Contents

1 An introduction to dimensional analysis

1.1 Introductory remarks about the course: Why these topics are hopefully interesting!

In introductory physics courses you learn basic ideas of *mechanics*, such as analyzing the motion of objects due to applied forces, the *conservation of (mechanical) energy* and the interconversion of kinetic and potential energies, and *thermodynamics*, which introduces the concepts of internal and thermal energy, entropy, etc. The concepts of force and energy balances are at the heart of the study of fluid mechanics. More generally, problems that involve **heat transfer**, e.g. thermal conduction, energy transfer by flow, and radiation, **mass transfer**, e.g. diffusion of a chemical or transport of a chemical by flow, and **momentum or stress** transfer, which is at the heart of the study of fluid motion, all can be analyzed using *conservation* or *balance laws*. Explaining, illustrating and applying the principles of heat, mass and momentum transfer will form the basis for the majority of this course.

Before we embark on the study of **transport phenomena**, however, we will introduce elementary ideas related to modelling and, in particular, we will emphasize the *dimensional* structure of equations and problem statements. The majority of this chapter focuses on developing a style of inquiry and thinking that takes advantage of the dimensions of the variables that appear in a problem statement. Many examples are provided and some important general principles are given, e.g. the significance of *dimensionless groups* and the Buckingham-Pi Theorem. In this course we will make frequent use of these ideas as we learn to characterize the kinds of problems that arise in transport processes.

In this chapter we will:

(1) In the remainder of this section we will motivate some of the physical kinds of problems for which we can use dimensional analysis to provide quantitative insight. The examples are described qualitatively and come from many different subject areas which should help highlight the manifold aspects of studying transport processes.

(2) In the second section we show the idea of dimensionless groups, we discuss the distinction between units and dimensions, and we give the classical example where the period of the simple pendulum is deduced via arguments based solely on dimensions.

(3) In the third section we give the Buckingham-Pi Theorem which states that the solution to any problem can be organized using a fewer number of dimensionless groups than the number of original variables in the problem statement.

(4) The fourth section gives many examples showing how dimensional argument lead to quantitative answers.

(5) The fifth section discusses the spring-mass-damper system, familiar from basic physics and mechanics courses, from several perspectives. In particular, we illustrate an important idea, which is how to make an equation dimensionless. The steps required are only elementary algebra and calculus but it is important to understand them, and to be able to carry them out easily, in order to get more insight into problems later in the semester.

(6) In Section 6 we simply give a detailed example for making a partial differential equation (from bending of a beam which you would have been introduced if you have taken ES120) dimensionless.

(7) A formal procedure for carrying out dimensional analysis using a *tabular method* for keeping track of dimensions is shown in Section 7.

(8) Finally, a brief discussion of dimensional analysis for problems involving electrical charge is given in Section 8.

(9) A long list of problems concludes the chapter.

An example of the potential importance of a result in scaling form: In some cases, simply recognizing the dimensional structure of an answer is important to draw significant conclusions; e.g. one example we will learn about when we study fluid flow in pipes concerns the relationship of the pressure drop Δp and the volumetric flow (volume/time) Q in a pipe of radius R. We will find

$$\Delta p \propto \frac{Q}{R^4} \tag{1}$$

Explain why an equation of the form $\Delta p \propto Q/R^4$ implies that a 10% reduction in radius, leads to approximately a 40% decrease in flow rate for the same Δp.

and the larger integer power on the radial dependence has important health implications: a 10% reduction in radius, perhaps produced by a diet high in fatty foods, leads to approximately a 40% decrease in flow rate (in this case, blood supply) for the same pressure drop maintained by your heart! By the end of this course we should understand why $\Delta p \propto Q/R^4$ for a laminar pipe flow of a viscous fluid.

There are many examples that help illustrate why thinking first about the *dimensional structure* of problems may prove useful. For example, why is it that the smallest cells are nearly spherical in shapes but large organisms, e.g. fish, mammals, etc. are more cylindrical, and some are shaped very much like aerofoils (e.g. an airplane wing)? Hint: The answer is linked to the surface to volume ratio which increases as the size of the object decreases. In a similar spirit, why are large soda bottles more commonly made out of plastic while smaller soda bottles are made out of thicker plastic or glass, and most canned beverages are made out of metal? Hint: The answer has to do with the transport of gases across the container's surface (this changes the properties of the liquid, in this case related to taste of the beverage) and so again involves thinking about implications of the surface to volume ratio.

Other examples involve modern technology. For example, we will study briefly how the operating speed of modern computers (gaming machines, laptops, the fastest processors) is a strong function of temperature that exhibits a decreasing speed with increasing temperature. Thus, the heat transfer characteristics of microelectronic components are crucial to their performance and in cases like this having lots of extra surface area to enhance heat transport is important. Again, there arises the significance of the surface area (where heat is lost) to volume (where heat is generated).

Some of the examples we will see during the course involve nature. For example, we will learn about beetles that when attacked excrete a liquid at the bottom of their feet because there is a substantial suction or adhesive force that arises when liquid is confined in a narrow gap between two surfaces. The origin of this macroscopic force is the surface tension that characterizes the forces or energy at a liquid-air interface; it is this same interfacial force that is responsible for moving fluid from the roots of plants and trees to the leaves (perhaps 100 meters high for the largest trees). Again, implicit in these examples will be the role of forces at surfaces. Thinking about these problems will be aided by having a good understanding of dimensional considerations when setting up problems.

The classic book on the topic of dimensional reasoning, appropriately titled "Dimensional Analysis" was written in the 1920s by P.W. Bridgman, a professor of physics at Harvard. The book was based on a series of lectures Bridgman gave, some of which, at least judging from the book, were probably like some of the lectures for this course! In addition, Lord Rayleigh, who is famous in mathematics and physics for many original research contributions, not the least of which was the discovery of the first noble gas, argon, for which he received the 1903 Nobel Prize in Physics, wrote a beautiful short article in

Nature in 1914 where he argues for the importance of dimensional reasoning. The article is given at the end of these notes and we will on occasion draw questions from Lord Rayleigh's list of examples. Along with Rayleigh's article, we include three other articles with a biological focus, which address the theme of *size and scale* as it impacts the mechanical understanding, and the relation of structure to function, in biological systems. These articles should make clear the power of using dimensional reasoning to help understand systems.

1.2 The basic ideas related to dimensional analysis

Here we introduce the idea of *dimensional analysis*, *rescaling* of variables and *nondimensional parameters*. We will use these ideas throughout the course. At the outset we will content ourselves with illustrating the spirit and style of the ideas and give a few examples. Then, in later sections we will provide a more systematic approach since there is a structure and a formal mathematical description that can be useful, especially in problems involving many variables.

In your previous science and mathematics courses, you have made mathematical models, developed equations, obtained analytical solutions, substituted lots of numbers into equations, etc. There was probably one idea that was overlooked when you had your first few courses and this concerns some *structural* features of any problem statement that generally *constrain* the form of the answer. Here we illustrate how the *dimensions* of the variables that appear when expressing a given problem or question provide valuable information about possible forms of the answer. In some cases, these ideas even, more or less, give the explicit quantitative answer!

Many students have the very (very) bad habit of always first substituting numbers into equations at the start of problem solving. You can never see the structure of any problem solution if you do this and you must learn to work symbolically.

1.2.1 A first example from introductory mechanics: a projectile

Example 1 - *Problem statement*: As illustrated in Figure 1 a projectile of mass m has initial height h above the ground. It is projected horizontally (x) with velocity v_0. *Find the horizontal displacement when the projectile hits the ground.* We neglect resistance from the air, but do note that later in this course we will learn how to calculate such resistance.

Before beginning the detailed solution presented next you should first ask: How many variables are involved in the problem description?

Here is how you would proceed in an introductory course. Label the horizontal and vertical coordinates as x and z, respectively, and let g indicate the gravitational acceleration (downward). Recall that in a uniform gravitational field the vertical motion $z(t)$ of the projectile follows from the algebraic equation

You should obviously be able to derive this result beginning with Newton's Second Law, $\mathbf{F} = m\mathbf{a}$.

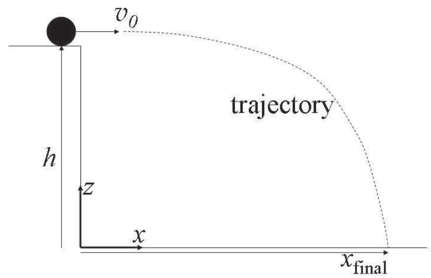


Figure 1: A ball projected horizontally with initial speed v_0. Find the final landing distance x_{final} as a function of the parameters in the problem, such as the initial height h, the initial speed v_0, and the gravitational acceleration g.

$$z(t) = -\frac{1}{2}gt^2 + h \ , \tag{2}$$

so that the object hits the ground ($z = 0$) at a time $t = \sqrt{2h/g}$. The corresponding final horizontal displacement x_{final} follows from $x(t) = v_0 t = v_0\sqrt{2h/g}$ or

$$x_{\text{final}} = \sqrt{\frac{2v_0^2 h}{g}}. \tag{3}$$

We observe that the projected distance increases with the square root of the initial height h and also that the mass of the object does not affect the result (provided we neglect air drag).

How is x_{final} changed if h increases by 30% or by a factor of 2?

These algebraic formulae, in one form or another, are typically how the answers are left in most books and courses. However, notice that given the phrasing of the question, a basic distance (or length scale) involved in the problem statement is the initial height h. So, how large is the final horizontal displacement x_{final} *relative* to the initial vertical displacement? Notice that the ratio of x_{final}/h is dimensionless since both the numerator and denominator are lengths. From the exact answer in equation (3) we can write

Recognize that h is a basic unit of length given in the problem statement!

$$\underbrace{\frac{x_{\text{final}}}{h}}_{\text{dimensionless distance}} = \sqrt{2} \underbrace{\sqrt{\frac{v_0^2}{gh}}}_{\substack{\text{dimensionless}\\ \text{parameter}}} . \tag{4}$$

This equation completely describes the solution in a compact and useful form. Notice that this result says that the projected distance relative to the initial height h is the **same** if the ratio v_0^2/gh is the same. We can then immediately observe that the ratio x_{final}/h is doubled if h is reduced by a factor of four or if v_0 is increased by a factor of 2. Such elementary "scaling" results carry a lot of quantitative information and do not necessarily require writing detailed equations to explain the prediction to a colleague.

Look again at equation (4). We refer to x_{final}/h as a *dimensionless parameter* since it is the ratio of two lengths, and v_0^2/gh is a second dimensionless parameter or *dimensionless group* since it is formed from a dimensionless combination of physical constants that appear in the problem statement. Since x_{final}/h is dimensionless we see that the final answer is a relationship between **two** dimensionless groups, while the original problem statement contained **five** dimensional objects $(x_{\text{final}}, h, g, v_0, m)$, which certainly seems more complicated. Thus, we conclude that (4) represents a simplification, at least in keeping track of information. We will emphasize such dimensionless representations, and how they present more meaningful characterizations of a problem statement and solution, throughout this course.

Verify that v_0^2/gh is dimensionless.

We will return to this problem shortly and explain how the basic functional form of equation (4), i.e. the dependence on two dimensionless parameters, though not the explicit formula involving a square root form, could have been predicted knowing only the *dimensions* of the quantities given in the problem statement.

A final take-away message about the functional dependence in problems: It is important to recognize that when the projectile problem was stated, you could have immediately concluded that the quantity you sought x_{final} was a function of the other quantities that appear, either explicitly or implicitly, in the problem statement. In this case, these quantities are h, g, v_0, and m. All of the quantities have dimensions and no matter what the form of the mathematical relationship between these variables, the final result must be consistent with respect to the dimensions. This idea of functional dependence is crucial to discussing the idea of dimensional analysis further.

1.2.2 Dimensions and units

It is useful to distinguish the *dimensions* of a given quantity from the *units* in which it is measured. In mechanics, which is concerned with force and the corresponding motion (e.g. velocity, acceleration), it is common to take

variable	dimensions[1]	systems of units
mass	M	kg, g, lb_m
time	T	sec, min, hr, years
length	L	m, cm, microns, ft, yard
area	L^2	m^2, ...
volume	L^3	m^3, ...
angle (=length/length)	-	radians, degrees
density	M/L^3	kg/m^3, lb_m/ft^3
velocity	L/T	m/s, cm/s, ft/min
acceleration	L/T^2	m/s^2, ft/s^2
force	ML/T^2	kg m $/\text{s}^2$=N, $\text{lb}_f = 32.2\ \text{lb}_m\ \text{ft/s}^2$
torque	ML^2/T^2	kg m^2/s^2
energy	ML^2/T^2	J = Nm; erg = g cm^2/s^2
power=energy/time	ML^2/T^3	J/s
pressure or stress	$M/L/T^2$	Pascal = N/m^2, psi = lb_f/in^2

Table 1: [1]Here we are considering mass (M), length (L) and time (T) as the fundamental dimensions. N = Newtons. You are probably familiar with the metric or SI sytem (the Systeme Internationale), MKS and CGS units, but not so familiar with the English system, which is basically only used today in the US and United Kingdom (and one or two other places). We will occasionally use the English system just to give you some exposure to it.

as *fundamental* dimensions mass (M), length (L) and time (T). Then, for example, by Newton's second law, $\mathbf{F} = m\mathbf{a}$, we know that acceleration has dimensions $[a] = L/T^2$ and so we speak of force (F) as a derived quantity with dimensions ML/T^2. It is convenient to introduce the bracket notation [.] to represent the "dimensions" of a quantity. For example, we would write $[F] = ML/T^2$.

It is helpful to be familiar with the many variables that arise in common science and engineering problems. Table 1 indicates in the first column different variables, in the second column the dimensions of the variable, and in the third column the different expressions for the variable using different systems of units. For example, in the MKS system, we report velocity in the units m/s, in the CGS system we report velocity in the units cm/s, and in the English, or British, system we report velocity in the units $feet/s$. It is a simple matter to convert between these different systems of units.

Conversion of units from one system to another system of units is important. For a recent (expensive) error of this type, see the next article about a NASA project.

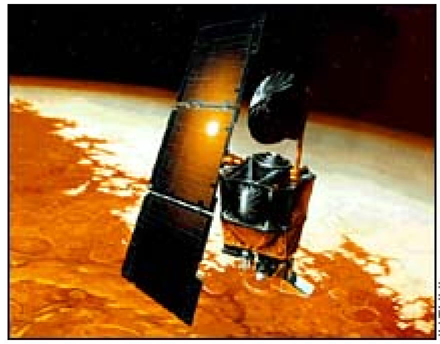
Metric mishap caused loss of NASA orbiter

September 30, 1999
In this story:

[Metric system used by NASA for many years]

[Error points to nation's conversion lag]

By Robin Lloyd
CNN Interactive Senior Writer



NASA's Climate Orbiter was lost September 23, 1999

(CNN) -- NASA lost a $125 million Mars orbiter because a Lockheed Martin engineering team used English units of measurement while the agency's team used the more conventional metric system for a key spacecraft operation, according to a review finding released Thursday.

The units mismatch prevented navigation information from transferring between the Mars Climate Orbiter spacecraft team in at Lockheed Martin in Denver and the flight team at NASA's Jet Propulsion Laboratory in Pasadena, California. Lockheed Martin helped build, develop and operate the spacecraft for NASA. Its engineers provided navigation commands for Climate Orbiter's thrusters in English units although NASA has been using the metric system predominantly since at least 1990. No one is pointing fingers at Lockheed Martin, said Tom Gavin, the JPL administrator to whom all project managers report. "This is an end-to-end process problem," he said. "A single error like this should not have caused the loss of Climate Orbiter. Something went wrong in our system processes in checks and balances that we have that should have caught this and fixed it."

The finding came from an internal review panel at JPL that reported the cause to Gavin on Wednesday. The group included about 10 navigation specialists, many of whom recently retired from JPL. "They have been looking at this since Friday morning following the loss," Gavin said. The navigation mishap killed the mission on a day when engineers had expected to celebrate the craft's entry into Mars' orbit.

