity function approaches a constant value, $F(\phi_o \rightarrow 0) \ \rightarrow \ 2\pi$. From dimensional analysis, we can expect that this will hold for all simple pendulums that are described by the VPlist (22).

2.4.2 Tension in the line

For the tension, T, a force, we could write

$$T = F(t, M, L, g, \phi_0) , \tag{24}$$

which at this stage is not different from (14) (aside from the dependent variable).

- A VPlist for the tension in the line of a simple pendulum: (25)

1. the tension, $T \doteq [1 \quad 1 \quad \text{-}2]$, the one and only dependent variable,
2. time, $t \doteq [0 \quad 0 \quad 1]$, the independent variable,
3. mass of the bob, $M \doteq [1 \quad 0 \quad 0]$, a parameter,
4. length of the line, $L \doteq [0 \quad 1 \quad 0]$, a parameter,
5. acceleration of gravity, $g \doteq [0 \quad 1 \quad \text{-}2]$, a parameter,
6. the initial angle, $\phi_0 \doteq [0 \quad 0 \quad 0]$, a parameter.

The mathematical model for tension, Eq. (10), includes the angular velocity squared, $(d\phi/dt)^2$, which is proportional to the centrifugal acceleration. Even if we knew that, $d\phi/dt$ should be omitted from Eq. (25), since including $d\phi/dt$ would amount to having a second dependent variable that must itself depend upon t, M, L, g, and ϕ_0.

An informal analysis of this VPlist is much the same as that for $\phi(t)$ and leads quickly to $K = 6 - 3 = 3$ nondimensional variables,

$$\frac{T}{Mg} = F(\frac{t}{\sqrt{L/g}}, \phi_o), \tag{26}$$

the form used in Fig. (1), lower. The mass of the bob, M, *does* make an appearance in the nondimensional tension, since the dependent variable, T, is a force that has a dimension of mass. That the mass must appear is evident since the tension in the line will equal the weight of the bob, Mg, in the absence of motion, which is an important, limiting case. The maximum tension in the line of a swinging bob will exceed the weight of the bob by a factor that increases with the amplitude of the motion defined by ϕ_o.

The tension may also be characterized by a time-independent value, say $\hat{T}$, for example the minimum or maximum value during an oscillation, T_{min} and T_{max}, and so

- A VPlist for a time-independent tension: (27)

1. a time-independent tension, $\hat{T} \doteq$ [1 1 -2], the dependent variable,
2. the mass of the bob, $M \doteq$ [1 0 0], a parameter,
3. length of the line, $L \doteq$ [0 1 0], a parameter,
4. acceleration of gravity, $g \doteq$ [0 1 -2], a parameter,
5. the initial angle, $\phi_0 \doteq$ [0 0 0], a parameter.

The rule of thumb estimate of the number of nondimensional variables is $K = 5 - 3 = 2$, or one less than required for the time-dependent tension, and they are related by

$$\frac{\hat{T}}{Mg} = F(\phi_o). \tag{28}$$

This leads to a legible format of the numerical solutions, Fig. (4), lower. Notice that the length, L, has dropped out going from (26) to (28).

But which $\hat{T}$ is this, T_{min} or T_{max}? Evidently (27) and (28) are adequate for either, though of course there is a different similarity function $F(\phi_o)$ for T_{min} and T_{max}. This illustrates an important and perhaps unexpected property of VPlists generally: VPlists are likely to encompass a greater range of phenomena i.e., they are likely to be much less specific, than the phenomenon you intend to describe and analyze. And thus the VPlist (27) will suffice just as well for the time-averaged tension as it does for the minimum and maximum values. This is consistent with the comment made following Eq. (15) that a VPlist generally makes a rather spare description of a problem.

2.5 Parameters

One of the interesting aspects of dimensional analysis is that it treats parameters on an equal footing with dependent and independent variables. (The algorithms implemented here in Sec. 3 put the dependent and independent variables in a special place, just for convenience.)

Unnecessary parameters. Dimensional analysis has revealed that the period of a simple, inviscid pendulum does not depend upon the mass of the bob, M. This result might seem to suggest that the inclusion of extra or superfluous variables in a VPlist will not spoil the result. However, in most cases an extra variable will not be detected by a dimensional analysis, and will lead to an extra or at least an unnecessary nondimensional variable. For example, if we had included the bob diameter, D_b, in the VPlist (15), it would have been carried through to a nondimensional variable, D_b/L. Access to experimental data would soon show that D_b/L was of no great significance in determining the period of a nearly conservative pendulum, and so it would have been dropped from the final result.

Omitted parameters. Will the omission of a relevant variable be detected? Yes, but rarely, and only if the omission makes it impossible to nondimensionalize the dependent variable. For example, if we analyzed tension under the assumption that the mass of the bob M would be irrelevant (as it was for the oscillation period) then it would be impossible to construct a nondimensional tension. That would be a clear signal that something important had been left out of the VPlist. However, if the dependent variable can be nondimensionalized with the variables that are included, and in practice that is much more likely, then the purely formal procedure of dimensional analysis will not be able to identify an incomplete VPlist.

Locally and globally constant parameters. In the discussion of a simple pendulum, it was implicit that our interest included all possible values of the initial displacement, $-\pi \leq \phi_o \leq \pi$. In that event, the parameter ϕ_o would be deemed a 'locally constant' parameter, meaning that it is constant for a given experiment or realization, but that it may vary from one realization to the next. This is the default meaning of 'parameter'.

Suppose instead that the scope of a study is limited to a single value of the initial displacement, say $\phi_o = \beta$, and no other values are relevant. In that case, ϕ_o will be termed a 'globally constant' parameter, meaning that it is the same for every realization. The similarity function $F(\phi_o)$ would then be a constant specific to that β, say $F(\phi_o = \beta) = {}^{\beta}C$. The dependence of $\phi(t/\sqrt{L/g})$ upon nondimensional parameters that are made up solely from parameters that are constants (common examples could be heat capacity, fluid density and other thermodynamic parameters) may thus be ascribed to the similarity constant, and so

$$\phi(t/\sqrt{L/g}) = {}^{\beta}C$$

would suffice in this case. The globally constant nondimensional parameter ϕ_o may then be dropped from the list of nondimensional variables, though it leaves a mark on the similarity constant. If there are two such globally constant nondimensional parameters, say $\phi_o = \beta$ and $\phi_1 = \gamma$, then this may be extended,

$$\phi(t/\sqrt{L/g}) = {}^{\beta,\gamma}C$$

and ϕ_o and ϕ_1 both dropped. The happy result[13] is that the number of relevant (or active) nondimensional parameters may be less than the number of nondimensional parameters that first appears.

2.6 Summary: Why use nondimensional variables?

The informal dimensional analysis of this section has shown two significant benefits that you can expect to follow from the use of nondimensional variables.

More efficient representation. There will be a reduced number of nondimensional variables compared with dimensional variables. This will often lead to a more legible presentation of complex data sets, as in Figs. (1) and (3). It will also streamline model equations.

Physical interpretation. When dimensional variables are scaled (nondimensionalized) with the natural scales of a phenomenon, as they will be during a dimensional analysis, they are likely to have a significant physical interpretation. For example, the nondimensional reduced period of a simple pendulum oscillation is found to be $\sqrt{L/g}$ for *all* simple pendulums in small amplitude motion. This kind of result could come from a mathematical model, if it is available, or from dimensional analysis, even if not.

A different and wider perspective. There is a third possible benefit of dimensional analysis that is not as immediate, but that you may come to appreciate with time and experience. The method and the goals of dimensional analysis are quite different from the intensely focussed analysis methods that characterize most research, e.g., solving an equation of motion as in Sec. 2.1. The change in perspective that comes with dimensional analysis, e.g., thinking about the range of the relevant parameters of a problem, may help you develop insights on your problem that wouldn't necessarily come as readily from a more focussed approach to problem solving.

2.7 Pendulum problems

1) Compare the tension and the angular displacement of Figs. (1) and (3). Why is the frequency of oscillation different for these two variables? Notice that the range of the (nondimensional) maximum tension seen in Fig. 4, lower, is 1 to 5, while the minimum tension has a range -1 to 1. Can you explain these values using energy conservation and the radial equation of motion, Eqs. (9) and (10) of Sec. 2.1?

2) The approximate isochronism of simple pendulums in small amplitude motion is an important result going all the way back to Galileo's systematic investigations. It is not immediately obvious why this should hold for simple pendulums, and indeed, it does not hold at large amplitudes. From what you can see in Fig. (4), how would you characterize the limit $P(\phi_o \rightarrow 0)$? Can you explain the approximate isochronicity at low amplitudes? Why is the departure at larger amplitude towards longer period oscillations? (Hint: What does this imply about the restoring force?) What is the similarity constant for ϕ_o = 0.01, 0.1, 1.0 and π radians?

3) The simple-minded observational methods used here did not include a means to measure the angular velocity, $\omega = d\phi/dt$, and so the initial ω was taken to be zero. The effect of this is to exclude an entire family of simple pendulums. What new phenomenon and nondimensional variables would appear if the scope of the study had included an initial ω?

4) It is amusing to think of making pendulum measurements on Mars or the Moon where gravity is different from Earth's, but it is likely that our pendulum experiments will remain Earthbound. If that is so, then is it really necessary to carry along a constant parameter g in the dimensional analysis, as in Eq. (23)? Whether g is viewed as a locally or globally constant parameter must depend upon the application and the required precision. Can you envision exploiting the g-dependence of the pendulum period to measure the rather small, O(10^{-3}), variations of g along Earth's surface? [17]

5) Lead is an appropriate material for the bob of a pendulum[18] since its comparatively high density results in less drag with the air. However, lead can be toxic if handled carelessly, and so instead of a lead bob, consider a wooden bob or perhaps a foam rubber bob. Should the density of such a bob appear in the VPlist? Why did we ever leave it out? Would the similarity constant 2π of the small amplitude period then change?

6) In the analysis of tension in the line, Sec. 2.4.2, it was noticed that the line length L dropped out when going from the time-dependent tension to the time-independent tension, Eqs. (26) to (28). Can you explain why or how this happened?

7) Free fall in constant gravity is closely related to the (constrained) motion of a simple pendulum. Suppose the dependent variable is the vertical coordinate, z, and that the equation of motion is just

$$\frac{d^2z}{dt^2} = -g. \tag{29}$$

The ICs are

$$z = z_o(t=0) \;\text{ and }\; \frac{dz}{dt} = 0 \;\text{ at }\; t = 0. \tag{30}$$

What is the dimensional analysis version of this problem? How does your dimensional analysis result for $z(t)$ compare to the result from integrating (29) and (30)?

---

[17] An interesting discussion of applications is by https://microglacoste.com/product/fg5-x-absolute-gravimeter/

[18] Lead's chemical symbol is Pb, from the Latin *plumbum*, from which is derived 'plumb bob', a lead ball suspended on a string and used to observe the direction of gravitational acceleration, i.e., the vertical. If the bob is set into motion, the plum bob is a simple pendulum that may be used to infer the magnitude of the gravitational acceleration.

3 A basis set of nondimensional variables

Once a VPlist has been defined, the second and mathematical step of a dimensional analysis is to find a complete set of nondimensional variables for that VPlist. With a little experience and for small problems such as the simple pendulum of Sec. 2, the nondimensional variables can be constructed by inspection. For larger problems it may be helpful to use a computational method described in Sec. 3.2 that has its roots in linear algebra.[6] This method differs from most others[5] in that it does not rely on the Buckingham Pi theorem, although it comes to the same result, and instead computes the null space basis of the dimensional matrix (defined below) to find a basis set of nondimensional variables. The value of this method lies less in the calculation, which is straightforward, but in the perspective it provides on the result.

3.1 The mathematical problem

How to calculate a set of nondimensional variables from a VPlist? The nondimensional variables, called Pi variables, Π, are presumed to be the products of the dimensional variables, say X_1, X_2 and X_3 that are raised to some power,

$$\Pi \ = \ X_1^{S_1} \times X_2^{S_2} \times X_3^{S_3}, \tag{31}$$

where the exponents S_1, S_2, S_3 are unknowns. This is consistent with the nondimensional variables found by the informal analyses of Sec. 2. For a very abbreviated example consider

- A VPlist for the period of a simple pendulum in small amplitude motion: (32)

1. period, $P \ \doteq$ [0 0 1], the dependent variable,
2. length of the line, $L \ \doteq$ [0 1 0], a parameter,
3. acceleration of gravity, $g \ \doteq$ [0 1 -2], a parameter.

A Π variable formed as the product Eq. (31) will then have physical dimensions

$$\Pi \ \doteq \ S_1 \, [0\;0\;1] \ + \ S_2 [0\;1\;0] \ + \ S_3 [0\;1\;-2] \tag{33}$$

and will be nondimensional

$$\Pi \ \doteq \ [0\;\;0\;\;0]$$

if the exponents satisfy

$$S_2 \ + \ S_3 \ = \ 0, \qquad \text{and} \qquad S_1 \ - 2S_3 \ = \ 0. \tag{34}$$

This is two independent equations in three unknowns, and so this system is underdetermined. We may as well choose $S_1 = 1$ so that P will appear to the first power, and then $S_3 = 1/2$ and

$S_2 = -1/2$. Since this system is linear and homogeneous, $S_1 = \alpha$ and then $S_2 = -\alpha/2$ and $S_3 = \alpha/2$ with α real is also a solution. In this case (and assuming $\alpha = 1$) there is just one nondimensional variable,

$$\Pi_1 \;=\; P^{S_1} \times L^{S_2} \times g^{S_3} \;=\; P^1 \times L^{-1/2} \times g^{1/2} \;=\; P/\sqrt{L/g}, \;\; = constant.$$

From experimental or theoretical analysis (Fig. 4) we know that this $constant \;=\; 2\pi$ provided that the amplitude of the motion is not relevant (and not included in the VPlist (32)).

3.2 Computing the null space

These are the kind of solutions we are looking for. However, we need *all* of the possible solutions and for systems that may be much larger (more variables). These complete solutions will come from solving

$$\mathbb{D}\,\mathbb{S} \;=\; 0, \tag{35}$$

where $\mathbb{D}$ is the dimension matrix, and $\mathbb{S}$ is the matrix of solution vectors that constitute the null space of $\mathbb{D}$. This is the matrix equivalent of Eqs. (31) - (33) (Problem 1 below).

$\mathbb{D}$ may be read directly from the VPlist. For the VPlist of the simple pendulum, Eq. (15), the dimension matrix is

$$\mathbb{D} \;=\; \begin{matrix} \\ m \\ l \\ t \end{matrix} \begin{matrix} \;\;\phi \;\; t \;\; M \;\; L \;\;\; g \;\;\; \phi_0 \\ \begin{bmatrix} 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 & -2 & 0 \end{bmatrix} \end{matrix}. \tag{36}$$

The columns of $\mathbb{D}$ correspond to the dimensional variables. The order of including the dimensional variables is noteworthy only insofar as the algorithm will seek to make the first few dimensional variables appear in the nondimensional variables with an exponent = 1. Hence it is helpful to have the dependent variable ϕ come first, the independent variable t come next, and after that there is no obvious, preferred ordering. The rows of $\mathbb{D}$ correspond to the physical dimension of those variables: row 1 is mass, row 2 is length and row 3 is time. This order is arbitrary, but has to be maintained once chosen.[19] Thus for the acceleration of gravity, which is the fifth dimensional variable in the VPlist, $g \;\doteq\; [0\;1\;-2]$ and so the 5th column of $\mathbb{D}$ is $D_{1\,5} \;=\; 0, \;\; D_{2\,5} \;=\; 1$, and $D_{3\,5} = -2$.

The matrix $\mathbb{S}$ contains the unknown exponents. It is convenient to think of the columns of $\mathbb{S}$ as

[19] Within the main text the physical dimensions on a variable are written as a row vector to save space. These data are the column vectors of the matrix $\mathbb{D}$.

vectors, $\mathbf{S}$,

$$\mathbf{S_1} = \begin{bmatrix} S_{11} \\ S_{21} \\ S_{31} \\ S_{41} \\ S_{51} \\ S_{61} \end{bmatrix}, \quad \mathbf{S_2} = \begin{bmatrix} S_{12} \\ S_{22} \\ S_{32} \\ S_{42} \\ S_{52} \\ S_{62} \end{bmatrix} \text{ and } \ldots\ldots \quad \mathbf{S_K} = \begin{bmatrix} S_{1K} \\ S_{2K} \\ S_{3K} \\ S_{4K} \\ S_{5K} \\ S_{6K} \end{bmatrix},$$

to which nondimensional variables will correspond one-to-one.

The system of equations (35) is underdetermined in the usual case that there are more unknown exponents (the number of dimensional variables) than there are independent equations (usually the number of physical dimensions and also the number of rows of $\mathbb{D}$). However, it can be less (see Problem 4 of Sec. 3.5). The number of columns in $\mathbb{S}$, and thus the number of vectors, $\mathbf{S}$, is K, and is found as a part of the calculation of the null space. To represent or span the null space requires a basis set of K vectors, from which any solution vector can be constructed (more on this below). This K has the same important interpretation as the K of Sec. 2.3.1; *viz.*, K is the number of nondimensional variables in a basis set.

The calculation of a null space basis proceeds along the lines of Eqs. (31) to (34); not difficult, but tedious if done by hand. The calculation of a null space basis is readily automated, as will be assumed from here on. Links to relevant scripts are in Sec. 3.4.

3.2.1 A basis set of nondimensional variables of the angular displacement

The null space basis for the time-dependent and finite amplitude angular displacement defined by the $\mathbb{D}$ of Eq. (36) has $K = 3$ solution vectors,

$$\mathbf{S_1} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \quad \mathbf{S_2} = \begin{bmatrix} 0 \\ 1 \\ 0 \\ -1/2 \\ 1/2 \\ 0 \end{bmatrix}, \text{ and } \quad \mathbf{S_3} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}.$$

The elements of $\mathbf{S}$ are required to be rational numbers. Here and typically, they are small rational numbers.

The corresponding basis set of K nondimensional variables is constructed from the solution vectors as

$$\Pi_1 = \mathbf{X}^{\mathbf{S_1}} = \phi^1\, t^0\, M^0\, L^0\, g^0\, \phi_0^0, = \phi \tag{37}$$

$$\Pi_2 = \mathbf{X}^{\mathbf{S_2}} = \phi^0\, t^1\, M^0\, L^{-1/2}\, g^{1/2}\, \phi_0^0 = t/\sqrt{L/g}, \tag{38}$$

$$\Pi_3 = \mathbf{X}^{\mathbf{S_3}} = \phi^0\, t^0\, M^0\, L^0\, g^0\, \phi_0^1 = \phi_0, \tag{39}$$

where notice that the elements of the $\mathbf{S}$ vectors are exponents on the elements of the $\mathbf{X}$ vectors. The dependent variable ϕ appears in the first nondimensional variable only, and to the first power, a convenient property built in to the algorithm.

The functional relationship among the nondimensional Πs may be written as

$$\Pi_1 = F(\Pi_2, \Pi_3),$$

or, once again,

$$\phi(t) = F(\frac{t}{\sqrt{L/g}}, \phi_0)\,. \tag{40}$$

Since the maximum of $\phi(t)$ will be equal to ϕ_o, it makes sense to normalize ϕ as

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

This amounts to dividing Π_1 by Π_3, which is allowed by properties of the null space (discussed in Sec. 3.3). But we can not then omit ϕ_o from the righthand side of (40) as that would reduce the basis set of nondimensional variables.

