SIAM Review
Vol. 61, No. 1, pp. 159–184

Dimensional and Scaling Analysis*

Gjerrit Meinsma†

Abstract. A complete theory of dimensional and scaling analysis is presented and its power is demonstrated through a series of examples. A vector-matrix exponentiation is introduced to simplify notation and calculus.

Key words. dimensional analysis, mathematical modeling, units, scaling, vector-matrix exponentiation

AMS subject classifications. 00A71, 00A73, 76M55, 15A03, 15A04, 97F70, 97M50

DOI. 10.1137/16M1107127

1. Introduction. Why is it immediate that the equality $\int_0^\infty \mathrm{e}^{-\alpha t}\,\mathrm{d}t = \alpha^p$ can only be correct if $p=-1$? Why is it easy to see that the movement of a pendulum does not depend on its mass? And how is it possible that we can solve complicated minimization problems such as

$$\min_{x:\mathbb{R}\to\mathbb{R}} \int_0^\infty x^4(t) + \left(\tfrac{\mathrm{d}}{\mathrm{d}t}x(t)\right)^2 \mathrm{d}t \quad \text{given } x(0)>0$$

easily with just a few rewritings? This is because of dimensional and scaling analysis. This analysis, and especially the Buckingham π-theorem, is easy to explain and fantastically useful in mathematical modeling, and yet it is not well appreciated in the mathematical community. Sure, it originated in physics [4, 19, 7, 15, 9, 5, 6] and uses "units" and "dimensions" which are normally absent in mathematical formulations, and because of that one might think that mathematics can do without these notions. On the contrary, it is a great tool that enables us to quickly spot errors in formulas, interpret and simplify mathematical models, and more. I teach this topic in a three-hour lecture to second-year students with diverse backgrounds and they are always amazed by the power of dimensional arguments.

On the basis of many short examples I hope to convince you of the beauty and usefulness of dimensional analysis, which goes beyond its physical origins. Now, complete proofs of the π-theorem are hard to come by, and many suffer from implicit assumptions. In this paper I shall be explicit and I believe that with the introduction of *vector-matrix exponentiation* one can be precise and efficient in the calculus and formulation of the π-theorems and their proofs. I also want to make the point that it is important to distinguish between a "mathematical" and a "physical" version of

*Received by the editors December 8, 2016; accepted for publication (in revised form) March 27, 2018; published electronically February 7, 2019.
http://www.siam.org/journals/sirev/61-1/M110712.html
†Department of Applied Mathematics, University of Twente, Enschede, The Netherlands (g.meinsma@utwente.nl).

the π-theorem. Loosely speaking, they form two sides of the same coin. With the introduction of the *scaling matrix* and its link with the common *dimensional matrix*, the connection between the two becomes explicit.

2. Features of Dimensional Analysis. In the following sections we develop and demonstrate the theory of dimensional analysis, but to get an idea of where this is heading we first present its main features, in particular, how to use the π-theorem. This theorem is the central result in dimensional analysis.

All that we assume in this section is a basic understanding of dimensions such as "length" and "time." The dimension of a quantity x is denoted as $[x]$. A dot on top means derivative, e.g., $\dot{x}(t) = \frac{\mathrm{d}}{\mathrm{d}t}x(t)$, and then the dimension of $\dot{x}(t)$ is that of x over that of t: $[\dot{x}] = [x][t]^{-1}$.

Example 2.1 (a first-order differential equation). Let $x : \mathbb{R} \to \mathbb{R}$ and $a, x_0 \in \mathbb{R}$, and suppose that x satisfies the differential equation

$$\dot{x}(t) = -ax(t), \qquad x(0) = x_0. \tag{2.1}$$

Instead of trying to solve this equation in its full generality (which you probably can) the idea is to first exploit dimensions so as to simplify the problem. To make it more concrete, suppose that t denotes time and $x(t)$ denotes the temperature at time t of some object sitting in a medium of constant zero degrees. Clearly, the hotter the object, the faster the cooling of the object. Newton's law of cooling says that the rate of cooling is proportional to the temperature, that is, (2.1) holds for some constant $a > 0$. Since $\dot{x}(t)$ is a cooling *rate*, the dimension of a is $[a] = \text{time}^{-1} = [t]^{-1}$. The temperature $x(t)$ is completely determined by x_0, t, and a, so we have a relation between four variables involving two dimensions:

$$\begin{aligned} [x(t)] &= \text{temperature}, \\ [x_0] &= \text{temperature}, \\ [t] &= \text{time}, \\ [a] &= [t]^{-1} = \text{time}^{-1}. \end{aligned}$$

The π-theorem then claims the following (just believe us for now—we are going to prove it in the next section): since there are $k = 2$ dimensions (temperature, time) in a relation between $n = 4$ quantities $(x(t), x_0, t, a)$, there exist $n - k = 2$ quantities that are dimensionless with respect to temperature and time, for instance,

$$\frac{x(t)}{x_0} \quad \text{and} \quad at,$$

and then the first of these must be a function of the second, that is,

$$\frac{x(t)}{x_0} = f(at) \tag{2.2}$$

for some function f. To be more precise, the π-theorem claims that there is a unique function $f : \mathbb{R} \to \mathbb{R}$ (not depending on $x(t), x_0, t, a$) such that for *every* x_0, t, a the solution of (2.1) equals (2.2). It is quite impressive that we may draw this conclusion without having to manipulate the equations. Also, it shows that we may as well choose $x_0 = 1$ and, $a = 1$ and, once we have solved that particular problem, substitute arbitrary x_0, t, a in (2.2) to obtain the general solution. For this example it means

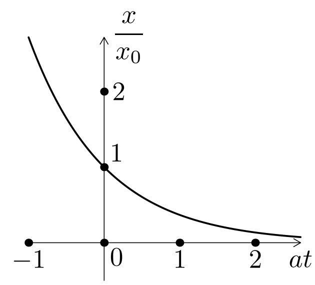


Fig. 1 *All solutions of* $\dot{x}(t) = -ax(t), x(0) = x_0$ *in a single graph; see Example* 2.1.

that we can represent the complete solution of (2.1) in a *single* graph by scaling the axes; see Figure 1. (The cases $x_0 = 0$ and $a = 0$ should be done separately, but these are easy.)

Dimensional analysis does not help us in finding the function f; it just says that such a function exists. However, you probably know that the solution of (2.1) is

$$x(t) = x_0\,\mathrm{e}^{-at}, \tag{2.3}$$

and clearly this equals (2.2) for $f(z) = \mathrm{e}^{-z}$.

In the next example we can simplify the model even before having equations. It is based on Example 1.3b from [20].

Example 2.2 (pendulum). Consider the pendulum shown in Figure 2. The figure depicts a mass m hanging from a ceiling on a thin flexible cable of fixed length ℓ. The angle between the cable and the vertical position is denoted by ϕ. Now suppose that at time $t = 0$ we release the mass at an initial angle ϕ_0 and with zero speed. We are interested in ϕ as a function of time. In the absence of friction it seems reasonable to assume that $\phi(t)$ depends on

initial angle ϕ_0,
time t,
mass m,
cable length ℓ,
gravitational acceleration g.

The standard model of this system under the above assumptions is the second-order differential equation

$$m\ell\ddot{\phi}(t) + mg\sin(\phi(t)) = 0, \qquad \phi(0) = \phi_0,\ \dot{\phi}(0) = 0.$$

(We are not going to derive it here, since the derivation is not important for the discussion.) Dividing the differential equation by $m\ell$ shows that in this model the angle $\phi(t)$ only depends on t, ϕ_0 and the ratio g/ℓ, but not on mass m:

$$\ddot{\phi}(t) + \frac{g}{\ell}\sin(\phi(t)) = 0, \qquad \phi(0) = \phi_0,\ \dot{\phi}(0) = 0. \tag{2.4}$$

To further simplify the analysis we replace $\sin(\phi(t))$ by its linear approximation $\phi(t)$, which is reasonable if $\phi(t)$ is small. This turns the model into

$$\ddot{\phi}(t) + \frac{g}{\ell}\phi(t) = 0, \qquad \phi(0) = \phi_0,\ \dot{\phi}(0) = 0. \tag{2.5}$$

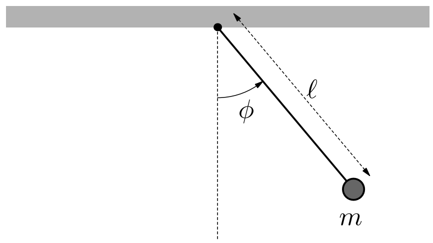


Fig. 2 *Pendulum; see Example* 2.2.

This final model is a standard second-order differential equation and it has the unique solution

$$\phi(t) = \phi_0 \cos\left(\sqrt{\tfrac{g}{\ell}}t\right). \tag{2.6}$$

This $\phi(t)$ is periodic with period $T = 2\pi\sqrt{\ell/g}$.

The beauty of the π-theorem is that we can arrive at almost the same conclusions but without the need to derive equations, let alone solutions of those equations! The reasoning is as follows: there is a relation between $n = 6$ quantities $(\phi(t), t, \phi_0, m, \ell, g)$ and it involves $k = 3$ dimensions (time, length, mass). Then $n - k = 3$ quantities exist that are dimensionless with respect to (time, length, mass), for instance,[1]

$$\phi(t), \quad \sqrt{\tfrac{g}{\ell}}t, \quad \phi_0.$$

(In the next section we make precise which dimensionless quantities we are allowed to take.) The π-theorem then goes on to claim that these three completely specify the problem; in fact, it claims that $\phi(t)$ must be a function of the other two dimensionless quantities, that is, must be of the form

$$\phi(t) = f\left(\sqrt{\tfrac{g}{\ell}}t, \phi_0\right)$$

for some function f that is dimensionless with respect to time, length, and mass. This means that f does not depend on any "hidden" quantity that involves one or more of the dimensions time, length, and mass, so again, we see that $\phi(t)$ necessarily does not depend on m and that *if* the angle $\phi(t)$ is periodic, the period is proportional to $\sqrt{\ell/g}$. According to the π-theorem *this holds for every model in these quantities* $(\phi(t), t, \phi_0, m, \ell, g)$, so not just the linearized model (2.5), but also the "true" nonlinear model (2.4). In *every* such model. How about that!

