THE UNITS TELL YOU WHAT TO DO

Simon Brown

Abstract: Many students have some difficulty with calculations. Simple dimensional analysis provides a systematic means of checking for errors and inconsistencies and for developing both new insight and new relationships between variables. Teaching dimensional analysis at even the most basic level strengthens the insight and confidence of students, and provides them with a tool that can be applied in almost any context.

Key words: activity coefficient; dimensional analysis; equilibrium coefficient; pH

1. Introduction

Every science teacher has been confronted with a student unable to work out how to do certain calculations. Sometimes it is a simple problem, like calculating the number of moles in a given volume of solution of a specified concentration, but it can be more complicated. The decline in the mathematical skills [1, 33] and insight of students has increased the frequency with which this problem is encountered. Given this, strategies for addressing such problems have become increasingly important. One approach is to use the units as a guide.

I often repeat the title of this paper to my students in the hope that they will remember this idea if nothing else. The assertion is based on the concepts of dimensional analysis (DA), a technique often taught to chemists, physicists and engineers, but less frequently to biologists. At its simplest, DA can be used to check for errors in calculations or equations, but with very little more sophistication it can also enable one to work out the general form of an equation rather than having to remember it or look it up. The most sophisticated application facilitates the formulation of new relationships between variables: for example the relationship between particle size and the wavelength of light in Rayleigh scattering was determined using DA [32]; and it has been used by many other people, including Maxwell [24], Fourier [16] and Einstein [14].

If the correct use of units can be informative, an incorrect use can be catastrophic. Perhaps the best known example of the latter is the Mars Climate Orbiter ‘lost’ by NASA in 1999 because two groups of engineers used different units preventing correct data transfer [21]. Aviation is especially rich in such confusion, for example in 1983 an Air Canada flight ran out of fuel on the way to Edmonton because a non-metric fuel density (1.77 lb L⁻¹) was used to convert a mass of fuel in kilogrammes to a volume in litres [2]; in 1997 a Korean Air flight flew into a hillside as it approached the airport at Guam, apparently because the flight crew were unclear whether they were working in feet or metres [27]. Medical professionals sometimes have difficulty calculating drug doses because they have to interconvert units leading to the accidental administration of incorrect doses of medication [36, 37]. On at least two occasions American long and triple jumpers have encountered problems in international competition because metric units are used, rather than the imperial units (feet and inches) used in the United States. For example, Melvin Lister did not reach the triple jump final at the 2004 Olympics. Later he said “They made me use a metric tape for my run-up and I'm used to imperial. It ruined my trust.” [29] and, perhaps more understandably, “Nobody told me they were only going to have metric out there. I couldn't figure out what my mark was.” [12].

Here I outline the basis of DA and provide some examples of its application. In the first instance, I give some very simple illustrations of the fundamental ideas needed to develop and analyse technical
problems (sections 2 and 3). Finally, I outline a more formal approach and give some more examples (section 4), and then describe the value added from teaching DA (section 5).

2. The fundamentals of dimensional analysis

The general rule is very simple: the dimensions of all parameters related by addition, subtraction and an equals sign have to agree. Specifically, the dimensions on either side of ‘=’ must be the same (obviously, it makes no sense to write something like 1 mL = 1 mm); if numbers are added or subtracted they must have the same dimensions (for example, it makes no sense to write 5 mL + 2 mm); and if numbers are multiplied (or divided) the dimensions are just combined according to the mathematical operations (2 mm × 2 mm = 4 mm² or 20 km h⁻¹ / 2 km = 1 h⁻¹). Moreover, it is not possible to apply some mathematical functions to a unit (for example, it is not reasonable to write log(1 mm), sin(1 mm) or exp(1 mm)) so the arguments of those functions have to be dimensionless. A useful practice is to replace all SI prefixes with the appropriate multiplier (for example 1.5 mm = 1.5 × 10⁻³ m or 1 ML = 1 × 10⁶ L), which not only maintains dimensional consistency, but provides an explicit connection with ‘scientific notation’. Maxwell [24] introduced the convention that the dimensions of a variable (x) are indicated using square brackets ([x]). In chemistry such square brackets usually denote concentration, but I will use Maxwell’s convention here and write the concentration of x as cₓ.