After a 286-day journey, the probe fired its engine on September 23 to push itself into orbit. The engine fired but the spacecraft came within 60 km (36 miles) of the planet -- about 100 km closer than planned and about 25 km (15 miles) beneath the level at which the it could function properly, mission members said.

The latest findings show that the spacecraft's propulsion system overheated and was disabled as Climate Orbiter dipped deeply into the atmosphere, JPL spokesman Frank

O'Donnell said. That probably stopped the engine from completing its burn, so Climate Orbiter likely plowed through the atmosphere, continued out beyond Mars and now could be orbiting the sun, he said.

Climate Orbiter was to relay data from an upcoming partner mission called Mars Polar Lander, scheduled to set down on Mars in December. Now mission planners are working out how to relay its data via its own radio and another orbiter now circling the red planet. Climate Orbiter and Polar Lander were designed to help scientists understand Mars' water history and the potential for life in the planet's past. There is strong evidence that Mars was once awash with water, but scientists have no clear answers to where the water went and what drove it away.

NASA has convened two panels to look into what led to the loss of the orbiter, including the internal peer review panel that released the Thursday finding. NASA also plans to form a third board -- an independent review panel -- to look into the accident.

Metric system used by NASA for many years

A NASA document came out several years ago, when the Cassini mission to Saturn was under development, establishing the metric system for all units of measurement, Gavin said. The metric system is used for the Polar Lander mission, as well as upcoming missions to Mars, he said. That review panel's findings now are being studied by a second group -- a special review board headed up by John Casani, which will search for the processes that failed to find the metric to English mismatch. Casani retired from JPL two months ago from the position of chief engineer for the Lab.

"We're going to look at how was the data transferred," Gavin said. "How did it originally get into system in English units? How was it transferred? When we were doing navigation and Doppler (distance and speed) checks, how come we didn't find it?"
"People make errors," Gavin said. "The problem here was not the error. It was the failure of us to look at it end-to-end and find it. It's unfair to rely on any one person."
Lockheed Martin, which failed to immediately return a telephone call for comment, is building orbiters and landers for future Mars missions, including one set to launch in 2001 and a mission that will return some Mars rocks to Earth a few years down the line. It also has helped with the Polar Lander mission, set to land on Mars on December 3 and conduct a 90-day mission studying martian weather. It also is designed to extend a robotic arm that will dig into the nearby martian soil and search for signs of water. NASA managers have said the Polar Lander mission will go on as planned and return answers to the same scientific questions originally planned -- even though the lander will have to relay its data to Earth without help from Climate Orbiter.

Error points to nation's conversion lag

Lorelle Young, president of the U.S. Metric Association, said the loss of Climate Orbiter brings up the "untenable" position of the United States in relation to most other countries, which rely on the metric system for measurement. She was not surprised at the error that arose. "In this day and age when the metric system is the measurement language of all sophisticated science, two measurements systems should not be used," Young said. "Only the metric system should be used because that is the system science uses," she said. She put blame at the feet of Congress that she said has squeezed NASA's budget to the point that it has no funds to completely convert its operations to metric. "This should be a loud wake-up call to Congress that being first in technology requires funding," she said, "and it's a very important area for the country."

Remark 1: It is also useful on occasion to view the fundamental dimensions as force (F), length (L) and time (T). In this case, mass is a derived quantity with (do this as an exercise) dimensions FT^2/L.

Remark 2: Some properties involve temperature. In this case we use degrees Kelvin (K), as the appropriate dimension.

Remark 3: The gravitational acceleration g has dimensions L/T^2.

Exercise: What are the dimensions of power? heat capacity?

Exercise: In the quantum description of nature, Planck's constant h enters the quantization of different energy levels according to the equation $h\nu =$ energy, where ν is a frequency. What are the dimensions of h?

Exercise: For understanding the molecular character of matter and properties of materials, Boltzmann's constant k_B enters. This constant is defined so that k_BT is an energy. What are the dimensions of k_B?

1.2.3 Mathematical description of nature: dimensional consistency

Consider your entire experience studying quantitative subjects. When making comparisons, it is necessary to compare "like" with "like". Thus, we take it for granted that

(i) Equations must be dimensionally consistent. For example,

The equation could not be $x(t) = \frac{1}{2}gt^3$.

$$\underbrace{x(t)}_{L} = \frac{1}{2}\underbrace{gt^2}_{\frac{L}{T^2}T^2} \quad \Rightarrow \text{dimensions are: } L = L \tag{5}$$

(ii) A complete statement of a physical law must be independent of the **units** used for representation of quantities and measurement. If this were not true, then two observers using two different measurement systems could arrive at different conclusions to the same question. It is this fact that forces us to the conclusion that the solution to all problems must be expressible in dimensionless form, since it then follows that the results that are obtained will not depend on the choice of coordinate system.

Notice how we distinguish dimensions and units.

1.2.4 A second example from introductory mechanics: the pendulum

Example 2: A second mechanical example we learn as physics and engineering students (and maybe even the first one that makes a lasting impression) is the description of the oscillations of a pendulum consisting of a mass m at the end of a (massless) connector of length ℓ (see Figure 2). We seek the period of the oscillations, assume that the pendulum is initially inclined at an angle θ_0 from the vertical, and neglect friction with the air. The mathematical problem statement is to solve for $\theta(t)$ according to the ordinary differential equation

Can you begin with Newton's 2nd law and derive equation (6)?

$$m\ell\frac{d^2\theta}{dt^2} + mg\sin\theta = 0 \ , \qquad \theta(0) = \theta_0, \quad \frac{d\theta}{dt}(0) = 0 \ . \tag{6}$$

Without solving the problem we can ask about the period τ of the motion as a function of the parameters that appear in the problem statement: m, g, ℓ, θ_0 or

$$\tau = f(m, g, \ell, \theta_0) \qquad \text{... a relation involving 5 objects,} \tag{7}$$

where $f(\cdot)$ indicates "a function of". The basic premise of *dimensional analysis* is that any physical response must be expressible in a form that is independent

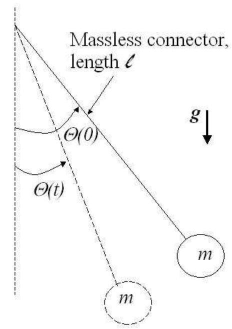


Figure 2: A simple pendulum.

of the system of units[1] used to make measurements or report results. Hence, we expect that the period τ must depend on a functional form of the parameters that has dimensions of time. Since the gravitational acceleration g, which has dimensions L/T^2, is the only variable on the right-hand side of (7) that involves time, then we are forced to conclude that the only possibility for the formula for the period is $\tau \propto \sqrt{\ell/g} f(\theta_0)$, where $f(\theta_0)$ is some function of the initial angle (f is not to be confused with the same symbol that appears in equation (7)). This result is a relationship among TWO *dimensionless groups*, which we emphasize by writing:

$$\frac{\tau}{\sqrt{\ell/g}} = f(\theta_0) \ . \tag{8}$$

Note that we do not require a detailed solution to deduce the dependence of the period on ℓ and g nor to determine that the period is *independent* of the mass m of the pendulum bob. On the other hand, also recognize that we cannot determine from these kinds of arguments the functional form $f(\theta_0)$ for the dependence of the period on the initial inclination angle θ_0.

A detailed calculation shows that the dependence of the period on θ_0 is very weak.

Important remark about graphical data: Suppose you were to obtain **experimental data** to test the pendulum result, equation (8). You would measure the period τ for different lengths ℓ, while maintaining the same initial angle θ_0. You would then plot τ versus ℓ. The *scaling* result (8) predicts that on a **log-log** plot the data will be linear! – this is a simple and useful result. Also,

Often the simple idea of making log-log plots of experimental results yields valuable information! From (8) then $\log \tau = \frac{1}{2} \log \ell + \text{constant}$, which means that a plot of $\log \tau$ versus $\log \ell$ is linear with slope 1/2.

[1]Remember to distinguish dimensions such as length (L), mass (M) and time (T) from specific units used for their measurement, such as lengths being reported in inches, feet, centimeters or meters. An angle is defined as the ratio of arclength to the radius of a circular arc, so angles (e.g. radians) have no dimensions.

it is one of the reasons that the mathematical form of power-law responses is relatively easy to check and seek in experimental data. For one real example, see problem 28 at the end of these notes.

> *Exercise*: Solve (6) using the small angle approximation.
>
> *Exercise*: If the initial condition has $\frac{d\theta}{dt}(0) = \Omega_0$, then using dimensional reasoning, what can you conclude about the functional form of the period? Using the small angle approximation, solve this new problem and confirm your answer.

1.3 The Buckingham-Pi (Π) Theorem

We might now expect that it is always possible to introduce *dimensionless variables* so that the description of a problem is *independent* of the specific units of measurement. A quantitative statement of the number of dimensionless variables (or groups) is expressed by the Buckingham-Pi Theorem:

> **Buckingham-Pi Theorem**: Given n variables that are expressible in terms of r independent dimensions, then there are no more than $n - r$ independent dimensionless variables.

Remark: It is common when discussing the subject in general terms to denote the different dimensionless variables as Π_1, Π_2, etc., hence the name of the theorem given by Buckingham. It is doubtful I will ever use this Π notation since I have never found it really helpful.

Typically in mechanics problems, the *independent dimensions are mass (M), length (L) and time (T)* so $r = 3$. If temperature is involved in the problem formulation, then $r = 4$ and if electrical effects (i.e. charge) are included then $r = 5$.

Let's reconsider one of the example above. For the oscillations of the simple pendulum, the period τ involves $\{\ell, g, m, \theta_0\}$. Hence, because an angle (being defined as the ratio of two lengths) is dimensionless, we see that $n = 5$ and $r = 3$, so there are two dimensionless groups. This conclusion is consistent with the answer in the form $\tau/\sqrt{\ell/g} = f(\theta_0)$.

Exercise: Reconsider the example discussed in section 1.2.1. What quantities are involved in the problem statement? Explain why the form of the answer given is consistent with the Buckingham-Pi Theorem.[2]

1.4 A series of examples illustrating (informally) dimensional analysis for arriving at the structure of answers

We now provide a series of examples where it is possible to arrive at the form of the answer *without* writing down any detailed equations. Rather, all that is required is to make a list of the variables and physical constants that appear in the problem statement and problem description and then demand that the final answer be dimensionally related to this set of parameters. We will arrive at most of the answers to within a multiplicative (dimensionless) constant, which hopefully will impress you since we will not have written down any detailed equations. Nevertheless, the only way to determine the unknown constants in the examples below is to either solve some detailed equation or do a single experiment. We will later explain more systematically how to obtain the dimensionless characterization implied by the Buckingham-Pi Theorem in a more systematic way.

(i) A ball of mass m is dropped from a height h and falls due to gravity g. Determine the time τ to contact the ground.

Answer: We expect that the time τ can at most depend on m, h and g. It is common to express the potential functional dependence as

$$\begin{aligned} \tau &= f(m, h, g) && (9a) \\ \text{dimensions: } T & \qquad M \;\; L \;\; L/T^2 && (9b) \end{aligned}$$

Note that the Buckingham-Pi result yields $4-3=1$ dimensionless parameter for the problem. What is this parameter? Since the time τ does not have dimensions involving mass, then m can't appear in the answer since it is the only parameter in the above list with dimensions involving mass. Because the only parameters remaining are h and g, we note that the *only* way to obtain a result with dimensions of *time* is $\sqrt{h/g}$. Hence, we see that the answer must have the form

NOTE: This result is ONE dimensionless parameter.

[2] Answer: Recall the first example problem that involved the final horizontal distance obtained by a projectile released with velocity v_0 from a height h above the ground. In that problem there were 5 variables and 3 independent dimensions (mass, length and time), so from the Buckingham-Pi Theorem we should expect that the problem solution should involve $5-3=2$ dimensionless groups. Indeed, that is the case as we see in equations (3) or (4).

$$\tau = c_1\sqrt{h/g} \qquad \text{or} \qquad \frac{\tau}{\sqrt{h/g}} = c_1 \ , \tag{10}$$

where c_1 is a (dimensionless) constant. The ratio $\tau/\sqrt{h/g}$ is dimensionless and represent the ratio of the actual fall time to the "natural time scale", $\sqrt{h/g}$ in the problem statement.

(ii) A ball of mass m is shot vertically in the air with initial velocity v_0. The gravitational acceleration is g. Determine the maximum height h to which the ball can rise.

Answer: The solution form is potentially

$$h = f(m, v_0, g) \tag{11}$$

Again, the Buckingham-Pi Theorem yields $4 - 3 = 1$ dimensionless parameter. Then, with dimensional considerations we conclude there can be no dependence on m. Since the only way to obtain dimensions of length is v_0^2/g (note that g has dimensions L/T^2, so we need the v_0^2 in the numerator to eliminate dependence on time), we conclude

$$h = c_2 v_0^2/g \qquad \text{or} \qquad \frac{h}{v_0^2/g} = c_2 \ , \tag{12}$$

where c_2 is a (dimensionless) constant.

(iii) Drive at speed v along a circle of radius R. Determine the acceleration a.

Answer: Since $a = f(v, R)$, then elementary dimensional inspection yields $a = c_3 v^2/R$. Also, we should reiterate that this form of argument does not determine the constant c_3, nor the sign of the acceleration (whether it is directed inwards or outwards).

(iv) Reconsider the example from section 1.2.1. We are interested in the final horizontal distance x_{final} achieved by the object projected horizontally from height h. Read the problem statement. Clearly, if the only variables involved in the solution are those mentioned in the problem statement, then at most we can expect

$$x_{\text{final}} = f(\underbrace{m}_{M}, \underbrace{h}_{L}, \underbrace{v_0}_{\frac{L}{T}}, \underbrace{g}_{\frac{L}{T^2}}) \tag{13}$$

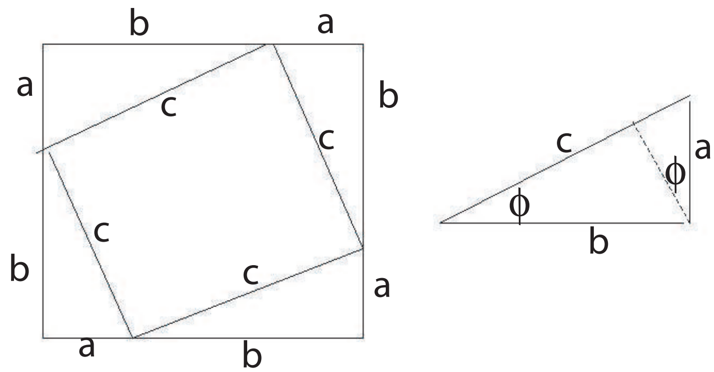


Figure 3: The Theorem of Pythagoras.

The Buckingham-Pi result yields the expectation that there should be $5-3=2$ dimensionless parameters, as we in fact obtained by a direct calculation earlier. To continue with a purely dimensional characterization, in this case, we can first conclude that the final result must be independent of m, since x_{final} has dimensions of length (L) and there are no other variables with dimensions of mass (M). So, we can expect the dependence of x_{final} on the other variables to be

$$x_{\text{final}} = f(h, v_0, g) \tag{14}$$

Now if we measure x_{final} relative to h then the only nondimensional description is

$$\frac{x_{\text{final}}}{h} = f_1\left(\frac{v_0^2}{gh}\right) , \tag{15}$$

where f_1 is some function (unknown at this stage). This result is precisely the functional form of equation (4) that was obtained by obtaining the exact solution of the trajectory equations. Nevertheless, even without knowing what is the precise functional form of the solution (i.e what is f_1 in equation 15), we note that simply arriving at the dimensionless form (15) is a significant step forward. For example, we can conclude that doubling v_0 has the same effect as decreasing g by a factor of four.

(v) *The Theorem of Pythagoras*: We have all learned that for a right triangle with hypotenuse c and sides a and b that $a^2 + b^2 = c^2$. One elementary proof is indicated in Figure 3 and takes advantage of a clever decomposition of a large square (area=$(a = b)^2$) into a smaller square (area=c^2) and four triangles (area=$4.\frac{1}{2}ab$), where the area of a triangle is half the product of the base and the height. Apparently, this simple proof was known to the Babylonians long ago. On the other hand we can prove this theorem using dimensional analysis (and a little bit of geometry), which is clever and surprising!