The above two examples hint at another startling property: observe that the argument $-at$ of the exponential function in (2.3) is dimensionless and also that the argument $\sqrt{g/\ell}t$ of the cosine in (2.6) is dimensionless. This suggests that

the argument x of "every" function $f(x)$ is dimensionless.

It turns out that this is essentially correct and the only exceptions to this rule are the power functions $f(x) = cx^\alpha$. A precise statement is given in Theorem 5.8. With

[1] The dimension of g is length$/[t]^2$ and the dimension of ℓ is length, so $\sqrt{g/\ell}$ has dimension $\sqrt{1/[t]^2} = 1/[t]$ and therefore $\sqrt{g/\ell}t$ indeed is dimensionless.

this rule many expressions become easier to grasp and memorize. For instance, if in a course on Laplace transformation you encounter the integral

$$\int_{-\infty}^{\infty} f(t)\,\mathrm{e}^{-st}\ \mathrm{d}t,$$

then it will be instantly clear to you that s has dimension $[t]^{-1}$ and you will understand why they call s a "frequency."

3. The Buckingham π-Theorem.

That's great, but why would x need to be dimensionless in expressions like e^x and $\cos(x)$, and why would the π-theorem be correct, and how is it possible that we can conclude all that without equations? Also, in the examples in the previous section we were vague about the construction of the dimensionless quantities (we just stated them). To understand all that, we need to dig into

- *coordinate calculus* and *vector-matrix exponentiation*,
- *dimensional matrices*,
- *physical laws* and *unit independence*.

These we will explore now, and then we use them to formulate and prove the Buckingham π-theorem (or just π-theorem). This is the classic physical version. In section 5 we cover a more mathematical version of the π-theorem.

3.1. Coordinate Calculus and Vector-Matrix Exponentiation.

In the pendulum example we constructed a new quantity $\sqrt{g/\ell}t$ from given t, g, ℓ. We need to understand this type of construction and now we set up a calculus for it. As a toy example, consider the two positive numbers x_1, x_2 and with them construct two new numbers,

$$y_1 = x_1 x_2^3, \qquad y_2 = x_2^{-2}.$$

Can we recover x_1, x_2 from y_1, y_2? Yes, with a bit of trial and error we find that it is possible with

$$y_1 y_2^{3/2} = (x_1 x_2^3)(x_2^{-3}) = x_1 \quad \text{and} \quad y_2^{-1/2} = (x_2^{-2})^{-1/2} = x_2. \tag{3.1}$$

For the π-theorem to develop we need to generalize this type of calculus to arbitrary row vectors $x = (x_1, \ldots, x_n)$ in the positive orthant $\mathbb{R}^n_+ = \{x \in \mathbb{R}^n \mid x_i > 0\ \forall i = 1, \ldots, n\}$.

DEFINITION 3.1 (vector-matrix exponentiation). *Let* $x = \begin{bmatrix} x_1 & x_2 & \cdots & x_n \end{bmatrix} \in \mathbb{R}^n_+$ *and* $C \in \mathbb{R}^{n\times k}$. *The* vector-matrix exponentiation *is the operation that maps* x *and* C *to its* vector-matrix power x^C *defined as*

$$x^C = \begin{bmatrix} x_1^{c_{11}} \cdots x_n^{c_{n1}} & x_1^{c_{12}} \cdots x_n^{c_{n2}} & \cdots\cdots & x_1^{c_{1k}} \cdots x_n^{c_{nk}} \end{bmatrix}.$$

It is a row vector with k *entries.*

This is less complicated than it looks. It is very similar to regular vector-matrix multiplication, except that multiplication is now exponentiation and addition is multiplication. For example, $5^{[0\ 1\ -2]} = \begin{bmatrix} 1 & 5 & \frac{1}{25} \end{bmatrix}$ and $[x_1\ \ x_2]^{\left[\begin{smallmatrix} 1 & 0 \\ 3 & -2 \end{smallmatrix}\right]} = [x_1 x_2^3\ \ x_2^{-2}]$. Vector-matrix exponentiation has the neat property

$$(x^A)^B = x^{(AB)} \text{ with } AB \text{ the normal matrix product.} \tag{3.2}$$

This is a perfect generalization of what we know of exponentiation with numbers. The proof is a bit of a bore, but is essentially easy. Rule (3.2) comes in really handy;

in particular, it shows that $(x^C)^{C^{-1}} = x^I = x$. Now we can systematically invert the mapping from x to $y = x^C$:

Example 3.2 (vector-matrix exponentiation and coordinate change). Let $C = \begin{bmatrix} 1 & 0 \\ 3 & -2 \end{bmatrix}$ and $x = \begin{bmatrix} x_1 & x_2 \end{bmatrix}$. Define y as

$$y = x^C = \begin{bmatrix} x_1 & x_2 \end{bmatrix}^{\begin{bmatrix} 1 & 0 \\ 3 & -2 \end{bmatrix}} = \begin{bmatrix} x_1 x_2^3 & x_2^{-2} \end{bmatrix}.$$

Since C is invertible we can switch back to x with

$$x = y^{C^{-1}} = \begin{bmatrix} y_1 & y_2 \end{bmatrix}^{\begin{bmatrix} 1 & 0 \\ 3/2 & -1/2 \end{bmatrix}} = \begin{bmatrix} y_1 y_2^{3/2} & y_2^{-1/2} \end{bmatrix}.$$

Indeed, it agrees with what we found in (3.1).

The mapping from a vector of coordinates x to $y = x^C$ is called a *coordinate change* if C is invertible. Thus, with a coordinate change we lose no information if we switch from x to $y = x^C$.

3.2. Dimensional Matrix. In the π-theorem we need a special coordinate change. To determine this coordinate change we first have to determine the *dimensional matrix.* This is best explained using an example.

Example 3.3 (pendulum dimensional matrix). We continue with the pendulum example. The dimensions of the $n = 6$ quantities $x = (\phi(t), t, \phi_0, m, \ell, g)$ in terms of the $k = 3$ dimensions (time, length, mass) are

$\phi(t)$	dimensionless,
t	time,
ϕ_0	dimensionless,
m	mass,
ℓ	length,
g	length/time^2.

We collect these dimensions in a *dimensional matrix* $M \in \mathbb{R}^{k \times n}$ with as many rows as we have dimensions and as many columns as we have quantities:

x	$\phi(t)$	t	ϕ_0	m	ℓ	g
time	0	1	0	0	0	−2
length	0	0	0	0	1	1
mass	0	0	0	1	0	0

The so-defined $k \times n = 3 \times 6$ dimensional matrix M is a convenient representation of the dimensions because if we want to know the dimension of, say, $\ell^{10} g^3$, we need only multiply the dimensional matrix from the right with the corresponding vector,

$$\begin{bmatrix} \phi(t) & t & \phi_0 & m & \ell & g \\ \hline 0 & 1 & 0 & 0 & 0 & -2 \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 10 \\ 3 \end{bmatrix} = \begin{bmatrix} \ell^{10} g^3 \\ \hline -6 \\ 13 \\ 0 \end{bmatrix}.$$

The dimension of $\ell^{10} g^3$ is hence length^{13}/time^6. This identity demonstrates that *the dimensional matrix of x^C is MC if M is the dimensional matrix of x*. We use this a lot in what follows.

Now for the next step: consider the coordinate change

$$\begin{bmatrix} \phi(t) & t & \phi_0 & m & \ell & g \\ \hline 0 & 1 & 0 & 0 & 0 & -2 \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 \end{bmatrix} \underbrace{\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & -1/2 & 0 & 1 & 0 \\ 0 & 0 & 1/2 & 0 & 0 & 0 \end{bmatrix}}_{C} = \begin{bmatrix} \phi(t) & \phi_0 & \sqrt{\frac{g}{\ell}}t & t & \ell & m \\ \hline 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

This C is invertible, and hence x^C is indeed a coordinate change. It is special in that the first three columns of MC contain zeros only. That is, $M^\perp$ defined as the first three columns of C,

$$M^\perp = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1/2 \\ 0 & 0 & 1/2 \end{bmatrix},$$

is in the null space of M. As a result, the first three $x^{M^\perp} = (\phi(t), \phi_0, \sqrt{g/\ell}t)$ of the new coordinates x^C are *dimensionless* with respect to (time, length, mass). The example continues later.

3.3. Physical Laws and the Buckingham π-Theorem. From a physical perspective every quantity $x_1, \dots, x_n$ is assumed from the outset to have a dimension, and a relation between quantities is considered valid only if it is a *physical law*. This is defined as follows.

DEFINITION 3.4 (relation and physical law). *A* relation $\mathbb{X}$ *between quantities* $x_1, \dots, x_n$ *is the set of all possible quantities*

$$\mathbb{X} = \{x = (x_1, \dots, x_n) \mid x \textit{ satisfies the model equations}\}.$$

A relation is a physical law *(with respect to a finite set of given dimensions) if*

- *each x_i has a dimension and this dimension is a product of powers of the given dimensions (hence x has a dimensional matrix with respect to the given dimensions), and*
- *the relation is* unit independent *with respect to the given dimensions. Unit independence means that if we fix the unit of all but one of the given dimensions, then varying the unit of the remaining dimension does not change the relation.*

Thus, a physical law is relative to a set of given dimensions. These dimensions are given (chosen) by us. Typically, we choose "length" and "time," etc., but we may also choose, say, "length/mass."

Example 3.5 (physical laws and unit independence). Consider Newton's second law

$$F = ma,$$

which defines a physical law with respect to the dimensions (length, mass, time). Indeed, it has a dimensional matrix with respect to these dimensions and the relation $\{(m, a, F) \mid F = ma\}$ is unit independent. If, for instance, $F = ma$ holds in the

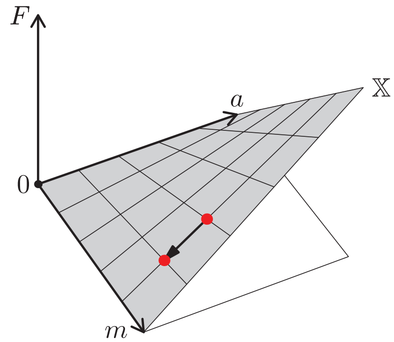


Fig. 3 *The relation* $\mathbb{X} = \{(m, a, F) \in \mathbb{R}^3_+ \mid F = ma\}$*; see Example* 3.5.

units $m\,(\mathrm{kg})$, $a\,(\mathrm{m/s^2})$, and $F\,(\mathrm{kg\,m/s^2})$, then it also holds in any other unit of the dimensions (length, mass, time), such as $m\,(\mathrm{kg})$, $a\,(\mathrm{km/s^2})$, and $F\,(\mathrm{kg\,km/s^2})$. The numerical values of a and F change (both times 0.001), but the identity $F = ma$ remains unchanged; that is, unit-independence holds. Another look at unit independence goes as follows: consider the relation $\mathbb{X} = \{(m, a, F) \in \mathbb{R}^3_+ \mid F = ma\}$ as a two-dimensional surface in $\mathbb{R}^3$; see Figure 3. Doubling the length unit from, say, 1 meter to 2 meters, halves the *values* of a and F (see the two red dots in the figure), but it does not change the identity $F = ma$, i.e., it does not change the two-dimensional surface $\mathbb{X}$. This literally means that the relation $\mathbb{X}$ is unit independent.