Naturally, dimensional consistency and dimension interconversion can be analysed using this idea [34], but one can also use it to work out how to combine variables correctly. For example, the common unit of concentration (c) is molarity ([c] = mol L⁻¹), but molality (m) is also used ([m] = mol kg⁻¹) and these units are connected by the density of the solvent ([ρ] = kg L⁻¹). To convert a molarity to a molality assume that

\[ [m] = [c]^a \times [\rho]^b \]

or

\[ \text{mol kg}^{-1} = (\text{mol L}^{-1})^a \times (\text{kg L}^{-1})^b = \text{mol}^a \times L^{-a-b} \times \text{kg}^b \]

and, by equating the exponents of the units on each side

\[ \text{mol}^1 = \text{mol}^a \Rightarrow a = 1 \text{ and kg}^{-1} = \text{kg}^b \Rightarrow b = -1 \]

but, obviously, since L does not appear on the left \( L^{-a-b} = L^0 \), so \(-a - b = 0\), consistent with these values of \( a \) and \( b \). Of course, this means that \( m = c \times \rho^{-1} \). While this is a trivial example, the same approach works with much more complicated problems and I shall make use of it below.

The same approach works equally well if the dimensions are reduced to their essential components: length (L), mass (M), time (T) and temperature (Θ). More complex units are just combinations of these fundamental dimensions, for example

1. speed is just distance (L) moved in a given time (T), so [speed] = L T⁻¹;
2. density is just a mass (M) in a given volume (L³), so [density] = M L⁻³;
3. energy is measured in joules (J) and 1 J = 1 kg m² s⁻² (the energy needed to accelerate 1 kg at 1 m s⁻² over 1 m), so [energy] = M L² T⁻²; and
4. work (measured in watts, W) is the rate at which energy is expended, so [work] = [energy] T⁻¹ = M L² T⁻³.

Not having to worry about all the details of the specific units can make it easier to work out what calculation is actually required.

3. Devising new relationships and gaining insight

A simple application of DA can lead to apparently surprising insights that challenge students to question their assumptions, leading to deeper understanding. For example, one often sees
The units tell you what to do

\[ \text{pH} = -\log_{10}(c_{\text{H}^+}) \]  

(1)

or, equivalently,

\[ c_{\text{H}^+} = 10^{-\text{pH}}, \]

(2)

relating pH to the concentration of protons \((c_{\text{H}^+})\). By convention \([c_{\text{H}^+}] = \text{mole L}^{-1}\), rather than any of the other concentration units that might be employed, and, since \([\text{pH}] = 1\) and therefore \([10^{-\text{pH}}] = 1\), (2) is clearly dimensionally inconsistent (as is (1), since (1) and (2) are equivalent). As the argument of the logarithm must be dimensionless (because \(\log(\text{mole L}^{-1})\) is a nonsense, as Sørenson [31] recognised), there must be a factor that converts \(c_{\text{H}^+}\) into a dimensionless equivalent \((a_{\text{H}^+})\), so that (1) is actually

\[ \text{pH} = -\log_{10}(a_{\text{H}^+}). \]

(3)

In practice, \(a_{\text{H}^+}\) is the thermodynamic activity of protons, to which \(c_{\text{H}^+}\) is related by the apparent activity coefficient \((\gamma_{\text{H}^+}')\), so \(a_{\text{H}^+} = \gamma_{\text{H}^+}'c_{\text{H}^+}\), and \([\gamma_{\text{H}^+}'] = \text{L mol}^{-1}\). It is more usual [7] to assume a dimensionless activity coefficient \(\gamma_{\text{H}^+} = c_0\gamma_{\text{H}^+}'\), where \(c_0\) is a standard concentration (= 1 mol L\(^{-1}\)), so (1) becomes

\[ \text{pH} = -\log_{10}(a_{\text{H}^+}) = -\log_{10}(\gamma_{\text{H}^+}'c_{\text{H}^+}) = -\log_{10}(\gamma_{\text{H}^+}c_{\text{H}^+}/c_0) = -\log_{10}(\gamma_{\text{H}^+}) - \log_{10}(c_{\text{H}^+}/c_0). \]

Only in an ideal solution does \(\gamma_{\text{H}^+} = 1\), in which case

\[ \text{pH} = -\log_{10}(c_{\text{H}^+}/c_0), \]

so values of \(\gamma_{\text{H}^+} \neq 1\) indicate deviation from ideality. The realisation that (1) is dimensionally inconsistent, and the resolution of that inconsistency, prompts students to question and helps them to develop considerable insight. The simple reasoning leads to a deepening of student understanding.