The mass M has an exponent of zero in all of the solution vectors, Eqs. (37) to (39), consistent with the informal analysis of Sec. 2 which showed that there was no way to construct a nondimensional variable from a single parameter having physical dimensions of mass. To say it a little differently, for the mass of the bob to be retained, there would have to be another parameter with dimension mass in the VPlist (as there will be in the more comprehensive models of Sec. 4). Notice that the angles ϕ and ϕ_0 sailed into the basis set of nondimensional variables untouched, since they were already nondimensional.

3.2.2 A basis set for tension in the line

Tension can be analyzed in the same manner, starting with the dimensional matrix read from the VPlist, (25),

$$\mathbb{D} = \begin{matrix} \\ m \\ l \\ t \end{matrix} \begin{matrix} T & t & M & L & g & \phi_0 \\ \end{matrix} \left[\begin{matrix} 1 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 1 & 0 \\ -2 & 1 & 0 & 0 & -2 & 0 \end{matrix} \right]. \tag{41}$$

The null space basis vectors are

$$\mathbf{S_1} = \begin{bmatrix} 1 \\ 0 \\ -1 \\ 0 \\ -1 \\ 0 \end{bmatrix}, \quad \mathbf{S_2} = \begin{bmatrix} 0 \\ 1 \\ 0 \\ -1/2 \\ 1/2 \\ 0 \end{bmatrix}, \quad \text{and} \quad \mathbf{S_3} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}.$$

A basis set of nondimensional variables is thus

$$\Pi_1 = \mathbf{X}^{\mathbf{S_1}} = T^1\, t^0\, M^{-1}\, L^0\, g^{-1}\, \phi_0^0 = T/Mg \tag{42}$$

$$\Pi_2 = \mathbf{X}^{\mathbf{S_2}} = T^0\, t^1\, M^0\, L^{-1/2}\, g^{1/2}\, \phi_0^0 = t/\sqrt{L/g}, \tag{43}$$

$$\Pi_3 = \mathbf{X}^{\mathbf{S_3}} = T^0\, t^0\, M^0\, L^0\, g^0\, \phi_0^1 = \phi_0. \tag{44}$$

The second and third of these are identical to Π_2 and Π_3 of the displacement noted above. The functional relationship for the time-dependent tension can then be written

$$\frac{T}{Mg} = F(\frac{t}{\sqrt{L/g}}, \phi_0). \tag{45}$$

Notice that the dependent variable T has been scaled by the weight of the bob, Mg, a fairly obvious natural scale in this problem.

A time-independent tension, say the minimum or the maximum during a complete oscillation, can be analyzed by omitting the time, t, to find

$$\frac{T_{min}}{Mg} = F(\phi_0), \quad \text{and} \quad \frac{T_{max}}{Mg} = F(\phi_0), \tag{46}$$

the forms used in Fig. (4), lower. Notice that the line length, L, has disappeared going from (45) to (46). It is a little jarring to see $F(\phi_o)$ on the right side of two different equations in (46), but keep in mind that this F is merely holding a place for a similarity function that has to be determined by some method beyond dimensional analysis.

3.3 Properties of a null space basis

It is important to understand the properties of a null space basis, and especially that a given null space basis is generally *not* a unique solution to the underdetermined problem of the sort Eq. (35) encountered in dimensional analysis. There are usually many possible solutions, and just as many null space bases. In fact, the specific null space basis that comes from a solution algorithm will depend upon the order in which the dimensional variables are listed in the dimensional matrix. The construction of the best (most useful, most insightful) basis set of nondimensional variables is facilitated by the following three properties of a null space basis.

> *Property 1, P1: The number of solution vectors in a null space basis is $K = I - R$, where I is the number of dimensional variables and R is the rank of the dimensional matrix, $\mathbb{D}$.*

It is highly desirable to have the smallest possible K. However, the rank, R, is determined strictly by the dimension matrix $\mathbb{D}$ and thus by the VPlist; there is nothing that can be done to change K, except to amend the VPlist.[20]

> *P2: Nondimensional variables correspond one-to-one with solution vectors of the null space, and are K in number. These nondimensional variables are themselves a basis set for the problem defined by the VPlist.*

The K is the same for all basis sets and in that important regard, all basis sets are equally efficient, i.e., they all achieve the same reduction in the number of variables. Nevertheless, one particular basis set may be more useful than the others. A transformation from one basis set to another is readily accomplished because

> *P3: The vectors of a null space basis are linearly independent and they span the null space. Any vector that is a solution of the homogeneous system can thus be formed as a linear combination of the vectors that make up any basis set.*

For example, if $\mathbf{S_1}$ and $\mathbf{S_2}$ are vectors in 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 also a vector in the null space, and hence $\mathbf{S_3}$ is also a solution. Let the nondimensional variable corresponding to this new vector be Π_3. If Π_3 and Π_1 are preferred over the initial Π_1 and Π_2, then a revised basis set can be taken as Π_1 and Π_3, while omitting Π_2. The revised basis set has the same number of vectors and nondimensional variables as the initial basis set, and will also span the null space.

[20] The rank, R, of the dimension matrix is computed and reported as part of the automated calculation. It is useful to know how R may change. In brief, the rank of a matrix is the number of linearly independent rows or columns (the same for either). Suppose that a VPlist and thus the corresponding $\mathbb{D}$ includes a time scale, ω_1, and now add another time scale, ω_2. The number of dimensional variables increases by 1, but the rank of $\mathbb{D}$ does not increase because the new column associated with ω_2 is identical to the column that represents ω_1 and already present. The number K of nondimensional variables in the null space basis thus goes up by 1 with the inclusion of ω_2 in the VPlist. A special case considered in Problem 4) below is $\mathbb{D} = 0$, the zero matrix, which has rank of zero. For much more on the rank of a matrix and on the null space generally, see https://ocw.mit.edu/courses/18-06-linear-algebra-spring-2010/ and G. Strang, 1998, *Introduction to Linear Algebra*, Wellesley-Cambridge Press, Wellesley, MA.

Given these three properties of a null space basis, the initial basis set of nondimensional variables that comes directly from a calculation may be transformed to some other preferred basis set by multiplying or dividing the Πs in any order. Defining a preferred basis set of nondimensional variables by this kind of transformation is the third step of a dimensional analysis that will be emphasized in Secs. 4, 6 and 7.

3.4 Automated calculation

The calculation of a null space basis is best done by machine, say with Matlab

https://www2.whoi.edu/staff/jprice/wp-content/uploads/sites/199/2024/06/DanalysisA2.zip

or Python,

https://www2.whoi.edu/staff/jprice/wp-content/uploads/sites/199/2024/10/DA2Python.zip

To use the Matlab, unpack a script, DanalysisA2.m, and three functions, Dclear.m, Din.m and Dnullspace.m, and enter the text below in red into a command window. The blue text is returned by Matlab.

DanalysisA2 starts a script that will run several cases automatically. You will be asked to reply to two questions that may be defaulted with n. To define a new problem, either edit this script, or, call the functions that make up the algorithm.

To use the functions, first declare the following variables to be global:

global D VPname nVP S Pi, where

D is the dimension matrix,
VPname is the name of a variable in the VPlist (character data),
nVP is the number of variables and parameters in the VPlist,
S is the null space vectors that we seek, and
Pi is a symbolic math representation of the nondimensional variables.

Dclear should be called at the start of a problem to reset these global variables.

Then enter the data that define the VPlist by a sequence of calls to Din('text', [mass length time]), once for each variable in the VPlist. The 'text' is the name of the variable, and the row vector defines the physical dimensions of 'text' as [mass length time]. Enter the single dependent variable first, then the independent variables, then the parameters, e.g., for the VPlist Eq. (15),

Din('phi', [0 0 0]) angular displacement, the single dependent variable
Din('t', [0 0 1]) time, an independent variable

Din('M', [1 0 0])	mass of the bob
Din('L', [0 1 0])	length of the line
Din('g', [0 1 -2])	acceleration of gravity
Din('phio', [0 0 0])	the initial angle.

That completes the VPlist for the oscillations of a simple pendulum.

And now the calculation,

Dnullspace

The number of dimensional variables, I = 6
The rank of the dimension matrix, D, is R = 3
The number of non-d variables is K = I - R = 3
The dimensional variables and solution vectors (exponents) that
make one possible basis set of non-d variables are:

phi	t	M	L	g	phio
1	0	0	0	0	0
0	1	0	-0.5	0.5	0
0	0	0	0	0	1

A basis set of non-dimensional variables is thus:

(phi, t * sqrt(g) / sqrt(L), phio)

This is consistent with Eq. (40), though written in an implicit form.[11]

3.5 Null space problems

1) Use the null space basis method to analyze the period of a simple pendulum given the VPlist (32). How many nondimensional variables are there? What does this imply for the similarity function, or is it a similarity constant? Given that there is only solution vector, verify by direct substitution that the matrix relation Eq. (35) is equivalent to Eqs. (31) - (33). What if an angle, say ϕ_o, is included in the analysis above? To add a new variable to an existing VPlist you need only call Din('phio', [0 0 0]), and then call Dnullspace to see the new result. What happens if a second length, L_2 is also included?

2) Suppose that you scramble the order of the VPlist of Eq. (15)? How about changing the order of the physical dimensions, i.e., in place of [mass length time] used here, how about [length mass time]?

3) In most problems the number of nondimensional variables, K, can be estimated by the rule of thumb noted in Sec. 2.3.1, K = the number of variables in the VPlist minus the number of fundamental units. But that doesn't always work, as the following example, which is contrived but not impossible, shows. Suppose that a VPlist consists of four velocities. The simple rule of thumb: there are two fundamental units, length and time, and so the expectation is $K = 4 - 2 = 2$ nondimensional variables. However, an application of the null space method indicates 3 nondimensional variables because the dimension matrix has rank $R = 1$ (the minimum, aside from the special case of a zero matrix; discussed below) since the variables all have the same dimensionality. Some questions for you: What are the nondimensional variables in this case, and which estimate of their number is correct? Suppose a fifth variable having dimension length, $L \doteq [\,0\;1\;0\,]$ is added into the VPlist; what changes? Suppose a sixth variable having dimension time is added?

4) A good check on a model or algorithm is that it behaves appropriately at the limits of its domain. In this context, an interesting limit is a case in which all of the variables of a VPlist are nondimensional, as if the output of a null space calculation was used as the input of a second calculation. The corresponding dimension matrix is the zero matrix, $\mathbb{D} = 0$, i.e., all entries are zero, and which is the only matrix that has a rank $R = 0$.[20] What do you expect from a null space calculation? Form a hypothesis, and then run such a calculation beginning with a VPlist comprising three angles, say.

5) The dimensional analysis of tension in the line, Sec. 2.4.2, and Eq. (46), concluded that the time-independent tension, i.e., the minimum or maximum over an oscillation, is independent of the line length, L. This was suggested by a handful of numerical experiments seen in Fig. (1), and dimensional analysis assures that this holds rigorously for the VPlist that describes the most important properties of a simple pendulum. Something for you to think about: does dimensional analysis provide a satisfactory *explanation* of this (small and singular) fact?

4 Damping a simple pendulum

The motion of a free-running pendulum will decay in time as energy is dissipated into the surrounding air.[21] This section will analyze the mechanisms that cause this damping.

4.1 Observations of the decaying amplitude

The amplitude of the oscillation, $\Phi(t)$, was observed by measuring the maximum of the cord length at intervals of 30 seconds to 2 minutes. To minimize the measurement noise associated with the rather coarse least count on the measured cord, 10^{-3} m, it was found helpful to study a longer pendulum, $L = 3.70$ m, than used before. The line was smooth monofilament having a diameter $D_l = 0.40 \times 10^{-3}$ m, and the bob was a nearly spherical, smooth lead fishing sinker with a diameter $D_b = 0.0211$ m and mass $M = 0.055$ kg. The line was supported on a needle bearing (a fishhook with the point impinging on a hard, metal surface) to minimize interactions between the swinging pendulum and its support.

The observed amplitude, $\Phi(t)$, (the red crosses of Fig. (6), upper) was found to be highly repeatable, and can be characterized roughly as an e-folding in about 10 minutes. During this time the pendulum made hundreds of oscillations, and hence the damping of this pendulum is weak in comparison to gravity; more on this below.

The decay rate of the oscillation amplitude is defined by

$$\Gamma = \frac{1}{\Phi}\frac{d\Phi}{dt}, \tag{47}$$

where Φ is the amplitude of the angular displacement (and $\phi(t)$ is the time-dependent, instantaneous angle). This Γ is an inverse time. It is expected that Γ will itself be amplitude-dependent and thus time-dependent.

4.2 A preliminary VPlist for the decay rate of a simple pendulum

The primary damping process is presumed to be aerodynamic drag between the swinging pendulum and the surrounding, viscous air.[22] If that is the case, then the diameter of the bob, D_b, and of the

[21] A pendulum-based clock requires an escapement mechanism to couple the pendulum into the clock mechanism, and, to do a very small positive work on each swing of the pendulum to offset frictional losses. One of the first successful mechanisms for a pendulum clock was designed by Galileo Galilei, late in life and after he had lost his sight. Pendulum-based clocks were later developed into the most precise means of time-keeping available for almost three centuries. https://en.wikipedia.org/wiki/Galileo_escapement

[22] An excellent introduction to viscosity in fluid flow is by http://galileo.phys.virginia.edu/classes/152.mf1i.spring02/ and see the chapter on Fluids. The pioneering work on this topic (and much else) is the remarkable study by Stokes, G. G.,

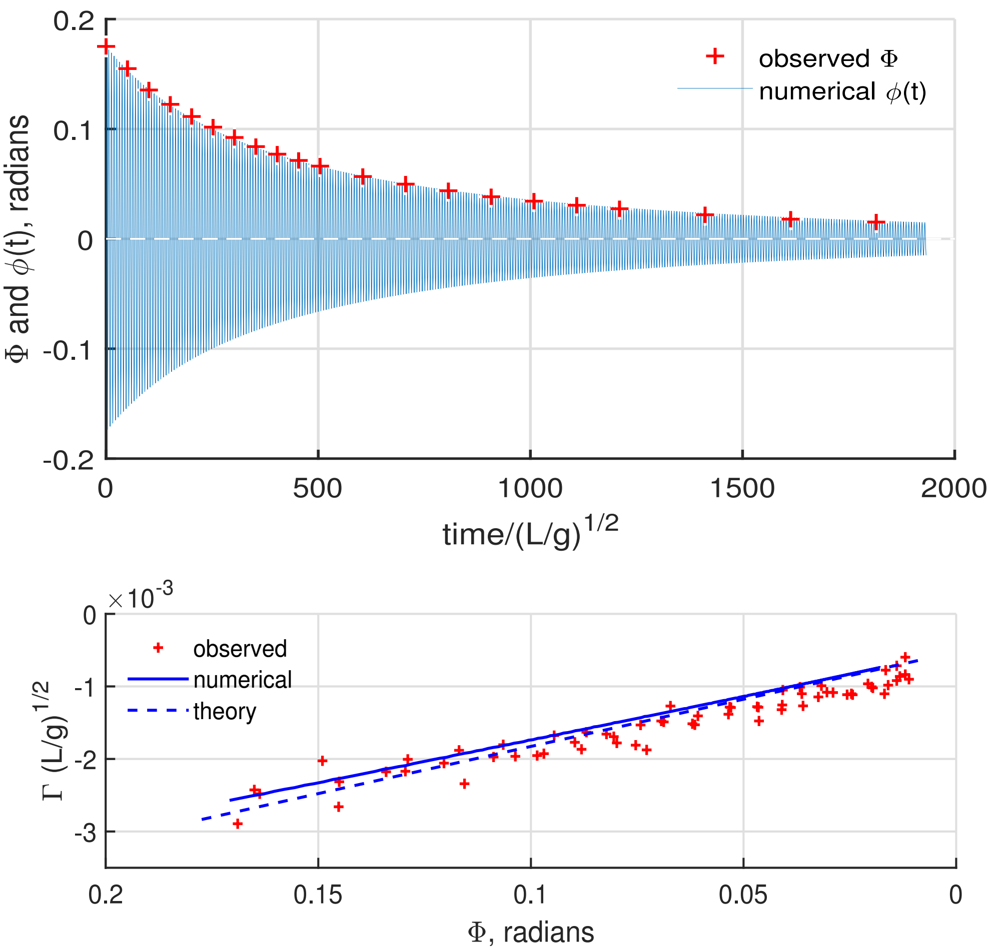


Figure 6: **(upper)** Observations of the amplitude, Φ, of a simple pendulum at intervals of 30 seconds to 2 minutes (the red crosses). Also shown here is a numerical solution for $\phi(t)$ (the thin, solid blue line), described in Sec. 4.4.1. **(lower)** The decay rate computed directly from Eq. (47) with data from three repetitions of the experiment (the red crosses), and from a numerical solution (blue, solid line). An approximate analytic solution, Eq. (70), is the blue, dashed line. Drag that is linear in the angular velocity produces exponential decay of the amplitude in time and thus a constant nondimensional decay rate. Drag that is quadratic in the angular velocity produces a nondimensional decay rate that decreases with decreasing amplitude, Φ. Most of what can be seen here is the latter.

line, D_l, would seem to be relevant, along with the density and kinematic viscosity of air, ρ and ν.

- A preliminary VPlist for the decay rate of a simple pendulum with viscosity: (48)

1. the decay rate, $\Gamma \doteq [0\ \ 0\ \ -1]$, the dependent variable;
2. mass of the bob, $M \doteq [1\ \ 0\ \ 0]$, a parameter,
3. length of the line, $L \doteq [0\ \ 1\ \ 0]$, a parameter,
4. acceleration of gravity, $g \doteq [0\ \ 1\ \ -2]$, a parameter,
5. the amplitude of the motion, $\phi_o \doteq [0\ \ 0\ \ 0]$, a parameter,
6. diameter of the line, $D_l \doteq [0\ \ 1\ \ 0]$, a parameter,
7. diameter of the bob, $D_b \doteq [0\ \ 1\ \ 0]$, a parameter,
8. density of air, $\rho \doteq [1\ \ -3\ \ 0]$, a parameter (1.2 kg m^{-3}, nominal),
9. kinematic viscosity of air, $\nu \doteq [0\ \ 2\ \ -1]$, a parameter ($1.5\times10^{-5}$ m^2 s^{-1}, nominal).

Dimensional analysis via the method of Sec. 3.2 (from here on omitting all of the intermediate steps) indicates a basis set of $K = 9 - 3 = 6$ nondimensional variables,

$$\Gamma\sqrt{L/g} = F(\phi_o, \frac{D_b}{L}, \frac{D_l}{L}, \frac{\rho D_b^3}{M}, \frac{g^{1/2}L^{3/2}}{\nu}). \tag{49}$$

The dependent variable, Γ is scaled with the fast, oscillatory time scale, $\propto \sqrt{L/g}$, the natural time scale of a simple pendulum, and is O(10^{-3}) for the nearly conservative pendulum observed here, Fig. (6). This is the sense in which this pendulum is weakly damped. The first nondimensional parameter, ϕ_o, is familiar as the amplitude of the motion. The next two nondimensional parameters have straightforward geometric interpretations, the third parameter is proportional to the buoyancy force on the bob (not significant since the density of lead is much greater than the density of air). The last nondimensional parameter involving the viscosity can be recast as a Reynolds number, which has an important physical interpretation discussed below. While the nondimensional variables and parameters are interpretable individually, we are not ready to make use of such a comprehensive model.[23] We may still be thinking of the nearly conservative pendulum of Sec. 2.1, but this nine-variable VPlist includes all simple pendulums and fluid mediums, including bobs that float and fluid that is sufficiently viscous that there is no oscillation. Before we can expect a useful result from dimensional analysis we will have to first identify the most relevant parameters for the kind of nearly conservative (weakly damped) pendulum described in Sec. 4.1.