If we accept that our quantities satisfy a physical law, then the π-theorem follows:

Example 3.6 (pendulum example continued). Consider again the pendulum example and let us continue with the transformed coordinates x^C and corresponding special dimensional matrix

x^C	$\phi(t)$	ϕ_0	$\sqrt{\frac{g}{\ell}}t$	t	ℓ	m
time	0	0	0	1	0	0
length	0	0	0	0	1	0
mass	0	0	0	0	0	1

.

Recall that the first three of these are $x^{M^\perp} = (\phi(t), \phi_0, \sqrt{g/\ell}t)$ and that they are dimensionless. Assuming that the relation is a physical law means that they satisfy the relation iff they do so in any other unit of (time, length, mass). Changing these three units only affects the value of the final three (t, ℓ, m) because the first three are dimensionless. In fact, t only depends on the time unit, ℓ only depends on the length unit, and m on the mass unit, thus they can be independently scaled at will by the choice of unit,

$$\left(\phi(t), \phi_0, \sqrt{\tfrac{g}{\ell}}t, \frac{t}{\alpha}, \frac{\ell}{\beta}, \frac{m}{\gamma}\right). \tag{3.3}$$

In particular, for each t, ℓ, m *we may choose* $\alpha = t$*,* $\beta = \ell$*, and* $\gamma = m$*.* Thus, it satisfies the relation iff the six numbers

$$\left(\phi(t), \phi_0, \sqrt{\tfrac{g}{\ell}}t, 1, 1, 1\right)$$

satisfy the relation, which depends on just the first three (the dimensionless quantities). Thus, the values of t, ℓ, m cannot play a role in the relation (once the dimensionless $x^{M^\perp}$ is known) and may, therefore, just as well be discarded.

This last example should sink in. It effectively finishes the proof of the general π-theorem as formulated below.

Notice that changing units means that quantities are scaled with *positive* numbers. For example, the scalings α, β, γ in (3.3) are positive numbers. If in the pendulum example we are interested in positive and negative time t, then we cannot achieve $t/\gamma = 1$ by choice of $\gamma > 0$. In that case we can achieve $t/\gamma = \pm 1$ by choice of γ, i.e., the relation depends on the three dimensionless quantities $x^{M^\perp}$ *together with the signs of* t, ℓ, m. In principle, then, for each orthant we might have a different model. In the pendulum example that is not the case, but for other applications of the π-theorem it can happen (see Example 5.7). We formulate the π-theorem for the positive orthant $\mathbb{R}^n_+ = \{(x_1, \ldots, x_n) \mid x_i > 0 \ \forall i = 1, \ldots, n\}$, but it applies to any orthant.

THEOREM 3.7 (Buckingham π-theorem). *Suppose the relation*

$$\mathbb{X} = \{x \in \mathbb{R}^n_+ \mid x \textit{ satisfies the model equations}\}$$

is a physical law with respect to given dimensions, and let M be the corresponding dimensional matrix. If the columns of some matrix $M^\perp$ span the null space of M, then

$$x \in \mathbb{X} \quad \iff \quad x^{M^\perp} \in \mathbb{X}^{M^\perp},$$

where $\mathbb{X}^{M^\perp} = \{x^{M^\perp} \mid x \in \mathbb{X}\}$. In words this means that x satisfies the physical law iff $x^{M^\perp}$ satisfies a certain physical law.

Moreover, $x^{M^\perp}$ is dimensionless with respect to the given dimensions.

Proof. Given a matrix Q, define $\mathbb{X}^Q = \{x^Q \mid x \in \mathbb{X}\}$. Let k be the rank of M. If k is less than the number of rows of M, then one by one remove dimensions (i.e., remove rows from M) that do not affect the null space of M until the number of rows is reduced to k. With respect to this reduced number of dimensions, the relation is still unit independent. Now, there is a matrix A such that $\left[\begin{smallmatrix} A \\ M \end{smallmatrix}\right]$ is square invertible. Define C as $C = \left[\begin{smallmatrix} A \\ M \end{smallmatrix}\right]^{-1}$. This is invertible and has the property that $MC = \left[\begin{smallmatrix} 0 & I_k \end{smallmatrix}\right]$. Then $y = x^C$ is a coordinate change and the dimensional matrix of x^C is $MC = \left[\begin{smallmatrix} 0 & I_k \end{smallmatrix}\right]$. The final k entries of x^C can hence be assigned to any positive number by choice of unit. By definition of a physical law, the relation does not depend on the choice of unit so these final entries do not play a role in the relation. That is, the relation is completely determined by x^{C_0}, where C_0 is the matrix formed by the first $n-k$ columns of C. We have thus proved that

$$x \in \mathbb{X} \quad \iff \quad x^C \in \mathbb{X}^C \quad \iff \quad x^{C_0} \in \mathbb{X}^{C_0}.$$

Since both C_0 and $M^\perp$ span the null space of M we have that $M^\perp = C_0 P$ for some matrix P, but also that $C_0 = M^\perp Q$ for some matrix Q. From that it is easy to see that $x^{C_0} \in \mathbb{X}^{C_0}$ iff $x^{M^\perp} \in \mathbb{X}^{M^\perp}$.

$x^{M^\perp}$ is dimensionless because its dimensional matrix is $MM^\perp = 0$. ∎

That's it. Notice that any $M^\perp$ will do as long as its columns span the null space of M. Linear algebra tells us that the minimal number of such columns is $n-k$, where k is the rank of M. Typically, k equals the number of given dimensions. Thus, the more dimensions we have the higher the rank of M normally is and the fewer entries $n-k$ we end up with in $x^{M^\perp}$. But if we miss a dimension, then the π-theorem still applies. As an aside, it is called "π-theorem" because Edgar Buckingham back in 1914 happened to denote the dimensionless quantities as $\pi_1, \pi_2, \ldots$ [5].

4. Applications of the π-Theorem.

Now that we understand the π-theorem we can be more precise in our statements, if needed. We limit the number of examples from physics, though, because these can be found elsewhere.

Example 4.1 (differential equation). In this example we redo Example 2.1, but now with all the steps of the π-theorem spelled out. Let $x : \mathbb{R} \to \mathbb{R}$ and $a, x_0 \in \mathbb{R}$. What does the π-theorem have to say about solutions of the differential equation

$$\dot{x}(t) = -ax(t), \qquad x(0) = x_0?$$

The differential equation defines a relation between $x(t)$ and x_0, a, t:

$$\mathbb{X} = \{(x(t), x_0, a, t) \mid \dot{x}(\tau) = -ax(\tau), x(0) = x_0\}.$$

With respect to the two dimensions $[x], [t]$ we have that $[a] = [t]^{-1}$ and so the dimensional matrix of this relation becomes

	$x(t)$	x_0	a	t
$[x]$	1	1	0	0
$[t]$	0	0	−1	1

The dimensional matrix has two rows and four columns and its null space is spanned by, for example,

$$M^\perp = \begin{bmatrix} 1 & 0 \\ -1 & 0 \\ 0 & 1 \\ 0 & 1 \end{bmatrix}.$$

Now the π-theorem says that $\begin{bmatrix} x(t) & x_0 & a & t \end{bmatrix}^{M^\perp} = \begin{bmatrix} x(t)/x_0 & at \end{bmatrix}$ completely specifies the relation. Any first course on differential equations will tell you that $x(t)$ follows uniquely from (x_0, a, t), but then, given x_0, the ratio $x(t)/x_0$ also follows uniquely from at. Hence, $x(t)/x_0 = f(at)$ for some function f.

Once you get the hang of it there is no need to write down the relation $\mathbb{X}$. The equations suffice.

Example 4.2 (Newton's second law). Suppose you forgot Newton's second law $F = ma$. All you remember is that the force F is some function of mass m and acceleration a and that it involves the dimensions (length, time, mass). With respect to these dimensions the dimensional matrix M is

	F	m	a
length	1	0	1
time	−2	0	−2
mass	1	1	0

Notice that the rank of M is 2 (not 3). So there is $n-2 = 3-2 = 1$ column that spans the null space, e.g., $(1, -1, -1)$. The relation is thus equivalent to a relation in

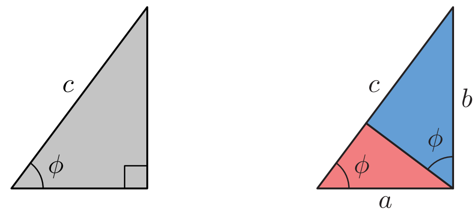


Fig. 4 *Pythagorean theorem. See Example 4.3.*

the single quantity

$$\begin{bmatrix} F & m & a \end{bmatrix} \begin{bmatrix} 1 \\ -1 \\ -1 \end{bmatrix} = \frac{F}{ma}.$$

Since F by assumption is unique once that m and a are known, $F/(ma)$ is also unique. This unique dimensionless number we denote as c, and this finally shows that

$$F = cma.$$

Fun. The π-theorem does not help us in finding the value of c.