In addition to dimensional inconsistency, there are at least three other problems with (1) and (2). First, there need be no unique relationship between \(c_{\text{H}^+}\) and pH, because \(\gamma_{\text{H}^+}\) need not be constant (for example the mean activity coefficients of various acids depend on the concentration of the acid [19]). McCarty and Vitz [25] give several examples for which (1) is incorrect in this respect, and suggest that students should be taught (3) rather than (1), but concentration is a challenge to many students [17] and activity is even more difficult. The second question is the significance of the ‘p’ in pH. If pH is related to the concentration of \(H^+\), then what has the ‘p’ got to do with it and why does it mean that pH decreases as \(c_{\text{H}^+}\) increases? It is of little help to most students to relate the ‘p’ to \textit{potenz, puissance} or \textit{power} [22, 28] and suggesting that ‘p’ is related to ‘–log10’ is often only marginally better and reduces the whole issue to one of calculation. The third problem, related to the second, is that the significance of ‘log10’ is often unclear to students because of their relatively poor mathematical literacy. In many classrooms the use of pH has been reduced to the question of which button to press on a calculator [35]. Whether (1) or (3) is taught to students, this fundamental difficulty remains and the question of whether the argument of the logarithm should be \(c_{\text{H}^+}\) or \(a_{\text{H}^+}\) is actually secondary to the more important consideration of mathematical competence.

A simple, and well known, extension of this example relates to the dissociation of an acid (HA)

\[ \text{HA} \rightleftharpoons \text{H}^+ + \Lambda^-, \]

(4)

for which the mass action ratio \((Q)\) is usually written

\[ Q = c_{\text{H}^+}c_\Lambda^-/c_{\text{HA}}, \]

(5)

where \(c_\Lambda^-\) and \(c_{\text{HA}}\) are the concentrations of \(\Lambda^-\) and HA, respectively, and the equilibrium constant is \(K_a\). At equilibrium, \(Q = K_a\) and one often sees this written as the Henderson-Hasselbalch equation [11, 30]

\[ \text{pH} = pK_a + \log_{10}(c_\Lambda^-/c_{\text{HA}}), \]

(6)

where \(pK_a = -\log_{10}(K_a)\) and pH is given by (1). Since \([\text{pH}] = 1\) and because the argument of the logarithm is also dimensionless, \([pK_a] = 1\). Consistent with the argument leading to (3), \([K_a] = 1\) and so \([Q] = 1\) [10], but it appears from (5) that \([Q] = \text{mol L}^{-1}\), indeed \(K_a\) values are often reported as
having some specific dimension [9], despite recommendations to the contrary [10]. For an electrolyte that dissociates into $\nu_+ cations$ and $\nu_--anions$, the activity is defined as $a_{\text{solute}} = a^+\nu^- a^-\nu^+$, where $\nu = \nu_+ + \nu_-$. Since the activities of individual ions can not be measured, the (geometric) mean activity coefficient $\gamma = (\gamma_+\gamma_-)^{\nu}$ is used. For the acid dissociation reaction (4), $\nu_+ = \nu_- = 1$ and $\gamma = (\gamma_H\gamma_A)^{\nu}$, which varies with the concentration of the acid (Figure 1). Based on this, the Henderson-Hasselbalch equation (6) becomes

$$\text{pH} = pK_a + \log_{10}(c_A/c_{HA}) + \log_{10}(\gamma),$$

where the third term on the right can be thought of as representing deviation from the behaviour of an ideal solution.

---

**Figure 1.** Activity coefficients ($\gamma$) of some hydrogen halides (HCl, HI, HBr, HF) [19] as a function of molality ($m$) and the limiting (7) and extended (8) Debye-Hückel approximations (dashed and solid curves, respectively).