In order to proceed with a dimensional argument we consider a generic right triangle. It is well known from elementary geometry that given one sides and one of the two acute angles, then everything is known about the right triangle. Let us always choose the hypotentuse c and the smallest of the two acute angles, call it ϕ, as the basic variables. Now consider the area A of the triangle. We expect $A(c, \phi)$. Clearly, by dimensional analysis

A critical examination will show that this is a key assumption.

$$\text{area} \;=\; A = c^2 f(\phi), \tag{16}$$

where the function $f(\phi)$ is not known at this time (indeed, we do not need to know it as you will next see).

Next, subdivide the right triangle into two smaller right triangles, one with hypotenuse a and smallest acute angle ϕ, and the other with hypotenuse b and smallest acute angle ϕ (verify these statements). But we just learned that the area of these two triangles are, respectively, $a^2 f(\phi) + b^2 f(\phi)$. Summing the areas we find $A = c^2 f(\phi) = a^2 f(\phi) + b^2 f(\phi)$ or $c^2 = a^2 + b^2$. We have proved the Theorem of Pythagoras using dimensional analysis!

(vi) The energy of the atomic bomb: There is a famous, but true, story of an English physicist named G.I. Taylor, who in about 1947 deduced the approximate energy of the atomic bomb, which at the time was considered a secret known only to top-level officials in the US government. The simplified version of this deduction can be explained using dimensional analysis!

As the story goes, Taylor used a publicly available photograph of the nuclear explosion, published in Life Magazine, at the time a leading international publication, to complete his estimate. His letter to US officials inquiring as to the accuracy of his estimate was undoubtedly received with shock. Apparently, he was admonished for publishing his

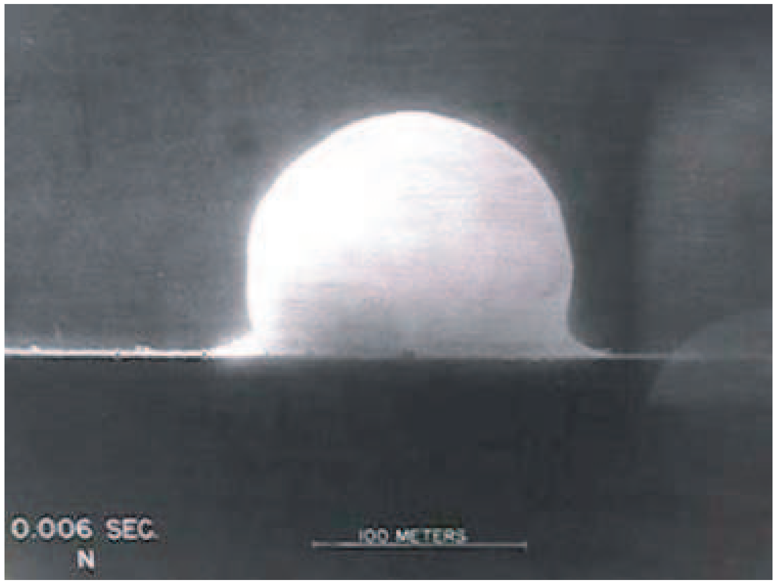


Figure 4: A photograph showing the nearly spherical explosion cloud from an atomic bomb.

results even though they were based on publically available information![3]

As for the analysis, we consider a spherical explosion of energy E into a medium (air) of density ρ (see figure 4). We ask for the radius r as a function of time t. Dimensional considerations (based on this admittedly simplistic problem characterization) then yield

$$r = f(t, E, \rho). \tag{17}$$

Since the dimensions of $[E] = ML^2/T^2$ and $[\rho] = M/L^3$, then $[E/\rho] = L^5/T^2$. We are then lead to the conclusion

$$r = c\left(Et^2/\rho\right)^{1/5}. \tag{18}$$

The scaling with the two-fifths power of time is in near perfect agreement with the data and the constant c can be deduced from graphing the experimental data. From this result, Taylor was able to estimate the energy E of the bomb.

A modern day applications: A few years ago there was a massive explosion, which killed 22 people, at a fireworks warehouse in the city of

[3] G.K. Batchelor, *The Life and Legacy of G.I. Taylor*, Cambridge University Press, 1996.

Figure 5: Left: Mass on a spring. Right: Mass-spring-damper system.

Enschede in the Netherlands. Apparently, in the resulting court case there was a question of the quantity of fireworks in the warehouse. It turns out that an amateur photographer had been videoing the city at the time of the accident, and the film and the "Taylor bomb calculation" were used to help determine who was actually telling the truth!

1.5 Mass-spring systems studied using dimensional reasoning

1.5.1 One more view of elementary dimensional analysis: Motion of a mass on a spring

Let us examine in different ways the solution of the simple harmonic oscillator of a mass m connected to a spring (spring constant k).

What are the units of k?

Again, using M, L, T as indicating dimensions.

Answer: force $= -kx \quad \Rightarrow \quad [\text{force}] = \frac{ML}{T^2} \quad \Rightarrow \quad [k] = \frac{[F]}{[x]} = \frac{M}{T^2}$.

Now we know that this system undergoes (possibly damped) oscillations. What is the typical period of the oscillation? In the case of damped vibrations, what is the typical time over which the damping occurs? We refer to both questions as asking about *time scales*. In the previous sentences we use the word "typical" to emphasize that we are obtaining a good estimate of the time of a physical process. We now attempt to answer the questions just posed about time scales.

The differential equation for $x(t)$: Let $x(t)$ denote the displacement of the spring from its equilibrium length. $x(t)$ satisfies the differential equation

$$m\frac{d^2x}{dt^2} = \sum \text{forces} = -kx \qquad \Rightarrow \qquad m\frac{d^2x}{dt^2} + kx = 0 \ . \tag{19}$$

We typically wish to solve this equation subject to initial conditions

$$\text{displacement :} \qquad x(0) = x_0 \qquad \text{and} \qquad \text{velocity :} \qquad \frac{dx}{dt}(0) = v_0 \ . \tag{20}$$

For the discussion that follows, we will choose $v_0 = 0$, so the spring begins at rest with displacement (from equilibrium) x_0.

Four ways to look at the idea of a *characteristic time*:

(i) What is the approximate period of the oscillation?

Answer by dimensional arguments alone: Let's denote the period as τ. Clearly the period τ must depend on the parameters in the problem. Therefore, if we first list the constants involved, we expect that $\tau = f(m, k, x_0)$ as m, k and x_0 are the only three physical constants that appear in the introduction above. Below we note the dimensions of each of the quantities:

$$\begin{array}{ccc} \tau & = & f(m, \ k, \ x_0) \\ T & & \ \ M \ \ \frac{M}{T^2} \ \ L \end{array} \tag{21}$$

In order to obtain the period, which has dimensions of time, we see that we must eliminate the mass and so $\sqrt{m/k}$ has dimensions of time. Thus we write

$$\tau \propto (m/k)^{1/2} \qquad \text{or} \qquad \tau = (\text{const})\,(m/k)^{1/2} \ . \tag{22}$$

We have really done very little work but have arrived at a useful result. *Important reminder*: This type of inspectional (dimensional) analysis can not determine the multiplicative constant (in this case 2π if we refer to the period of the motion) in the equation.

(ii) Of course, for the initial conditions $x(0) = x_0$ and $\dot{x}(0) = 0$, we can solve equation (19) exactly. You should verify that the solution is

$$x(t) = x_0 \cos\left[(k/m)^{1/2}\, t\right] \; . \tag{23}$$

From this exact solution we deduce that the exact period of the motion is $2\pi\,(m/k)^{1/2}$. It is thus natural to speak of the (undamped) spring mass system as having a 'characteristic time scale' $\sqrt{m/k}$ since this is the typical order of magnitude over which $x(t)$ changes (e.g. from its maximum to its minimum).

(iii) Another look at scaling by beginning with the governing equation. Consider the mass-spring differential equation:

$$m\frac{d^2x}{dt^2} + kx = 0 \; . \tag{24}$$

Let's only think about the order of magnitude of the terms in equation (24), and indicate their dimensions. We observe that when typical displacements, say Δx, occur on a time scale τ, then the dimensional estimates of each of the terms in equation(24) are

$$\underbrace{\frac{m\Delta x}{\tau^2} \;, \quad k\Delta x}_{\text{compare}} \qquad = 0 \quad \Rightarrow \quad \tau^2 \approx m/k \quad \Rightarrow \quad \text{again} \;\; \tau \approx (m/k)^{1/2} \tag{25}$$

Since equation (24) is *linear* in the displacement $x(t)$, then Δx does not enter the final estimate in (25).

(iv) *Nondimensionalize the equation*: It is possible to write (19) so that all the dimensional variables (and even the multiplicative constants) are eliminated. Let us define new, dimensionless variables according to

$$X = x/x_0 \;, \quad T = t/\sqrt{m/k} \tag{26}$$

Then, substituting into (24) and (20) we obtain an equation for $X(T)$:

$$\frac{d^2X}{dT^2} + X = 0 \;, \quad X(0) = 1 \;, \quad \frac{dX}{dT}(0) = 0 \; . \tag{27}$$

If you are to solve or plot the solution to this equation you will see that the magnitude of (changes in) X is less than or equal to 1 (we typically write $O(1)$, and pronounce the "big Oh" as "order"), when changes of T are $O(1)$.

It is important that you feel comfortable with beginning with equation (24) and arriving at equation (27). Please do not be alarmed by this language. It is simply useful in more complicated problems for expressing ideas related to approximate answers of complex problems.

1.5.2 A mass-spring system with damping

We have emphasized above that it is useful to express solutions in dimensionless form. We will give one additional example of this style of thinking applied to a differential equation familiar from your previous mathematics and physics courses.

Many basic models (biological, chemical, electrical, mechanical, thermal, physiological, acoustic) involve second-order, constant coefficient ordinary differential equations (ODEs). For example, a mass-spring oscillator with linear damping (coefficient ζ) has a displacement from equilibrium $x(t)$ that is described by the equation

$$m\frac{d^2x}{dt^2} + \zeta\frac{dx}{dt} + kx = 0 \ . \tag{28}$$

Let us assume that this equation is to be solved with the two initial conditions $x(0) = x_0$, $\frac{dx}{dt}(0) = 0$.

Upon a first inspection of this problem you might think that understanding the possible solutions requires independent specification of four physical constants: m, ζ, k, and x_0. In fact, this is **not** true: the structure of the solution only involves one (dimensionless) parameter which is the ratio of the physical constants.

Making a differential equation dimensionless – the steps involved in rescaling variables: Let $X = x/x_0$ and $T = t/t_c$, where t_c is a constant, with dimensions of time, which is to be determined. Substituting into equation (28) and rearranging slightly gives

Verify that you can follow this change of variables in the differential equation. It is a very common step.

$$\frac{m}{kt_c^2}\frac{d^2X}{dT^2} + \frac{\zeta}{kt_c}\frac{dX}{dT} + X = 0 \ , \qquad \text{with } X(0) = 1, \ \frac{dX}{dT}(0) = 0. \tag{29}$$

To make the coefficients of the equation simple, we see that it is convenient to choose $t_c = \sqrt{m/k}$ in which case the equation is

$$\frac{d^2X}{dT^2} + \underbrace{\frac{\zeta}{\sqrt{mk}}}_{=\Lambda,\ \text{dimensionless parameter}}\frac{dX}{dT} + X = 0 \ , \qquad \text{with } X(0) = 1, \ \frac{dX}{dT}(0) = 0. \tag{30}$$

Therefore, the basic physical problem for determining $X(T)$ only involves ONE ratio of physical constants, $\Lambda = \zeta/\sqrt{mk}$. The form of this parameter makes clear that increasing ζ by a factor of two is equivalent to increasing m or k by a factor of four.

In this example we considered the oscillations of a damped spring-mass system. Dimensional considerations have that the solution for $x(t; m, \zeta, k, x_0)$, which involves 6 variables, should require knowledge of $6-3=3$ dimensionless groups. This is exactly the meaning of obtaining the dimensionless solution, expressed functionally as $X(T; \Lambda)$, by solving equation (28).

Exercise: Consider instead equation (28) subject to the initial conditions $x(0)=0$, $\frac{dx}{dt}(0)=v_0$. Put the problem statement in dimensionless form.

1.6 Bending of a beam: An example of rescaling variables and equations (you do not have to know the physical situation to appreciate the steps involved)

As a further example of rescaling equations, consider the equation for the transverse oscillations of a thin elastic beam of length ℓ. If $u(x,t)$ is the transverse displacement and ρ the mass/length (assumed constant), with x the distance along the beam and t time, then you may take it as given that $u(x,t)$ satisfies the equation

$$\rho\frac{\partial^2 u}{\partial t^2} + B\frac{\partial^4 u}{\partial x^4} = 0 \ . \tag{31}$$

The constant B is known as the bending modulus. By inspection you should note that $[B/\rho] = L^4/T$. In fact, equation (31) is only valid in the limit that the deflections are small, which here means that $u_m/\ell \ll 1$. You can consider u_m/ℓ as a dimensionless parameter.

Typical boundary conditions for $t > 0$ are

Since the equation has four spatial derivatives, we need to specify four boundary conditions.

$$\text{fixed end:} \qquad u(0,t) = 0;\ \text{zero slope:}\ \frac{\partial u}{\partial x}(0,t) = 0 \tag{32a}$$

$$\text{zero force and torque at } x = \ell: \qquad \frac{\partial^2 u}{\partial x^2}(\ell,t) = 0 \ ,\ \frac{\partial^3 u}{\partial x^3}(\ell,t) = 0 \ . \tag{32b}$$

Also, we need two initial conditions and we consider an initial deformation:

Two initial conditions are needed since the equation involves two time derivatives.

$$\text{parabolic initial shape: } u(x,0) = u_m(x/\ell)^2 \ ,\ \text{zero initial velocity: } \frac{\partial u}{\partial t}(x,0) = 0 \ . \tag{33}$$

With this description u_m is the maximum displacement of the free end (see sketch). Let's now answer a few questions.

(a) Using the Buckingham-Pi theorem, how many dimensionless groups do you expect when you solve for $u(x,t)$.

Answer: We seek $u(x,t;\rho,B,u_m,\ell)$ so we expect $n-r = 7-3 = 4$ dimensionless variables. Note that if you had initially remarked that the equation only involves B/ρ, whose dimensions only involve length and time, then you would write $u(x,t;B/\rho,u_m,\ell)$, and as the list of variables does NOT involve mass, we would conclude that here there are only two independent dimensions (length and time; $r = 2$) in the variable set, so $n-r = 6-2 = 4$, as before.

(b) Nondimensionalize (31) and the boundary and initial conditions. In this direct way, verify your answer to (a).

Answer: Let $X = x/\ell$, $T = t/t_c$, and $U = u/u_m$. Here it seems natural to "measure" x and u relative to the "typical" scales in the problem statement, which, in this case are, respectively, the length of the beam ℓ and the maximum initial displacement u_m. The "time scale" t_c we shall next determine. Substitute these new variables into the problem statement to obtain an equation for $U(T)$. After some algebra we find

$$\frac{\partial^2 U}{\partial T^2} + \frac{B t_c^2}{\rho \ell^4} \frac{\partial^4 U}{\partial X^4} = 0 \tag{34}$$

Therefore, it is natural to choose $t_c = \ell^2 \left(\rho/B\right)^{1/2}$. This time t_c is essentially the period for the oscillation of the beam.We have learned from these simple steps that if you double the length of the beam the period is four times as long. The complete dimensionless problem statement is then:

$$\frac{\partial^2 U}{\partial T^2} + \frac{\partial^4 U}{\partial X^4} = 0, \tag{35a}$$

$$U(0,T) = 0; \ \frac{\partial U}{\partial X}(0,t) = 0, \ \frac{\partial^2 U}{\partial X^2}(1,T) = 0, \ \frac{\partial^3 U}{\partial X^3}(1,T) = 0 \tag{35b}$$

$$U(X,0) = X^2, \quad \frac{\partial U}{\partial T}(X,0) = 0. \tag{35c}$$

Indeed, we now see that ALL the variables have been scaled out of the problem, and we need now determine $U(X,T)$. This result corresponds to a relationship among THREE dimensionless variables, U, X and T. Since equation (31) required $u_m/\ell \ll 1$ (see brief discussion above) then u_m/ℓ is one dimensionless parameter and U, X, and T are three others, giving a total of four dimensionless parameters, which is what we expected based on the answer to (a).