Example 4.3 (Pythagorean theorem). This example appears to be popular on the internet. It is a proof of the Pythagorean theorem using the π-theorem. The idea is to explore the relation between the area A of a right triangle and its hypotenuse c and acute angle ϕ; see Figure 4 (left). The area A has dimension length^2, c has dimension length, ϕ is dimensionless. With respect to the length dimension the dimensional matrix of this relation is

	A	c	ϕ
length	2	1	0

The null space of this dimensional matrix is, for instance, spanned by the columns of

$$M^{\perp} = \begin{bmatrix} 0 & 1 \\ 0 & -2 \\ 1 & 0 \end{bmatrix},$$

so the π-theorem says that the relation between (A, c, ϕ) is equivalent to a relation between the dimensionless quantities $[A \quad c \quad \phi]^{M^{\perp}} = [\phi \quad A/c^2]$. Given c, the area A follows uniquely from ϕ, but then A/c^2 also follows uniquely from ϕ, so

$$A = c^2 f(\phi)$$

for some (unique) dimensionless function $f > 0$. Consequently, the area of the red part of Figure 4 (right) is $a^2 f(\phi)$, and the area of the blue part is $b^2 f(\phi)$, which add up to the total area of the triangle, that is,

$$a^2 f(\phi) + b^2 f(\phi) = c^2 f(\phi).$$

Cancelling $f(\phi)$ leaves us with the famous $a^2 + b^2 = c^2$.

Example 4.4 (melodies). A melody can be seen as an ordered sequence of notes, each with its own frequency. Here we neglect rhythm, type of instruments, and duration and volume of the notes. We only care about the sequence of frequencies $(f_1, f_2, \dots, f_n)$. In this case there is only one dimension (that of frequency) and so, according to the π-theorem, what sets one melody apart from another is determined solely by the null space of the dimensional matrix

	f_1	f_2	$\cdots$	f_n
frequency	1	1	$\cdots$	1

.

One appealing choice of $M^\perp$ is

$$M^\perp = \begin{bmatrix} -1 & 0 & \cdots & 0 \\ 1 & -1 & \ddots & \vdots \\ 0 & 1 & \ddots & 0 \\ \vdots & \ddots & \ddots & -1 \\ 0 & \cdots & 0 & 1 \end{bmatrix},$$

for then the dimensionless quantities are the $n-1$ ratios of successive frequencies

$$\begin{bmatrix} f_1 & f_2 & \cdots & f_n \end{bmatrix}^{M^\perp} = \begin{bmatrix} f_2/f_1 & f_3/f_2 & \cdots & f_n/f_{n-1} \end{bmatrix}.$$

The π-theorem hence implies that the absolute frequency (pitch) is not relevant for the characterization of a melody. More explicitly, it says that two melodies $(f_1^A, f_2^A, \dots, f_n^A)$ and $(f_1^B, f_2^B, \dots, f_n^B)$ are considered equal if one is a scaled version of the other,

$$(f_1^A, f_2^A, \dots, f_n^A) = (\mu f_1^B, \mu f_2^B, \dots, \mu f_n^B) \qquad \text{for some } \mu > 0.$$

On a logarithmic scale these two sequences differ by a constant $\log(\mu) \in \mathbb{R}$. Not surprisingly, perhaps, standard musical notation adopts a more or less logarithmic scale.

Realize that there is a freedom in the choice of $M^\perp$: the π-theorem says that any matrix $M^\perp$ whose columns span the null space of M is fine. It is worthwhile to think about this choice because choosing the "wrong" one might easily mess up $x^{M^\perp}$. This is particularly true for the next example.

Example 4.5 (zeros of polynomials). What about the zeros $\hat{x} \in \mathbb{R}$ of a polynomial

$$ax^2 + bx + c$$

for some given $a, b, c \in \mathbb{R}$? Of course, we know how to solve it, but let us apply the π-theorem once more and see what it does. Since there are no natural dimensions, we simply assign dimensions ourselves. We choose to assign the dimensions $[x]$ and $[c]$ and then for consistency we need $[a] = [c]/[x]^2$, $[b] = [c]/[x]$. There is a relation between the zeros $\hat{x}$ and the numbers a, b, c. The dimensional matrix for this relation and choice of dimensions is

	$\hat{x}$	a	b	c
$[x]$	1	−2	−1	0
$[c]$	0	1	1	1

,

which has rank 2 so $4 - 2 = 2$ dimensionless quantities suffice. One choice of dimensionless quantities is z, d defined as

$$z = \frac{a\hat{x}}{b}, \qquad d = \frac{ac}{b^2}.$$

Verify this yourself. (There are many other pairs of dimensionless quantities, but this pair z, d appears to work nicely.) Then by the π-theorem the relation is equivalent to a relation between z and d. By substitution of $a = b = 1$ we see that the relation between z and d is

$$z^2 + z + d = 0, \tag{4.1}$$

that is, $d = -z^2 - z$; see Figure 5. Now the general relation follows by simply scaling the axes, also shown in Figure 5. Given the values of a, b, c, this figure represents all possible real solutions $\hat{x}$ of $ax^2 + bx + c = 0$ and it demonstrates, for instance, that no real solutions $\hat{x}$ exist iff $d = ac/b^2 > 1/4$.

Of course, (4.1) is the same as $(z+1/2)^2 = 1/4-d$ so the solution is $z = f(d)$ with f the dimensionless multivalued function $f(d) = -1/2 \pm \sqrt{1/4 - d}$. Then according to the π-theorem the general solution of the original problem is $\hat{x} = \frac{b}{a} f\left(\frac{ac}{b^2}\right)$.

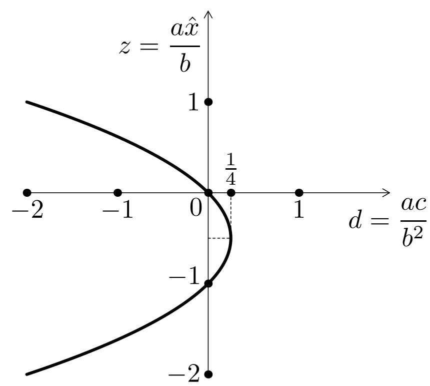


Fig. 5 *Given a, b, c, the graph determines all real zeros $\hat{x}$ of $ax^2 + bx + c$. There are two zeros if $d < 1/4$, one zero if $d = 1/4$, and there is no zero if $d > 1/4$. See Example* 4.5.

Now for some serious business. The next example deals with a complicated mathematical optimization problem. Such problems are usually solved with methods from calculus of variations or optimal control, but here we tackle it with basic dimensional arguments and some general insight.

Example 4.6 (minimization of integrals). Let x_0 be some positive number. We want to determine the differentiable function $x : \mathbb{R} \to \mathbb{R}$ that solves the minimization problem

$$\min_{x:\mathbb{R}\to\mathbb{R}} \int_0^\infty x^4(t) + \dot{x}^2(t)\,\mathrm{d}t \qquad \text{given } x(0) = x_0 > 0. \tag{4.2}$$

This looks like a complicated problem. First notice that it is dimensionally weird because the dimension of the derivative $\dot{x}$ is $[x]/[t]$ so $\dot{x}^2$ has dimension $[x]^2/[t]^2$, which differs from the dimension of x^4 unless $[t] = [x]^{-1}$. To remedy this we include an extra parameter α in our problem:

$$\min_{x:\mathbb{R}\to\mathbb{R}} \int_0^\infty x^4(t) + \alpha^2\dot{x}^2(t)\,\mathrm{d}t \qquad \text{given } x(0) = x_0 > 0, \alpha > 0. \tag{4.3}$$

Now we can assign two dimensions $[x]$ and $[t]$ and then for consistency set $[\alpha] = [x][t]$. Given x_0, α we expect that a function x that solves the problem exists and is unique.

Consequently, at any t the solution $x(t)$ follows uniquely from x_0, α, t. In particular, there is a relation between $\dot{x}(0)$ and (x_0, α). With respect to the dimensions $[x]$ and $[t]$ the dimensional matrix of this relation is

	$\dot{x}(0)$	x_0	α
$[x]$	1	1	1
$[t]$	−1	0	1

The null space of this dimensional matrix is spanned by $(1, -2, 1)$ and so the corresponding dimensionless quantity is

$$\begin{bmatrix}\dot{x}(0) & x_0 & \alpha\end{bmatrix}\begin{bmatrix}1\\-2\\1\end{bmatrix} = \frac{\dot{x}(0)\alpha}{x_0^2}.$$

Similar to Example 4.2, this implies that $\dot{x}(0) = -\frac{c}{\alpha}x_0^2$ for some unique dimensionless constant c. The inclusion of a minus sign will be explained shortly. Now by Bellman's *principle of optimality* we have in fact that[2]

$$\dot{x}(t) = -\frac{c}{\alpha}x^2(t) \quad \text{given } x(0) = x_0 > 0, \alpha > 0 \tag{4.4}$$

for all time. We have included a minus sign because then the solution $x(t)$ of this differential equation converges to zero only if $c > 0$. This convergence is needed to ensure convergence of the integral in (4.3). With this differential equation, the minimization problem (4.3) reduces to a minimization over a single number,

$$\min_{c>0} \int_0^\infty (1+c^2)x^4(t)\,\mathrm{d}t \quad \text{given } \dot{x}(t) = -\tfrac{c}{\alpha}x^2(t) \text{ and } x(0) = x_0 > 0, \alpha > 0.$$

We can apply the π-theorem once more: since the solution $x(t)$ of the differential equation $\dot{x}(t) = -\frac{c}{\alpha}x^2(t)$ with $x(0) = x_0$ depends on $(x_0, t, \frac{c}{\alpha})$, we also have the dimensional matrix

	$x(t)$	x_0	t	c/α
$[x]$	1	1	0	−1
$[t]$	0	0	1	−1

The differential equation is hence equivalent to one in $4 - 2 = 2$ dimensionless quantities, for instance, $(\frac{x(t)}{x_0}, \frac{c}{\alpha}x_0 t)$. Therefore,

$$\frac{x(t)}{x_0} = f(\tau) \tag{4.5}$$

for some dimensionless function f and dimensionless "time" $\tau = \frac{c}{\alpha}x_0 t$. In this form the integral that we want to minimize over positive c becomes

(4.6)
$$\int_0^\infty (1+c^2)x^4(t)\,\mathrm{d}t = \int_{\tau=0}^{\tau=\infty} (1+c^2)x_0^4 f^4(\tau)\,\mathrm{d}\big(\tfrac{\alpha}{cx_0}\tau\big) = \left(\frac{1}{c}+c\right)x_0^3\alpha\int_0^\infty f^4(\tau)\,\mathrm{d}\tau.$$

[2] Bellman's principle of optimality is a famous result in optimal control. Proof: Take an arbitrary $T > 0$ and split the integral in (4.3) into two parts, one from 0 to T and one from T to infinity: $\int_0^T x^4(t) + \alpha^2\dot{x}^2(t)\mathrm{d}t + \int_T^\infty x^4(t) + \alpha^2\dot{x}^2(t)\mathrm{d}t$. If x with given initial condition $x(0) = x_0$ minimizes the combined integral, then it also minimizes the second integral $\int_T^\infty x^4(t) + \alpha^2\dot{x}^2(t)\mathrm{d}t$ with given initial condition $x(T)$. This similarly gives that $\dot{x}(T) = -\frac{c}{\alpha}x^2(T)$. Since this holds for every $T > 0$, the result follows.