For molalities ($m$) of less than about 0.05 mol kg$^{-1}$ (Figure 1), $\gamma$ can be estimated using the Debye-Hückel limiting equation

$$\ln \gamma = \frac{-z_+z_-e\alpha}{\rho s}, \quad (7)$$

where $z_+$ and $z_-$ are the valences of the cation and anion, respectively, (so $z_+ = 1$ and $z_- = -1$ for the dissociation of an acid) and

$$\alpha = \sqrt{2\pi N_A\rho s^3},$$

where $N_A$ is Avogadro’s constant, $\rho$ is the density of the solvent and

$$s = \frac{e^2}{4\pi\varepsilon_0\varepsilon r_{\text{es}}kT}$$

is a distance characteristic of the solvent which depends on the dielectric constant ($\varepsilon_r = 78.54$ for water), the charge of an electron ($e = 1.6021 \times 10^{-19}$ C), Boltzmann’s constant ($k = 1.38 \times 10^{-23}$ J K$^{-1}$), the absolute temperature ($T = 298.15$ K) and the permittivity of free space ($\varepsilon_0 = 8.85 \times 10^{-12}$ C V$^{-1}$ m$^{-1}$). Using these values, $s = 0.7136$ nm and $\alpha = 1.171$ (kg mol$^{-1}$)$^{\frac{1}{2}}$. At higher molalities ($m < 0.1$ mol kg$^{-1}$), the extended Debye-Hückel equation [13]

$$\ln \gamma = -\frac{|z_+z_-|e\alpha\sqrt{I}}{1 + 2\alpha(a/s)\sqrt{I}}, \quad (8)$$
is more reliable (Figure 1). In (8), \(a\) is the average smallest distance between two ions \((a = 0.30-0.35\, \text{nm})\). Both (7) and (8) depend on the ionic strength \((I)\) of the solution

\[I = 0.5\sum_{j} m_j a_j^2\]

[23] where \(m_j\) is the molality of the ion \(j\) which has valence \(z_j\) and so \([I] = \text{mol}\,\text{kg}^{-1}\). It is clear from Figure 1 that neither (7) nor (8) is adequate at high concentrations, and also that they do not work for hydrofluoric acid (HF), which behaves quite differently from other hydrogen halides (HCl, HBr and HI) because of the solvation of \(\text{F}^-\) by HF at even quite low concentrations [26]. Indeed, whereas HCl, HBr and HI are strong acids \((pK_a = -7, -9\) and \(-11\), respectively), HF is a weak acid \((pK_a = 3.45)\) at low concentrations.

The usual means of deriving an expression for the activity coefficient is based on the observation that for a solute \(i\) the chemical potential \((\mu)\) can be viewed as the sum of the chemical potential of the ideal solute \((\mu_{i,\text{ideal}})\) and a term representing non-ideality \((\mu_{i,\text{non-ideal}})\) which is

\[\mu_{i,\text{non-ideal}} = R T \ln \gamma_i,\]

where \(R\) is the gas constant. If the non-ideality arises from electrostatic interactions, then

\[
\ln \gamma_i = -\frac{\beta}{8\pi e_0 e_i^{2} kT} \sum_{i} n_i z_i^2
\]

where \(\beta\) is the Debye length, which is a measure of the approximate extent of the electrostatic field of the ion,

\[\beta^2 = \frac{e^2}{e_0 e_i^{2} kT} \sum_{i} n_i z_i^2\]

and to convert \([n_i] = \text{m}^3\) to \([m_i] = \text{mol}\,\text{kg}^{-1}\) we use \(m_i = N_A \rho c_i / (10^{-3}\,\text{m}^3\,\text{L}^{-1})\) to obtain

\[\beta^2 = \frac{N_A \rho c_i^{2}}{10^{-3} e_0 e_i^{2} kT} \sum_{i} n_i z_i^2 = \frac{2 N_A \rho c_i^{2}}{10^{-3} e_0 e_i^{2} kT} I.\]