1.7 A formal approach to dimensional analysis

As emphasized above, the basic feature of *dimensional consistency* of equations places constraints on the possible functional forms allowed for the answer. Although in the simplest problems, the functional form can be guessed, problems with many variables generally require a more systematic approach, which we outline here. There are two main steps: First we identify (list) all variables that appear in the problem statement, and second we organize these variables into dimensionless ratios by sequential elimination of dimensions (M, L and T in the case of the usual mechanics problem).

We give two examples here using the projectile problem and mass-spring-damper problem mentioned above. The approach shown here is purely formal, and not really very physical, so we will often find that while it is a useful first step, in more complicated problems some rearrangement of the final results can be helpful.

1.7.1 The projectile problem revisited

We recall the example of a projectile of mass m fired horizontally at speed v_0 from an initial vertical height h. We seek the final horizontal displacement when the projectile has hit the ground and expect the functional dependence

$$x_{\text{final}} = f(m, h, v_0, g) \ . \tag{36}$$

We next make a table of all of the variables distinguishing the variable in which we are interested from the other independent variables or physical constants or other parameters that appear in the problem statement. By each variable we provide in parentheses the dimensions. We then systematically eliminate one dimension at a time until all variables are dimensionless.

Thus, we have

x (L)	m (M)	h (L)	v_0 (L/T)	g (L/T^2)	remark
x (L)	-	h (L)	v_0 (L/T)	g (L/T^2)	mass can be eliminated
x/h	-	-	v_0/h (T^{-1})	g/h (T^{-2})	scale out h
x/h	-	-	-	$\frac{gh}{v_0^2}$	simplify above as $[(v_0/h)] = T^{-1}]$

The last line indicate that there are TWO dimensionless groups: x/h and gh/v_0^2. This approach does NOT provide the functional relation between these two independent dimensionless groups, so we can simply conclude

$$\frac{x}{h} = f\left(\frac{gh}{v_0^2}\right) \ . \tag{37}$$

It is VERY IMPORTANT to recognize that you do **not** know the function f from this kind of approach, so we do not know everything. However, you should recognize that the problem has been significantly simplified and you have learned some quantitative information.

1.7.2 The mass-spring-damper problem revisited

This example is similar to the previous one. We first have

$$x = f(t, x_0, k, m, \zeta) \ . \tag{38}$$

The Buckingham-Pi theorem allows us to expect $6 - 3 = 3$ dimensionless variables.

x (L)	t (T)	x_0 (L)	k (M/T^2)	m (M)	ζ (M/T)	remark
x/x_0	t (T)	-	k (M/T^2)	m (M)	ζ (M/T)	length can be scaled out
x/x_0	t (T)	-	k/m (T^{-2})	-	ζ/m (T^{-1})	scale out m
x/x_0	$t/\sqrt{m/k}$	-	-	-	$\frac{\zeta}{\sqrt{mk}}$	simplify above as $[\sqrt{k/m}] = T^{-1}]$

The last line indicates that there are THREE dimensionless parameters. We do not know the functional relationships among these three objects, so we conclude

$$X = \frac{x}{x_0} = f\left(\frac{t}{\sqrt{m/k}}, \frac{\zeta}{\sqrt{mk}}\right) \ . \tag{39}$$

1.8 Extensions to account for electrical charge and thermal effects

There are of course other physical effects to take into account when studying nature. For example, there are electrical and magnetic interactions. Lets consider electrical effects. In introductory physics, you would have seen the basic ideas introduced in two ways:

(i) The force F between two charges q_1 and q_2 separated by a distance r, with $\mathbf{e}_r$ the unit vector along the line of separation, is

$$\mathbf{F} = \frac{q_1 q_2 \mathbf{e}_r}{4\pi\epsilon_0 r^2} \tag{40}$$

The physical quantity ϵ_0 is referred to as the permittivity of free space.

$\epsilon_0 = 8.85\ C^2/Jm$

(ii) Maxwell's equation for electrostatics (no moving charges) states that the electric field $\mathbf{E}$ and the charge density ρ_e (charge per volume) are related by

$$\nabla \cdot \mathbf{E} = \frac{\rho_e}{\epsilon_0} \tag{41}$$

You should be familiar with the notation of the divergence of a vector field, $\nabla \cdot \mathbf{E}$.

In order to discuss electrical phenomenon we must then introduce a dimension of *charge*, which we shall denote Q. In the Systeme Internationale (SI) the fundamental unit of charge is the Coulomb and for example the charge on an electron is 1.6×10^{-19} C. Using (i) and the dimensions of force $[F] = ML/T^2$, we observe that the dimensions of the electrical permittivity are

$$[\epsilon_0] = Q^2T^2/ML^3\ . \tag{42}$$

In order to discuss many physical phenomenon it is then necessary to have as fundamental dimensions the list M, L, T and Q.

Exercise: Recall that the force $\mathbf{F}$ on a charge q in an electrical field $\mathbf{E}$ is $\mathbf{F} = q\mathbf{E}$. Use this fact and equation (41) to show that ϵ_0 has the dimensions given in (42).

1.8.1 An example with charge as an independent dimension

Consider the energy E of an electron in orbit about an atomic nucleus. The electron has mass m and charge e. For such subatomic particles it is necessary to use ideas from quantum mechanics which introduces Planck's constant h.

Note that the dimensions of Planck's constant can be easily deduced from the fundamental relation for electromagnetic effects that $E = h\nu$, where ν is the frequency (you should recall this formula from introductory chemistry and physics); so $[h] = ML^2/T$.

The above problem statement can now be expressed functionally in terms of the given variables:

$$E = f(m, e, \epsilon_0, h) \tag{43}$$

which is a relation among 5 variables. As there are 4 independent dimensions (M, L, T, Q), by the Buckingham-Pi theorem we expect $5-4=1$ dimensionless group. We can now determine the form of this single dimensionless group using systematic elimination of the basic dimensions, as in the matrix method introduced earlier.

Since the only variables involving charge are ϵ_0 and e we see that the final result must involve ϵ_0/e^2. Note that $[\epsilon_0/e^2] = T^2/ML^3$. So, in order to eliminate mass as a variable we can note that the functional form must be

$$\underbrace{\frac{E}{m}}_{L^2/T^2} = f\left(\underbrace{\frac{m\epsilon_0}{e^2}}_{T^2/L^3}, \underbrace{\frac{h}{m}}_{L^2/T}\right) . \tag{44}$$

Elimination of time as a variable then allows us to conclude

$$\underbrace{\frac{Em}{h^2}}_{L^{-2}} = f\left(\underbrace{\frac{\epsilon_0 h^2}{me^2}}_{L}\right) . \tag{45}$$

Both groups above have dimensions related only to length so we conclude

$$\frac{Em}{h^2} = \frac{c_1}{\left(\epsilon_0 h^2/me^2\right)^2} \qquad \Rightarrow \qquad E = c_1 \frac{me^4}{\epsilon_0^2 h^2} , \tag{46}$$

where c_1 is a pure number. This result only involves one dimensionless parameter because we can express it as $\frac{E}{me^4/\epsilon_0^2 h^2} = \text{constant}$.

1.8.2 Temperature

As a final remark if problems are considered that involve temperature or thermal energy, then it is necessary to introduce a dimension for temperature. It is common to denote this dimension as Θ. The SI unit for temperature is the Kelvin (K). We shall discuss this topic more when we begin heat and mass transfer.

1.9 Exercises

Using dimensional analysis answer the following questions.

0. In statistical physics and kinetic theory Boltzmann's constant k_B is the fundamental parameter that enters the discussion via the equation $k_B T$ = energy, where T is the absolute temperature. What are the dimensions of k_B?

1. A ball of mass m is thrown into the air with initial velocity v at an angle θ_0 from the horizontal. Find the functional form for the horizontal distance travelled and the time the ball is in the air.

2. Given the coupled ODEs, involving $x(t)$, $y(t)$ and the constants a, b, c and x_0,

$$\frac{dx}{dt} = ax - bx^3 y \quad \text{and} \quad \frac{dy}{dt} = cx \qquad \text{with} \quad x(0) = x_0 \ , \ \ y(0) = 0 \ , \tag{47}$$

nondimensionalize these equations by appropriate choices of scales for t, x and y. Clearly state the dimensionless equations and the dimensionless parameters that appear. You do not have to solve these equations.

3. The it Casimir effect is a quantum mechanical phenomenon that refers to a force between two uncharged adjacent surfaces in vacuum.[4] This effect arises owing to so-called vacuum fluctuations that produce fluctuations in the electric field. In order to explain this phenomenon it is necessary to use quantum mechanics (which introduces Planck's constant $\hbar$) and electrodynamics (which introduces the speed of light c). Now, consider two parallel plates separated by a distance d. Determine the pressure p in this system (i.e. it is common to refer to the pressure or force/area acting on the plates rather than the force itself). In other words, using dimensional analysis we consider $p = f(\hbar, c, d)$. Find a formula for p.

4. Consider an undamped spring-mass system where the restoring force is nonlinear in displacement x according to $F = -kx^3$. If the initial amplitude of the system is a_0, what is a typical time for the system's response?

[4]This effect was predicted by the Dutch physicist Hendrick Casimir in 1948 and was measured in 1986 by Steven Lamoreaux. For a recent article in *Physics World* on the Casimir effect see **http://physicsweb.org/article/world/15/9/6**.

5. Bouncing of an elastic ball: A ball of radius R, density ρ and elastic modulus E bounces off a table. For simplicity we assume the table is made of a material of the same density and elastic modulus as the ball. Assume that the sphere has velocity V when it contacts the table. Determine the functional form of the contact time τ, i.e. the time the ball remains in contact with the table. State clearly how the contact time varies with the radius of the ball. *Hint*: According to the Buckingham-Pi Theorem, how many dimensionless variables should you expect?

6. *Superstring theory*: A modern trend in theoretical physics has been the development of superstring theory, which seeks to provide a description of the elementary particles (quarks, etc.) and the forces that act between them (e.g. J.H. Schwarz *Physics Today*, 1987). Therefore, it is necessary to bring together electromagnetism, quantum mechanics and gravitation. These three aspects of the universe are characterized by the physical constants c, $\hbar$ and G, i.e. the speed of light, Planck's constant and the gravitational constant, respectively. Look up or otherwise determine the dimensions of these three constants. Using dimensional analysis determine the characteristic length, mass and time scales at which these three physical phenomena act together.

6. An object of mass m is on a string of length L and travels in a horizontal plane at speed v in a circle of radius R. Determine the angle θ the string makes with the vertical and the tension T (a force) in the string. Note that $\sin\theta = R/L$.

   Solution: $\theta = f_1(m, g, L, v)$ and $T = f_2(m, v, L, g)$. Now use dimensional analysis.

7. A particle of mass m falls from a height h along a curved path. At the end of the path is a spring with spring constant k. Determine the form for the compression of the spring Δx.

8. Newton's law of gravitational attraction of two masses m_1 and m_2, with center to center distance r, states that the attractive force $F = Gm_1m_2/r^2$. What are the dimensions of G?

9. A planet of mass m moves around the sun (mass $M \gg m$) along an elliptical path with average radius R (in practice, the trajectory is only a little different from a circle). How is period T for one rotation related to R, G and any other variables (the dimensions of G are determined in problem 8)? This is one of Kepler's three laws.

10. N springs, some with spring constant k_1 and the rest with spring constant k_2, form a complicated network. Determine the functional form of the effective spring constant k_{eff} for the network.

11. How much potential energy is stored in a spring (spring constant k) that has been compressed a distance δx?

12. One important applications area of fluid mechanics is the efficient mixing of materials. For example, imagine trying to mix a blob of blue dye in water. An important question to then address is the increase of the *interfacial area* between the dyed material and the liquid as the time for mixing increases.

    So, consider a sphere of radius a, which is being stretched at a rate $\dot{\gamma}$; note that $[\dot{\gamma}] = \text{time}^{-1}$. It should be familiar from your calculus course that a sphere has the smallest surface area for a given volume. So, stretching will increase the surface area. After time t estimate the change in the surface area. Hint: List the variables in the problem.

13. A spring has a nonlinear damping term:

$$m\ddot{x} + \zeta|\dot{x}|\dot{x} + kx = 0 \tag{48}$$

    Find the dimensions of ζ and estimate the damping rate for the spring-mass system.

14. *Euler's disk*: Place a coin on edge and spin it. After some time it is spinning with a very small angle α relative to the surface and you can observe the rotation frequency (or vibration frequency) Ω increase as the angle decreases until contact occurs. The motion depends on the mass m of the disk, gravity g, the disk radius a and the inclination angle α.

    (a) In the absence of any frictional effects (damping), Ω is related to α. Find the form of the relationship using dimensional arguments.

    (b) The motion eventually stops because of damping. Assume that the origin of the damping is from the viscosity μ of air. Viscosity has dimensions M/L/T. Using dimensional arguments determine the typical damping time.

15. *Coiling rope*:[5] A rope of linear density ρ (mass/length), radius a and stiffness or Young's modulus E (dimensions of force/area) is fed at speed

[5]Professor L. Mahadevan of DEAS has written a very nice paper on this subject!

v from a distance h above the ground. It is a common experience that such a rope tends to coil into a pile with a typical radius R. Denote the associated rate of coiling as ω. Suppose that you seek R or ω as a function of the other parameters in the problem statement (remember to include the gravitational constant g).

Use the Buckingham-Pi theorem to determine the number of independent dimensionless parameters. Then, find the complete dimensionless charaterization of the problem.

17. Hooke's law for a linear elastic solid states that stress σ is proportional to strain ϵ. For a bar pulled along its axis, the strain is defined as the change in length relative to the initial length, so $\sigma = E\epsilon$, where E is called Young's modulus. Determine the dimensions of Young's modulus E.

18. The sound speed c in a material arises due to compressibility, or how the density ρ changes with pressure p. In a gas we expect $c = f(p, \rho, \gamma)$ where $\gamma = c_p/c_v$, which is the ratio of specific heats. Using dimensional analysis find the functional form of c.

18. The Casimir effect was discussed briefly in problem 3. There is analogous effect when a colloidal suspension is placed between two plates a distance d apart. In a colloidal system, the particles undergo Brownian motion where the characteristic thermal energy is k_BT, where k_B is Boltzmann's constant. The two plates experience a force which is characterized in terms of the pressure p. Find $p(k_BT, d)$ using dimensional analysis.

19. From playing with soap films and soap bubbles, you are familiar with surface tension even if you have not tried to do problem solving with it. Surface tension γ characterizes the energy/area required to deform an interface. Equivalently, it can be considered the force/length that acts along a boundary. Determine the dimensions of γ.

For problems 20 and 21, read Rayleigh's article titled "The Principle of Similitude", which is reproduced in the notes following the appendix on dimensional analysis.

20. Justify Rayleigh's remark (page 1, second column) "The frequency of vibration of a globe of liquid, vibrating in any of its modes under gravitation, is independent of the diameter and directly as the square root of the density."

21. Justify Rayleigh's remark "The frequency of vibration of a drop of liquid, vibrating under capillary force, is directly as the square root of the capillary tension and inversely as the square root of the density and as the 3/2 power of the diameter."

    Remark: The dimensions of surface tension are discussed in problem 19.

    Note: For problems (22-25), which involve charge, we need the dimension of the electric charge [Q].

22. Use Gauss's law $\nabla \cdot \mathbf{E} = \rho_e/\epsilon_0$, where ρ_e is the charge/volume and the expression for the force on a charge to determine the dimensions of the permittivity ϵ_0.

23. Consider the typical frequency of radiation f emitted when an electron passes from one energy state to another. Note that $[f] = T^{-1}$ and so seek $f = F(m, e, \epsilon_0, h)$ where m is the mass of the electron, which has charge e and h is Planck's constant.