1851, 'On the effect of the internal friction of fluids on the motion of pendulums'. Trans. of the Cambridge Philosophical Soc., Ninth Vol., No. 10, pp. 8 - 106. Bibcode: 1851TCaPS...9....8S More recent is P. T. Squire, 1986, 'Pendulum damping', Am. J. Phys. **54**, 984–991, and R. A. Nelson, and M. G. Olsson, 1985, 'The pendulum: Rich physics from a simple system', Am. J. Phys. **54**, 112–121.

[23] This looks like a big problem compared to those encountered up to here. However, it is small compared to the problems that arise in some real-world applications, e.g., food processing.[2]

4.3 Aerodynamic drag on a moving sphere and cylinder

A piecewise approach is tried next. Consider in isolation the aerodynamic drag on a smooth sphere (the bob) that is in steady motion through a viscous fluid (air) that is at rest. Despite the idealized configuration of this problem, it is challenging to compute the drag from first principles in the common case that the flow around the sphere is turbulent. However, dimensional analysis combined with laboratory measurement leads to a useful result.

- A VPlist for drag on a sphere moving steadily through viscous fluid: (50)

1. drag (a force), $H \doteq [\,1 \quad 1 \quad \text{-}2\,]$, the dependent variable,
2. speed of the sphere, $U \doteq [\,0 \quad 1 \; \text{-}1\,]$, a parameter,
3. diameter of the sphere, $D_b \doteq [\,0 \quad 1 \quad 0\,]$, a parameter,
4. density of the fluid, $\rho \doteq [\,1 \quad \text{-}3 \quad 0\,]$, a parameter,
5. kinematic viscosity of the fluid, $\nu \doteq [\,0 \quad 2 \; \text{-}1\,]$, a parameter.

There are five variables and parameters having three physical dimensions and hence there are two nondimensional variables in the preliminary basis set,

$$\Pi_1 = \frac{H}{\rho\, D_b^2\, U^2} \quad \text{and} \quad \Pi_2 = \frac{\nu}{U\, D_b}. \tag{51}$$

This Π_2 is the inverse of a very widely occurring nondimensional variable called the Reynolds number,

$$Re = \frac{U\, D_b}{\nu}, \tag{52}$$

which is preferred. From P2 of the null space (Sec. 3.3) we know that the Π_1 and Π_2 of Eq. (51) are not uniquely determined by dimensional analysis, and that we are free to reform the initial basis set by adding or subtracting the solution vectors to one another in any order, which amounts to multiplying or dividing the Πs in any order. The first change will be to simply invert the initial Π_2 of Eq. (51), and use instead the conventional Reynolds number,

$$\Pi_2 = \frac{U\, D_b}{\nu}.$$

Thus to here we have the slightly revised basis set,

$$\Pi_1 = \frac{H}{\rho\, D_b^2\, U^2}, \quad \Pi_2 = R_e = \frac{U\, D_b}{\nu}, \tag{53}$$

and so

$$\frac{H}{\rho\, D_b^2\, U^2} = C(Re), \tag{54}$$

which is close to a useful result. Notice that in place of the F used to indicate an unknown similarity function in Secs. 2 and 3, this Eq. (54) and the rest of this section will use instead C, which is said to be a 'drag coefficient'.

The drag coefficient on an immersed object is generally defined to be

$$C_d = \frac{H}{\frac{1}{2}\rho A U^2}, \tag{55}$$

with A the frontal area, $A = \pi D^2/4$ for a sphere. The factor $1/2$ in the denominator of (55) is almost always included. With these numerical factors acknowledged, the basis set (54) becomes[24]

$$\Pi_1 = \frac{H}{(\pi/8)\rho D_b^2 U^2} \quad \text{and} \quad \Pi_2 = Re = \frac{U D_b}{\nu}. \tag{56}$$

Before accepting (56) as final, we should consider other possibilities, and specifically that Π_1 may also be multiplied by some power of the Reynolds number, Π_2. A somewhat general basis set is then

$$\frac{H}{(\pi/8)\rho D_b^2 U^2} R_e{}^n = C_n(Re), \tag{57}$$

where n is any real number. This assumes that H and $\Pi_2 = Re$ may as well remain to the first power. The subscript n on C_n is to acknowledge that the drag coefficient (the similarity function in this problem), will be dependent upon n. In this case, the criterion for deciding the best n will be the n that yields the least variable drag coefficient, $C_n(Re)$.

Regardless of the n finally chosen, an essential result from Eq. (57) is that the nondimensional drag, the left side of (57), is expected to be a function of Re alone. Laboratory measurements of drag, speed, etc., can thus be used to define a similarity function $C_n(Re)$ that should hold for all steadily moving spheres or cylinders, just the way that the function $F(\phi_0)$ (Sec. 2.4) sufficed to define the period for all inviscid, simple pendulums. What is most important is that drag coefficients estimated from experiments conducted over a very wide range of speeds and through different Newtonian viscous fluids do indeed collapse to a well-defined, though not necessarily simple function of the Reynolds number alone, just as dimensional analysis had lead us to expect. Experimental estimates for spheres are in Fig. (7), left, and the result for cylinders is Fig. (7), right.

This is a result, characteristic of dimensional analysis, that is at once profound and trivial. One might say trivial because, after all, dimensional analysis told us that the drag coefficient

[24] The inclusion of these two numerical factors doesn't change anything fundamental about the drag coefficient (54). However, if we want to make contact with the huge body of historical studies on aerodynamic drag, then we should account for them. The factor 1/2 that multiplies ρU^2 can be viewed as the dynamic pressure of the Bernoulli equation, $P_d = \frac{1}{2}\rho U^2$, where U is the free stream velocity brought to zero on the surface of an object that is immersed in the flow. Dynamic pressure is the immediate mechanism of inertial drag, and can be measured directly.

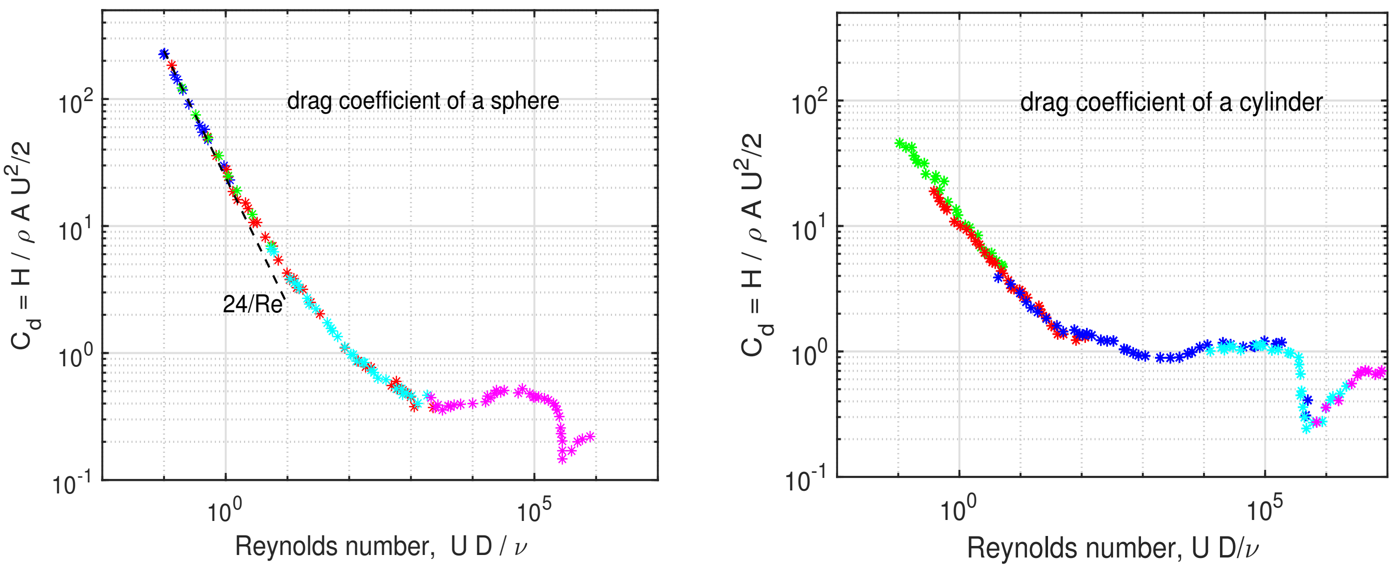


Figure 7: Experimental estimates of the conventional drag coefficient, $C_d = H/\frac{1}{2}\rho AU^2$, of a smooth sphere **(left)** and a cylinder **(right)** moving steadily through a viscous, Newtonian fluid, air, water and various oils. This is a representative subset of experimental data extracted from Rouse (1946) Sec. 4.[5]

should depend upon Re alone. From that perspective, an effective collapse of the experimental data to a single curve $C(Re)$ merely verifies that carefully controlled laboratory conditions can approximate the idealized VPlist (50). But it is also profound in that dimensional analysis has shown the way to a portable, useful result, the $C(Re)$ of Fig. (8), where there would otherwise have been be an unwieldy mass of highly specific data, as in going from Fig. (1), upper to lower.

4.3.1 A zero order solution as a scale for the dependent variable

The most important (and in this case the only) choice is how to nondimensionalize, or scale, the dependent variable, H, which then defines Π_1. The null space basis method has framed the issue in a convenient way, but the next step requires that something be added. One strategy is to choose a scale that reflects a physically meaningful 'zero order' solution for the dependent variable. A zero order solution is a highly simplified, approximate solution that has a physically plausible behavior in an appropriate limit (examples coming just below). If such a zero order solution can be found, then the corresponding nondimensional variable, Π_1, should have a magnitude of O(1) in that limit.

Constructing a zero order solution requires some sense of the physics of the problem. Visual observations of the flow around a steadily moving sphere suggest that drag can arise from two distinct processes.

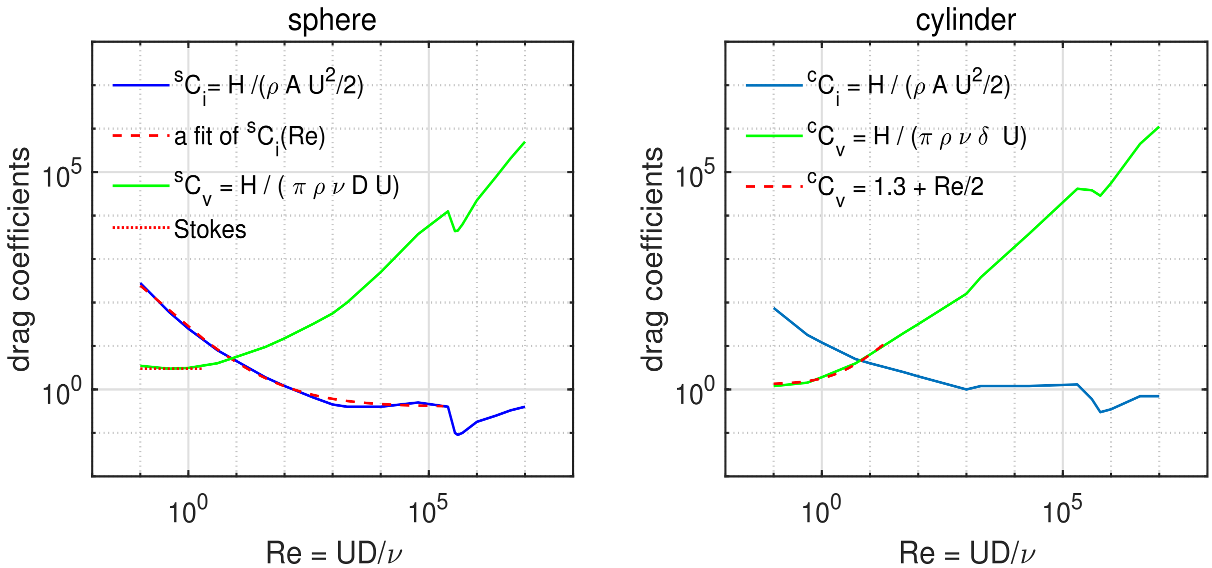


Figure 8: Empirical drag coefficients for a sphere, **(left)**, and a cylinder, **(right)**, moving at a steady speed U through a Newtonian viscous fluid. Two forms of a drag coefficient are shown here, the inertial drag coefficient denoted by C_i (the blue lines), and the so-called viscous drag coefficient denoted by C_v (the green lines). The inertial drag coefficients were read from Munson et al. (1998)[5], Fig. (7.7) and Rouse (1946),[5] Figs. (125) and (126) and the viscous drag coefficients were then computed by Eq. (57) with $n = 1$. Note that ^sC_v and ^cC_v are O(1) if Re is small, $Re \leq 1$, and that ^sC_i and ^cC_i are O(1) if the Reynolds number is much larger, $Re \geq 1000$. The dashed red lines are approximate fits to the drag coefficients in the Re range appropriate to the bob, Eq. (62), at left, or the line, Eq. (60), at right. The dotted red line at left and very small Re is Stokes' solution for creeping flow discussed in Sec. 4.5, Problem 1.

Viscous drag. If the sphere is moving very slowly so that the flow around the sphere is laminar and symmetric fore and aft with little or no wake, then the drag will be mainly viscous stress acting on the surface of the sphere. In the common case of a Newtonian fluid, the viscous, surface drag is proportional to the dynamic viscosity of the fluid, $\rho\,\nu$, times the shear of the flow around the sphere, $\rho\nu U/D_b$. Assuming that this viscous stress acts over a surface area proportional to D_b^2, then a zero order solution for drag on the sphere is $H \propto \rho\nu D_b U$. If this viscous, surface drag is the dominant drag-producing process, then it would be appropriate to nondimensionalize (scale) the aerodynamic drag as

$$\underline{v}\text{iscous drag coefficient for a } \underline{s}\text{phere:} \quad \frac{H}{\rho\nu D_b U} = {}^sC_v(Re), \tag{58}$$

on the expectation that the Re-dependence of ^sC_v would then be minimized compared to other choices. This is the $n = 1$ version of Eq. (57). This form of a drag coefficient will be used to estimate drag on the line of the pendulum in Sec. 4.4.

Inertial drag. Even if the fluid was nearly inviscid,[25] there would still be drag on a moving sphere because the fluid that is displaced by a moving object must be accelerated up to a speed $\propto U$. If the displaced fluid forms a turbulent wake, as is observed behind a rapidly moving sphere, then the drag would be expected to be roughly proportional to the rate of change of the momentum of the displaced fluid, $\propto \rho U$, times the rate of change of the volume of the wake, frontal area $A = \pi D_b^2/4$ times the speed, U. Thus the drag would be estimated as $H \propto \rho A U^2$. The drag coefficient that corresponds to this inertial drag process, often called 'form drag', is

$$\underline{i}\text{nertial drag coefficient for a }\underline{s}\text{phere:} \quad \frac{H}{\frac{1}{2}\rho A U^2} = {}^sC_i(Re), \tag{59}$$

which is the $n = 0$ version of Eq. (57). This is the form of the drag coefficient that is almost always encountered, as in Fig. (7), and is usually written C_d and said to be the 'drag coefficient' without mentioning inertial. The Re-dependence of ${}^sC_i$ shows the departures from inertial drag due to viscous effects, presumably. Either form of the drag coefficient effectively conveys the laboratory data and in that respect they are equally useful (consistent with P1 and P2 of a basis set, Sec. 3.2). However, if the range of the relevant Re is restricted, then it would be appropriate to use the form of the drag coefficient – inertial or viscous, $n = 0$ or $n = 1$ – that has the least variation over that subrange of Re. In this case, the inertial drag coefficient for a sphere will be used to estimate the aerodynamic drag on the bob.

4.3.2 Nondimensional parameters; the Reynolds number

Once the form of the dependent nondimensional variable, Π_1, has been set, the remaining nondimensional variables can be formed in ways that define the geometry of the problem, or that reflect a balance of terms in a governing equation, or that follow the established norms of your field. This is necessarily vague because the possibilities are almost limitless. However the task is eased by the null space basis method which ensures that you start with a complete basis set of nondimensional variables. All that you have to do (!) is rearrange the basis set to make the most useful basis set for your purpose.

In the example of aerodynamic drag on a moving sphere, there is only one remaining nondimensional variable, the Reynolds number, $Re = \frac{UD}{\nu}$ (or any power of the Reynolds number), where U is the speed of the current, usually in a free-stream or undisturbed region, D is the spatial scale over which U changes by O(1), 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 often in problems of fluid mechanics.

Recall that for the purpose of modeling drag, a slowly moving sphere has a nearly

[25] If the fluid is a superfluid having literally zero viscosity, then the viscous drag will vanish. Such superfluids are in effect a new state of matter for which quantum physics supplants classical physics.

undisturbed, laminar wake. Observational evidence shows that laminar flow occurs when Re is small, $Re \leq 1$, regardless of speed *per se*; a small diameter or large viscosity are equally important to Re. Dimensional analysis tells us as much in that the drag coefficient of a given class of object (e.g., spheres or cylinders) depends only upon Re. The small Re range is that of a small bug swimming through water, or the thin line of a pendulum swinging slowly through air. In the small Re range, the viscous drag coefficient ${}^sC_v$ is O(1), both numerically and in the sense that ${}^sC_v$ is nearly independent of Re for $Re \leq 1$ (Fig. 8).

For creatures and objects anywhere near our size, e.g., golf balls, bikers or automobiles, Reynolds numbers of $O(10^4)$ and greater are the norm, and inertial drag (form drag) is generally more important than is viscous drag. Notice that for moderately large Re, $Re \geq 10^3$, the inertial drag coefficient C_i is O(1) in magnitude and very roughly constant within subranges of Re (Fig. 8).[26]

The pendulum studied here has Reynolds numbers for the line and the bob that fall in an intermediate range in which both viscous and inertial drag are likely to be important, discussed further in the next section, 4.4.

4.3.3 Three-dimensional flow effects upon drag

Before leaving the subject of aerodynamic drag, it should be noted explicitly that the drag on an immersed object depends very sensitively upon the three-dimensional flow from front to back around the object, and not just the frontal area, which is the only geometric property accounted for in the drag coefficients (58) and (59). For example, a flat disk of a given area that is face-on to a flow will have a drag coefficient of about 2 at large Re. A tear-drop-shaped, streamlined object having the same frontal area may have a drag coefficient as low as 0.04, or a factor 50 less at the same Re.[27] As a consequence of this sensitive dependence of drag upon the three-dimensional flow around an object, each distinct shape will require its own drag coefficient denoted by a superscript, here a sphere or a cylinder, ${}^sC$ or ${}^cC$ in Fig. (8), left and right. (For a more ambitious approach to this, see Problem 6 in Sec. 4.5 below.)

[26] However, even at a very large Re it does not follow that viscosity is entirely irrelevant. Significant changes in the drag coefficient of a sphere or a cylinder occur in the vicinity $Re \approx 4 \times 10^5$ due to changes in the boundary layer flow around the object and the width of the turbulent wake (sometimes referred to as the 'drag crisis'). This is the Re range of a well-hit golf ball or tennis ball, and is consistent with the observation that aerodynamic drag on these objects has a surprising sensitivity to small scale surface roughness and to spin. For more on these phenomena see S. Vogel, 1994, *Life in Moving Fluids*, Princeton Univ. Press, and P. Timmerman, and J. P van der Weele, 1999, 'On the rise and fall of a ball with linear and quadratic drag', *Am. J. Phys.*, **67**, 538–546.