Substituting this into (9) yields

\[
\ln \gamma_i = -\frac{\beta}{(4\pi e_0 e_i^{2} kT)^{3/2}} \sqrt{\frac{2\pi N_A \rho}{10^{-3}}} \sqrt{I} = -z_i^2 \alpha \sqrt{I},
\]

but, for the acid dissociation reaction (4), \(\gamma_e = (\gamma_+ \gamma_-)^{1/2}\) or \(\ln \gamma_e = \frac{1}{2}(\ln \gamma_+ + \ln \gamma_-)\) and \(z_+ = 1\) and \(z_- = -1\), so the Debye-Hückel limiting equation (7) is

\[
\ln \gamma_e = -\alpha \sqrt{I}
\]

which, in water at 298 K, is \(\ln \gamma_e = 1.174 I^{1/6}\) or \(\log_{10} \gamma_e = 0.509 I^{1/6}\) (Figure 1).

Even this derivation is relatively demanding, but an approximation of the Debye-Hückel limiting equation can be obtained quite simply using DA, as will be demonstrated in the next section.

4. A more formal version of dimensional analysis

Buckingham [8] described a more formal approach to DA that can be stated in a few steps: (1) count the number of variables \((N_v)\) and the number of independent dimensions \((N_d)\); (2) form \(N_v - N_d\) dimensionless groups \((\theta_i, i = 1, 2, \ldots, N_v - N_d)\); and (3) the most general relationship between the variables is

\[\theta_i = F(\theta_i, i = 2, 3, \ldots, N_v - N_d),\]

where \(F\) is a dimensionless function. Each dimensionless group is formed by multiplying the variables involved (say \(x\) and \(y\)) raised to separate powers (say \(a\) and \(b\), as illustrated in section 2) to form \(\theta =\)

Volume 2, Number 4, 2009
\( x^\theta \) such that \( [\theta] = 1 \). There may be several possible ways of forming the \( \theta_i \), and one can choose from the options.

In the case of the Debye-Hückel limiting expression (7), if \( \ln \gamma \) depends on electrostatic interactions between ions, then it involves \( e, \varepsilon_0, \varepsilon_r, k, T, \rho, N_e, \) and \( I \), for which

\[
[\ln \gamma] = 1, \quad [e] = Q, \quad [\varepsilon_0 \varepsilon_r] = Q^2 (M L^3 T^{-2})^{-1}, \quad [k] = M L^2 T^{-2} \Theta^{-1},
\]

\[
[T] = \Theta, \quad [\rho] = M L^{-3}, \quad [N_e] = n^{-1}, \quad [I] = n M^{-1}
\]

and \( N_e = 8 \) and \( N_d = 6 \), so two dimensionless group can be formed. The first is just \( \theta_1 = \ln \gamma \) and the other is

\[
[\theta_2] = 1 = Q^a \times (Q^2 (M L^3 T^{-2})^{-1})^b \times (M L^2 T^{-2} \Theta^{-1})^c \times (M L^{-3} T^{-1})^d = n^{g} \times (n M^{-1})^h.
\]

Collecting terms yields

\[
Q^{a + 2b} \times M^{b + c + f - h} \times L^{-3b + 2c - 3f} \times T^{2b - 2c} \times \Theta^{-d} = 1,
\]

but since all of the exponents must equal zero, then

\[
\begin{pmatrix}
1 & 2 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & -1 & 1 & 0 & 1 & 0 & -1 & 0 \\
0 & -3 & 2 & 0 & -3 & 0 & 0 & 0 \\
0 & 2 & -2 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & -1 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & -1 & 1 & 0
\end{pmatrix}
\begin{pmatrix}
a \\
b \\
c \\
d \\
f \\
g \\
h
\end{pmatrix} = 0.
\]

This system of equations cannot be solved explicitly for all the variables because there are fewer expressions than variables, but it is clear that \( a = -2b, \; b = c = d = f = g = h \). Taking \( f = g = h = 1 \), then \( b = c = d = -3 \) and \( a = 6 \) from which

\[
\theta_2 = \frac{e^2}{\varepsilon_0 \varepsilon_r kT} N_A \rho I
\]

and the general form of the expression is

\[
\ln \gamma_\pm = F\left(\frac{e^2}{\varepsilon_0 \varepsilon_r kT} N_A \rho I\right),
\]

which is clearly similar to (10), but it required no complex mathematics. If it had been assumed that \( f = g = h = 0.5 \), then the Debye-Hückel limiting equation (7) would have been approximated even more closely.