24. Niels Bohr proposed a model for the atom where electrons orbit the nucleus and are confined to specific radial positions, which depend on the energy of the electron. For the hydrogen atom, which has only one electron, the radial distance at which the electron orbits the nucleus is referred to as the Bohr radius (often denoted a_0). We expect a_0 to depend on the mass (m_e) and charge (e) of the electron as well as Planck's constant (h) and ϵ. Using dimensional analysis, determine the functional form for a_0.

25. In some applications, such as versions of ink jet printing, small droplets are charged so that they can be manipulated with an external electric field. However, it is also known that if the total charge, call it Q_D, on the drop is too large, surface tension γ cannot maintain the drop in a spherical form – instead the droplet shatters (it basically explodes). Determine the functional form for the maximum stable radius of a liquid drop R of charge Q_D.

    Hint: What variable list should you start with?

    Note: This critical radius (or the associated charge) is known as the Rayleigh limit.

26. Consider an elastic circular column of radius a. Denote the shear modulus (same dimensions as the elastic modulus E) as G. When a torque

τ is applied, the column twists an amount angular amount θ per unit length; denote this value per length by θ_ℓ. What are the dimensions of torque? What can you conclude by dimensional analysis?

Hint: Using the Buckingham-Pi theorem, how many dimensionless variables do you expect?

27. Read the handout"Gulliver was a bad biologist" by Francis Moog (*Scientific American*, 1948), which is attached with the articles on size and scale. The issue of "scale" is crucially important in many areas of biology since if an organism has typical linear dimension L, then we expect the surface area of the object to increase proportional to L^2 and the volume (or mass) to increase as L^3.

    Identify as many scaling type considerations as can in Moog's article. Do you know any more examples in "Gulliver's Travels"?

    Consider the example of the frequency f of your voice. In fact, the vocal cords work similar to a stretched string. So consider a string of density ρ, length ℓ, and cross-sectional area A, which is pulled by a force F. The linear density ρA is the important physical quantity when the string oscillates and the frequency of the transverse oscillations clearly depend on F (e.g. think of a guitar string). Show that the frequency of oscillation then must have the form $f \propto \ell^{-1}$. Is your result consistent or inconsistent with the discussion in Moog's article, where (p. 3) it is stated that "pitch ... varies inversely with the linear dimensions".

28. Consider the case of a bouncing drop of liquid. A contribution to the journal *Nature* illustrates that a liquid drop, of initial radius R, can "bounce" when fired at a surface with a velocity V (the article is attached). The pictures make clear that large deformations take place during and after the contact with the surface. Suppose that the droplet is characterized by the surface tension γ and the density ρ (we neglect viscous influences for this short time motion). (i) Use dimensional analysis to determine the variation of the contact time τ with the variables in the problem. (ii) The data shows that τ is independent of velocity. Hence, conclude why figure 2b in the *Nature* article, which is a log-log plot of the contact time versus radius, has a slope of 3/2.

29. A block of mass m is pushed with initial velocity v_0 along a horizontal plane. The frictional force resisting sliding is proportional to speed, $F_{\text{friction}} = -\zeta v$, where v is the instantaneous speed of the block and ζ

Am. J. Crit. Care Med. **149,** 1327–1334 (1994).
11. Clements, J. A., Hustead, R. F., Johnson, R. T. & Gribetz, I. *J. Appl. Physiol.* **16,** 444–450 (1961).
12. Stamenovic, D. & Wilson, T. A. *J. Appl. Physiol.* **73,** 596–602 (1992).

Competing financial interests: declared none.

Surface phenomena

Contact time of a bouncing drop

When a liquid drop lands on a solid surface without wetting it, it bounces with remarkable elasticity[1–3]. Here we measure how long the drop remains in contact with the solid during the shock, a problem that was considered by Hertz[4] for a bouncing ball. Our findings could help to quantify the efficiency of water-repellent surfaces (super-hydrophobic solids[5]) and to improve water-cooling of hot solids, which is limited by the rebounding of drops[6] as well as by temperature effects.

The way in which a water drop of radius R deforms during its impact with a highly hydrophobic solid depends mainly on its impinging velocity, V. The Weber number, $W=\rho V^2R/\gamma$, compares the kinetic and surface energies of the drop, where ρ and γ are the liquid density and surface tension, respectively. The greater the value of W, the larger are the deformations that occur during the impact (Fig. 1).

High-speed photography (Fig. 1) enabled us to measure the drop's contact time, τ. The frame rate could be greater than 10^4 Hz, allowing precise measurements of τ, which we found to be in the range 1–10 ms. As the impact is mainly inertial (with a restitution coefficient[2] as great as 0.91), τ is expected to be a function of only R, V, ρ and γ, and thus to vary as $R/V.f(W)$. For a Hertz shock, for example, the maximum vertical deformation, δ, scales as $R(\rho^2V^4/E^2)^{1/5}$, where E is the Young's modulus of the ball[7]. Taking a drop's Laplace pressure, $E\approx\gamma/R$, as an equivalent modulus and noting that $\tau\approx\delta/V$, we find for a Hertz drop that $f(W)\sim W^{2/5}$ and that the contact time varies as $V^{-1/5}$ and $R^{7/5}$.

Figure 2a shows that the contact time does not depend on the impact velocity over a wide range of velocities (20–230 cm s^{-1}), although both the deformation amplitude and the details of the intermediate stages largely depend on it. This is similar to the case of a harmonic spring, although oscillations in the drop are far from being linear. Moreover, this finding confirms that viscosity is not important here.

Figure 2b shows that τ is mainly fixed by the drop radius, because it is well fitted by $R^{3/2}$ over a wide range of radii (0.1–4.0 mm). Both this result and the finding shown in Fig. 2a can be understood simply by balancing inertia (of the order $\rho R/\tau^2$) with capillarity (γ/R^2), which yields $\tau\approx(\rho R^3/\gamma)^{1/2}$, of the form already stated with $f(W)\sim W^{1/2}$. This time is slightly different from the Hertz time because the kinetic energy for a solid is stored during the impact in a localized region, whereas in our case it forces an overall deformation of the drop (Fig. 1).

The scaling for τ is the same as for the period of vibration of a drop derived by Rayleigh[8], and is consistent with a previous postulation[9], although the motion here is asymmetric in time, forced against a solid, and of very large amplitude. Absolute values are indeed found to be different: the prefactor deduced from Fig. 2b is 2.6 ± 0.1, which is significantly greater than $\pi/\sqrt{2}\approx2.2$ for an oscillating drop[8]. Another difference between the two systems is the behaviour in the linear regime ($W<<1$): for speeds less than those shown in Fig. 2, we found that τ depends on V, and typically doubles when V is reduced from 20 to 5 cm s^{-1}, which could be due to the drop's weight[10].

The brevity of the contact means that a drop that contains surfactants, which will spread when gently deposited onto the solid, can bounce when thrown onto it; this is because the contact time is too short to allow the adsorption of the surfactants onto the fresh interface generated by the shock. Conversely, the contact time should provide a measurement of the dynamic surface tension of the drop.

Denis Richard*, Christophe Clanet†, David Quéré*

**Laboratoire de Physique de la Matière Condensée, URA 792 du CNRS, Collège de France, 75231 Paris Cedex 05, France*

e-mail: quere@ext.jussieu.fr

†*Institut de Recherche sur les Phénomènes Hors Équilibre, UMR 6594 du CNRS, BP 146, 13384 Marseille Cedex, France*

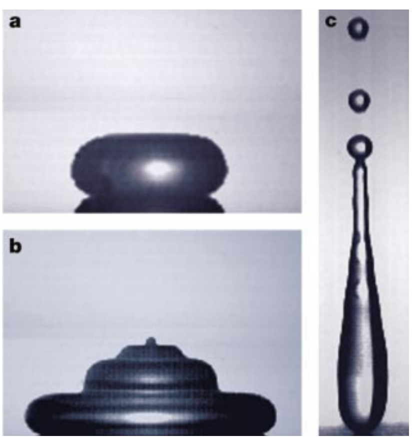


Figure 1 Millimetre-sized water drops with different Weber numbers (W) hitting a super-hydrophobic solid. W compares the kinetic and surface energies of the drop ($W=\rho V^2R/\gamma$, where R is the drop radius, V is the impact velocity, and ρ and γ are the density and surface tension, respectively, of the liquid). **a,** When W is close to unity, the maximum deformation during contact becomes significant. **b,** When $W\approx4$, waves develop along the surface and structure the drop. **c,** When $W\approx18$, the drop becomes highly elongated before detaching and gives rise to droplets; however, the contact time is independent of the details of the impact (see Fig. 2a).

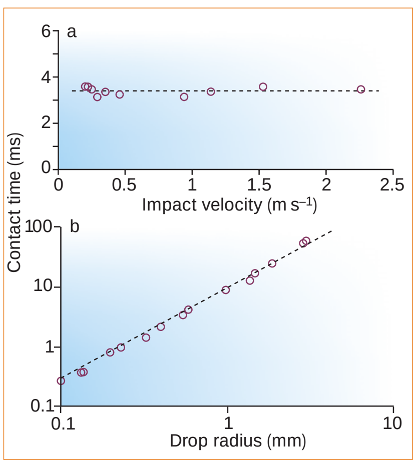


Figure 2 Contact time of a bouncing drop as a function of impact velocity and drop radius. **a, b,** In the explored interval (Weber number, W, between 0.3 and 37), the contact time is **a**, independent of the impact velocity, V, but **b**, depends on the drop radius, R. Dotted lines indicate slopes of 0 (**a**) and 3/2 (**b**).

1. Hartley, G. S. & Brunskill, R. T. in *Surface Phenomena in Chemistry and Biology* (ed. Danielli, J. F.) 214 (Pergamon, Oxford, 1958).
2. Richard, D. & Quéré, D. *Europhys. Lett.* **50,** 769–775 (2000).
3. Aussillous, P. & Quéré, D. *Nature* **411,** 924–927 (2001).
4. Hertz, H. *J. Reine Angew. Math.* **92,** 156–171 (1881).
5. Nakajima, A., Hashimoto, K. & Watanabe, T. *Monatshefte Chim.* **132,** 31–41 (2001).
6. Frohn, A. & Roth, R. *Dynamics of Droplets* (Springer, Berlin, 2000).
7. Landau, L. D. & Lifschitz, E. M. *Theory of Elasticity* 3rd edn (Pergamon, Oxford, 1986).
8. Rayleigh, Lord *Proc. R. Soc. Lond. A* **29,** 71–97 (1879).
9. Wachters, L. H. J. & Westerling, N. A. J. *Chem. Eng. Sci.* **21,** 1047–1056 (1966).
10. Perez, M. *et al. Europhys Lett.* **47,** 189–195 (1999).

Competing financial interests: declared none.

Evolutionary biology

Hedgehog crosses the snail's midline

According to the dorsoventral axis-inversion theory[1], protostomes (such as insects, snails and worms) are organized upside-down by comparison with deuterostomes (vertebrates)[2–5], in which case their respective ventrally (belly-side) and dorsally (back-side) located nervous systems, as well as their midline regions, should all be derived from a common ancestor[5]. Here we provide experimental evidence for such homology by showing that an orthologue of *hedgehog*, an important gene in midline patterning in vertebrates, is expressed along the belly of the larva of the limpet *Patella vulgata*. This is a constant; this model is appropriate when there is a thin lubricating film between the block and the plane.

(a) What are the dimensions of ζ?

(b) Use dimensional arguments to determine the horizontal distance ℓ the block moves before coming to rest.

(c) Consider instead the friction rule $F_{\text{friction}} = -\zeta v^{1/2}$. Again, use dimensional arguments to determine the horizontal distance ℓ the block moves before coming to rest.

30. Consider a mass m attached to a spring (spring constant k). Let the mass be given an initial velocity v_0.

(a) Use dimensional analysis to determine the time τ for the mass to reach its maximum displacement.

(b) Suppose instead that the spring is nonlinear with a force (F) versus displacement (x) relationship $F = -k_1 x^3$. Now, use dimensional analysis to determine the time τ for the mass to reach its maximum displacement.

(c) Return to the problem statement (a). Suppose that there is damping as well, such that $F_{\text{damping}} = -\zeta \dot{x}$, where ζ is a constant. Use dimensional analysis to determine the time τ for the mass to reach its maximum displacement. *Hint*: How many dimensionless groups are there?

31. A linear elastic brittle material is characterized by Young's modulus E. The material is used to construct a long narrow rod of length ℓ that is connected to a mass m. The rod is then is rotated at angular frequency ω. Use dimensional analysis to determine an expression for the angular frequency at which the rod will be observed to fracture. (State physically why such a response might be expected.)

ES123: *Introduction to Fluid Mechanics and Transport Processes*
Spring 2008

This handout includes four articles that focus on the issues of *size*, *scale*, and *dimensional characterization*. The first article uses the classic story *Gulliver's Travel* as a way to think about the plausibility, in terms of basic physics, of Gulliver and his various encounters. Haldane and Gould were two leading biologists who argue in their respective articles about the importance of size and shape, and the basic principles of physics, for understanding natural biological systems. The final article is by Lord Rayleigh, one of the greatest scientists, who provides a terse summary of dimensional reasoning and its use in recognizing the quantitative form of an answer.

The articles are

1. "Gulliver was a bad biologist" by F. Moog, *Scientific American* 1948.

2. "On being the right size" by J.B.S. Haldane

3. "Size and shape: The immutable laws of design set limits on all organisms" by S.J. Gould, *Natural History* 1974.*

    a. Note: Steven J. Gould (1941-2002) was one of Harvard's most distinguished and best known professors. He wrote widely about the (biological) sciences, both for scientists and non-scientists. For more information on Gould, see http://www.stephenjaygould.org/original.html

4. "The principle of similitude" by Lord Rayleigh, *Nature* 1915.

Scientific American, 1948

GULLIVER WAS A BAD BIOLOGIST

Jonathan Swift's famous fantasy gives the modern biologist an opportunity to reflect upon the way living things are tailored to their environment

by Florence Moog

WHEN Jonathan Swift's Captain Lemuel Gulliver first published his account of his remarkable adventures in undiscovered Pacific lands, his contemporaries appear to have responded with some skepticism. Their reluctance to believe in six-inch men, floating islands and educated horses is mirrored in Gulliver's overprotesting preface to the second edition of his now-famous *Travels*. Whether his contemporaries were impressed by his insistence that from consorting with the Houyhnhnms he had been able to rid himself of "that infernal habit of lying" common to Yahoos is doubtful. In any case the two centuries that have elapsed since then have seen the growth of a body of knowledge by which the improbability of the creatures of Gulliver-land may be translated into impossibility.

Much of this knowledge has been the direct concern of biologists, those present-day kindred of Gulliver's academicians of Lagado. Indeed, for a student of comparative biology Gulliver's book may serve as an unpremeditated textbook on biological absurdities and, as a corollary, on the nicety with which all living organisms are tailored to the physical conditions of their existence.

The most unlikely of Gulliver's inventions, the 60-foot Brobdingnagians, actually could have been explained away, long before the biologists got around to it, by a principle of physics first developed by Galileo almost 100 years before Gulliver's odyssey appeared. According to the principle of dimensional analysis, the weight of a system increases as the cube of its linear dimensions. The principle seems to have been well known to Gulliver's Lilliputians, for it was the means they used in calculating that Gulliver equaled 1,728 Lilliputians. Since six-foot Gulliver was 12 times as tall as a six-inch Lilliputian, they computed that he weighed as much as one Lilliputian times 12^3 ($12 \times 12 \times 12 = 1{,}728$). The weight of a 60-foot Brobdingnagian may be similarly calculated as 10^3 times that of a six-foot man, let us say 180 pounds times 1,000, which is 180,000 pounds or 90 tons!