Assuming that the integral is convergent, the above is minimal over $c > 0$ iff $\frac{1}{c} + c$ is minimal. That is the case iff $c = 1$. The differential equation that determines the optimal solution is thus

$$\dot{x}(t) = -\tfrac{1}{\alpha}x^2(t) \qquad \text{given } x(0) = x_0 > 0, \alpha > 0.$$

Here the contribution of dimensional arguments ends. (One can solve (4.4) and the answer is $x(t) = \frac{x_0}{tx_0c/\alpha+1}$, which equals (4.5) for $f(\tau) = 1/(\tau + 1)$. Substituting $c = \alpha = 1$ we get the solution of the original minimization problem (4.2).)

By the way, aren't you surprised that the right-most side of (4.6) contains a factor x^3, while the integral on the left-most side contains a factor x^4? Did we make a mistake? You know what needs to be done now: we need to check the dimensions. Since c and f and τ are all dimensionless, the dimension of the right-most side of (4.6) is $[x]^3[\alpha] = [x]^3[x][t] = [x]^4[t]$. This does agree with the dimension of the integral on the left-most side. Equation (4.6) is probably correct after all.

In the above example we tacitly assumed that the minimization problem has a solution. In the final example of this section we introduce a possible crowd model, or at least one component of it. It is inspired by dimensional arguments and it appears that these models do not suffer from the somewhat unnatural bouncing effects that some crowd models exhibit, e.g., [10]. In any event it demonstrates once more the power and beauty of dimensional analysis.

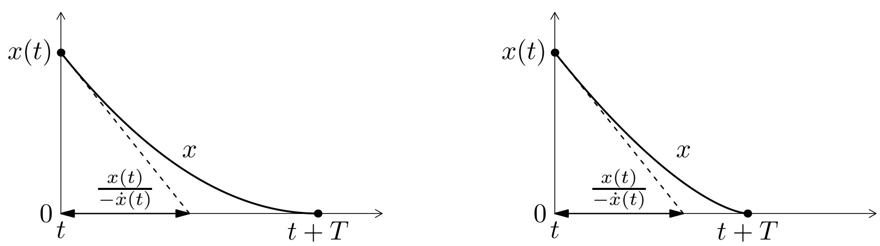


Fig. 6 *Crowd modeling—collision avoidance. Given $x(t) > 0$ and $\dot{x}(t) < 0$, the figure on the left shows the solution x for constant acceleration. The figure on the right shows the solution x for constant power. See Example* 4.7.

Example 4.7 (crowd modeling—collision avoidance). Suppose that at some time t you find yourself at a distance $x(t) > 0$ from a wall and running toward this wall with some speed $\dot{x}(t) < 0$. What would you do? If you want to avoid a collision, but without much effort, then one option might be to apply a constant acceleration $\ddot{x}(t)$ that brings you, at some time in the future, to a standstill at the wall; see Figure 6. That is, you might want to choose the smallest constant acceleration that prevents you from running into the wall. This constant acceleration $\ddot{x}(t)$ is a function of $x(t)$ and $\dot{x}(t)$. With respect to the dimensions $[x]$ and $[t]$ the dimensional matrix of this relation is

	$\ddot{x}(t)$	$\dot{x}(t)$	$x(t)$
$[x]$	1	1	1
$[t]$	−2	−1	0

.

The null space is spanned by $(1, -2, 1)$ so the π-theorem dictates that the relation is equivalent to a relation in the single dimensionless quantity

$$\begin{bmatrix}\ddot{x}(t) & \dot{x}(t) & x(t)\end{bmatrix}\begin{bmatrix}1\\-2\\1\end{bmatrix} = \frac{\ddot{x}(t)x(t)}{\dot{x}^2(t)}.$$

This $\ddot{x}(t)x(t)/\dot{x}^2(t)$ is unique because $\ddot{x}(t)$ follows uniquely from $\dot{x}(t), x(t)$. Hence,

$$\ddot{x}(t) = c\frac{\dot{x}^2(t)}{x(t)} \tag{4.7}$$

for some dimensionless constant c. A constant acceleration means that $x(t)$ is a parabola and from that fact it is not too hard to show that $c = 1/2$; see Figure 6 (left). In fact, we will prove it soon.

This differential equation (4.7) is preferred over an explicit formula for the parabola because the differential equation solves the problem at any moment anew. If, for instance, someone pushes you to another state $x(t), \dot{x}(t)$, then the differential equation (4.7) immediately provides the correct acceleration from that new state.

The next question might be: how long does it take for you to reach the wall? This time T is a function of $x(t), \dot{x}(t)$ and then again it is a matter of standard dimensional arguments to prove that T must be of the form

$$T = b\frac{x(t)}{-\dot{x}(t)} \tag{4.8}$$

for some dimensionless constant b. As Figure 6 (left) suggests, the value of this constant is $b = 2$. You can do the derivation yourself.

But why would you apply a constant acceleration? Perhaps it is better to apply a constant power. Power is the derivative of energy with respect to time, $\frac{\mathrm{d}}{\mathrm{d}t}(\frac{1}{2}m\dot{x}^2(t)) = m\ddot{x}(t)\dot{x}(t)$. So, up to a constant, the power equals $\ddot{x}(t)\dot{x}(t)$. The smallest constant power $\ddot{x}(t)\dot{x}(t)$ that prevents you from hitting the wall is a function of $x(t), \dot{x}(t)$ and hence this again defines a relation between $\ddot{x}(t), \dot{x}(t), x(t)$, but then the π-theorem dictates that it must again be of the form (4.7) for some dimensionless constant c. It can be shown[3] that now $c = 1/3$; see Figure 6 (right).

In fact, the π-theorem says that *every* policy—whether it be constant acceleration, constant power, or satisfying whatever rule—that at every moment in time t anew determines the acceleration $\ddot{x}(t)$ as a function of just $x(t), \dot{x}(t)$, must be of the form (4.7), and then the "hit time" T must be of the form (4.8). It is interesting at this point to interpret the parameter c and to allow for other values of $c > 0$. It is dimensionless (nice) and it expresses how cautious you are: the larger c is, the more cautious you are. A deeper analysis shows that for $0 < c < 1$ you hit the wall in finite time $T = \frac{1}{1-c}\frac{x(t)}{-\dot{x}(t)}$ and that for $c \geq 1$ you approach the wall closer and closer but never hit it.

There are many more applications of dimensional analysis. G.I. Taylor's accurate estimation of the energy released by the first atomic blast in New Mexico is often recounted as an application of the π-theorem. However, Taylor's approach [17, 18] is in

[3] Equation $\ddot{x} = c\dot{x}^2/x$ is equivalent to $\ddot{x}/\dot{x} = c\dot{x}/x$. Integrating both sides of the equality over time shows that it is equivalent to $\dot{x} = \alpha x^c$ for some constant α. From this it follows that $\ddot{x} = \alpha c x^{c-1}\dot{x} = \alpha^2 c x^{2c-1}$ and $\ddot{x}\dot{x} = \alpha^3 c x^{3c-1}$. Hence, $\ddot{x}\dot{x}$ is constant iff $c = 1/3$ (and $\ddot{x}$ is constant iff $c = 1/2$).

fact quite different [8]. Still, useful applications are abundant. A classic dimensionless quantity is the *Froude number*, which is defined as $v/\sqrt{gh}$ where v is a speed, h is a height, and g is the gravitational acceleration. It plays an import role in fluid mechanics, but it has also been used in models of walking animals where now v denotes the speed of the animal, h its hip height, and g, still, the gravitational acceleration. To some degree, every type of walking (such as normal walking or running) appears to have its own Froude number [2]. An extension of this model takes into account the stride length s, which is the distance between two footsteps. In that case the π-theorem says that the relation must be of the form

$$\frac{v}{\sqrt{gh}} = f\left(\frac{s}{h}\right)$$

for some function f. Based on this form and walking experiments with all sorts of animals, speeds, and surfaces it has been argued that

$$v = 0.25\sqrt{gh}\,\left(\frac{s}{h}\right)^{1.67}$$

is a reasonable model for the walking speed of all types of animals and all types of walking [1]. This is interesting because footprints of walking dinosaurs have been found, so once we know their hip heights we might have a way of figuring out how fast dinosaurs used to walk on our planet millions of years ago [1].

5. Scaling Matrices and a Scaling π-Theorem. The classic π-theorem, presented in previous sections, *assumes* that we are dealing with a physical law, so it *assumes* that all quantities have a certain dimension composed of predefined dimensions that we somehow are supposed to know, and it *assumes* that the relation is unit independent with respect to these dimensions. All in all this is quite a few assumptions.

From a mathematical perspective there are no predefined dimensions and unit-independence properties, and this begs the question of whether there is a π-theorem that does not require dimensions and units. There is! It turns out that we need to look at the scaling properties of the relation. The scaling approach that we present in this section assumes much less than the physical approach and in that sense is easier and more general. But it is also more abstract. In the end, here too we like to work with "dimensions," but then only as a tool to collect the scaling properties.