[27] An especially clear illustration of the flow around such objects is Plate XVIII of Rouse (1946)[5]

4.4 Evaluating models of a damped pendulum

4.4.1 A numerical solution

To model the decay process requires including the aerodynamic 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 assumed implicitly that the instantaneous speed of the bob or line gives the same drag as would a steady motion of the same speed. Whether this assumption is appropriate remains to be seen.

The main task is to account for the Re-dependence of the drag coefficients. Because the line is comparatively thin, $D_l = 0.4 \times 10^{-3}$m, the Reynolds numbers of the line $Re_l = UD_l/\nu$ are comparatively small, $Re_l \leq 20$, 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 may be approximated by

$$ {}^{c}C_v(Re_l) = 1.3 + \frac{Re_l}{2}, \tag{60} $$

(the short red dashed line of Fig. 7, right). The drag per unit length of the line, $\delta = dr$, can then be computed by the drag law corresponding to Eq. (58) as

$$ H = \pi \rho \nu \, {}^{c}C_v(Re_l) \, U \, dr $$

and the (dimensional) torque due to drag over the length of the line is

$$ \tau_l = \int_0^L r \, H \, dr = \rho \left(1.3\frac{\pi}{2}\nu L^3 + \frac{1}{12} D_l L^4 \mid \frac{d\phi}{dt} \mid \right) \frac{d\phi}{dt}. \tag{61} $$

The absolute value operator ensures that the drag opposes the motion.

The bob has a comparatively large diameter, $D_b = 0.021$ m, and thus a much larger Reynolds number; $Re_b = L\frac{d\phi}{dt}D_b/\nu$ is O(1000). In this range of Re, a convenient fit to the inertial drag coefficient of a sphere is[26]

$$ {}^{s}C_i(Re_b) = \frac{24}{Re_b} + \frac{6}{1 + \sqrt{Re_b}} + 0.4, \tag{62} $$

the dashed red line of Fig. 7, left. The drag-induced torque on the bob is then estimated as

$$ \tau_b = \frac{\pi \rho}{8} \, {}^{s}C_i(Re_b) \, D_b^2 \, L^3 \mid \frac{d\phi}{dt} \mid \frac{d\phi}{dt}, \tag{63} $$

where Re_b and ${}^{s}C_i$ are evaluated at each time step of the numerical integration (with care to avoid division by zero when evaluating (62)).

The revised angular momentum balance is (in dimensional variables),

$$\frac{d^2\phi}{dt^2} = -\left(\frac{g}{L}\sin(\phi) + \frac{\tau_l + \tau_b}{L^2 M}\right). \tag{64}$$

Together with Eqs. (60) - (63) and the initial condition Eq. (3), this makes a complete if inelegant model that can be integrated numerically. With these drag terms included, the period of the oscillation is nearly unchanged, but the amplitude decays with an e-folding in about 10 minutes (Fig. 6, upper, the envelope of the thin blue line). The amplitude simulated by the numerical model looks plausible when compared with the observations, suggesting that this application of the steady drag laws for a sphere and cylinder is appropriate.

4.4.2 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. (6), upper, that sinϕ of Eq. (55) can be approximated as ϕ. Second, the torque is due mostly, $\approx 80\%$, to the line, and so it should be acceptable to make the approximation that the inertial drag coefficient for the bob is a constant, $C_i = 0.7$, an average for the high Re_b range of the bob where most of the drag occurs. With these approximations, a solvable model for the simple, viscous pendulum (now in nondimensional variables)

$$\frac{d^2\phi}{dt^{*2}} = -\left(\phi + a\frac{d\phi}{dt^*} + b\,|\frac{d\phi}{dt^*}|\,\frac{d\phi}{dt^*}\right), \tag{65}$$

where the coefficient in the linear drag term is

$$a = \frac{\pi}{2}\frac{\rho\,\nu\,L^{3/2}}{M\,g^{1/2}} \tag{66}$$

and the coefficient in the quadratic term is

$$b = \frac{\rho}{8M}(0.7D_b^2\,L + \frac{2\pi}{3}D_l\,L^2). \tag{67}$$

The linear term $\propto a$ comes from the low Re limit of viscous drag on the line, and the quadratic term $\propto b$ comes from the higher Re approximations of inertial drag on the line and the bob.

Approximate solutions for small damping are given by Timmerman and van der Weele.[26] Linear drag causes the amplitude to decay at a nondimensional rate

$$\frac{1}{\Phi}\frac{d\Phi}{dt^*} = -\frac{a}{2} \tag{68}$$

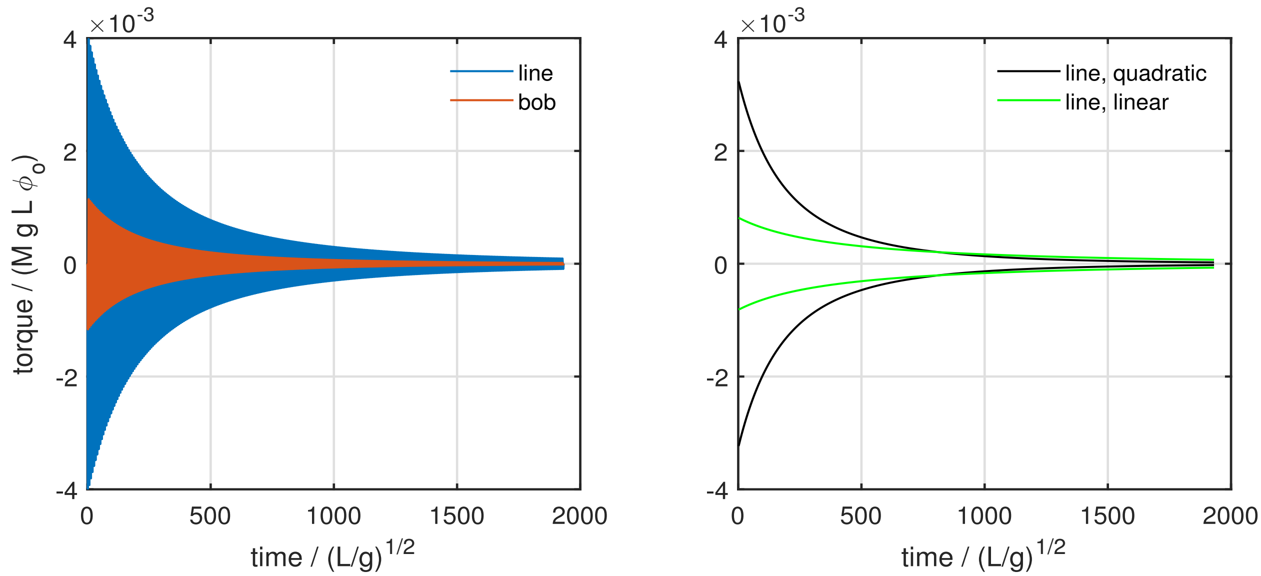


Figure 9: **(left)** The drag-induced torque on the line and the bob (blue and red lines) nondimensionalized with the gravitational-induced torque on the bob (which is much larger). The drag-induced torque on the line contributes about 80% of the total drag-induced torque. **(right)** The torque on the line due to quadratic and linear terms of Eqs. (65) - (67) (black and green lines; the envelope only is shown here). The sum is equal to the total shown as the blue line at left. For short times, when the amplitude of the displacement is largest, the torque on the line is due mainly to the quadratic term, and thus the inertial drag process. For longer time and smaller amplitude, the torque on the line is due about equally to the quadratic and linear terms, and thus about equally to inertial and viscous drag.

and the quadratic term causes decay at a rate

$$\frac{1}{\Phi}\frac{d\Phi}{dt^*} = -\frac{8\,b}{6\pi}\Phi\,. \tag{69}$$

For small damping, these two may be added together, and when evaluated for the specific pendulum observed here (Sec. 4.1),

$$\Gamma\sqrt{L/g} = \frac{1}{\Phi}\frac{d\Phi}{dt^*} \approx -\left(5.2\times10^{-4} + 1.6\times10^{-2}\Phi\right) \tag{70}$$

shown as the blue, dashed line (theory) of Fig. (6), lower. This approximate solution gives a good account of the damping found in numerical simulations even for much stronger damping, e.g., by imposing a larger viscosity. It also shows clearly how the decay rate is expected to vary with the parameters that characterize the pendulum and the surrounding fluid. All of the pieces of this model were present in the first attempt at dimensional analysis of a damped, viscous pendulum, Eq. (49), though we were not prepared to interpret them at that time.

The decay rate, $\Gamma = \Phi^{-1}d\Phi/dt$, can be estimated from the observations and from the numerical solution by first differencing the observed and model-simulated amplitudes, Fig. (6), lower. This makes a much more sensitive test of the drag formulations than does a comparison of the amplitude cf., Fig. (6), upper. The first thing to note is that the decay is not a simple exponential in time as it first appears on inspection of $\Phi(t)$. Instead, there is a significant dependence of the decay rate upon amplitude, with a greater decay at the beginning of an experiment when the amplitude was largest, $\Phi \geq 0.1$ radians, than at long times when the amplitude was small, ≤ 0.05 radians.

Within the numerical solution, the drag-induced torque on the pendulum was due mainly to drag on the line rather than drag on the bob. The drag on the line was due mainly to inertial drag at short times, when the amplitude of the motion was larger, and then due mainly to viscous drag at long times, when the amplitude was small, $\Phi \leq 0.05$ radians, cf. Figs. (6) and (9).

4.5 Damped pendulum problems

1) Stokes' landmark 1851 study of viscous fluid flow[22] included the detailed calculation of drag on a smooth sphere moving at a steady speed, U, through a viscous fluid. The speed was presumed to be slow enough that the flow remained laminar (no turbulence and no inertial effects). Such a slow flow is sometimes called 'creeping flow', and in practice requires that the Reynolds number must be very small, $Re = UR/\nu \leq O(1)$. Stokes found that the viscous drag on a sphere of radius R is then

$$H = 6\pi\nu RU. \tag{71}$$

How does this compare with the inferred (empirical) viscous drag coefficient of Fig. (8), left, at very low Re? And how about the inertial drag coefficient of Fig. (7), left, the dashed black line, $24/Re$?

2) Under the plausible assumption that the terminal (steady) velocity, U, of a dense sphere falling through viscous fluid reflects a balance of the viscous drag from (71) and buoyancy force, $\propto R^3 g\delta\rho$, show that the Stokes settling velocity, U_{set}, is

$$U_{set} = \frac{2}{9}\frac{g\,\delta\rho}{\nu}R^2, \tag{72}$$

where $\delta\rho$ is the density difference $\rho_{sphere} - \rho_{fluid}$. This has proven to be an accurate and widely useful estimate of the settling velocity of small particles in air and water (dust and sediment). When the velocity may be observed and the density and radius of the particle are known, then (72) is a handy means for estimating the viscosity of the surrounding fluid. Now suppose that you do not know (72), but that you do know the VPlist for the corresponding dimensional analysis of velocity. What can you infer about the dependence of settling velocity U_{set} upon $\delta\rho$ and particle radius?

3) The decay rate (Fig. 6, lower) decreases significantly as the amplitude of the pendulum's motion dies away. At very small amplitude, $\Phi \leq 0.05$ radians, the visual observations of cord length made here were rather noisy (poor resolution). However, an extrapolation of the observed and modeled decay rates to $\Phi \rightarrow 0$ looks plausible. How does the inferred decay rate compare with the analytic solution for linear drag, Eq. (68)? It is surprising that (68) does not depend upon the line diameter, D_l. Can you explain why this model of linear, viscous drag-induced torque does not depend upon D_l, while the (generally more important) quadratic, inertial drag and torque do?

4) The empirical formula for the inertial drag coefficient $C_i(Re)$ Eq. (62) has terms that are proportional to Re^{-1}, $Re^{-1/2}$, and Re^0, and that are important at low, medium and high Re. We have considered the low Re limit in Problem 1 above, and now consider the high(er) Re limit. What does dimensional analysis tell us about the drag coefficient in the limit that Re is very large, say because $\nu \rightarrow 0$? (We won't know the value of C_i in that limit without help from laboratory data or sophisticated and intensive numerical computations, but we might expect it to be O(1), as indeed it is.)

5) Looking back at the first try at dimensional analysis for a damped pendulum, Sec. 4.1, it seems that the scope of the initial VPlist (48) was too broad. Let's try the opposite: narrow the scope to the one specific pendulum that was observed. Thus while L, D_l, D_b, and M are all relevant to the damping rate, they can also be declared 'globally constant' parameters (Sec. 2.7), if we choose to narrow the scope to just one specific pendulum. The same holds for g and the viscosity of air, ν. What are we left with? How does the new, greatly pared-down analysis compare with Fig. (6), lower?

6) Aerodynamic drag on an object is a property of the three-dimensional flow around the object, and not just its frontal area as the usual drag coefficient allows (7). Consider objects that are azimuthally symmetric in the direction parallel to the flow, e.g., a disk or a sphere or an ellipse, and suppose that we allow in just one length to define the object's geometry: the length from front to back, B. The ratio of this length compared to the diameter of the object is sometimes called fineness; $B/D \rightarrow 0$ for a thin disk, $B/D = 1$ for a sphere, and so on. Would you prefer a drag coefficient that varies with both Re and B/D, and so encompasses the entire family of azimuthally symmetric objects, or, would you prefer a separate drag coefficient for each shape, i.e., a Cd for each specific B/D. Discuss this in terms of locally and globally constant parameters, Sec. 2.7. Qualitatively, how would you expect the viscous drag coefficient to vary with B/D? How about the inertial drag coefficient?

7) It is surprising that the formulations for steady state flow and drag give accurate estimates of the aerodynamic drag on the bob and line of an oscillating, time-dependent pendulum. A possible rationalization: presume that the line interacts with the surrounding air on the scale of the line width, D_l. Can you show that the velocity change on this space scale is $O(D_l/L) \propto 10^{-3}$, i.e., very small? Hence, the flow around the line is quasi-steady. Where is this in the first try at dimensional analysis of a viscous pendulum, Eq. (48)?

5 A similarity solution for diffusion in one dimension

Dimensional analysis can help simplify and reduce a model system and help find solutions that might otherwise have been missed. A good example is afforded by Stokes' first problem (also called the Rayleigh boundary layer problem) in which a fluid column of uniform density is driven from rest by a speed, V_o, imposed on a boundary, here the upper boundary, $z = 0$, Fig. (10). The problem is to find the resulting current, $U(z,t)$. A key physical assumption is that the momentum supplied at the boundary will diffuse into the fluid at a rate set by the kinematic viscosity of the fluid, ν, times the velocity shear,

$$\tau \;=\; \nu\frac{\partial U}{\partial z}, \tag{73}$$

which is characteristic of Newtonian fluids (air and water). The quantity τ is proportional to the diffusive flux of momentum, and, it may also be interpreted as a shear stress, a horizontal force per unit area exerted by the fluid on a boundary or an adjacent fluid parcel. The kinematic viscosity is presumed to be a given constant, and not dependent upon the flow. In that case the governing equation for the current, $U(z,t)$, is the elementary one-dimensional diffusion equation,

$$\frac{\partial U}{\partial t} \;=\; \frac{\partial \tau}{\partial z} \;=\; \nu\frac{\partial^2 U}{\partial z^2}, \tag{74}$$

a partial differential equation (PDE) in z and t. The initial condition is a state of rest,

$$U(z,t=0) \;=\; 0. \tag{75}$$

The boundary conditions are that the fluid sticks to an upper surface that is set into motion with the speed Vo at $t = 0$, and the fluid also sticks to a lower boundary at $z = -L$ that remains at rest. (Stokes first problem presumes that $L \rightarrow \infty$, discussed more below.) The two required boundary conditions on U are then

$$U(z=0,t\geq 0) \;=\; Vo, \quad \text{and} \quad U(z=-L,t) = 0. \tag{76}$$

This linear PDE system is not difficult to solve; Fourier transform leads to an infinite sine series that can be summed to high accuracy, and numerical integration is straightforward (Fig. 10, left).[28]

5.1 Dimensional analysis to simplify or reduce a model system

Dimensional analysis can help point the way to yet another possibility, a similarity solution, that may yield more insight than is likely to come from an infinite series or from numerical data. As usual, a start can be made with a straightforward list of the important variables and parameters,

[28] Matlab code: https://www2.whoi.edu/staff/jprice/wp-content/uploads/sites/199/2024/08/Stokes1.txt

(77)

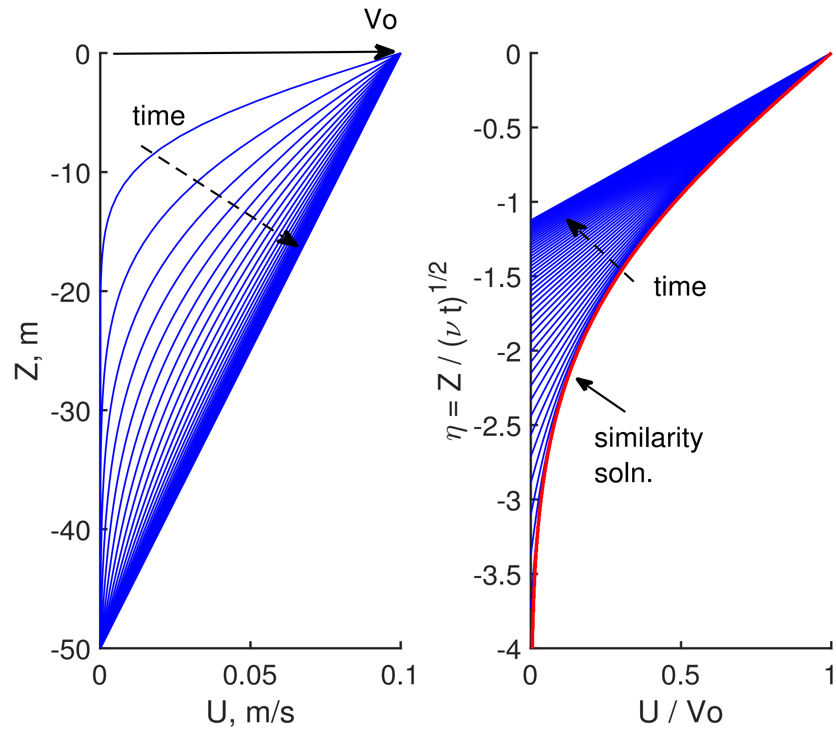


Figure 10: Numerical solution of the elementary diffusion model, Eqs. (74) - (76). **(left)** The solution in dimensional coordinates. The parameters are those of an upper ocean boundary layer over shallow water, $\nu = 10^{-2}$ m^2 s^{-1} and $L = 50$ m. Current profiles are plotted at intervals of 10^4 seconds. **(right)** The solution in nondimensional coordinates (blue lines) and the similarity solution, Eq. (86), (red line). The similarity solution presumes an infinitely deep fluid and does not recognize the no-slip boundary condition at $z = -L = -50$ m imposed here. For small time, only about the first three profiles, the numerical solutions are coincident with the similarity solution. At longer times the effect of finite depth and the no-slip lower boundary condition becomes appreciable, and the numerical solution departs from the similarity form.