No wonder Gulliver neglected to mention the Brobdingnagians' weight! No very acute insight into structural principles is needed to see that such a tremendous bulk could not be borne in a frame of human proportions. The upper limit of weight which a body built on the human pattern will carry is perhaps no more than the 500 pounds reached by an occasional eight- or nine-foot rarity. A greater bulk would necessitate a truly ponderous skeleton. The long bones of the legs would be shortened relatively to prevent their bending under their great burden; the head would become comparatively small, for reasons we shall look into later, but its larger absolute size would entail shortening and thickening of the neck; much of the increased weight would be taken up in the trunk, for the internal organs would have to undergo relative enlargement to provide adequate power to move the huge machine.

GIANT Brobdingnagians, here talking with tiny Gulliver, can be shown to be an engineering impossibility.

Examination of a few hoofed mammals will neatly illustrate this adaptation of form to mass. A gazelle of 150 pounds, for example, has a rather long neck on which is mounted a head which, though large in relation to the slight body, is small as tall animals' heads go; the heavier head of a 1,200-pound horse, though smaller in proportion to body size, requires a shorter and more powerful neck to support it; while the great head of a five-ton elephant, though not large in relation to the gargantuan body, is too heavy to afford the luxury of any noticeable neck at all. Similarly, the slender legs of the gazelle may constitute two thirds of the height at the shoulder, whereas the sturdy limbs of a plow horse are only about half the height, and the pillarlike props of the elephant not much more than one-third. The mind shrinks from picturing the broad-beamed corpulence of a Brobdingnagian. In fact we need no more than the zoo-keepers' rule that once around the forefoot of an elephant is half the height of the body to make it clear that the delicacy of the feminine ankle must have been a matter of no interest in Brobdingnag.

We need have no fear of ever finding such a neckless, short-legged monster peering into our sixth-story windows, for no 90-ton animal could ever walk on dry land. Certainly such bulk could not walk on the arched structure of the human foot, which is too ready to flatten under a little additional strain in normal-sized people. A flesh-and-bone foot ten times longer than the normal human one and a thousand times heavier would have as much difficulty supporting even itself as would a covered bridge enlarged to span the Mississippi. Mount 90 tons on it and such a foot would require bones of steel bound by ligaments of wire cable.

The limiting strength of living tissues, especially muscle and connective tissues, is probably the reason why nature, in millions of years of experimentation, only once succeeded in designing a land animal even half as ponderous as a Brobdingnagian; this animal was a now long-extinct rhinoceros. The tremendous dinosaurs, in their vain attempt to make muscles outweigh brains, may have approached a Brobdingnagian weight, but they were not strictly land animals; they lived in swamps, sharing their burden with the buoyant water as the whale does today.

If the Brobdingnagians were too big to exist, the mouse-sized Lilliputians were too small to be human. So long as the laws of physics and chemistry obtain, living cells cannot vary much in size. Hence large animals must be built of more cells than small ones. In many organs cell number is not important, but the brain is in this respect like a telephone system (*see page 14*): a small private telephone system is of limited usefulness compared with one in which a great number of individual units, with their connecting wires and central switchboard equipment, make it possible for any person or institution to get in quick touch with any other. The human cortex, which is the portion of the brain that receives sensory information and deals intelligently with it, has an estimated 14 billion cells. On the inconceivably numerous interconnections which keep this vast assemblage of units in touch with one another depend the adaptability and educability of the human being. Were this tremendous number of cortical cells to be much reduced, the apparently inexhaustible capacity of good human minds for learning, remembering, perceiving and thinking would wither; it would shrink perhaps to the low level of defectives in whom the brain is cramped by a "pinhead" skull or by the abnormal presence of fluid in the cranium.

Now if we allow to a Lilliputian nerve cells as large as those of a mouse (which have about one-fourth the volume of human nerve cells), his tiny cranium could accommodate something like 35 million cortical cells—a large number indeed, but only a small fraction of what even a chimpanzee has at his disposal. On such a small allotment of intellectual equipment the Lilliputians could never have devised their delightful court routine, which yielded nothing in intricacy or absurdity to the best that Augustan England had to offer.

The Lilliputians also would have needed disproportionately large heads to carry useful eyes. Anyone who has ever quizzically scanned an elephant, trying to determine just where the enormous beast has its diminutive eye, must realize that eye size varies far less than body size. A small animal seems to have too much eye in relation to the expanse of its head, and this is because the limits of eye magnitude are dictated by the physical properties of light, which must enter through a pupil not too small, and must impinge on a sufficient number of seeing elements—the rod- and cone-shaped cells of the retina

TINY Lilliputians have various logical drawbacks, among them the difficulty of constructing an eye that would fit their small heads. The drawings on this and the opposite page are from a 1768 edition of the works of Swift.

—of almost invariable size. The eye is thus in a sense a doorway which must admit a certain minimal amount of light, but need not admit much more; just as an architectural doorway, whether of a cottage or a mansion, must be big enough to admit a man, but need not be much bigger.

So if a Lilliputian had had a head large enough to hold an intelligent brain and serviceable eyes, he would have needed a heftier body to hold up the head. The smallest known human race, the African pygmies, stand four and a half feet high. Even the tiniest human dwarfs, who never achieve quite correct physical proportions, are almost always more than two feet tall.

But overlooking for the moment the matter of unaccountable intelligence in unreasonably small heads, let us take note of another Lilliputian character that casts doubt on their creator's veracity. The voices of the miniature people, we read, were shrill—a shrewd guess, but not shrewd enough. Had Gulliver considered the difference in pitch which a small difference in size makes among the members of the viol family, he might have been more cautious about assigning any audibility to the Lilliputian voice. Pitch, measured in cycles per second, varies inversely with the square of the linear dimensions of the vibrating surface. So the vocal cords of a six-inch human being would vibrate 144 times faster than those of a six-footer; allowing our voices to center comfortably at 256 cycles per second (middle C), the small voice would vibrate at about 37,000 cycles per second—more than seven octaves higher! This would not inconvenience a Lilliputian, whose ears would probably have a sensitivity proportioned to his voice, but Gulliver's ears must have been practically deaf to sound of more than 10,000 cycles per second. Even had the captain been a prodigy of aural acuity he could not have heard a Lilliputian voice any more than we can hear the cries that bats seem to utter constantly as they fly; to such sounds only the ears of their small confreres are attuned.

With Brobdingnagians the case is no better, for their voices must have been at the lower limit of audibility—averaging perhaps three cycles per second. At such a rate the vibrations, though they might be heard, would not merge into a continuous sound, but would seem like the sad undulations of a phonograph record dragging to a stop. Now and then a Brobdingnagian soprano or piccolo player might have produced some notes that Gulliver could hear normally, but the sensation could hardly have been pleasant, for the lowest (as well as the highest) tones, to which our ears are sensitive can be heard only when they are so loud that they can also be felt. Thus Gulliver was doubly wrong in claiming that he improved the giants' music by retreating from its loudness: most of it would not have sounded like music at all, and those few notes that might have had the earmarks of music would have become inaudible as soon as they became painless.

The apparently modest appetite of the Lilliputians is another reason why we may doubt their existence. Small, warm-blooded animals must take in more calories per unit of body weight than large ones, for the rate of living varies with size; small lungs breathe faster, small hearts beat faster, small bodies consume oxygen and turn out waste products faster. So a six-inch man might be expected to have about the same food requirements as a mouse,

PYGMIES of Africa are good evolutionary approximation of the smallest practical size for human being.

that is, approximately eight times as many calories per ounce of body weight as a full-scale man needs—or 24 meals a day instead of three. Gulliver missed on two counts here; he failed to realize that the creatures of his invention would have spent the larger part of their time stuffing themselves with food, and by the same token he did not see that by allowing him 1,728 times their dietary they were giving him as much food in a day as he could conveniently eat in a week.

Indeed, had Gulliver known anything of differential metabolism, his concept of a Lilliputian humanity would have been altered in every respect. For it is not difficult to see that an animal that has to provide itself with the equivalent of 24 of our meals a day would not have time enough left over for developing the nicer aspects of civilization. Worse than that, the very duration of life is related to size; an elephant may see a century pass, but fast-burning, voracious little mammals run through their lives in a space of time too brief to allow for the sort of education on which civilized society depends.

It must be concluded that the author of the Houyhnhnm hoax could have had no adequate appreciation of the physical characteristics from which human life has sprung. From the evolutionary point of view man is in essence a tall anthropoid whose big head accommodates a sizable brain and is provided with forward-looking eyes that can be used stereoscopically; he stands easily erect on flat, supple feet and carries at the end of his long, free-swinging arms a pair of instruments so beautifully designed, so perfectly adapted to uses without number, that even the products of his clever brain have never equaled them.

Our remotest apelike ancestors may not have been much more attractive than the Yahoos of Houyhnhnmland, but they were able to take the road to civilization because they were physically equipped for the journey. Their equipment did not include manual dexterity or conceptual thought, for these were goals farther along the road. Primitive man, their descendant, could enlarge the use of his brain partly because his hands, shaped by the tree-living habit of his forebears and freed by the erect posture derived from the same habit, could pick up and handle what his excellent eyes—another product of arboreal life—wished to examine; to this day the tendency to touch what we see is almost instinctive in most of us. Eyes and hands thus produced material on which the brain could work. The ideas that began to come in their turn provided work for the hands and eyes: the rock that was used as a weapon would be shaped into a blade and then mounted on a shaft; the thick-growing boughs that provided shelter would be used to construct a hut; the skins that served for clothing would be fashioned with bone needles and gut threads. Through innumerable such cycles of action and thought and action, civilization rolled slowly on toward atom bombs and prefabricated houses and the garment workers' union.

If the indispensability of our apelike (or Yahoo-like) appointments in the development of civilization is not at once apparent, a suitable appreciation of them can be readily derived from a critical examination of Gulliver's educated horses. Even if we overlook the horse's low-browed construction, with a head so weighted down with jaws that it can hardly indulge in the luxury of a big intellectual brain, and even if we credit horse-lovers' accounts of the remarkable intelligence of the much-admired animal, we must still draw the line at Gulliver's stories of horses cooking oats, grinding flints and building houses.

An unlettered stableboy would snort at the notion of a horse threading a needle

—even if the limbs could do the trick, both eyes would be looking the wrong way! For the horse, having neither offensive weapons nor the hands to make them with, must tend first of all to his own safety; his eyes are placed to sweep a wide area for signs of danger; his limbs are designed to carry him swiftly away should danger materialize. To this end the limbs are so fixed that they swing freely in only one direction, and the body is so mounted on them that the horse, unlike certain small-clawed quadrupeds, cannot disengage his forefeet from their normal task even by sitting down. Nor is this any real disadvantage now, for the five fingers that might have led to the ability to grasp and manipulate objects were long since paid by the ancestors of the horse as the price of the stout single digit that makes such a superb running instrument. This exchange is typical of the law of compensation that nature rigidly enforces; only by refraining from specialization that might enable it to do anything perfectly has the flexible five-fingered human hand on its loose-jointed arm retained the ability to do everything after a fashion. A similar rule holds for the mind.

This brief catalogue by no means exhausts Gulliver's crimes against nature. The alterations which an increase or diminution of body size would enforce on the heart, lungs, liver and intestines could be made the subject of a large treatise. As for larks so small that they were provided with invisible feathers, eagles so big that they could not possibly pack enough power to get themselves into the air, insects far beyond the size limits imposed by the simple respiratory mechanisms of their kind—the very blades of grass in Gulliver's fantastic kingdoms cry out their impossibility. No biological principle is clearer than that every living thing—from man with his rapaciously expanding control over the environment, to the patient, insensible slime mold—lives in harmonious adjustment to the conditions of its life.

But, after all, we must not be too hard on Gulliver for failing to understand the biological conditions that made him a man—and an implausible liar. His talents, like those of his friend and teacher, the unhappy Dean Swift of Saint Patrick's, were in the psychological realm. The etymology of the name Houyhnhnm, his master horse tells us, is "the perfection of nature." Gulliver may not have understood biology, but he did understand biologists, who after his time were to endow their own species with the somewhat wishful name, *Homo sapiens*.

Florence Moog, winner of the AAAS-George Westinghouse Award for the best magazine science article of 1948, is assistant professor of zoology at Washington University in St. Louis.

WATUSI of Uganda, in East Africa, are among the world's tallest people. Many of them exceed a height of seven feet. Although there have been other kinds of big human beings, probably few were much taller than the Watusi.

On Being the Right Size[†]

by J. B. S. Haldane

The most obvious differences between different animals are differences of size, but for some reason the zoologists have paid singularly little attention to them. In a large textbook of zoology before me I find no indication that the eagle is larger than the sparrow, or the hippopotamus bigger than the hare, though some grudging admissions are made in the case of the mouse and the whale. But yet it is easy to show that a hare could not be as large as a hippopotamus or a whale as small as a herring. For every type of animal there is a most convenient size, and a large change in size inevitably carries with it a change of form.

Let us take the most obvious of possible cases, and consider a giant man sixty feet high - about the height of Giant Pope and Giant Pagan in the illustrated Pilgrim's progress of my childhood. These monsters were not only ten times as high as Christian, but ten times as wide and ten times as thick, so that their total weight was a thousand times his, or about eighty to ninety tons. Unfortunately the cross sections of their bones were only a hundred times those of Christian, so that every square inch of giant bone had to support ten times the weight borne by a square inch of human bone. As the human thigh-bone breaks under about ten times the human weight, Pope and Pagan would have broken their thighs every time they took a step. This was doubtless why they were sitting down in the picture I remember. But it lessens ones respect for Christian and Jack the Giant Killer.

To turn to zoology, suppose that a gazelle, a graceful little creature with long thin legs, is to become large, it will break its bones unless it does one of two things. It may make its legs short and thick, like the rhinoceros, so that every pound of weight has still about the same area of bone to support it. Or it can compress its body and stretch out its legs obliquely to gain stability, like the giraffe. I mention these two beasts because they happen to belong to the same order as the gazelle, and both are quite successful mechanically, being remarkably fast runners.

Gravity, a mere nuisance to Christian, was a terror to Pope, Pagan, and Despair. To the mouse and any smaller animal it presents practically no dangers. You can drop a mouse down a thousand-yard mine shaft; and, on arriving at the bottom it gets a slight shock and walks away, provided that the ground is fairly soft. A rat is killed, a man is broken, a horse splashes. For the resistance presented to movement by the air is proportional to the surface of the moving object. Divide an animal's length, breadth, and height each by ten; its weight is reduced to a thousandth, but its surface only a hundredth. So the resistance to falling in the case of the small animal is relatively ten times greater than the driving force.

An insect, therefore, is not afraid of gravity; it can fall without danger, and can cling to the ceiling with remarkably little trouble. It can go in for elegant and fantastic forms of support like that of the daddy-longlegs. But there is a force which is as formidable to an insect as gravitation to a mammal. This is surface tension. A man coming out of a bath carries with him a film of water about one-fiftieth of an inch in thickness. This weighs roughly a pound. A wet mouse has to carry about its own weight of water. A wet fly has to lift many times its own weight and, as everyone knows, a fly once wetted by water or any other liquid is in a very serious position indeed. An insect going for a drink is in a great danger as man leaning out over a precipice in search of food. If it once falls into the grip of the surface tension of the water – that is to say, gets wet – it is likely to remain so until it drowns. A few insects, such as water-beetles, contrive to be unwettable; the majority keep well away from their drink by means of a long proboscis.

Of course tall land animals have other difficulties. They have to pump their blood to greater heights than a man, and, therefore, require a larger blood pressure and tougher blood-vessels. A great many men die from burst arteries, greater for an elephant or a giraffe. But animals of all kinds find difficulties in size for the following reason. A typical small animal, say a microscopic worm of rotifer, has a smooth skin through which all the oxygen it requires can soak in, a straight gut with sufficient surface to absorb its food, and a single kidney. Increase its dimensions tenfold in every direction, and its weight is increased a thousand times, so that if it to use its muscles as efficiently as its miniature counterpart, it will need a thousand times as much food and oxygen per day and will excrete a thousand times as much of waste products.