For the analysis to come we need to extend the rules of calculus of vector-matrix exponentiation. In accordance with this we define the *elementwise product* $x \odot z$ of two vectors $x = \begin{bmatrix} x_1 & x_2 & \cdots & x_n \end{bmatrix}$ and $z = \begin{bmatrix} z_1 & z_2 & \cdots & z_n \end{bmatrix}$ as

$$x \odot z = \begin{bmatrix} x_1 z_1 & x_2 z_2 & \cdots & x_n z_n \end{bmatrix}.$$

Vector-matrix exponentiation, together with this type of multiplication, has the neat properties

$$(x^A)^B = x^{(AB)} \text{ with } AB \text{ the normal matrix product,} \tag{5.1a}$$

$$(x \odot z)^C = (x^C) \odot (z^C). \tag{5.1b}$$

These properties are neat because they are perfect generalizations of what we know of exponentiation with numbers. The first property we have used already, and the second property we need now. Presenting the proofs of these properties is just a matter of writing them out. The proof of (5.1b) for vectors with two entries is

$$\begin{aligned}\left(\begin{bmatrix}x_1 & x_2\end{bmatrix}\odot\begin{bmatrix}z_1 & z_2\end{bmatrix}\right)^{\left[\begin{smallmatrix}a & b\\ c & d\end{smallmatrix}\right]} &= \begin{bmatrix}x_1z_1 & x_2z_2\end{bmatrix}^{\left[\begin{smallmatrix}a & b\\ c & d\end{smallmatrix}\right]}\\ &= \begin{bmatrix}(x_1z_1)^a(x_2z_2)^c & (x_1z_1)^b(x_2z_2)^d\end{bmatrix}\\ &= \begin{bmatrix}x_1^ax_2^c & x_1^bx_2^d\end{bmatrix}\odot\begin{bmatrix}z_1^az_2^c & z_1^bz_2^d\end{bmatrix}\\ &= \begin{bmatrix}x_1 & x_2\end{bmatrix}^{\left[\begin{smallmatrix}a & b\\ c & d\end{smallmatrix}\right]}\odot\begin{bmatrix}z_1 & z_2\end{bmatrix}^{\left[\begin{smallmatrix}a & b\\ c & d\end{smallmatrix}\right]}.\end{aligned}$$

In the scaling approach, the role of the dimensional matrix is taken over by what is called "scaling matrix."

Example 5.1 (scaling matrix). Consider the relation

$$\mathbb{X} = \{x\in\mathbb{R}_+^3 \mid x_3 = \tfrac{1}{6}x_1^2/x_2\}.$$

It has two clear scaling properties: for every $\mu_1, \mu_2 > 0$, if we independently scale x_1, x_2 as

$$x_1 \to \mu_1 x_1 \qquad \text{and} \qquad x_2 \to \mu_2 x_2, \tag{5.2}$$

then $x_3 \to \mu_1^2/\mu_2 x_3$. We collect these two scaling properties in a *scaling matrix* $S \in \mathbb{R}^{2\times 3}$ with as many rows as we have scalings and as many columns as we have quantities,

(5.3)

	x_1	x_2	x_3
$(x_1,x_2)\to(\mu_1x_1,x_2)$	1	0	2
$(x_1,x_2)\to(x_1,\mu_2x_2)$	0	1	−1

As with dimensional matrices it now pays to switch to new coordinates $y = x^C$, where the first column(s) of C span the null space of S. The null space of the above scaling matrix S is spanned by $S^\perp = (-2, 1, 1)$. Based on this we choose

$$C = \begin{bmatrix}-2 & 1 & 0\\ 1 & 0 & 1\\ 1 & 0 & 0\end{bmatrix},$$

and then the transformed coordinates become $y = x^C = \begin{bmatrix}x_2x_3/x_1^2 & x_1 & x_2\end{bmatrix}$. Obviously, the two scalings of (5.2) have no effect on $x_2x_3/x_1^2 = 1/6$ so the scaling matrix of the transformed coordinates x^C is

(5.4)

	x_2x_3/x_1^2	x_1	x_2
$(x_1,x_2)\to(\mu_1x_1,x_2)$	0	1	0
$(x_1,x_2)\to(x_1,\mu_2x_2)$	0	0	1

In these new coordinates, the first entry $x^{S^\perp} = x_2x_3/x_1^2$ is scale-invariant and it is also immediate that the relation $\mathbb{X}$ is completely specified by this first entry: $x^{S^\perp} = \frac{1}{6}$.

The notation and calculus of vector-matrix exponentiation is of use now: let μ be the row vector $\mu = \begin{bmatrix}\mu_1 & \mu_2\end{bmatrix}$ and let S be the above scaling matrix. Then we have

$$\begin{aligned}\mu^S\odot x &= \begin{bmatrix}\mu_1 & \mu_2\end{bmatrix}^{\left[\begin{smallmatrix}1 & 0 & 2\\ 0 & 1 & -1\end{smallmatrix}\right]}\odot\begin{bmatrix}x_1 & x_2 & x_3\end{bmatrix}\\ &= \begin{bmatrix}\mu_1 & \mu_2 & \mu_1^2/\mu_2\end{bmatrix}\odot\begin{bmatrix}x_1 & x_2 & x_3\end{bmatrix} = \begin{bmatrix}\mu_1x_1 & \mu_2x_2 & \mu_1^2/\mu_2x_3\end{bmatrix}.\end{aligned}$$

So the scaling property that says that certain scaled versions of x also satisfy the relation can be succinctly expressed as

$$x \in \mathbb{X} \quad \Longrightarrow \quad \mu^S \odot x \in \mathbb{X} \quad \forall \mu > 0.$$

Here the inequality $\mu > 0$ means that every entry of μ is positive.

Motivated by the last example we define scaling matrices as follows.

DEFINITION 5.2 (scaling matrix). *A matrix $S \in \mathbb{R}^{k\times n}$ is a* scaling matrix *of a relation $\mathbb{X} \subseteq \mathbb{R}^n_+$ if $x \in \mathbb{X}$ implies $\mu^S \odot x \in \mathbb{X}$ for all $\mu \in \mathbb{R}^k_+$.*

Given a scaling matrix, we say that x and $\hat{x}$ are *equivalent* if one is a scaled version of the other: $\hat{x} = \hat{\mu}^S \odot x$ for some $\hat{\mu} > 0$. It is important to realize that x is in the relation iff one of its equivalent $\hat{x}$ is in the relation. Indeed, if $x \in \mathbb{X}$, then by definition $\hat{\mu}^S \odot x \in \mathbb{X}$. Conversely, if $\hat{\mu}^S \odot x \in \mathbb{X}$, then $\mu^S \odot (\hat{\mu}^S \odot x)$ is in $\mathbb{X}$ for every $\mu > 0$ which for the elementwise inverse $\mu = 1/\hat{\mu}$ gives $x \in \mathbb{X}$. So x satisfies the relation *iff* a conveniently scaled $\hat{\mu}^S \odot x$ does so.

In Example 5.1 it was quite easy to figure out the scaling matrix (5.4) of the transformed $y = x^C$. In the general case this is also easy.

LEMMA 5.3 (transformed scaling matrix). *Let $\mathbb{X} \subseteq \mathbb{R}^n_+$ and $\mathbb{Y} = \{x^C \mid x \in \mathbb{X}\}$. If S is a scaling matrix of $\mathbb{X}$, then SC is a scaling matrix of $\mathbb{Y}$. If C is invertible, then S is a scaling matrix of $\mathbb{X}$ if and only if SC is scaling matrix of $\mathbb{Y}$.*

Proof. If S is a scaling matrix of $\mathbb{X}$ and $y = x^C$ for some $x \in \mathbb{X}$, then $\mu^{SC} \odot y = \mu^{SC} \odot (x^C) = (\mu^S \odot x)^C$, which is in $\mathbb{Y}$ because $\mu^S \odot x \in \mathbb{X}$. So SC is a scaling matrix of $\mathbb{Y}$. If C is invertible, then, likewise, any scaling matrix $\tilde{S}$ of $\mathbb{Y}$ makes $\tilde{S}C^{-1}$ a scaling matrix of $\mathbb{X}$. ∎

In Example 5.1 the scaling matrix (5.4) of x^C indeed equals SC. Now the scaling π-theorem is a piece of cake.

THEOREM 5.4 (scaling π-theorem). *Let $S \in \mathbb{R}^{k\times n}$ be a scaling matrix of a relation $\mathbb{X} \subseteq \mathbb{R}^n_+$. If the columns of some matrix $S^\perp$ span the null space of S, then*

$$x \in \mathbb{X} \quad \Longleftrightarrow \quad x^{S^\perp} \in \mathbb{X}^{S^\perp},$$

where $\mathbb{X}^{S^\perp} = \{x^{S^\perp} \mid x \in \mathbb{X}\}$. Moreover, $x^{S^\perp}$ is scale-invariant with respect to S.

Proof. The proof is very similar to that of Theorem 3.7. Suppose for now that the rank of S equals the number of rows of S. Then there is a matrix A such that $\left[\begin{smallmatrix} A \\ S \end{smallmatrix}\right]$ is square invertible. Define C as $C = \left[\begin{smallmatrix} A \\ S \end{smallmatrix}\right]^{-1}$. It is invertible and has the property that $SC = \left[\begin{smallmatrix} 0 & I_k \end{smallmatrix}\right]$. Then $y = x^C$ is a coordinate change and the scaling matrix becomes $SC = \left[\begin{smallmatrix} 0 & I_k \end{smallmatrix}\right]$. Now $y = x^C$ is in the relation iff $\hat{\mu}^{SC} \odot y$ is for some convenient $\hat{\mu}$. For the elementwise inverse $\hat{\mu} = 1/y$ we find $\hat{\mu}^{SC} \odot y = (y_1, \ldots, y_{n-k}, 1, \ldots, 1)$; that is, the relation is completely determined by $(y_1, \ldots, y_{n-k})$. We have that $(y_1, \ldots, y_{n-k}) = x^{C_0}$, where C_0 is the matrix formed by the first $n-k$ columns of C. We have thus proved that $x \in \mathbb{X} \iff x^C \in \mathbb{X}^C \iff x^{C_0} \in \mathbb{X}^{C_0}$. Since both $S^\perp$ and C_0 span the null space of S, we have that $x^{C_0} \in \mathbb{X}^{C_0} \iff x^{S^\perp} \in \mathbb{X}^{S^\perp}$.

If the rank of S is less than the number of rows of S, then one by one remove rows from S that do not affect the null space of S until the number of rows equals the rank. The reduced matrix S_0 obtained is still a scaling matrix. Now apply the above arguments with S replaced by S_0.

The $x^{S^\perp}$ is scale-invariant because its scaling matrix is $SS^\perp = 0$. ∎

The theorem says that the relation in x is equivalent to a relation in $x^{S^\perp}$. Normally, $x^{S^\perp}$ has fewer entries than x. As before, linear algebra tells us that the minimal number of entries of $x^{S^\perp}$ is $n - \text{rank}(S)$. Therefore, the more "independent" scaling properties we have, the higher the rank of the scaling matrix S is and, hence, the fewer entries we end up with in $x^{S^\perp}$.