- A VPlist for one-dimensional diffusion:

1. current, $U \doteq [\,0 \;\; 1 \;\; \text{-}1\,]$, the dependent variable,
2. time, $t \doteq [\,0 \;\; 0 \;\; 1\,]$, an independent variable,
3. depth, $z \doteq [\,0 \;\; 1 \;\; 0\,]$, a second independent variable,
4. surface boundary value, $Vo \doteq [\,0 \;\; 1 \;\; \text{-}1\,]$, a parameter,
5. the kinematic viscosity, $\nu \doteq [\,0 \;\; 2 \;\; \text{-}1\,]$, a parameter,
6. the depth of the fluid column, $L \doteq [\,0 \;\; 1 \;\; 0\,]$, a parameter.

Given this VPlist, a basis set of nondimensional variables has four members, and a relation among them may be written

$$\frac{U}{Vo} = F\left(\frac{zVo}{\nu}, \frac{tVo^2}{\nu}, \frac{z}{L}\right), \tag{78}$$

as one possibility.

5.2 Tuning the VPlist

The model equations, (74) - (76), are linear. In consequence, the numerical solutions show that the current at a given depth and time is directly proportional to the boundary value, Vo. Thus, the sole effect of changing Vo is to change the current in direct proportion. This important property has not been built into the VPlist, nor is it reflected in the initial basis set of nondimensional variables, Eq. (65). Instead, this VPlist includes a much more general problem in which Vo appears in the nondimensional variables combined with z and t as if Vo affected the diffusion process. In general, it does! On physical grounds, this can be expected to hold when the diffusion process results from turbulence generated by the boundary forcing rather than by (the implicitly laminar) diffusion process represented by ν. Whether the flow is turbulent or laminar depends upon the distance from the boundary, the current speed, and the fluid viscosity, i.e., a Reynolds number, zVo/ν.

If we insist that the diffusion process must be represented by a constant viscosity (what is meant by elementary diffusion) then we are implicitly limiting the analysis to small Reynolds number flows that are laminar, and the dynamics are identical to the diffusion of heat in a solid. To assert this idealization in the VPlist requires a small but significant change — retain the dependent variable U/Vo, but remove Vo from the list of parameters. A basis set for this revised VPlist is found to be

$$\frac{U}{Vo} \;=\; F(\frac{z}{\sqrt{t\nu}}, \frac{z}{L}), \tag{79}$$

which has one fewer variables than the initial basis set, Eq. (78), a significant change.

One step further — suppose that L is very large compared to the depth that diffusion has reached up to the time, t. This will be plausible if the fluid is very deep, or, the time very short. In that case the depth L will be irrelevant, and the current at a given depth ($z << L$) will depend only upon time, depth, and the kinematic viscosity that are combined into the single independent, nondimensional variable, η,

$$\eta = \frac{z}{\sqrt{t\nu}}, \tag{80}$$

and thus (79) may be written

$$\frac{U}{Vo} \;=\; F(\eta). \tag{81}$$

The variable η is said to be a 'similarity' variable, and the function $F(\eta)$ is a similarity function, something we have encountered before in Sec. 4 as the Reynolds number and drag coefficients. If Eq. (81) is relevant, then the current profiles at various times will have a similar shape, being more or less stretched out in z depending upon $\eta = z/\sqrt{t\nu}$, Fig. (6), right.

It has been noted on other occasions that the power of nondimensional analysis comes in large part from a reduction in the number of variables that is required to define a solution. Often the variables that are compacted are parameters, but on this occasion they are the independent variables, z and t. Dimensional analysis *per se* doesn't recognize the difference between a parameter and an independent variable or even the dependent variable. The implementation of the null space method suggested here (Sec. 3.3) strives to keep the dependent variable to the first power, but beyond that, independent variables and parameters are treated as equal.

5.3 When there is only one independent, nondimensional variable

The analysis above suggests that the original 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 holds, substitute $U = V_o F(\eta)$ into the governing equation (74). The partial time derivative becomes

$$\frac{\partial U}{\partial t} = Vo\frac{\partial F}{\partial \eta}\frac{\partial \eta}{\partial t} = -VoF'\frac{\eta}{2t}, \tag{82}$$

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

$$\frac{\partial^2 U}{\partial z^2} = VoF''\frac{1}{t\nu}. \tag{83}$$

Substitution into the governing equation and noting that z and t appear only in the combination $z/\sqrt{t}$ shows that the second order, linear partial differential equation (74) may indeed be transformed into the second order, linear ordinary differential equation (ODE),

$$F'' + \frac{1}{2}\eta F' = 0. \tag{84}$$

Substitution of η into the upper and lower boundary conditions gives

$$F(\eta = 0) = 1, \quad \text{and} \quad F(\eta = -\infty) = 0. \tag{85}$$

So far as the current is concerned, small time and large z are equivalent, and the initial condition is identical to the lower boundary condition. The ODE model equations (84) and (85) may be integrated by an approach much like that used with the finite amplitude pendulum, Eq. (13). First, let $G(\eta) = F'$ and substitute into (84). Separate variables, G and η, and integrate to find $F'(\eta)$. Then separate F and η, and integrate using (85) to find the similarity solution

$$\frac{U(\eta)}{Vo} = 1 - \frac{2}{\sqrt{\pi}}\int_0^{\eta/2} exp(-y^2)dy. \tag{86}$$

The integral at right is the error function, and this can be considered a closed solution.

A similarity solution (86) 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 assess *a priori*. Because it is more or less transparent, there are qualitative features evident in this

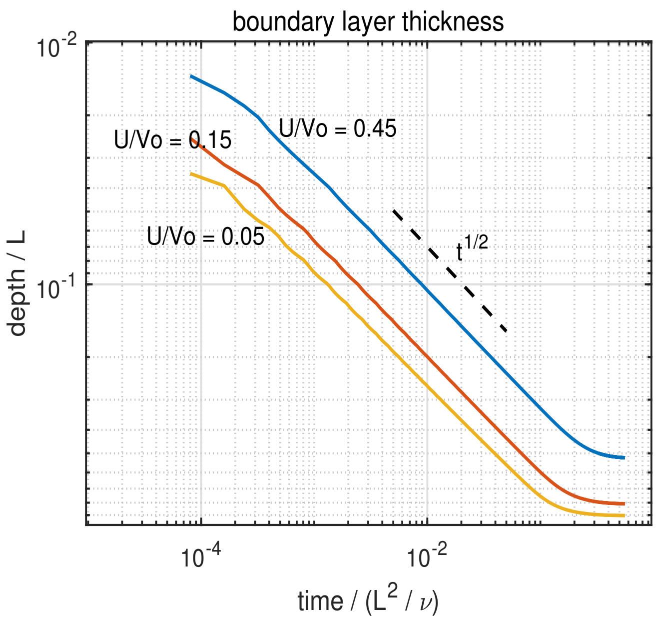


Figure 11: Boundary layer thickness, defined as the depth at which the current was 0.45, 0.15 or 0.05 of the imposed, surface value (three separate curves) diagnosed from the numerical solution of Fig. 10. Notice that there was an intermediate (nondimensional) time, $5 \times 10^{-4} \le t/(L^2/\nu) \le 10^{-1}$ during which the boundary layer thickness grew as $t^{1/2}$. The similarity solution Eq. (86) holds during this intermediate time, but not otherwise.

similarity solution that might have been missed in a mass of numerical data. In particular, note from (81) and (80) that a given value of η, say η_g, and thus a given value of $U/Vo = F(\eta_g)$, moves downward with time as

$$z_g = \eta_g \sqrt{t\,\nu}. \tag{87}$$

Thus the thickness of the layer that is directly affected by the boundary condition, i.e., the boundary layer, grows like the square root of time, Fig. (11), a characteristic of 1-dimensional elementary diffusion and random processes alike. This result can be turned around; the time t_g needed to diffuse a distance z_g away from the boundary is

$$t_g = \frac{z_g^2}{\nu\,\eta_g^2},$$

which forms a natural time scale for a one-dimensional diffusion problem.[29]

If the model domain includes a lower boundary at say $z = -L$ as here, then the similarity solution has the built-in limitation that it holds only for a limited, intermediate time interval. The elapsed 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 literal step function in time as is assumed. In a numerical model, the boundary layer must be thick enough that the current profile is not too strongly dependent upon the finite grid resolution. Judging from the curves of Fig. (10), this requires a time of about $t/(L^2/\nu) \ge 5 \times 10^{-4}$, in this numerical solution. However, the time can

[29] When we say that diffusion reaches a certain distance we mean that a given fraction of the boundary value amplitude is found at that distance from the boundary. The continuous diffusion equation has the unphysical property that every point in the domain is affected by the boundary condition instantaneously. If the point is very far away from the boundary in the sense that η is large, then the boundary effect will be correspondingly very small, though not zero.

not be too long either, or else diffusion will have caused an appreciable current near the lower boundary at $z = -L$. Once the current becomes appreciable near the lower boundary, judging from Fig. (10), $t/(L^2/\nu) \geq 10^{-1}$, the similarity solution will no longer hold accurately, and a more general solution of the form (79) will be required.

5.4 An oscillating upper boundary; Stokes second problem

Now consider another important diffusion problem that is in some respects the complement to the start-up problem discussed above. Suppose that the fluid is infinitely deep, so that there is no L to contend with, and that the upper boundary oscillates in time as $V_o cos(\omega t)$. If this has been going on for a very long time, the flow will become statistically steady though time-dependent, oscillating with the frequency ω imposed by the boundary. In this problem there is an important external time scale, ω, (an inverse time) that we cannot dispose of. Hence, there is no similarity solution, as in the case when the depth of the lower boundary was relevant, an imposed, external length scale.

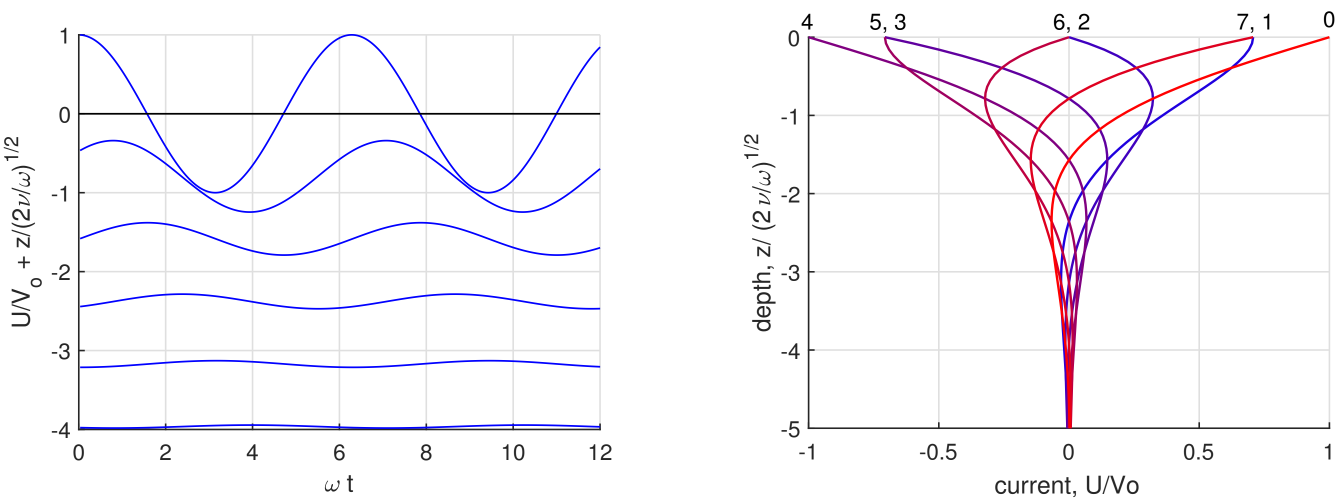


Figure 12: A solution for the time-dependent current under an imposed, oscillating surface current. **(left)** Time series of the current at six depths. **(right)** Eight profiles over one period of the surface current oscillation. The first profile is red, and successive profiles trend toward blue.

What to do? How about analyzing

- A VPlist for one-dimensional diffusion driven by an oscillating boundary: (88)

1. current, $U/Vo \doteq [\,0 \quad 0 \quad 0\,]$, the dependent variable,
2. depth, $z \doteq [\,0 \quad 1 \quad 0\,]$, an independent variable,
3. time, $t \doteq [\,0 \quad 0 \quad 1\,]$, an independent variable,
4. kinematic viscosity, $\nu \doteq [\,0 \quad 2 \quad \text{-}1\,]$, a parameter,
5. oscillation frequency of the boundary, $\omega \doteq [\,0 \quad 0 \quad \text{-}1\,]$, a parameter.

A basis set of nondimensional variables for this problem is then

$$\frac{U}{V_o} = F(\frac{z}{\sqrt{\nu/\omega}},\ \omega t). \tag{89}$$

This problem has a natural time scale, $\propto \omega^{-1}$, as well as a natural length scale, $\propto \sqrt{(}\nu/\omega)$, which depends upon the viscosity and the frequency, ω.

The solution can be presumed to be separable in depth and time,

$$U(z,t) = Real\,[A(z)\exp i\omega t]\,.$$

The depth dependence has to be a decay with increasing depth, and the time dependence will be an oscillation at the frequency ω. Substitution into the governing PDE (74) and minding the boundary conditions gives[30]

$$\frac{U}{V_o} = \exp(z/\sqrt{2\nu/\omega})\ cos(\omega t - z/\sqrt{2\nu/\omega}), \tag{90}$$

which we might have guessed, or certainly can recognize, up to the factor $\sqrt{2}$ that appears in the length scale of (90). The solution reveals that the boundary effect diffuses into the fluid, producing a sinusoidal current at depth, Fig. (12). The current at depth lags the surface current, and the amplitude decays over a comparable depth scale.

5.5 Diffusion problems

1) Consider the steady solution of Sec. 5.3. Suppose that the (water) column has a finite depth, L, and that the experiment is run out to a time that is long enough that the current becomes essentially steady. As before, assume that the boundary value of the current, V_o has the sole effect of changing the amplitude of the current. A VPlist for steady diffusion is then

[30] A thorough description of this solution method is by https://youtu.be/Jtg3WOTZfSw

- A VPlist for a steady state, one-dimensional diffusion:

(91)

1. current, $U/Vo \doteq$ [0 0 0], the dependent variable,
2. depth, $z \doteq$ [0 1 0], the independent variable,
3. the kinematic viscosity, $\nu \doteq$ [0 2 -1], a parameter,
4. the depth of the fluid column, $L \doteq$ [0 1 0], a parameter.

What do you find for U/V_o? What happened to the kinematic viscosity, ν? Where is the steady state solution evident in Fig. (10)? Can you interpret the steady solution in terms of the stress profile, $\tau(z)$, Eq. (74)? And specifically, how does the stress on the lower boundary (which is at rest) compare to the stress at the surface? Now suppose that the lower boundary is free-slip, which in practice means that there is no gradient normal to the surface, $\partial U/\partial z = 0$. What is the steady state $U(z/L)$ in that case?

2) The ocean surface layer. A classical boundary layer grows in thickness like $t^{1/2}$ or until the boundary layer reaches a lower boundary. On the other hand, we know that wind has been exerting a stress on the ocean surface for a very long time, and yet the surface boundary layer of the ocean, often called the Ekman layer, has a finite depth, very roughly, 50 m. The ocean surface layer is evidently not a classical boundary layer of the type considered here.

One important reason for this difference is that the ocean is rotating with respect to the distant stars at a rate $f = 2\Omega$sin(latitude), where Ω is Earth's rotation rate, 7.272 $\times 10^{-5}$ s^{-1}. At mid-latitudes, $f \approx 10^{-4}$ s^{-1}. The rotation rate f should thus appear in a VPlist of the ocean surface boundary layer. Under the assumption that the fluid thickness L is not relevant, can you use dimensional analysis to show that a steady boundary layer thickness could be possible (dimensionally)? Assuming that the diffusivity is roughly $\nu = 10^{-2}$ m^2 sec^{-1}, can you make a correspondingly rough estimate of the (rotating) upper ocean boundary layer thickness? (Hint: O(100 m).) This is often called the Ekman layer depth.[31]

3) A seasonal cycle of subsurface land temperature. The diurnal and seasonal variation of solar radiation warms the Earth's surface periodically. Heat loss from the surface occurs via infrared radiation and by sensible and evaporative heat fluxes so that over a day or a year there may be a more or less closed diurnal or annual cycle of the surface temperature. The periodic (but not necessarily sinusoidal or regular) variation of surface temperature diffuses downwards into the ground at a rate that depends upon the properties of the subsurface soil and rock; solid rock and water-saturated soil have a thermal diffusivity (reusing ν) $\nu \approx 1.2 \times 10^{-11}$ m^2 s^{-1}, while loose, dry soil is a comparative insulator, $\nu \approx 0.5 \times 10^{-11}$ m^2 s^{-1}. The amplitude, depth of penetration and phase of the resulting subsurface temperature response is of considerable importance for the design and operation of geothermal energy projects.[32]

---

[31] On the equator, $f = 0$, and so the rotation effect that produces a finite-depth Ekman layer vanishes. Is a classical (very thick) boundary layer observed on the equator? No. The notion of an Ekman layer depth discussed above is not applicable very near the equator, where instead a basin-scale, hydrostatic pressure gradient caused by a tilted sea surface and oceanic thermocline (subsurface density variation in the vertical) largely balances the mean wind stress. This basin-scale asymmetry is the background for the El Nino - Southern Oscillation phenomenon.

[32] Cuhac, C., A. Makiranta, P. Valisuo, E. Hiltunen and M. Elmusrati, 2020, 'Temperature measurements on a solar and low enthalpy geothermal open-air asphalt surface platform in a cold climate region', Energies, 13, 979 - 998. doi:10.3390/en13040979.

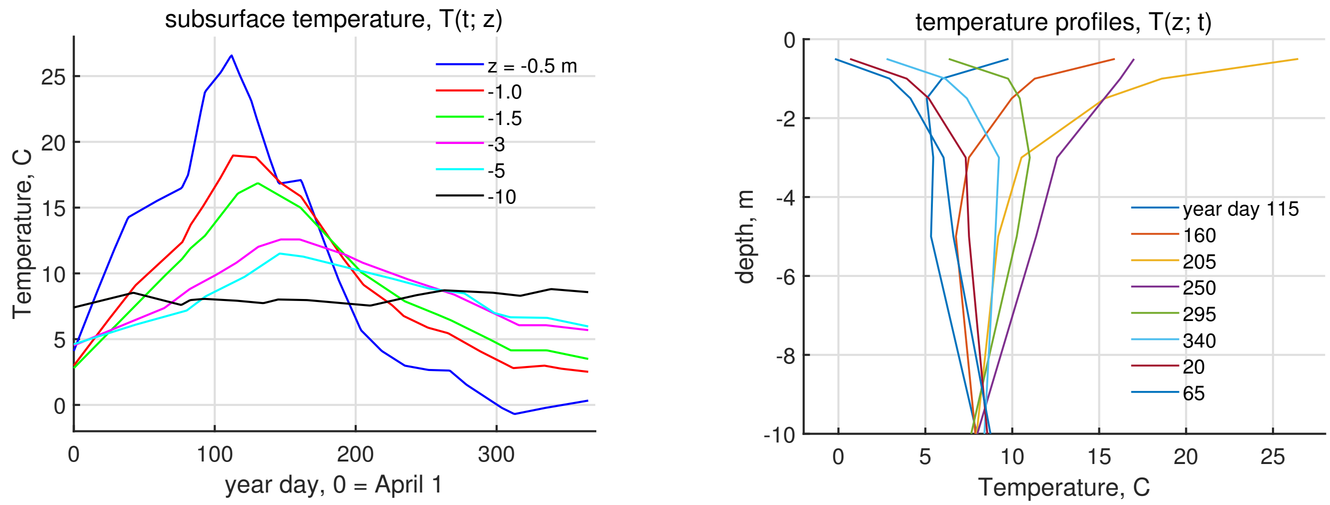


Figure 13: Subsurface (land) temperature observations made over a year as part of a geothermal energy study conducted in Finland by Cuhac et al. (2020).[32] These data were extracted from their Fig. (12). The data start in early April 2014 and continue until the following April. Notice that the beginning and ending temperatures do not coincide; evidently the following April was cooler than the first. Nevertheless, there is a clear signal of downward diffusing, summer warming.