Now if its shape is unaltered its surface will be increased only a hundredfold, and ten times as much oxygen must enter per minute through each square millimeter of skin, ten times as much food through each square millimeter of intestine. When a limit is reached to their absorptive powers their surface has to be increased by some special device. For example, a part of the skin may be drawn out into tufts to make gills or pushed in to make lungs, thus increasing the oxygen-absorbing surface in proportion to the animal's bulk. A man, for example, has a hundred square yards of lung. Similarly, the gut, instead of being smooth and straight, becomes coiled and develops a velvety surface, and other organs increase in complication. The higher animals are not larger than the lower because they are more complicated. They are more complicated because they are larger. Just the same is true of plants. The simplest plants, such as the green algae growing in stagnant water or on the bark of trees, are mere round cells. The higher plants increase their surface by putting out leaves and roots. Comparative anatomy is largely the story of the struggle to increase surface in proportion to volume.

Some of the methods of increasing the surface are useful up to a point, but not capable of a very wide adaptation. For example, while vertebrates carry the oxygen from the gills or lungs all over the body in the blood, insects take air directly to every part of their body by tiny blind tubes called tracheae which open to the surface at many different points. Now, although by their breathing movements they can renew the air in the outer part of the tracheal system, the oxygen has to penetrate the finer branches by means of diffusion. Gases can diffuse easily through very small distances, not many times larger than the average length traveled by a gas molecule between collisions with other molecules. But when such vast journeys – from the point of view of a molecule – as a quarter of an inch have to be made, the process becomes slow. So the portions of an insect's body more than a quarter of an inch from the air would always be short of oxygen. In consequence hardly any insects are much more than half an inch thick. Land crabs are built on the same general plan as insects, but are much clumsier. Yet like ourselves they carry oxygen around in their blood, and are therefore able to grow far larger than any insects. If the insects had hit on a plan for driving air through their tissues instead of letting it soak in, they might well have become as large as lobsters, though other considerations would have prevented them from becoming as large as man.

Exactly the same difficulties attach to flying. It is an elementary principle of aeronautics that the minimum speed needed to keep an aeroplane of a given shape in the air varies as the square root of its length. If its linear dimensions are increased four times, it must fly twice as fast. Now the power needed for the minimum speed increases more rapidly than the weight of the machine. So the larger aeroplane, which weighs sixty-four times as much as the smaller, needs one hundred and twenty-eight times its horsepower to keep up. Applying the same principle to the birds, we find that the limit to their size is soon reached. An angel whose muscles developed no more power weight for weight than those of an eagle or a pigeon would require a breast projecting for about four feet to house the muscles engaged in working its wings, while to economize in weight, its legs would have to be reduced to mere stilts. Actually a large bird such as an eagle or kite does not keep in the air mainly by moving its wings. It is generally to be seen soaring, that is to say balanced on a rising column of air. And even soaring becomes more and more difficult with increasing size. Were this not the case eagles might be as large as tigers and as formidable to man as hostile aeroplanes.

But it is time that we pass to some of the advantages of size. One of the most obvious is that it enables one to keep warm. All warm-blooded animals at rest lose the same amount of heat from a unit area of skin, for which purpose they need a food-supply proportional to their surface and not to their weight. Five thousand mice weigh as much as a man. Their combined surface and food or oxygen consumption are about seventeen times a man's. In fact a mouse eats about one quarter its own weight of food every day, which is mainly used in keeping it warm. For the same reason small animals cannot live in cold countries. In the arctic regions there are no reptiles or amphibians, and no small mammals. The smallest mammal in Spitzbergen is the fox. The small birds fly away in winter, while the insects die, though their eggs can survive six months or more of frost. The most successful mammals are bears, seals, and walruses.

Similarly, the eye is a rather inefficient organ until it reaches a large size. The back of the human eye on which an image of the outside world is thrown, and which corresponds to the film of a camera, is composed of a mosaic of "rods and cones" whose diameter is little more than a length of an average light wave. Each eye has about a half a million, and for two objects to be distinguishable their images must fall on separate rods or cones. It is obvious that with fewer but larger rods and cones we should see less distinctly. If they were twice as broad two points would have to be twice as far apart before we could distinguish them at a given distance. But if their size were diminished and their number increased we should see no better. For it is impossible to form a definite image smaller than a wave-length of light. Hence a mouse's eye is not a small-scale model of a human eye. Its rods and cones are not much smaller than ours, and therefore there are far fewer of them. A mouse could not distinguish one human face from another six feet away. In order that they should be of any use at all the eyes of small animals have to be much larger in proportion to their bodies than our own. Large animals on the other hand only require relatively small eyes, and those of the whale and elephant are little larger than our own. For rather more recondite reasons the same general principle holds true of the brain. If we compare the brain-weights of a set of very similar animals such as the cat, cheetah, leopard, and tiger, we find that as we quadruple the body-weight the brain-weight is only doubled. The larger animal with proportionately larger bones can economize on brain, eyes, and certain other organs.

Such are a very few of the considerations which show that for every type of animal there is an optimum size. Yet although Galileo demonstrated the contrary more than three hundred years ago, people still believe that if a flea were as large as a man it could jump a thousand feet into the air. As a matter of fact the height to which an animal can jump is more nearly independent of its size than proportional to it. A flea can jump about two feet, a man about five. To jump a given height, if we neglect the resistance of air, requires an expenditure of energy proportional to the jumper's weight. But if the jumping muscles form a constant fraction of the animal's body, the energy developed per ounce of muscle is independent of the size, provided it can be developed quickly enough in the small animal. As a matter of fact an insect's muscles, although they can contract more quickly than our own, appear to be less efficient; as otherwise a flea or grasshopper could rise six feet into the air.

And just as there is a best size for every animal, so the same is true for every human institution. In the Greek type of democracy all the citizens could listen to a series of orators and vote directly on questions of legislation. Hence their philosophers held that a small city was the largest possible democratic state. The English invention of representative government made a democratic nation possible, and the possibility was first realized in the United States, and later elsewhere. With the development of broadcasting it has once more become possible for every citizen to listen to the political views of representative orators, and the future may perhaps see the return of the national state to the Greek form of democracy. Even the referendum has been made possible only by the institution of daily newspapers.

To the biologist the problem of socialism appears largely as a problem of size. The extreme socialists desire to run every nation as a single business concern. I do not suppose that Henry Ford would find much difficulty in running Andorra or Luxembourg on a socialistic basis. He has already more men on his pay-roll than their population. It is conceivable that a syndicate of Fords, if we could find them, would make Belgium Ltd or Denmark Inc. pay their way. But while nationalization of certain industries is an obvious possibility in the largest of states, I find it no easier to picture a completely socialized British Empire or United States than an elephant turning somersaults or a hippopotamus jumping a hedge.

About the Author

John Burdon Sanderson Haldane (November 5, 1892 - December 1, 1964) was a geneticist born in Scotland and educated at Eton and Oxford University. He was one of the founders (along with Fisher and Wright) of population genetics. His famous book, The Causes of Evolution (1932), was the first major work of what came to be known as the "modern evolutionary synthesis", reestablishing natural selection as the premier mechanism of evolution by explaining it in terms of the mathematical consequences of Mendelian genetics. He was also a great science popularizer, and was perhaps the Stephen Jay Gould or Richard Dawkins of his day. His essay, Daedalus or Science and the Future (1923), was remarkable in predicting many scientific advances but has been criticized for presenting a too idealistic view of scientific progress.

Haldane was himself a very idealistic man, and in his youth was a devoted Communist and author of many articles in The Daily Worker. Events in the Soviet Union, such as the rise of the anti-Mendelian agronomist Trofim Lysenko and the crimes of Stalin, caused him to break with the Communist Party later in life.

He is also known for an observation from his essay, On Being the Right Size, which Jane Jacobs and others have since referred to as Haldane's principle. This is that sheer size very often defines what bodily equipment an animal must have: "Insects, being so small, do not have oxygen-carrying bloodstreams. What little oxygen their cells require can be absorbed by simple diffusion of air through their bodies. But being larger means an animal must take on complicated oxygen pumping and distributing systems to reach all the cells." The conceptual metaphor to animal body complexity has been of use in energy economics and secession ideas.

Haldane was friends with the author Aldous Huxley, and was the basis for the biologist Shearwater in Huxley's novel Antic Hay. Ideas from Haldane's Daedalus, such as ectogenesis (the development of fetuses in artificial wombs), also influenced Huxley's Brave New World. He had many students, the most famous of whom, John Maynard Smith, was perhaps also the one most like himself.

In one of the last speeches of his life, Biological Possibilities for the Human Species of the Next Ten Thousand Years (1963), Haldane coined the word "clone", from the Greek word for twig.

Pick from the Past
Natural History, January 1974
THIS VIEW OF LIFE

Size and Shape

The immutable laws of design set limits on all organisms.

Steven Jay Gould
Museum of Comparative Zoology, Harvard University

Who could believe an ant in theory?
A giraffe in blueprint?
Ten thousand doctors of what's possible
Could reason half the jungle out of being.

POET John Ciardi's lines reflect a belief that the exuberant diversity of life will forever frustrate man's arrogant claims to omniscience. Yet, however much we celebrate diversity and revel in the peculiarities of animals, we must also acknowledge a striking "lawfulness" in the basic design of organisms. This regularity is most strongly evident in the correlation of size and shape.

Animals are physical objects. They are shaped to their advantage by natural selection. Consequently, they must assume forms best adapted to their size. The relative strength of such forces as gravity varies with size in a regular way, and animals respond by systematically altering their shapes.

The geometry of space itself is the major reason for correlations between size and shape. Simply by growing larger, an object that keeps the same shape will suffer a continual decrease in relative surface area. The decrease occurs because volume increases as the cube of length (length x length x length), while surface increases only as the square (length x length): in other words, volume grows more rapidly than surface.

Why is this important to animals? Many functions that depend upon surface must serve the entire volume of the body. Digested food passes to the body through surfaces; oxygen is absorbed through surfaces in respiration; the strength of a leg bone depends upon the area of its cross section, but the legs must hold up a body increasing in weight by the cube of its length. Galileo first recognized this principle in his "Discorsi" of 1638, the masterpiece he wrote while under house arrest by the Inquisition. He argued that the bone of a large animal must thicken disproportionately to provide the same relative strength as the slender bone of a small creature.

One solution to decreasing surface has been particularly important in the progressive evolution of large and complex organisms: the development of internal organs. The lung is, essentially, a richly convoluted bag of surface area for the exchange of gases; the circulatory system distributes material to an internal space that cannot be reached by direct diffusion from the external surface of large organisms; the villi of our small intestine increase the surface area available for absorption of food (small mammals neither have nor need them).

Some simpler animals have never evolved internal organs; if they become large, they must alter their entire shape in ways so drastic that plasticity for further evolutionary change is sacrificed to extreme specialization. Thus, a tapeworm may be 20 feet long, but its thickness cannot exceed a fraction of an inch because food and oxygen must penetrate directly from the external surface to all parts of the body.

We are prisoners of the perceptions of our size, and rarely recognize how different the world must appear to small animals.

Other animals are constrained to remain small. Insects breathe through invaginations of the external surface. Since these invaginations must be more numerous and convoluted in larger bodies, they impose a size limit upon insect design: at the size of even a small mammal, an insect would be “all invagination” and have no room for internal parts.

We are prisoners of the perceptions of our size, and rarely recognize how different the world must appear to small animals. Since our relative surface area is so small at our large size, we are ruled by gravitational forces acting upon our weight. But gravity is negligible to very small animals with high surface to volume ratios; they live in a world dominated by surface forces and judge the pleasures and dangers of their surroundings in ways foreign to our experience.

An insect performs no miracle in walking up a wall or upon the surface of a pond; the small gravitational force pulling it down or under is easily counteracted by surface adhesion. Throw an insect off the roof and it floats gently down as frictional forces acting upon its surface overcome the weak influence of gravity.

The relative weakness of gravitational forces also permits a mode of growth that large animals could not maintain. Insects have an external skeleton and can only grow by discarding it and secreting a new one to accommodate the enlarged body. For a period between shedding and regrowth, the body must remain soft. A large mammal without any supporting structures would collapse to a formless mass under the influence of gravitational forces; a small insect can maintain its cohesion (related lobsters and crabs can grow much larger because they pass their “soft” stage in the nearly weightless buoyancy of water). We have here another reason for the small size of insects.

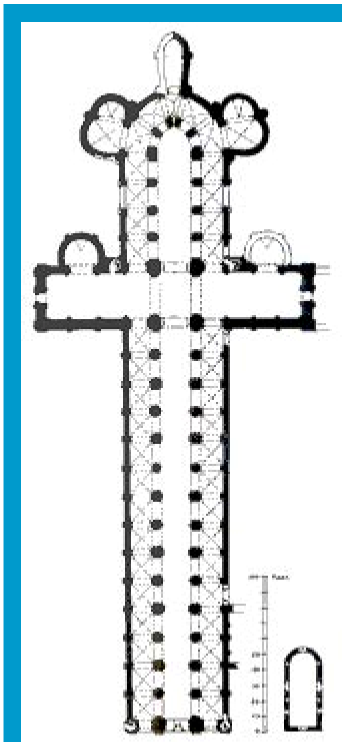
The creators of horror and science-fiction movies seem to have no inkling of the relationship between size and shape.

The creators of horror and science-fiction movies seem to have no inkling of the relationship between size and shape. These “expanders of the possible” cannot break free from the prejudices of their perceptions. The small people of *Dr. Cyclops*, *The Bride of Frankenstein*, *The Incredible Shrinking Man*, and *Fantastic Voyage* behave just like their counterparts of normal dimensions. They fall off cliffs or down stairs with resounding thuds; wield weapons and swim with olympic agility. The large insects of films too numerous to name continue to walk up walls or fly even at dinosaurian dimensions.

When the kindly entomologist of *Them* discovered that the giant queen ants had left for their nuptial flight, he quickly calculated this simple ratio: a normal ant is a fraction of an inch long and can fly hundreds of feet; these ants are many feet long and must be able to fly as much as 1,000 miles. Why, they could be as far away as Los Angeles! (Where, indeed, they were, lurking in the sewers.) But the ability to fly depends upon the surface area of the wings, while the weight that must be borne aloft increases as the cube of length. We may be sure that even if the giant ants had somehow circumvented the problems of breathing and growth by molting, their chances of getting off the ground would have been far worse than that of the proverbial snowball in hell.

Other essential features of organisms change even more rapidly with increasing size than the ratio of surface to volume. Kinetic energy, for example, increases as length raised to the fifth power. If a child half your height falls unsupported to the ground, its head will hit with not half, but only 1/32 the energy of yours in a similar fall. A child is protected more by its size than by a “soft” head. In return, we are protected from the physical force of its tantrums, for the child can strike with, not half, but only 1/32 of the energy we can muster. I have long had a special sympathy for the poor dwarfs who suffer under the whip of cruel Dr. Alberich in Wagner’s “Das Rheingold.” At their diminutive size, they haven’t a chance of extracting, with mining picks, the precious minerals that Alberich demands, despite the industrious and incessant leitmotif of their futile attempt.

This simple principle of differential scaling with increasing size may well be the most important determinant of organic shape. J.B.S. Haldane once wrote that “comparative anatomy is largely the story of the struggle to increase surface in proportion to volume.” Yet its generality extends beyond life, for the geometry of space constrains ships, buildings, and machines, as well as animals.



The great range of designs among medieval churches can be attributed partly to size. The twelfth-century parish church of Little Tey, Essex, England, was only 57 feet long and had a simple floor plan (lower right corner). The floor plan of Norwich Cathedral (left), also twelfth century, shows adaptations—transept, chapels—required for the 450-foot-long building. The need for light and support dictated complex cathedral layouts.

Click for a larger image.

Medieval churches present a good testing ground for the effects of size and shape, for they were built in an enormous range of sizes before the invention of steel girders, internal lighting, and air conditioning permitted modern architects to challenge the laws of size. The tiny, twelfth-century parish church of Little Tey, Essex, England, is a broad, simple rectangular building with a semicircular apse. Light reaches the interior through windows in the outer walls. If we were to build a cathedral simply by enlarging this design, then the periphery of the outer walls and windows would increase as length, while the area that light must reach would increase as length times length. In other words, the size of the windows would increase far more slowly than the area that requires illumination. Candles have limitations; the inside of such a cathedral would have been darker than the deed of Judas. Medieval churches, like tapeworms, lack internal systems and must alter their shape to produce more external surface as they are made larger.