We hope that you appreciate the fact that the scaling π-theorem assumes much less than the physical version. Indeed, the scaling properties *follow from* the relation, while in the physical version the dimensions and unit-independence properties are *imposed on* the relation. On the other hand, determining the scaling properties normally requires an explicit description of the relation (e.g., using equations), while the physical version can often be applied even before having equations.

5.1. Assigning Dimensions. In the scaling approach it is also convenient to work with dimensions. This is not strictly needed, but the advantage is that introducing suitable dimensions is easy in practice, while it can sometimes be tedious to work out the scaling properties.

Example 5.5 (interpreting scaling matrices as dimensional matrices). Consider again the relation of Example 5.1 and its two scalings (5.2), which we reproduce here:

$$x_1 \to \mu_1 x_1 \qquad \text{and} \qquad x_2 \to \mu_2 x_2.$$

Associated with these two scalings we introduce the two dimensions $[x_1]$ and $[x_2]$. To avoid comparing apples and oranges in the identity $x_3 = \frac{1}{6}x_1^2/x_2$ we should set $[x_3]$ equal to $[x_3] = [x_1]^2[x_2]$. With respect to the two dimensions $[x_1], [x_2]$ the dimensional matrix is

	x_1	x_2	x_3
$[x_1]$	1	0	2
$[x_2]$	0	1	-1

,

and this is precisely the scaling matrix (5.3). Is our relation a physical law with respect to these dimensions $[x_1], [x_2]$? Yes, because unit independence is equivalent to having the scaling property. For example—but you probably get it already—decreasing the unit of x_1 with a factor μ_1 corresponds to a scaling of the values $(x_1, x_2, x_3) \to (\mu_1 x_1, x_2, \mu_1^2 x_3)$.

From a mathematical/scaling perspective *every* relation is a physical law because we introduce only those dimensions that correspond to scaling properties. For instance, the weird looking relation

$$\mathbb{X} = \{x \in \mathbb{R}_+^3 \mid x_1 x_3^3 > (5 + x_1^2)x_2^2\}$$

can be turned into a physical law by introducing dimension $[x_2]$ and letting $[x_1] = 1$ (dimensionless) and $[x_3] = [x_2]^{2/3}$. This is a quick way to say that the relation has the scaling property

$$(x_1, x_2, x_3) \in \mathbb{X} \quad \Longrightarrow \quad (x_1, \mu x_2, \mu^{2/3} x_3) \in \mathbb{X} \quad \forall \mu > 0.$$

Introducing suitable dimensions is easy in most cases. In fact we did it already in Example 4.5 and some other examples in section 4. Intuition is fine, but if you are in doubt about your choice of dimensions, then a check of the implied scaling properties is needed. After all, the only thing that matters in the π-theorem is the scaling property.

5.2. Applications of the Scaling π-Theorem. The two versions of the π-theorem (physical and scaling) form two sides of the same coin and lead to the same conclusions in applications, although one might require more work than the other. To illustrate the equivalence we again consider the Pythagorean theorem but now using scaling arguments.

Example 5.6 (Pythagorean theorem). We explore the relation between the area A of a right triangle and its hypotenuse c and acute angle ϕ; see Figure 4 (left). It is clear that if the hypotenuse scales as $c \to \mu c$ and ϕ is kept the same, then the area scales as $A \to \mu^2 A$. The scaling matrix with respect to this scaling property is

	A	c	ϕ
$(c, \phi) \to (\mu c, \phi)$	2	1	0

and the null space of this scaling matrix $S = \begin{bmatrix} 2 & 1 & 0 \end{bmatrix}$ is spanned by the columns of

$$S^\perp = \begin{bmatrix} 0 & 1 \\ 0 & -2 \\ 1 & 0 \end{bmatrix}.$$

Thus, the scaling π-theorem says that the relation between (A, c, ϕ) is equivalent to a relation between $[A \;\; c \;\; \phi]^{S^\perp} = [\phi \;\; A/c^2]$. From here on the arguments are the same as those in Example 4.3 where we applied the physical version of the π-theorem.

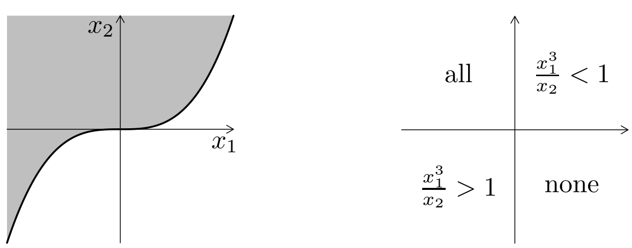


Fig. 7 *The relation* $\{x \in \mathbb{R}^2 \mid x_2 > x_1^3\}$ *(shown left) described in terms of* x_1^3/x_2 *is different in every quadrant (shown right). See Example* 5.7.

The two π-theorems require that the x_i are positive real numbers or, more generally, that they have a fixed sign in the relation. If the x_i are real numbers that may change sign, then we may have to split the domain into several orthants.

Example 5.7 (orthants and sets). In applications the relation in x is often a function, $x_1 = f(x_2, \ldots, x_n)$, but the π-theorem applies to general relations. Example 4.5 is a case in point. Here is another: Consider the relation

$$\mathbb{X} = \{x \in \mathbb{R}^2 \mid x_2 > x_1^3\};$$

see Figure 7 (left). It is defined on the entire two-dimensional plane so not just on one quadrant. The scalings μ are still positive. A possible scaling matrix is $S = \begin{bmatrix} 1 & 3 \end{bmatrix}$ and so we may choose $S^\perp = \left[\begin{smallmatrix} 3 \\ -1 \end{smallmatrix}\right]$. The π-theorem then says that *per quadrant* the $x^{S^\perp} = x_1^3/x_2$ completely specifies the relation. In this example the description of the relation in terms of x_1^3/x_2 in fact is different in every quadrant; see Figure 7 (right).

Now it is time to settle the claim put forward at the end of section 2: that x in expressions such as e^x and $\cos(x)$ is necessarily dimensionless. The scaling version is easier to formulate than the physical version.

THEOREM 5.8 (why function arguments and outcomes are dimensionless). *Let f be a function from $\mathbb{R}_+$ to $\mathbb{R}$ and consider the relation $\{(x_1, x_2) \mid x_2 = f(x_1), x_1 \in \mathbb{R}_+\}$.*

- Scaling version. *Either both x_1 and x_2 are scale-invariant with respect to every possible scaling of the relation, or $f(x_1) = cx_1^\alpha$ for some $c, \alpha \in \mathbb{R}$.*
- Physical version. *If the relation is a physical law in x_1, x_2 with respect to some dimension, and the relation does not involve any other dimensioned quantity, then either both x_1 and x_2 are dimensionless or $f(x_1) = cx_1^\alpha$ for some $c, \alpha \in \mathbb{R}$.*

Proof. Let $S = \begin{bmatrix} s_1 & s_2 \end{bmatrix}$ be a scaling matrix of the relation and note that every relation has at least the trivial scaling $S = \begin{bmatrix} 0 & 0 \end{bmatrix}$. Demanding that the relation is a physical law in x_1, x_2 and no other dimensioned quantity means that the relation has a dimensional matrix in x_1, x_2 alone:

	x_1	x_2
some dimension	m_1	m_2

The two versions are hence equivalent with $s_1 = m_1, s_2 = m_2$. We prove the first (the scaling version). Having a scaling property means that there are $s_1, s_2 \in \mathbb{R}$ such that

$$x_2 = f(x_1) \implies \mu^{s_2} x_2 = f(\mu^{s_1} x_1) \tag{5.5}$$

for every $\mu > 0$ and $x_1 > 0$. For the special choice $\mu = x_1^{-1/s_1}$ this says that $x_2 = f(x_1) \implies x_1^{-s_2/s_1} x_2 = f(1)$. That is, $x_2 = f(1)x_1^{s_2/s_1}$. Hence, the function is a power function,

$$f(x_1) = cx_1^\alpha, \tag{5.6}$$

with $c = f(1)$ and $\alpha = s_2/s_1$. Here we have assumed $s_1 \neq 0$. If $s_1 = 0$, then (5.5) implies that either f is the zero function (which is also of the form $f(x_1) = cx_1^\alpha$) or $s_2 = 0$, because otherwise x_2 is not a function of x_1. The latter means the zero scaling $s_1 = s_2 = 0$, i.e., x_1, x_2 are scale-invariant, i.e., are dimensionless. ∎

Functions of the form $f(x_1) = cx_1^\alpha$ are known as *power functions* and the above thus says that we can have nontrivial dimensions in $x_2 = f(x_1)$ only if it is a power function!

Example 5.9 (dimensionless functions). Let $\beta = \int_0^\infty e^{-\alpha t}\, dt$. Since both $e^{-\alpha t}$ and $-\alpha t$ must be dimensionless, we have that $[\beta] = [t]$ and $[t] = [\alpha]^{-1}$. Hence, $[\beta] = [\alpha]^{-1}$. Clearly, β is a function of α. Then the π-theorem dictates that $\beta = c/\alpha$ for some dimensionless constant c. More examples of this sort can be found in [11].

Knowing these dimensional arguments is really helpful in interpreting formulas and spotting errors. For instance, if you read a book on statistics and come across the probability density function of the exponential distribution,

$$f_\lambda(x) = \begin{cases} \lambda\, e^{-\lambda x} & \text{if } x \geq 0, \\ 0 & \text{if } x < 0, \end{cases} \tag{5.7}$$

then your immediate response should be that $[\lambda] = [x]^{-1}$. Indeed it is, and then you should not be surprised when the book calls λ a "rate." Now suppose that someone claims that the mean of the exponential distribution is λ; then you can immediately, and without any calculation, convince them otherwise because the mean must have the same dimension as x and so the mean cannot be λ. Incidentally, notice that this $f_\lambda(x)$ has dimension $[\lambda] = [x]^{-1}$, while we just claimed in Theorem 5.8 that only power functions can have a nontrivial dimension. The thing is that f_λ depends on more than one dimensioned quantity (namely, x and λ) and such functions are excluded in Theorem 5.8.