A typical annual cycle of subsurface land temperature is in Fig. (13). Some questions for you: How would you characterize these data, e.g., what is the depth of penetration of the summer warming? In terms of the idealized (perfectly sinusoidal) model solution (90), what is a plausible value for the thermal diffusivity at this site? What departures do you discern between these observations and the idealized model solution? What depth of penetration would you expect for an otherwise comparable diurnal cycle?

6 Choosing scales with a purpose

Dimensional analysis amounts to normalizing (or dividing) dependent and independent variables by an appropriate scale. Appropriate has two parts. First, the scale must have the same physical dimensions (e.g., mass, length, time) as the dependent variable, say. Second, if the scale is chosen thoughtfully, then normalizing will yield a nondimensional dependent variable that has a meaningful magnitude. This second aspect of choosing a scale is often called 'scaling analysis'. A dimensional analysis can be performed with no thought given to a scaling analysis. However, the results of a dimensional analysis will often be considerably more useful if aspects of a scaling analysis are considered either from the outset, or at the third, interpretive step of a dimensional analysis.

6.1 A nonlinear projectile problem

To illustrate the role of a scaling analysis, consider the following projectile problem (after Lin and Segel (1974)[5]). The problem is to calculate the height of a projectile of mass m that is launched upwards with speed V from the surface of a planet having a radius R and mass M. To make the problem interesting, the variation of the gravitational acceleration with height will be acknowledged. To keep the problem simple, the likely important effect of aerodynamic drag on the projectile will be ignored. With this simplification, the vertical component of the equation of motion for the projectile is

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

where G is the universal gravitational constant. The acceleration of gravity on the planet's surface $z = 0$ can be defined as $g = MG/R^2$, a parameter, and the equation of motion rewritten

$$\frac{d^2z}{dt^2} = \frac{-g}{(1+z/R)^2}. \tag{92}$$

Suitable initial conditions are

$$z(t=0) = 0 \quad \text{and} \quad \frac{dz}{dt}(t=0) = V. \tag{93}$$

The projectile has an initial positive upward velocity, V, a parameter, and thereafter is subject only to a downward acceleration due to gravitational attraction to the planet. As the height of the projectile increases, the gravitational acceleration decreases, and at some 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. To see just how, let's analyze a nondimensional form of the model equations (77) and (78), beginning with

(94)

- A VPlist of vertical motion in variable gravity:

1. height of the projectile above the planet surface, $z \doteq [0 \quad 1 \quad 0]$, the dependent variable,
2. time, $t \doteq [0 \quad 0 \quad 1]$, an independent variable,
3. the acceleration of gravity on the planet surface, $g \doteq [0 \quad 1 \quad -2]$, a parameter,
4. radius of the planet, $R \doteq [0 \quad 1 \quad 0]$, a parameter,
5. initial (vertical) speed, $V \doteq [0 \quad 1 \quad -1]$, a parameter.

The initial basis set of nondimensional variables for this VPlist is

$$\Pi_1 = \frac{z}{R}, \quad \Pi_2 = \frac{t}{R/V} \quad \text{and} \quad \Pi_3 = \frac{V^2}{gR}, \tag{95}$$

and the relation between these three nondimensional variables may be written

$$\frac{z}{R} = F(\frac{t}{R/V}, \frac{V^2}{gR}). \tag{96}$$

It will be helpful to denote these particular nondimensional variables with a prime,

$$z' = \frac{z}{R}, \quad \text{and} \quad t' = \frac{t}{R/V}$$

and the combination

$$\epsilon = \frac{V^2}{gR}.$$

With this notation, Eq. (96) is

$$z' = z'(t', \epsilon).$$

The maximum height that the projectile will reach, Z, is of particular interest, and the nondimensional form of Z corresponding to this initial basis set is

$$\frac{Z}{R} = F(\frac{V^2}{gR}),$$

or with the prime notation,

$$Z' = Z'(\epsilon). \tag{97}$$

Notice that $\epsilon = V^2/g\,R$ is the only nondimensional parameter in the relation for maximum height, i.e., g, R and V appear only in this combination. The parameter ϵ has a physical interpretation as the ratio of twice the initial kinetic energy of the projectile compared to the depth of the potential energy well of the planet evaluated on the planet's surface. A larger ϵ leads to a larger maximum height, and when $\epsilon \geq 2$, the kinetic energy exceeds the work required to climb out of the potential energy well. V is then said to be the escape velocity, approx. 12 km s^{-1} for Earth (ignoring drag

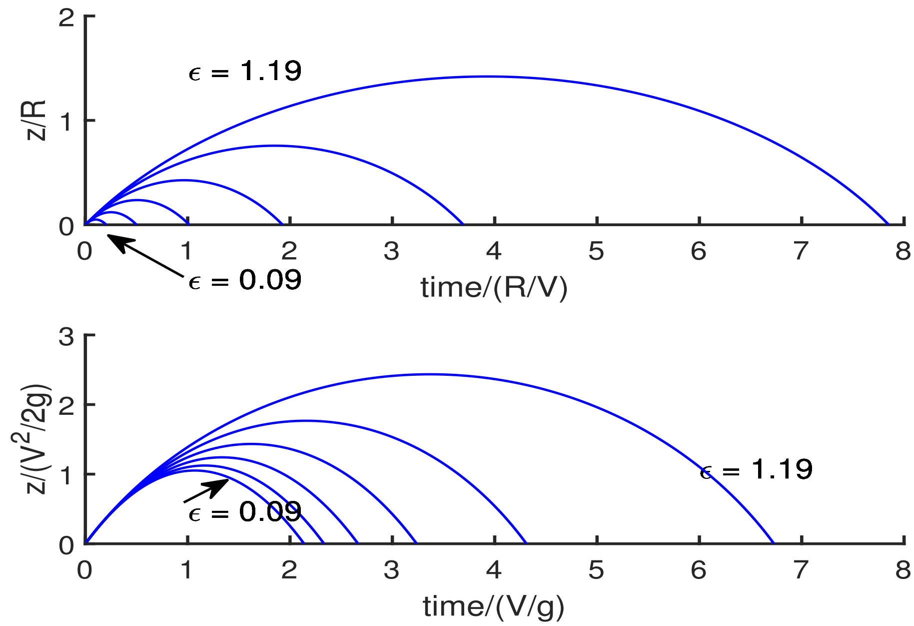


Figure 14: Projectile height computed numerically for several ϵ varying from 0.09 to 1.19. The solutions have been nondimensionalized in one of two ways; in the **(upper)** panel, using the initial basis set, Eqs. (96) and in the **(lower)** panel, using a second basis set built around a zero order solution described in Sec. 6.3. Notice that for small values of ϵ the set of curves in the lower panel appear to collapse toward one curve, while the set of curves in the upper panel do not.

with the atmosphere). Thus, in the usual (partial) way of dimensional analysis, we have already learned something useful about this problem.

Numerical solutions of Eqs. (92) and (93) nondimensionalized by this basis set look reasonable, Figs. (14), upper and (15), upper, in as much as Z' is a well-defined function of the nondimensional parameter ϵ. This is a mathematical certainty since these data are solutions of a numerical model whose parameters are known exactly and hence the VPlist and dimensional analysis should be consistent with the numerical solutions. This first basis set thus meets one of the main goals of dimensional analysis - to make a clear, concise 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).

6.2 Small parameter → small term?

An aspect of nondimensionalization that goes beyond previous discussions is the question of how or whether a nondimensional model equation can be used to develop an approximate solution. The solution method 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, the model equations can be amended to take account of the term dropped on the first pass, and solved again. Often the first such iteration will provide useful information regarding the consequences of the difficult term.

The initial basis set Eq. (97) amounts to scaling (and nondimensionalizing) the projectile height by the radius of the planet, R, and scaling time by the time interval required to move the distance R at a speed V. For the purpose of writing the model equation in nondimensional form, use that $z' = z/R$ and $t' = t/(R/V)$ and hence the nondimensional velocity of the projectile is

$$\frac{dz'}{dt'} = \frac{dz/R}{dt/(R/V)} = \frac{1}{V}\frac{dz}{dt}, \tag{98}$$

and the acceleration is

$$\frac{d^2z'}{dt'^2} = \frac{d^2z/R}{d(t/(R/V))^2} = \frac{R}{V^2}\frac{d^2z}{dt^2}. \tag{99}$$

When the equation of motion, Eq. (92), is written using these nondimensional variables the result is

$$\epsilon\frac{d^2z'}{dt'^2} = -\frac{1}{(1+z')^2} \tag{100}$$

and the ICs are just

$$z'(t=0) = 0 \quad \text{and} \quad \frac{dz'}{dt}(t=0) = 1. \tag{101}$$

The nondimensional equation of motion contains the single parameter, ϵ. For Earth-like values of R and g, and for V = 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 acceleration term of Eq. (92). It is plausible that the z-dependence could be ignored if $z \ll R$, and by extension, if $\epsilon \ll 1$. Thus, it should be possible 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, it might seem plausible that a first approximation to Eq. (100) could be obtained by dropping the term multiplied by the small parameter ϵ, the acceleration term. However, the solution to the reduced equation, $0 = 1/(1+z')^2$, is $z' = \infty$, which is contrary to the assumption of small z', and is nonsensical, generally. Either the idea that we could find a useful approximation by 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 justification for concluding that the acceleration term could be dropped simply because it is multiplied by the small parameter ϵ because we have no idea how big the nondimensional acceleration $\frac{d^2z'}{dt'^2}$ is compared with the terms kept, i.e., compared with 1. It turns out that $\frac{d^2z'}{dt'^2} \gg 1$ for $\epsilon \rightarrow 0$. As a result, when the acceleration term is dropped, it will not be possible to proceed toward an improved solution. It is important to understand that the nondimensional Eq. (100) is not at fault here; the terms of Eq. (100) still have the ratio one to another of the dimensional Eq. (92) since all that has been done is

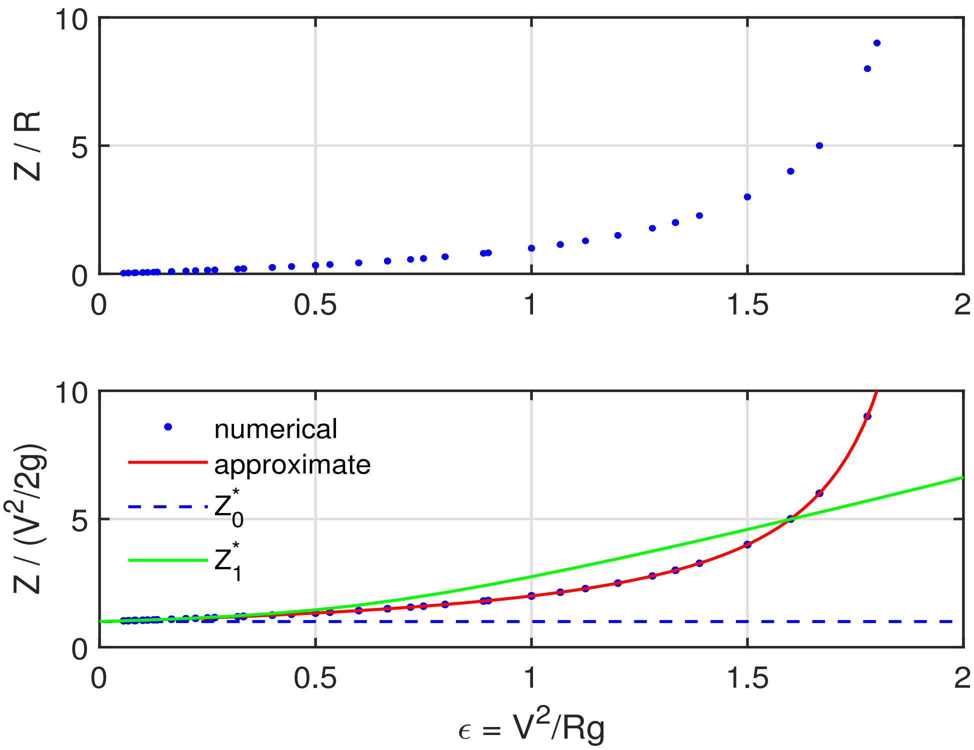


Figure 15: The maximum height of a projectile launched upwards at speed V from planets with a wide range of V, R and g. Each point is the maximum height diagnosed from a numerical solution. **(upper)** The maximum height nondimensionalized by R, consistent with the initial basis set of nondimensional variables, Eq. (96). **(lower)** Here the maximum height is nondimensionalized with a second basis set deduced from a scaling analysis (Sec. 6.3). Both basis sets show a clear-cut dependence of maximum height upon the single nondimensional parameter $\epsilon \quad = \quad V^2/gR$. Notice that as ϵ approaches 2 the nondimen sional height goes to infinity, indicating that the projectile has escaped the tug of gravity. In the lower panel, the dashed blue line is the maximum height from the zero order solution, Z_0^*, and the green line is the first order solution Z_1^* that takes some account of the reduced gravitational acceleration with height. The solid red line is an approximate solution Eq. (119) in which the effective gravitational acceleration is estimated as $g(1-\epsilon/2)$.

to divide by parameters. However, the inferences that might be drawn from Eq. (100) could be in error.

If the intent is to estimate the relative size of terms that are multiplied by nondimensional parameters, then the nondimensional dependent variables, in Eq.(100) z' and $\frac{d^2z'}{dt'^2}$, will have to be O(1) in the limit that the small parameter, ϵ, goes to zero, the relevant limit here. [33] In the first nondimensional basis set considered for this problem, the planet radius R was used as the length scale for the height of the projectile. While R is certainly an important length scale, it nevertheless has no direct relation to the maximum size of z, i.e., there is no basis for supposing that Z/R is O(1), generally. The imposition of this additional requirement on the choice of the scales is often said to comprise a 'scaling' analysis, implying a purposeful and thoughtful choice of the basis set of nondimensional variables.

6.3 Scaling the dependent variable

To ensure that the nondimensional z is O(1) as $\epsilon \rightarrow 0$ we have to choose a scale that is consistent with a physically motivated, even if highly simplified, model solution in that limit; in Sec. 4 this was said to be a 'zero order' solution. In this problem, a zero order model for the dimensional height z follows from ignoring the height dependence of the gravitational acceleration in Eq. (92),

$$\frac{d^2z_0}{dt^2} = -g, \tag{102}$$

where $(\ )_0$ refers to the order, which will be formalized shortly. The ICs are exactly as before

$$z_0(t=0) = 0 \quad \text{and} \quad \frac{dz_0}{dt}(t=0) = V, \tag{103}$$

and the solution is

$$z_0 = Vt - gt^2/2. \tag{104}$$

The zero order maximum height is then

$$Z_0 = V^2/2g.$$

[33] In Sec. 5 the 'big O' notation was used to indicate the numerical order of magnitude, e.g., O(1) or O(10^2), say. Here the meaning is extended to indicate the asymptotic behavior of a function. If a function $f(\epsilon)$ is O(1) in the limit $\epsilon \rightarrow 0$, then the function $f(\epsilon) \rightarrow constant$ in that limit. The constant need not be 1, though we have taken care that it will be here. If $f(\epsilon)/\epsilon^n \rightarrow C$, then $f(\epsilon)$ is said to be O(ϵ^n) in that limit. There are other possible gauges against which to measure asymptotic behavior besides the simple power law dependence that will suffice here. For more detail on the big O notation see http://encyclopedia.thefreedictionary.com/Big%20O%20notation A few questions: Based upon what you can see in Fig. 15, how would you characterize the order of the functions $Z'(\epsilon)$ and $Z^*(\epsilon)$, i.e., are they order ϵ^{-1}, ϵ^0 or ϵ? What is the order of d^2z'/dt'^2, and how does this impact the inferences we might draw from Eq. (100)?

When this is used as the scale for vertical distance, we can be sure that the maximum of the nondimensional height ≈ 1 in the range of small ϵ. An appropriate time scale is the time it takes the acceleration g to erase the initial velocity V, and is V/g. This new choice of scales amounts to a new basis set of nondimensional variables that is derivable from Eq. (96) when Π_1 and Π_2 are multiplied by Π_3^{-1}, or, reusing the Πs,

$$\Pi_1 = \frac{z}{V^2/2g}, \quad \Pi_2 = \frac{t}{V/g} \quad \text{and} \quad \Pi_3 = \frac{V^2}{gR}. \tag{105}$$

The relation between these three nondimensional variables can be written

$$\frac{z}{V^2/2g} = F(\frac{t}{V/g}, \frac{V^2}{gR}), \tag{106}$$

or using a $(\ )^*$ to denote these nondimensional variables,

$$z^* = F(t^*, \epsilon).$$

The nondimensional maximum height is then

$$\frac{Z}{V^2/2g} = F(\frac{V^2}{gR}), \qquad \text{or,} \tag{107}$$

$$\tag{108}$$

$$Z^* = Z^*(\epsilon), \tag{109}$$

with ϵ as before. This second basis set leads to a new interpretation, that $\epsilon \propto Z_0/R$, is the ratio of the zero order maximum height to the radius of the planet.

This new basis set also leads to a clearly defined functional dependence of maximum height upon the single parameter ϵ, Fig. (15), lower. This new form $Z^*(\epsilon)$ looks something like the initial form, $Z'(\epsilon)$, though with one significant difference. At small values of ϵ the new $Z^*(\epsilon)$ goes to a constant, 1, where the initial $Z'(\epsilon)$ decreased to zero as $\epsilon \rightarrow 0$. Thus, the new basis set is consistent with an underlying zero order solution in the limit of vanishing ϵ, where the initial basis set was not. This new basis set of nondimensional variables seems an obvious choice now that we have it in front of us, but then the first basis set seemed sensible as well (and was the form that happened to come from the computational algorithm).