As a final, slightly more complex illustration of the power of DA, consider Bjerrum’s [4] expression for \( K_a \). Bjerrum [4] combined the Debye-Hückel equation (7) with the notion of mass action (6) and obtained an expression for the equilibrium constant based on the electrostatic interaction between ions a distance \( r \) apart and a distance \( (R) \) representing association of ions. Assuming that the variables involved are

\[
[R] = L, \quad [r] = L, \quad [K_a] = 1, \quad [e] = Q, \quad [\varepsilon_0 \varepsilon_r] = Q^2 (M L^3 T^{-2})^{-1},
\]

\[
[k] = M L^2 T^{-2} \Theta^{-1}, \quad [T] = \Theta, \quad [\rho] = M L^{-3}, \quad [N_e] = n^{-1}, \quad [I] = n M^{-1}
\]

then \( N_e = 10 \) and \( N_d = 6 \), so four dimensionless groups can be formed

\[
\theta_1 = K_a, \quad \theta_2 = N_A \rho \Theta^3, \quad \theta_3 = e^2 / r \varepsilon_0 \varepsilon_r kT, \quad \theta_4 = r / R.
\]

Clearly \( \theta_2 \) and \( \theta_3 \) are linked, because (11) can be written \( \ln \gamma_\pm = F(\theta_2 \theta_3^3) \), so \( \theta_2 = \theta_3^{-3} F^{-1}(\ln \gamma_\pm) \) and

---

Acta Didactica Napocensia, ISSN 2065-1430
The units tell you what to do

\[ K_a = G(N_A \rho r^3, e^2/r_0 e_k T, r/R) = G(F^{-1}(\ln \gamma_\varepsilon) (r_0 \varepsilon_k T/\varepsilon)^3, e^2/r_0 e_k T, r/R). \]

Using these dimensionless groups, Bjerrum’s [4] expression can be written as

\[ K_a^{-1} = \frac{4\pi}{1000} \theta_2 \theta_3 \int_{\theta_2 \theta_4} \theta d y \exp(y) dy, \]

so even without solving the integral it is clear that \( K_a = G(\theta_2, \theta_3, \theta_4) \), as shown using DA.

These two examples illustrate the extraordinary power of DA to provide insight while using only simple algebra. Most significantly for a teacher, it facilitates the analysis of a problem from a scientific perspective without requiring the use of complex mathematics [5, 20], at least in the first instance. While DA will not always yield an answer, it can be relied on to generate insight into the problem. This is the basis of the assertion that the units tell you what to do.

5. The value of teaching dimensional analysis

Teaching even the simplest level of DA is powerful, because it develops (i) questioning; (ii) the habit of checking one’s own work, particularly in the sense of trying to assess whether an answer is reasonable; (iii) skills in the use of the SI prefixes and the link with scientific notation, neither of which is well understood by many students; (iv) fundamental algebraic skills; and (v) a powerful technique that can be applied in any discipline. A simple application of DA to familiar concepts, such as pH (section 3), elicits useful discussions and motivates students to think differently about what they think they know. Ultimately, the habits acquired by using DA contribute to the metacognitive development of students. Even if students do not appreciate DA initially, they are eventually persuaded of its value.

6. Conclusion

Dimensional analysis is a valuable tool that should be taught much more widely. Even used in a very simple way it is powerful and useful. Many examples of the use of DA can be found in the literature [3, 5, 6, 15, 20], but with a little practice it is not difficult to develop new examples. As with much in science, the ‘secret’ is allow yourself to think simply and let the units tell you what to do.

Literature


Volume 2, Number 4, 2009


[35] Watters, D. J. and J. J. Watters (2006), Student understanding of pH: "i don't know what the log actually is, i only know where the button is on my calculator", *Biochemistry and Molecular Biology Education*, 34, 4, 278-284.


**Author**

**Simon Brown**, School of Human Life Sciences, University of Tasmania, Tasmania, Australia, e-mail: Simon.Brown@utas.edu.au