The large cathedral of Norwich, as it appeared in the twelfth century, had a much narrower rectangular nave; chapels have been added to the apse and a transept runs perpendicular to the main axis. All these “adaptations” increase the ratio of external wall and window to internal area. It is often stated that transepts were added to produce the form of a Latin cross. Theological motives may have dictated the position of such “outpouchings,” but the laws of size required their presence. Very few small churches have transepts.

I have plotted periphery versus the square root of area for floor plans of all postconquest Romanesque churches depicted in Clapham’s monograph of English ecclesiastical architecture. As we would predict, periphery increases more rapidly than the square root of area. Medieval architects had their rules of thumb, but they had, so far as we know, no explicit knowledge of the laws of size.

Like large churches, large organisms have very few options open to them. Above a certain size, large terrestrial organisms look basically alike—they have thick legs and relatively short, stout bodies. Large Romanesque churches are all relatively long and have abundant outpouchings. The invention of the flying buttress strengthened later Gothic buildings and freed more wall space for windows. Churches could then become relatively wider and simpler in outline (as in the Cathedral of Bourges).

The “invention” of internal organs helped animals retain the highly successful shape of a simple exterior enclosing a large internal volume; and the invention of internal lighting and structural steel has helped modern architects design large buildings with simple exteriors. The limits are expanded, but the laws still operate. No large Gothic church is higher than it is long, no large animal has a sagging middle like a dachshund.

I once overheard a children’s conversation in a New York playground. Two young girls were discussing the size of dogs. One asked: “Can a dog be as large as an elephant?” Her friend responded: “No, if it were as big as an elephant, it would look like an elephant.” I wonder if she realized how truly she spoke.

http://www.naturalhistorymag.com/master.html?http://www.naturalhistorymag.com/editors_pick/1974_01_pick.html

ES123
Handout #13
5 April 2004

The geographer, speaking specially of the sandhill, says:—"The hill of sounding sand stretches 80 li east and west and 40 li north and south. It reaches a height of 500 ft. The whole mass is entirely constituted of pure sand. In the height of summer the sand gives out sounds of itself, and if trodden by men or horses, the noise is heard 10 li away. At festivals people clamber up and rush down again in a body, which causes the sand to give a loud rumbling sound like thunder. Yet when you look at it next morning the hill is just as steep as before."

Mr. Lionel Giles, from whose translation of the Tun-Huang-Lu these extracts are made, mentions that this sounding sandhill is referred to in another old Chinese book, the Wu Tai Shih.

JOSEPH OFFORD.

94 Gloucester Road, South Kensington, S.W.

The Green Flash.

I CAN confirm Dr. Schuster's observation of the green flash at sunrise, as in September last I saw a green segment herald the sun as it rose from the sea into a sky which was free from atmospheric glare (see the *Observatory*, December, 1914). Observations had previously been made at sunset, in one of which the eye was unquestionably fatigued, and the green flash was seen upon turning *away* from the sun at the instant after sunset. In a later sunset experiment precautions were taken to prevent retinal fatigue, and again the flash was seen.

My opinion is confirmed by Prof. Porter's experiment that "the reason why doubt has been cast upon records of the green flash is that the colour may arise in two different ways (complementary colour due to retinal fatigue, or dispersion by the atmosphere), and that the observer has not always been careful to avoid retinal fatigue, as was the case in my first (sunset) observation."

My observation, No. 2 (*loc. cit.*), is also in agreement with Dr. Schuster's experience, that with a very red sun no flash is to be seen.

W. GEOFFREY DUFFIELD.

University College, Reading, March 6.

Measurements of Medieval English Femora.

As the Editor of NATURE has insisted upon the great pressure at present upon his space I propose to reply to Dr. Parsons's letter, in the issue of March 11, adequately elsewhere. KARL PEARSON.

Galton Laboratory, March 15.

THE PRINCIPLE OF SIMILITUDE.

I HAVE often been impressed by the scanty attention paid even by original workers in physics to the great principle of similitude. It happens not infrequently that results in the form of "laws" are put forward as novelties on the basis of elaborate experiments, which might have been predicted *a priori* after a few minutes' consideration. However useful verification may be, whether to solve doubts or to exercise students, this seems to be an inversion of the natural order. One reason for the neglect of the principle may be that, at any rate in its applications to particular cases, it does not much interest mathematicians. On the other hand, engineers, who might make much more use of it than they have done, employ a notation which tends to obscure it. I refer to the manner in which gravity is treated. When the question under consideration depends essentially upon gravity, the symbol of gravity (g) makes no appearance, but when gravity does not enter into the question at all, g obtrudes itself conspicuously.

I have thought that a few examples, chosen almost at random from various fields, may help to direct the attention of workers and teachers to the great importance of the principle. The statement made is brief and in some cases inadequate, but may perhaps suffice for the purpose. Some foreign considerations of a more or less obvious character have been invoked in aid. In using the method practically, two cautions should be borne in mind. First, there is no prospect of determining a numerical coefficient from the principle of similarity alone; it must be found if at all, by further calculation, or experimentally. Secondly, it is necessary as a preliminary step to specify clearly *all* the quantities on which the desired result may reasonably be supposed to depend, after which it may be possible to drop one or more if further consideration shows that in the circumstances they cannot enter. The following, then, are some conclusions, which may be arrived at by this method:—

Geometrical similarity being presupposed here as always, how does the strength of a bridge depend upon the linear dimension and the force of gravity? In order to entail the same strains, the force of gravity must be inversely as the linear dimension. Under a given gravity the larger structure is the weaker.

The velocity of propagation of periodic waves on the surface of deep water is as the square root of the wave-length.

The periodic time of liquid vibration under gravity in a deep cylindrical vessel of any section is as the square root of the linear dimension.

The periodic time of a tuning-fork, or of a Helmholtz resonator, is directly as the linear dimension.

The intensity of light scattered in an otherwise uniform medium from a small particle of different refractive index is inversely as the fourth power of the wave-length.

The resolving power of an object-glass, measured by the reciprocal of the angle with which it can deal, is directly as the diameter and inversely as the wave-length of the light.

The frequency of vibration of a globe of liquid, vibrating in any of its modes under its own gravitation, is independent of the diameter and directly as the square root of the density.

The frequency of vibration of a drop of liquid, vibrating under capillary force, is directly as the square root of the capillary tension and inversely as the square root of the density and as the $1\frac{1}{2}$ power of the diameter.

The time-constant (*i.e.*, the time in which a current falls in the ratio $e:1$) of a linear conducting electric circuit is directly as the inductance and inversely as the resistance, measured in electro-magnetic measure.

The time-constant of circumferential electric currents in an infinite conducting cylinder is as the square of the diameter.

In a gaseous medium, of which the particles repel one another with a force inversely as the nth power of the distance, the viscosity is as the $(n+3)/(2n-2)$ power of the absolute temperature. Thus, if $n=5$, the viscosity is proportional to temperature.

Eiffel found that the resistance to a sphere moving through air changes its character somewhat suddenly at a certain velocity. The consideration of viscosity shows that the critical velocity is inversely proportional to the diameter of the sphere.

If viscosity may be neglected, the mass (M) of a drop of liquid, delivered slowly from a tube of diameter (a), depends further upon (T) the capillary tension, the density (σ), and the acceleration of gravity (g). If these data suffice, it follows from similarity that

$$M=\frac{Ta}{g}F\left(\frac{T}{g\sigma a^2}\right),$$

where F denotes an arbitrary function. Experiment shows that F varies but little and that within somewhat wide limits may be taken to be 3·8. Within these limits Tate's law that M varies as a holds good.

In the Æolian harp, if we may put out of account the compressibility and the viscosity of the air, the pitch (n) is a function of the velocity of the wind (v) and the diameter (d) of the wire. It then follows from similarity that the pitch is directly as v and inversely as d, as was found experimentally by Strouhal. If we include viscosity (ν), the form is

$$n=v/d.f(\nu/vd),$$

where f is arbitrary.

As a last example let us consider, somewhat in detail, Boussinesq's problem of the steady passage of heat from a good conductor immersed in a stream of fluid moving (at a distance from the solid) with velocity v. The fluid is treated as incompressible and for the present as inviscid, while the solid has always the same *shape* and presentation to the stream. In these circumstances the total heat (h) passing in unit time is a function of the linear dimension of the solid (a), the temperature-difference (θ), the stream-velocity (v), the capacity for heat of the fluid per unit volume (c), and the conductivity (κ). The density of the fluid clearly does not enter into the question. We have now to consider the "dimensions" of the various symbols.

Those of a are (Length)^1,
. . . . . . v . . (Length)^1 (Time)$^{-1}$,
. . . . . . θ . . (Temperature)^1,
. . . . . . c . . (Heat)^1 (Length)$^{-3}$ (Temp.)$^{-1}$,
. . . . . . κ . . (Heat)^1 (Length)$^{-1}$ (Temp.)$^{-1}$ (Time)$^{-1}$,
. . . . . . h . . (Heat)^1 (Time)$^{-1}$.

Hence if we assume

$$h=a^x\theta^y v^z c^u \kappa^v,$$

we have

by heat	$1=u+v$
by temperature	$0=y-u-v$,
by length	$0=x+z-3u-v$,
by time	$-1=-z-v$;

so that

$$h=\kappa a\theta\left(\frac{avc}{\kappa}\right)^z.$$

Since z is undetermined, any number of terms of this form may be combined, and all that we can conclude is that

$$h=\kappa a\theta.F(avc/\kappa),$$

where F is an arbitrary function of the one variable avc/κ. An important particular case arises when the solid takes the form of a cylindrical wire of any section, the length of which is perpendicular to the stream. In strictness similarity requires that the length l be proportional to the linear dimension of the section b; but when l is relatively very great h must become proportional to l and a under the functional symbol may be replaced by b. Thus

$$h=\kappa l\theta.F(bvc/\kappa).$$

We see that in all cases h is proportional to θ, and that for a given fluid F is constant provided v be taken inversely as a or b.

In an important class of cases Boussinesq has shown that it is possible to go further and actually to determine the form of F. When the layer of fluid which receives heat during its passage is very thin, the flow of heat is practically in one dimension and the circumstances are the same as when the plane boundary of a uniform conductor is suddenly raised in temperature and so maintained. From these considerations it follows that F varies as $v^{\frac{1}{2}}$, so that in the case of the wire

$$h\propto l\theta.\sqrt{(bvc/\kappa)},$$

the remaining constant factor being dependent upon the shape and purely numerical. But this development scarcely belongs to my present subject.

It will be remarked that since viscosity is neglected, the fluid is regarded as flowing past the surface of the solid with finite velocity, a serious departure from what happens in practice. If we include viscosity in our discussion, the question is of course complicated, but perhaps not so much as might be expected. We have merely to include another factor, ν^w, where ν is the kinematic viscosity of dimensions (Length)^2 (Time)$^{-1}$, and we find by the same process as before

$$h=\kappa a\theta.\left(\frac{avc}{\kappa}\right)^z.\left(\frac{c\nu}{\kappa}\right)^w.$$

Here z and w are both undetermined, and the conclusion is that

$$h=\kappa a\theta.F\left\{\frac{avc}{\kappa},\ \frac{c\nu}{\kappa}\right\},$$

where F is an arbitrary function of the *two* variables avc/κ and $c\nu/\kappa$. The latter of these, being the ratio of the two diffusivities (for momentum and for temperature), is of no dimensions; it appears to be constant for a given kind of gas, and to vary only moderately from one gas to another. If we may assume the accuracy and universality of this law, $c\nu/\kappa$ is a merely numerical constant, the same for all gases, and may be omitted, so that h reduces to the forms already given when viscosity is neglected altogether, F being again a function of a single variable, avc/κ or bvc/κ. In any case F is constant for a given fluid, provided v be taken inversely as a or b.

RAYLEIGH.

PERISCOPES.

WHILE the periscope of the submarine is developing in the direction of greater optical perfection and elaboration, there has been a return to the simplest and earliest types of periscope for use in land warfare. Some of these trench periscopes recall the polemoscope, described by Helvelius in the seventeenth century for military purposes; this polemoscope in its simplest form consisted of two mirrors with their reflecting surfaces parallel to each other, and

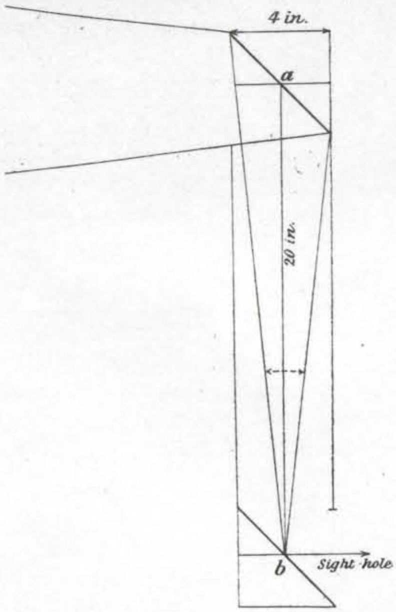


FIG. 1.

inclined at 45° to the direction of the incident light. These mirrors were mounted in a tube and separated a convenient distance (Fig. 1).

For modern trench warfare the convenient separation is about 18 to 24 in., and the mirrors are mounted in tubes, in boxes of square or oblong section, or attached to a long rod. In each case it is necessary that the mirrors should be fixed at the correct angle, and that there should be no doubling or distortion of the image.

The principal requirements of these trench periscopes are portability, lightness, small size and inconspicuous appearance, and large field of view. When there are no lenses the field of view is exactly the same as would be obtained by looking through a tube of the same length and diameter. Thus, with mirrors of 2 in. by 3 in. and a separation of about 22 in., a field of view of 5° would be obtained; and by moving the eye about, this field could be nearly doubled.

By using a box of oblong section the horizontal field of view can be increased without unduly increasing the size of the periscope. As the field of view is somewhat limited in any case, the principal objection to the use of a telescope or binocular, viz., the reduced field, no longer applies, and many periscopes are arranged to be used with a monocular or a binocular telescope.

Most periscopes can be used with a magnification of two or three, *i.e.*, with one tube of an ordinary opera glass; but when higher magnification is to be used the mirrors must be of better quality, both as regards flatness of surfaces and parallelism of the glass. When the mirrors are large enough—8 to 10 centimetres wide—both telescopes of the binocular may be used, but in this case the requirements for the mirrors are even more stringent, as the images formed by the two telescopes will not coincide unless the mirrors are plane. When suitable lenses are placed between the mirrors, the size of the mirrors can be reduced or the field of view increased; it is easy to provide a small magnification of the image or even to arrange for a variable magnification.

In such cases the lenses must be arranged to give an erect image, or mirrors or prisms employed to erect the image. An example of a periscope of this type is shown in Fig. 2, where the mirrors are replaced by reflecting prisms, and the prisms erect the image in much the same way as the prisms of a prism binocular.

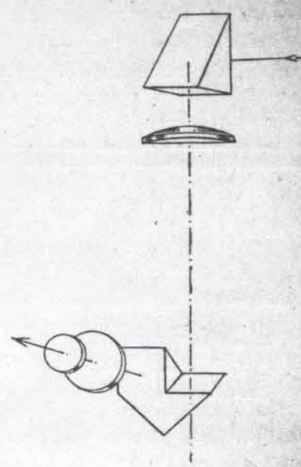


FIG. 2.

This arrangement is very suitable for a large magnification, but for larger fields the prism is unsuitable, unless it be silvered, and it is preferable to erect the image by means of lenses.

When longer tubes are used or larger fields are required, the design should approximate to that used in the submarine periscope.

This optical system has been steadily developed since its first introduction by Sir Howard Grubb in 1901.

The system consists of two telescopes, of which one is reversed, so that the image would be reduced in size, while the other magnifies this image, so that the final image is of the same size as the object, or is magnified one and a quarter or one and a half times. (As a very large angular field of view is required in these periscopes, the beam reflected into the tube must cover a large angle, and would soon fall on the sides of the tube; the reversed telescope, however, reduces the angle of the beam, and so enables it to pro-