6.4 Approximate and iterated solutions

We have now identified length and time scales that are appropriate specifically to the vertical motion of a free projectile in a constant gravitational field, as opposed to just any length and time scales that may happen to be present in the VPlist. Using this new basis set, and that

$$z^* = \frac{z}{V^2/2g}, \quad \text{and} \quad t' = \frac{t}{V/g}$$

are the nondimensional height and time, the nondimensional velocity of the projectile is

$$\frac{dz*}{dt*} = \frac{dz/(V^2/2g)}{dt/(V/g)} = \frac{2}{V}\frac{dz}{dt}, \tag{110}$$

and the acceleration is

$$\frac{d^2z^*}{dt^{*2}} = \frac{d^2z/(V^2/2g)}{d(t/(V/g))^2} = \frac{2}{g}\frac{d^2z}{dt^2}. \tag{111}$$

When ϵ is small, the dimensional acceleration is approximately equal to g, and this $\frac{d^2z^*}{dt^{*2}}$ is thus O(1). The same is true for the nondimensional velocity. Despite that the velocity goes to zero at the top of the trajectory we still say that dz^*/dt^* is O(1), because our concern is with the largest magnitude of a term, rather than its average or smallest value. The equation of motion Eq. (100) in these nondimensional variables is

$$\frac{d^2z^*}{dt^{*2}} = \frac{-2}{(1+\epsilon z^*/2)^2} \tag{112}$$

and the initial condition is

$$z^*(t=0) = 0 \quad \text{and} \quad \frac{dz^*}{dt}(t=0) = 2. \tag{113}$$

It is helpful to expand the right hand side of Eq. (112) using a binomial expansion,

$$\frac{d^2z^*}{dt^{*2}} = -2 + 2\epsilon z^* - \frac{3}{2}\epsilon^2 z^{*2} + \epsilon^3 z^{*3} + HOT, \tag{114}$$

that will converge for $\epsilon z^*/2 \leq 1$. Because we have taken care to nondimensionalize the height and time with appropriate scales, the size of each of the terms can be told by the exponent on the parameter ϵ. The acceleration term and the first term on the right hand side are independent of ϵ or O(ϵ^0), which is to say O(1) (a factor 2 notwithstanding). The second term on the right hand side is O(ϵ), the third term is O(ϵ^2), and so on, and the HOT is the sum of all the terms that are higher order in ϵ. The order of a model is said to be the highest exponent of ϵ in the terms retained in the expansion.

When the terms multiplied by ϵ are dropped, there follows a useful first approximation to the projectile problem,

$$\frac{d^2z_0^*}{dt^{*2}} = -2,$$

which is the so-called zero order model Eq. (104) written in nondimensional form. The ICs are exactly as Eq. (113) and the solution is

$$z_0^* = 2t^* - t^{*2}. \tag{115}$$

It seems that we have gone in a circle, but with Eq. (114) in hand we know how to proceed toward an improved solution; use this zero order solution as an estimate of z^* in the first order model, in which all terms higher order than ϵ^2 are omitted, to find,

$$\frac{d^2 z_1^*}{dt^{*2}} = -2 + 2\epsilon z^* \approx -2 + 2\epsilon z_0^* \tag{116}$$

$$= -2 + 2\epsilon(2t^* - t^{*2}), \tag{117}$$

which is easily integrable. The ICs do not involve ϵ, and the ICs of the first order model are Eq. (113) (the initial velocity is satisfied by the zero order solution alone). The first order solution is

$$z_1^* = 2t^* - t^{*2} + \epsilon(\frac{2}{3}t^{*3} - \frac{1}{6}t^{*4}). \tag{118}$$

Compared to the zero order solution, which appears here as the first and second terms on the right hand side, the new term is O(ϵ) and may be regarded as a correction to the zero order solution. The consequence of this new term for Z is that, for small ϵ so that t^* is O(1), there is an increased height compared to the zero order solution. Hence, this model and solution takes account (approximately) of the decrease in g with height. The maximum nondimensional height evaluated from Eq. (118), Z_1^*, shown as the green line in Fig. (15), lower, compares well with Z^* diagnosed from numerical solutions in the range $0 \leq \epsilon \leq 1/2$, and then begins to diverge from the numerical solution for larger ϵ, first above then below. Thus the first order solution represents approximately the effect of decreasing gravitational attraction with height above the planet's surface.[34]

An improved solution can be guessed by noting that the average value of z_0^* is $\approx 1/2$, and from Eq. (112) we might go on to guess that the height-dependent gravitational term in Eq. (117) could be estimated as the nominal gravity reduced by the factor $1 - \epsilon\frac{1}{2}$. This has the correct asymptotic behavior, i.e., it goes to 1 as ϵ vanishes, and it has the right qualitative behavior at $\epsilon = 2$ when the projectile should escape into deep space. The solution for the nondimensional maximum height is then

$$Z^\dagger = 1/(1 - \epsilon/2), \tag{119}$$

the solid red line of Fig. 15, lower, which compares well to Z^* diagnosed from the numerical solutions from the full model up to $\epsilon = 2$. Here and frequently, it happens that the first order model gives valuable insight into the parameter dependence (and one might say the physics) of a phenomenon in a way that numerical solutions, no matter how extensive and precise, may not.

[34] This iteration procedure is a simple and intuitive kind of perturbation analysis known as 'successive approximations' combined with an expansion in the small variable ϵ. Perturbation methods are powerful and very important techniques. Excellent references are Ch. 7 of Lin and Segel (1974),[3] and a very clear and concise text by J. G. Simmonds and J. E. Mann, 1986, *A First Look at Perturbation Theory*, Dover Publishing.

7 Applications / Homework projects

7.1 The Planck program, natural units for the universe

In Sec. 2 it was noted that the natural time scale of a simple pendulum is the reduced period, $P/2\pi = \sqrt{L/g}$, of small amplitude oscillation. When time is measured or normalized with this time scale, the parameters L and g disappear from the governing equation for a pendulum, a small convenience for calculations. This change of scale also points the way to a much more effective way to display a mass of data (Figs. (1) and (3)), and to think about pendulums generally. In 1899, Max Planck proposed something similar, but on a far, far grander scale — natural units for the universe.

Fundamental Physical Constants

constant	symbol	dimensions mass, length, time	SI value
speed of light	c	[0 1 -1]	299 792 458 m s^{-1}
Gravitational	G	[-1 3 -2]	6.674×10^{-11} kg^{-1} m^3 s^{-2}
Plancks const., reduced	$\hbar$	[1 2 -2]	1.054×10^{-34} kg m^2 s^{-2}

Table 1: Three of the Fundamental Physical Constants and their SI values. These important constants are known to considerably higher precision than implied here. To this short list could be added the charge of an electron and the Boltzmann constant, which would extend the scope to electro and thermodynamics.

7.1.1 A new fundamental constant with dimension mass

Even before the turn of the twentieth century it had been established that there were two fundamental physical constants that could be presumed to hold throughout virtually all of space and time, the speed of light in vacuum, c, and Newton's universal, gravitational constant, G (values in Table 1). Planck discovered a third fundamental constant, $\hbar \doteq [1 \quad 2 \quad -2]$, that represented the quantum of action appropriate to atomic scales, and it too could be presumed to be universal. Crucially, $\hbar$ includes the dimension mass. Given just these three universal constants, Planck recognized that one could construct the dimensions of any variable of mechanics.

In an address to the Prussian Academy of Sciences, he noted that units formed in this way would

> *necessarily retain their meaning for all times and for all civilizations, even extraterrestrial and non-human ones, and can therefore be designated as 'natural units.'*
>
> Max Planck, 1899

These natural units for the universe are called Planck units or Planck scales, Table 2.

Consider the Planck length, L_p to be the dependent variable of a VPlist; there are then four variables in the VPlist, L_p, plus the constants of Table (1), and the dimension matrix has rank $R = 3$. The inference from dimensional analysis is that there is one nondimensional variable,

$$\frac{L_p}{G\,h\,/\,c^4} \;=\; constant, \tag{120}$$

and that it is a constant. In other cases where the dimensional analysis came down to just one nondimensional variable, e.g., the period of a simple pendulum, the constant was found to be O(1) by comparison to observations or theory. Thus the nondimensional period, $P/\sqrt{L/g}$, is an appropriate representation of the period of a simple pendulum with respect to both the dependence upon L and g and the magnitude. Could that be the case also for the Planck length?

If the constant of Eq. (120) is O(1), then L_p is unimaginably small, $L_p \;=\; 1.616 \times 10^{-35}$ m, which is roughly 10^{-20} times the diameter of a proton. However, we do not know whether the Planck length has a direct physical interpretation, i.e., whether there is anything in nature having a length scale that is $\propto\ L_p$. This is a length scale that is well outside the reach of present experimental or observational methods, and so the suggestions for a physical interpretation are many and untestable: that it is the length scale of the nascent universe, that it is the smallest possible scale of a black hole, that it is the smallest possible, meaningful length, and so on. The physical interpretation of Planck length is an open question for physics and cosmology.

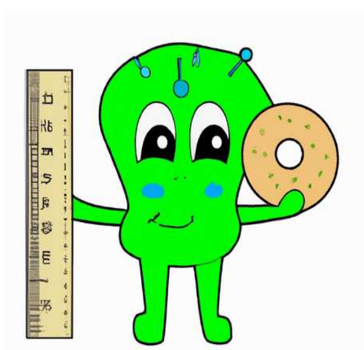
7.1.2 Planck problems

1) Use dimensional analysis to form a Planck time and mass from the fundamental constants c, G, and $\hbar$ of Table 1. How many nondimensional variables do you expect in the null space? You can check your method by computing the Planck length in SI units and comparing against Table 2.

Planck units

unit	symbol	construction	SI value
Planck mass	M_p	$\sqrt{\hbar G \,/\, c^3}$	2.176×10^{-8} kg
Planck length	L_p	$G\hbar/c^4$	1.616×10^{-35} m
Planck time	T_p	_?_	5.396×10^{-44} s
velocity		L_p/T_p = c	299 792 458 m s^{-1}
momentum		_?_	6.524 kg m s^{-1}
energy		$M_p c^2$	1.956×10^9 kg m^2 s^{-2}

Table 2: Planck units for mass, length and time formed from the Fundamental Constants of Table 1. The entries _?_ are left for you to fill in.



Come visit, Earthlings,
We have donuts, huge and tasty,
and well worth the trip.

2) Interstellar (mis)understandings. SETI[35] has at last detected and deciphered a radio signal coming from the neighborhood of Alpha Centauri that has been repeated over and over again for what may have been a very long time. The message is a short piece of text (see above), and then a long string of 1s and 0s, evidently a large binary number, $N \approx 10^{34}$. This appears to be a generous peace offering from a very advanced civilization. Their donuts may be good, but could they possibly be good enough and big enough to be worth such a lengthy trip — four years for a photon and much, much longer for mere Earthlings. The number N presumably defines the donut size in Planck units, the most likely system for interstellar communication. However, the clever ambassador neglected to indicate whether it was Planck volume or mass or energy or length. Assuming that their donuts are comparable to ours in terms of density, consistency, etc., and that their planet has a gravitational acceleration comparable to Earth's, what is a plausible interpretation of the number N, i.e., just how big are these donuts?

[35] Search for ExtraTerrestrial Intelligence: https://www.seti.org/

3) The dimensional form of gravitational mass attraction between two masses M_1 and M_2 is calculated by Newton's law to be

$$F = \frac{GM_1M_2}{r^2}. \tag{121}$$

Show that when the variables in this equation are written in terms of the Planck units for force, mass and length (Table 2), e.g., $r' = r/L_P \doteq nond$, (121) is simplified somewhat to

$$F' = \frac{M_1'M_2'}{r'^2}.$$

Since every variable has a prime, the prime may as well be dropped to arrive at

$$F = \frac{M_1M_2}{r^2}. \tag{122}$$

The interpretation of (122) could be unclear if not accompanied by a brief explanation, and specifically, what happened to the important gravitational constant, G, of (121)?

7.2 Open channel flow through a weir

Figure 16: **(left)** An adjustable weir that controls the flow of fresh water from Shawme Pond at left into Cape Cod Bay and the North Atlantic to the right. **(right)** A fish ladder that facilitates the upstream journey of river herring to their spawning grounds in Shawme Pond. Photos by the author in Sandwich, Massachusetts.

The transport of water within rivers and canals is of great practical interest, and difficult to monitor by direct measurements of the velocity. An indirect but simple and reasonably accurate method for monitoring the transport of open channel flows is to measure the height of the water surface at a location upstream of a weir, a barrier erected across the flow.[36] A weir has a well-defined notch, often rectangular, through which the water may flow and then fall freely into a lower pool, Fig. (16), left. The volume transport through the notch, Q, is expected to have a

[36] A very clear and appealing discussion of the physics and practice of open channel flows is by Caspar Hewett, https://www.youtube.com/channel/UCv2u6bFc7SSFsPHf98M50Mg

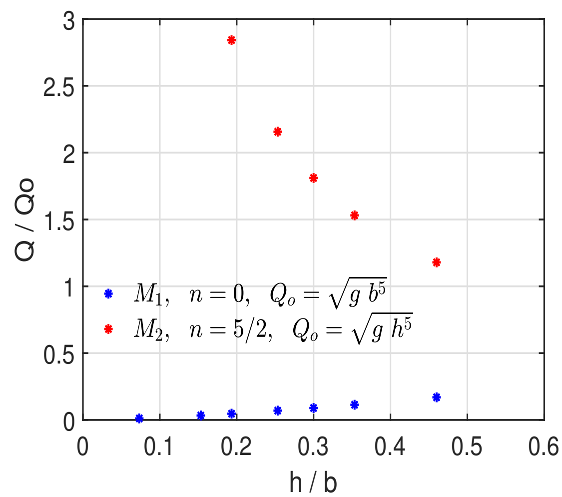


Figure 17: Two possible models of transport through a rectangular-notch weir evaluated using the transport data listed in the main text. Neither of these models indicates a satisfactory result insofar as the similarity functions have a very pronounced dependence upon h/b. These two models may be seen as end members of a family of models inferred from dimensional analysis (Eq. 125).

systematic dependence upon the width of the notch, here $b = 1.5$ m, and the upstream, still water surface height, h, measured with respect to the lower edge of the notch. The energy source in such a flow is gravitational potential energy released as the water falls through the notch, and hence the acceleration of gravity must be an important parameter. Water density would be relevant if the dependent variable was mass transport, but if the dependent variable is volume transport, then the water density may be omitted. That leaves a very concise VPlist for

- Volume transport through a rectangular-notch weir:

1. volume transport, $Q \doteq$ [0 3 -1], the dependent variable,
2. upstream water height, $h \doteq$ [0 1 0], the independent variable,
3. notch width, $b \doteq$ [0 1 0], a parameter,
4. acceleration of gravity, $g \doteq$ [0 1 -2], a parameter.

having only four variables and parameters and two physical dimensions, length and time. If attention is restricted to one specific weir, then two of these parameters, g and b, are 'globally' constant (Sec. 2.6).

The relationship $Q(h, b, g)$ may be inferred with help from dimensional analysis and the following data:[37]

upstream water level, h, m = [0.11 0.23 0.29 0.38 0.45 0.53 0.69]

volume transport, Q, m^3 s^{-1} = [0.097 0.286 0.403 0.601 0.773 0.982 1.46]

[37] Computed by a fairly comprehensive model that attempts to account for nominal friction at the weir as well as the finite size of the upstream basin: https://calcdevice.com/weir-spillway-id239.html

This VPlist is small enough that it can be analyzed by inspection, but let's use the null space basis method to illustrate an important point. If the variables and parameters of the VPlist are entered in the order above, then the nondimensional relationship among the variables comes out to be

$$M_1: \quad \frac{Q}{Q_o} = \frac{Q}{\sqrt{g\ b^5}} = F_1(\frac{h}{b}), \tag{123}$$

dubbed model M_1. This model is correct, (non)dimensionally, but notice the important term that scales the dependent variable, $Q_o = \sqrt{g\ b^5}$. If Q_o had come from a simplified model, then it could be a zero order solution with some physical significance. But in this case, Q_o is merely the first thing that came from the algorithm; it might be a good choice, or it might not be.

To find out, evaluate (123) using the data above, which gives the blue dots of Fig. (17). The result is not encouraging for $M1$ insofar as the h-dependence of the similarity function $F_1(h/b)$ is very substantial. Notice that this Q_o omits the upstream height, h, which we know is important. In fact, this Q_o is a constant for any one weir, and the crucially important dependence of Q upon h has to be relegated entirely to the similarity function of M_1. That isn't wrong, mathematically, but neither does this M_1 represent any real progress toward a useful model of weir transport.

Let's try the calculation again, but this time reversing the input order of h and b, to find a second model, M_2,

$$M_2: \quad \frac{Q}{\sqrt{g\ h^5}} = F_2(\frac{h}{b}), \tag{124}$$

This simply interchanges h with b, and now the zero order solution, $Q_o = \sqrt{g\ h^5}$, neglects b, the notch width, which is bound to be relevant. The argument h/b was retained in the similarity function F_2 so that M_1 and M_2 may be compared directly in Fig. (17). Model M_2 (the red dots of Fig. 17) seems no better than M_1 in that the similarity function varies significantly with h/b.

What can be learned from this? The calculation of a null space basis is a formal, logical (mathematical) procedure that takes no account of the physical meaning of the input variables. In this case, the calculation makes no distinction between two lengths, h and b, even while we know that they are quite different things.

You can improve on either of these first two models by combining the formal methods of dimensional analysis with physical insight, i.e., common sense plus a little experience. To start, show that a generalized relationship can be written

$$\frac{Q}{(g\ b^5)^{1/2} \times \left(\frac{h}{b}\right)^n} = {}^nF(\frac{h}{b}), \quad \text{where} \quad 0 \le n \le 5/2. \tag{125}$$

The first model M_1 is the case $n = 0$, and the second model M_2 is the case $n = 5/2$. Thus, dimensional analysis leads us to a family of models, depending upon the parameter, n, of which M_1 and M_2 are end members. The task is then to find the best n, the best member of the model family that has a physically sensible zero order solution, Q_o. A hint: Suppose a given h; what would you then expect for the dependence of Q upon b? You can test your new model by evaluating the similarity function as in Fig. (17) using the data listed above. If the result looks promising, i.e., if your new model yields a quasi-constant (or at least a much less variable) similarity function, then it could be plausible to represent the similarity function by a similarity constant.

A weir transport formula generally has to be calibrated for each specific weir to account for a variety of secondary effects including friction, that depends upon whether the weir is broad-crested, as here in Fig. (16), or sharp-crested, as well as the configuration and size of the upstream basin, i.e., its depth, width and length. For any specific weir these are globally constant parameters (Sec. 2.6) and hence their consequences may be lumped into (ascribed to) a similarity constant. What do you find for this weir?[38]

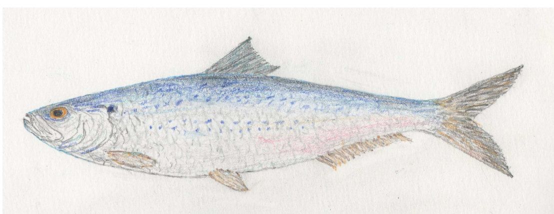


Figure 18: Blueback herring and the very similar alewife are known as river herring. An adult has a length up to 0.3 m and a mass of about 0.25 kg. The adult fish live at sea for four to five years, and then return to freshwater ponds to spawn in the spring. Drawing by the author.

The weir of Fig. (16) presents a formidable obstacle to the small river herring that migrate upstream from the North Atlantic to their spawning grounds in Shawme Pond. A springtime herring run was once an important food source for people and marine predators alike, but since the 1960s herring numbers have been greatly reduced by the damming of rivers and habitat degradation generally. The return of river herring to Shawme Pond is made much easier by a fish ladder, a bypass channel, Fig. (16), right, having a sequence of 12 identical small weirs and ponds. What would you expect for the variation of h from one weir to the next?