In fact, the dimension of $f(x)$ equals $[x]^{-1}$ for *every* probability density function f, because $\int_{-\infty}^{\infty} f(x)\,\mathrm{d}x = 1$ (think about it). So if some professor of mathematics would claim that the normal distribution with mean m and standard deviation $\sigma > 0$ has probability density function

$$f(x) = \frac{1}{\sqrt{2\pi\,\sigma}}\,\mathrm{e}^{-(x-m)^2/(2\sigma^2)},$$

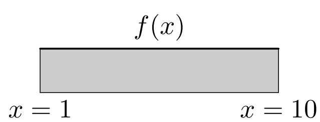
then you immediately know that he is mistaken,[4] because the standard deviation σ has the same dimension as x which means that $f(x)$ has the wrong dimension $1/\sqrt{[\sigma]} = 1/\sqrt{[x]}$. Based on this you probably even know how to correct the mistake.

The final example of this paper deals with an awesome application involving scale-invariance and probability density functions that are power functions (on a restricted domain).

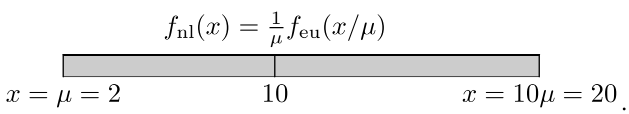
Example 5.10 (Benford's law via scaling). Suppose we are given the account balances of a large and diverse group of account holders. Obviously the leading digit of each balance lies in $\{1, 2, \ldots, 9\}$ and it seems safe to assume that all nine digits appear with equal probability. A simple "Nexit" thought experiment, however, demonstrates that this assumption is wrong! Nexit is the (hypothetical) process of the Netherlands leaving the European Union. To analyze this problem we express balances b in scientific notation:

$$b = x \times 10^m \quad \text{with } 1 \le x < 10 \text{ and } m \in \mathbb{N}.$$

Now it may *seem* natural to assume that every $x \in [1, 10)$ is equally likely, i.e., that its probability density function f is constant,



but if b is in euros, then back in guilders the probability density function is this f scaled and stretched by a factor of about $\mu = 2$ because 1 euro is about 2 guilders:



Now, in guilders, all balances from $x = 10$ up to $x = 19.99\ldots$ have leading digit 1. For $\mu = 2$ this is more than 50%. So the seemingly innocent assumption that f is constant for euros leads to the conclusion that in guilders more than half of the balances have leading digit 1. Surely not.

[4]There is a nice popular math book that displays this wrong density function several times.

We need a bit of modeling before we can proceed. What follows is based on [13]. Given a probability density function f_{bal} of balances b in, say, euros, the probability density function f of $x \in [1, 10)$ follows as

$$f(x) = \sum_{k=-\infty}^{\infty} 10^k f_{\text{bal}}(10^k x).$$

(You may need pen and paper to verify this equation.) This f is a probability density function over the interval $[1, 10)$, but it is defined for every $x > 0$. We have just seen that the uniform distribution is not scale-invariant. A natural question is: is there an f that *is* scale-invariant, i.e., that does not depend on the choice of unit (euro or guilder)? Under scaling of the balances, $b \to \mu b$, the probability density function of b becomes $f^{\text{new}}_{\text{bal}}(b) = \frac{1}{\mu} f_{\text{bal}}(b/\mu)$ and then the probability density function of x similarly becomes $f^{\text{new}}(x) = \frac{1}{\mu} f(x/\mu)$. Hence, the question is whether there is a probability density function f that satisfies the scaling property

$$f(x) = \tfrac{1}{\mu} f(x/\mu) \qquad \forall \mu > 0, \forall x \in [1, 10). \tag{5.8}$$

If so, then taking $\mu = x$ shows that f must be a power function of the form

$$f(x) = c/x$$

for some constant $c = f(1)$. Conversely, every such power function satisfies the scaling property (5.8). Hence, the functions that satisfy the scaling property (5.8) are *precisely* the power functions $f(x) = c/x$. Our function f also needs to be a probability density function on $[1, 10)$ so that $\int_1^{10} f(x)\,\mathrm{d}x = 1$. This gives $c = 1/\ln(10)$ and therefore

$$f_{\text{Benford}}(x) = \frac{1/\ln(10)}{x}$$

is the unique answer to our problem. This is the *Benford probability density function*, in honor of one of its inventors [3]. Now that f is known (see Figure 8), it is easy to figure out the distribution of the leading digits. From Figure 8 we can see that in about 30% of the cases the leading digit is 1, in about 18% the leading digit is 2, etc. This is not even close to a uniform distribution. It is an intriguing fact that many practical data sets follow the Benford distribution very well, so much so in fact that deviation from the Benford distribution in socio-economic data serves as an indicator of data manipulation and fraud [12, 14]. Very interesting material.

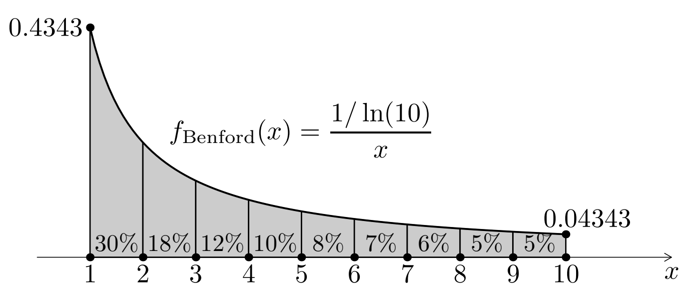


Fig. 8 *Benford probability density function; see Example* 5.10.

6. Concluding Remarks. I hope that by now you have developed the Pavlovian reaction that constantly makes you ask yourself, "what are the dimensions and is it dimensionally consistent?" In general this is a healthy reaction, but it might make you allergic to certain results where others see no problem. For instance, now that I am infected by the dimensional disease, I can no longer fully enjoy Shannon's entropy H_X for continuous processes. This entropy H_X is defined as

$$H_X = -\int_{-\infty}^{\infty} f(x) \ln(f(x))\,\mathrm{d}x,$$

where f a probability density function. Being a density function, it has dimension $[f(x)] = [x]^{-1}$. But then $\ln(f(x))$ is the logarithm of a dimensioned quantity. This is upsetting. Shannon himself was well aware of this problem and comments on it in [16]. It is not a real problem, he reasons, because in the end we only need *differences* of entropy and this cancels the dimensions. To me it still does not feel right, but—and this is an important point—in this case dimensional thinking was also valuable and because of it I understand entropy better now. That, in a nutshell, is what makes dimensional thinking so useful: it helps you to interpret each term in a formula and in that way you gain a better understanding of your formulas and mathematical models. On a lighter note, dimensional thinking is also a great way to impress your colleagues or fellow students because now you can grasp equations and spot errors much quicker.

Another goal of this paper was to make it self contained, with precise formulations and complete proofs of the π-theorems. Here, I believe, the notation and calculus of vector-matrix exponentiation was instrumental.

REFERENCES

[1] R. Alexander, *Estimates of speeds of dinosaurs*, Nature, 261 (1976), pp. 129–130. (Cited on p. 175)

[2] R. Alexander, *Optimization and gaits in the locomotion of vertebrates*, Physiol. Rev., 69 (1989), pp. 1199–1227. (Cited on p. 175)

[3] F. Benford, *The law of anomalous numbers*, Proc. Amer. Philos. Soc., 78 (1938), pp. 551–572. (Cited on p. 182)

[4] J. Bertrand, *Sur l'homogénéité dans les formules de physique*, Comptes Rendus, 86 (1878), pp. 916–920. (Cited on p. 159)

[5] E. Buckingham, *On physically similar systems: Illustrations of the use of dimensional equations*, Phys. Rev., 4 (1914), pp. 345–376. (Cited on pp. 159, 168)

[6] E. Buckingham, *The principle of similitude*, Nature, 96 (1915), pp. 396–397, https://doi.org/10.1038/096396d0. (Cited on p. 159)

[7] E. Carvallo, *Sur une similitude dans les fonctions des machines*, J. Phys. Theoret. Appl., 1 (1892), pp. 209–212. (Cited on p. 159)

[8] M. Deakin, *G.I. Taylor and the Trinity test*, Internat. J. Math. Ed. Sci. Tech., 42 (2011), pp. 1069–1079. (Cited on p. 175)

[9] A. Federman, *On some general methods of integration of first-order partial differential equations*, Ann. Inst. Polytech. Pierre le Grand à St. Pétersbourg, 16 (1911), pp. 1–59. (Cited on p. 159)

[10] I. Helbing, D. Farkas and T. Vicsek, *Simulating dynamical features of escape panic*, Nature, 407 (2000), pp. 487–490. (Cited on p. 173)

[11] S. Mahajan, *Street-Fighting Mathematics: The Art of Educated Guessing and Opportunistic Problem Solving*, MIT Press, Cambridge, MA, London, 2010. (Cited on p. 180)

[12] M. Nigrini, *I've got your number – how a mathematical phenomenon can help CPAs uncover fraud and other irregularities*, J. Accountancy, 187 (1999), pp. 79–83. (Cited on p. 182)

[13] R. Pinkham, *On the distribution of first significant digits*, Ann. Math. Statist., 32 (1961), pp. 1223–1230. (Cited on p. 182)

[14] B. Rauch, M. Göttsche, G. Brähler, and S. Engel, *Fact and fiction in EU-Governmental economic data*, German Econom. Rev., 12 (2011), pp. 243–255. (Cited on p. 182)

[15] D. Riabouchinsky, *Méthode des variables de dimension zéro et son application en aérodynamique*, l'Aérophile, (1911), pp. 407–408. (Cited on p. 159)
[16] C. Shannon, *A mathematical theory of communication*, Bell Syst. Tech. J., 27 (1948), pp. 379–423. (Cited on p. 183)
[17] G. Taylor, *The formation of a blast wave by a very intense explosion.* I. *Theoretical discussion*, Proc. Roy. Soc. Lond. Ser. A Math. Phys. Sci., 201 (1950), pp. 159–174. (Cited on p. 174)
[18] G. Taylor, *The formation of a blast wave by a very intense explosion.* II. *The atomic explosion of* 1945, Proc. Roy. Soc. Lond. Ser. A Math. Phys. Sci., 201 (1950), pp. 175–186, https://doi.org/10.1098/rspa.1950.0050. (Cited on p. 174)
[19] A. Vachy, *Sur les lois de similitude en physique*, Ann. Télégraphiques, 19 (1892), pp. 25–28. (Cited on p. 159)
[20] E. van Groesen and J. Molenaar, *Continuum Modeling in the Physical Sciences*, Math. Model. Comput. 13, SIAM, Philadelphia, 2007. (Cited on p. 161)