7.3 How fast did dinosaurs run?

The most tangible evidence of dinosaur locomotion comes from fossilized trackway data that shows the stride length, λ, the distance between successive prints of the same foot, as well as the foot length, f (Fig. 19).[39] It is often possible to estimate these parameters for a specific, distinguishable animal on a trackway, and it is easy to tell Theropods and Ornithopods (bipedal dinosaurs with bird-like feet) from Sauropods (quadrupedal, with elephant-like feet) though the species can not be told. Correlations made with fossil skeletons show that the hip height, h, is roughly proportional to the foot length, $h \approx 4f$. There is nothing directly evident regarding time or speed. However, an empirical relationship between these features of gross anatomy and the

[38] A typical similarity constant for weirs of the sort considered here is $\approx 1/2$.

[39] A remarkable example: https://www.fs.usda.gov/detail/gmug/news-events/?cid=FSEPRD1171521

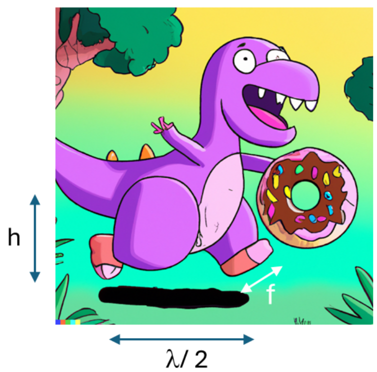


Figure 19: Evidence from fossil trackway data interpreted via dimensional analysis suggests that mighty Tyrannosaurus Rex was not a fast runner, typically. While the physical evidence and its analysis are fairly compelling, what do we know about her motivation and intention?

T Rex did not hurry,
Until, on a warm breeze, came the scent
Of sweet donut.

walking/running speeds for a very wide range of animals was developed by Alexander, 1976, A76,[40] on which this problem builds. How to infer dinosaur speed from trackway data?

- A VPlist for estimating running/walking speed from trackway observations:

1. speed, $V \doteq [\,0 \quad 1 \quad \text{-}1\,]$, the dependent variable,
2. stride length, $\lambda \doteq [\,0 \quad 1 \quad 0\,]$, a parameter,
3. height of the hip joint, $h \doteq [\,0 \quad 1 \quad 0\,]$, a parameter,
4. acceleration of gravity, $g \doteq [\,0 \quad 1 \quad \text{-}2\,]$, a parameter.

Gravity is highly likely to be relevant for any form of terrestrial locomotion, dinosaurs or not, and would seem to be the only plausible parameter that can provide the dimension time. There are four parameters having two dimensions, length and time, and so there are two nondimensional variables related by, as the first basis set,

$$\frac{V}{\sqrt{gh}} = F(\lambda/h). \tag{126}$$

where $F(\lambda/h)$ is an unknown similarity function. The dependent variable has been scaled with $\sqrt{gh}$, which is the phase speed of shallow water gravity waves on the sea surface. This is a natural scale from this VPlist, but is not the only possibility. The dependent nondimensional variable of (126) is the square root of an important nondimensional variable called the Froude number,

$$Fr = \frac{V^2}{gh},$$

[40] The pioneering study by Alexander, R. M., 1976, 'Estimates of speeds of dinosuars', Nature, Vol. 261, May 13. 129 - 130. See also R.. Alexander, 1991, 'How dinosaurs run', Sci. Am., 264, No. 4, 130 - 137. The first papers on this topic aimed to find correlations of the sort considered here, and discovered widespread, systematic relations that describe what would seem to be a very complex phenomenon — animal locomotion. Later papers have delved into the detailed structure of bones and muscles, e.g., W. I. Sellers and Manning, P. L., 2007, 'Estimating dinosaur maximum running speeds using evolutionary robotics'. Proc. Royal Soc. B, 274, 2711 - 2716. doi:10.1098/rsbp.2007.0846.

which often arises in contexts in which inertial forces $\propto V^2$ and gravitational forces $\propto g$ are at issue.

A76 used observations on the walking and running speed of living animals (including humans) to determine the similarity function $F(\lambda/h)$ of (126) and found that walking and running speed by a surprisingly wide variety of creatures — bipeds and quadrupeds, birds and mammals, creatures small and large — is consistent with

$$Fr = \frac{V^2}{gh} = .062\left(\frac{\lambda}{h}\right)^{3.33} \tag{127}$$

where the coefficient and the exponent were determined by fitting (126) to observations of U, λ and h.[41]

Given the following dinosaur trackway data of eight individual animals from A76,

stride, λ, m = [3.0 3.0 2.4 1.3 2.1 1.6 2.5 1.6]

hip height, h, m = [2.1 2.0 1.0 1.1 1.2 0.7 3.0 1.5]

estimate the speed of the responsible dinosaurs on the plausible assumption that (126) applies to these astounding creatures. The first six entries were made by bipedal dinosaurs, Theropods or Ornithopods, and the last two were made by quadrupedal Sauropods. Is there a systematic difference? An adult T Rex was a very large animal having a hip height up to $h = 4$ m. From what you can see in Fig. (19), how fast do you estimate T Rex was running on that occasion?

As we have seen in other problems, the detailed form of a solution is often not determined solely by dimensional analysis. Given λ and h, a speed requires a fourth parameter with dimension of time, and g is the obvious candidate. However, the way g enters does not have to follow (126), which uses $\sqrt{gh}$ as the velocity scale. Suppose instead that the dynamics of a swinging dinosaur leg is similar to a swinging pendulum; what do you get if the velocity scale is taken to be stride length / stride period $\propto \lambda/\sqrt{h/g}$? How can this be realized from the original form, Eq. (126)?

7.4 The expanding blast wave of an intense explosion

An interesting (although very grim) problem that yields to dimensional analysis is the analysis of the expanding blast wave that follows from the energy released into the atmosphere by an intense, localized explosion. The Trinity test of the first atomic bomb is an example of enormous historical

[41] Excluded from this analysis are animals that have a sprawled posture and support their weight mainly on their bellies, e.g., crocodilians and lizards. Recent and more comprehensive data collections reveal that there is, not unexpectedly, some systematic species dependence in the function (127). You may be pleased to know that humans are a little faster than the grand average reflected by (127).

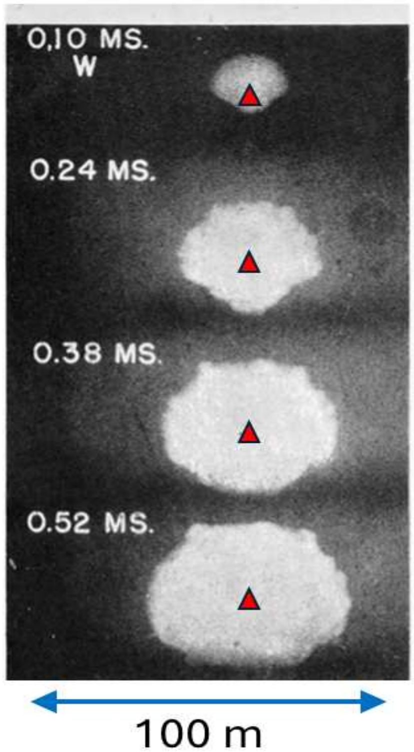


Figure 20: The 1945 Trinity test of the first atomic bomb as captured by very high-speed photography. Of this occasion, the project scientific director J. Robert Oppenheimer later recalled: '*W*e knew the world would not be the same. A few people laughed, a few people cried. Most people were silent.'

import, Fig. (20), documented in high-speed photographs from Taylor (1950).[42] These rather coarse images allow estimates of the radius, $r(t)$, of the rapidly expanding blast wave (for short times, coincident with the fireball visible in the available photographs):

time, t, 10^{-3} s, = [0.1 0.24 0.38 0.80 1.36 1.93 15.0 127.0]

radius, r, m, = [12 18 25 34 40 45 110 170]

Taylor's in-depth analysis of the phenomenon calculated the increase in pressure caused by a point-like release of thermal energy, and the subsequent expansion of the heated, high pressure air against the much smaller pressure of the ambient atmosphere. He started with the equations for momentum, mass and energy balance and boiled the problem down to a similarity solution for the expanding radius, r,

$$r(t) \;=\; F(\kappa)\left(\frac{E_0\;t^2}{\rho_a}\right)^{1/5} \tag{128}$$

where $\rho_a = 1.25$ kg m^{-3} is the nominal density of air. The parameter κ is the nondimensional ratio of the constant pressure specific heat of air and the constant volume specific heat of air, $\kappa \;=\; C_p/C_v$, which commonly arises in the analysis of acoustic phenomenon. For nominal

[42] Taylor, G. I., 1950, 'The formation of a blast wave by a very intense explosion: II'. The atomic explosion of 1945. Proc. R. Soc., London, A201, 175 - 190.

conditions, $\kappa \approx 1.4$, but of course an atomic blast is an extreme case that may involve fairly exotic effects. The bottom line is that Taylor's analysis found $F(\kappa) \approx 1$.

Here's a significant shortcut for this problem that omits the thermodynamic $F(\kappa)$ dependence and yet arrives at Taylor's essential result: solve for the (nondimensional) radius of the blast wave as a function of time, assuming that the only relevant parameters are the density of air, ρ_a, and the energy of the expanding blast wave, E_0. The key result is the power-law time-dependence of the radius. Does your result match (128)? And does this power law conform with the observations listed above? What do you infer for the energy parameter, E_0? [43]

Suppose that you did include the specific heat at constant pressure in the VPlist, $C_p \doteq$ [mass length time temperature] = [0 2 -2 -1], which has a dimension inverse temperature, and necessarily the specific heat at constant volume, C_v, same dimensions. What changes? [44]

7.5 Planetary orbits

> *It is one of the great tasks of science to learn what are accidents and what are principles, and about this we cannot always know in advance.*[45]
>
> Steven Weinberg, 2003.

The collaboration of the Dutch astronomer Tycho Brahe (1546 - 1601) and German polymath Johannes Kepler (1571 - 1630) was a milestone of the early Scientific Revolution. Tycho was an energetic and meticulous observer of the heavens who built the most precise observing instruments of the time, while also compiling historical, astronomical data from many sources. His protege, Kepler, was an accomplished mathematician whose worldview owed much to the classical Greeks and medieval scholastics. Kepler was strongly motivated by the mystical belief that the heavens must reflect a harmony and a geometric simplicity that evidenced creation by the Deity.[46]

---

[43] You wouldn't be confident in this analysis absent the much more detailed and comprehensive investigation of Taylor (1950), and, the opportunity to test against observations. Since this phenomenon is very far from common experience, it will be helpful to know that the E_0 estimated in this way is roughly 8×10^{13} Joules, which is equivalent to the energy released in the detonation of about 15 kilotons of TNT. This greatly simplified model ignores the work against the ambient atmosphere and hence may underestimate the actual thermal energy released. It also ignores the energy released as radiation that would have escaped the blast wave instantaneously, another likely bias toward a low total energy estimate.

[44] If you choose, you can see some of the results of this problem computed by the script DanalysisA2.m, Sec. 3.4.

[45] Weinberg, S., 2003, 'Can science explain everything? Can science explain anything?', Ch. 2 of *Explanations*, J. Crowell, Ed., Oxford Univ. Press. 23 - 38.

[46] Holton, G., 1956, 'Johannes Kepler's universe: its physics and metaphysics', *Am. J. Phys.*, Vol. XXIV, No. 5, 340 - 351. The literature on Kepler's remarkable life and accomplishments is voluminous and still growing. One of the better online references is https://www.aps.org/archives/publications/apsnews/201407/physicshistory.cfm and also https://plato.stanford.edu/entries/kepler/.

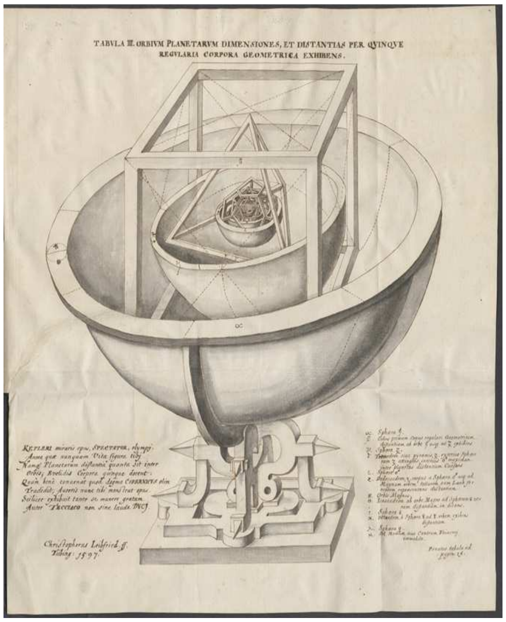


Figure 21: Kepler's Platonic solids model of the solar system. The outermost sphere represents the orbit of Saturn, and the next smaller sphere represents that of Jupiter. These are related by the inscribed cube which fits just inside the outer sphere, and encloses the inner sphere. Pure geometry thus defines the ratio of the radii of these spheres and, it was thought by Kepler, the orbital radii of Jupiter/Saturn. Given the freedom to choose the order of the remaining four solids, this scheme can be extended with some success to Mars and the inner planets. This model was in several respects a landmark in the Scientific Revolution; it brought the heavens down to Earth, and it supported a sun-centered solar system. Taken literally, it explained the number of planets that were then known, and it made a definite, quantitative prediction of the relationship among the planetary orbits that could be tested against observations.

Kepler came to prominence as a young mathematics professor when he published a geometric model of the solar system based upon the notion that the planetary orbital radii increased in proportion to the circumference of the five regular, Platonic solids nested on the Sun (Fig. 21).[47] The inspiration for this model seems to have come from the coincidence of there being exactly five Platonic solids, the number needed to account for the relative radii of the six known planets. The quantitative predictions of this model were within about 10% of the orbital estimates of Nicholas Copernicus — not convincingly right, but not clearly wrong.

7.5.1 Discovering a harmony in the heavens

The ambiguous match between Kepler's Platonic solids model and Copernicus' orbital estimates was no doubt part of Kepler's motivation to compute refined orbital estimates from Tycho's precise and extensive observations. The result was a significant advance in the precision of orbital estimates, with values close to modern (Table 3).[48] However, aside from Jupiter's orbital radius, the refined estimates were not very different from those made by Copernicus. Kepler's writings leave the impression that he had a clear understanding of what would constitute a meaningful departure of a model prediction from observed reality. And specifically, he found that the Platonic solids model was not a satisfactory fit to the more precise orbital estimates, as he must have hoped.

[47] This image is thanks to the Yale University Library https://search.library.yale.edu/catalog/1886333

[48] Kepler's data has been called to question for being a bit more precise than the observations should have allowed, https://www.nytimes.com/1990/01/23/science/after-400-years-a-challenge-to-kepler-he-fabricated-his-data-scholar-says.html

Historical and Modern Orbital Parameters

variable	Mercury	Venus	Earth	Mars	Jupiter	Saturn
Tycho and Kepler mean radius, Rp, AU	0.388	0.724	1	1.524	5.200	9.510
Copernican mean radius, Rp, AU	0.38	0.72	1	1.52	5.22	9.17
Modern mean radius, Rp, AU	0.387	0.723	1	1.52	5.200	9.54
Modern period, Tp, years	0.242	0.616	1	1.881	11.86	29.33
Modern planetary mass, Mp/Me	0.054	0.815	1	0.108	317.8	95.2
Modern eccentricity	0.205	0.006	0.017	0.093	0.048	0.054

Table 3: Estimates of orbital parameters by Tycho and Kepler, Copernicus, and modern values. These include the mean radius, Rp, and the period, Tp. Note that Tp is normalized with Earth's period, a year, and Rp is in units of Earth's mean radius, an Astronomical Unit, or AU. The absolute distance was not known in Kepler's era. Also included are modern estimates of planetary mass, Mp, in units of Earth's mass, Me, and the orbital eccentricity.[49]

Harmony among the planets of our solar system, if any exists, was not to be found in the relationship of the orbital radii.

A few questions for you. 1) Considering the six planetary orbital radii, Rp, of Table 3 available to Kepler, do you see a harmony i.e., some kind of regularity or pattern? 2) Evaluate the Platonic solids model given that a cube just touches the orbital sphere of Saturn at its eight corners, and encloses snugly the orbital sphere of Saturn (see Fig. 21). What is the ratio of the radii of the inner and outer spheres, and how does that compare with the ratio of the observed radii in Table (3)? Use both Copernican data and the Tycho/Kepler data. 3) Now consider the orbital radii Rp, and the orbital periods Tp, taken planet by planet; do you see a relationship Tp(Rp) that is in common to all six planets? (You may recognize your result as Kepler's third law of planetary motion.)

Kepler never gave up completely on the Platonic solids model, but his interests extended to modeling the detailed orbits of the individual planets, a key ingredient for forecasting astrology (a task that came with his post as a court astronomer). It was recognized that a circular, sun-centered orbit did not suffice for the eccentric orbit of Mars, and Kepler found that the *ad hoc* fix of adding one or a few additional suborbits (epicycles) was only a marginal improvement. After years of effort, Kepler discovered that an ellipse with the sun at one focus was a markedly better model of Mars' orbit, in terms of accuracy and simplicity. Moreover, he found that *all* of the planets moved on elliptical orbits, to within the high precision of the Tycho/Kepler analysis (Kepler's first law). Here, at last, Kepler found a genuine harmony, though of a different kind than he had sought with the Platonic solids model. Today we expect that this same law will hold for planets and asteroids in bound orbits found anywhere in the universe.

Kepler's analysis confirmed a crucial role for the Sun, and even more that the Sun acted to 'gather' planets to itself, i.e., that the Sun was the seat of an attractive force that decreased with distance. Despite that crucial insight, Kepler did not discover a physical mechanism for planetary motion before his death in 1630. Kepler's grave site has been lost, but his moving, self-written epitaph has survived:

I measured the skies, now the shadows I measure.
Skybound was the mind, Earthbound the body rests.

Johannes Kepler

7.5.2 An orbital mechanism

The mechanism behind elliptical orbits was clarified some fifty years after Kepler's death by Newton's theory of universal gravitation. You could just as well say that Kepler's laws of planetary motion imply the central force and the inverse square dependence on radius that is the heart of Newton's universal gravity. Newton's theory is amenable to a dimensional analysis. Suppose that the orbital period, Tp, depends upon gravitational interaction between the sun with mass Ms and a planet with mass Mp, their separation distance, Rp, and Newton's gravitational constant, $G \doteq$ [-1 3 -2]. What can a dimensional analysis tell us about the orbital period, Tp? Given the following numerical data, AU = 1.49 $\times 10^{11}$ m, G = 6.67 $\times 10^{-11}$ kg^{-1} m^3 s^{-2} and Ms = 1.9880 $\times 10^{30}$ kg, what is the similarity constant? (Hint; a common geometric factor.) There will be two nondimensional variables in this analysis. What is the dependence of the (nondimensional) period upon Mp/Ms? [50] If your result is indeed universal, then it should conform also with the orbital parameters of the first four moons of Jupiter discovered by Galileo, Table (4). Jupiter has a mass Mj = 1.8973 $\times 10^{27}$ kg.

Table 4: The Galilean moons of Jupiter

	Io	Europa	Ganymede	Callisto
Orbital radius, Rm, 10^3 km	421	671	1070	1882
Period, Tm, days	1.77	3.55	7.15	16.68

[49] Kepler's data are from C. M. Linton, 2004, *From Eudoxus to Einstein: A history of mathematical astronomy*, Cambridge U. Press, page 198. Copernicus' data is from https://pages.uoregon.edu/imamura/121/lecture-3/ Modern data are from https://nssdc.gsfc.nasa.gov/planetary/factsheet/joviansatfact.html

[50] If you choose, you can see some of the results of this problem computed by DanalysisA2.m (Sec. 3.4).

Index