This article was downloaded from https://rsc.li/3ClgcC2

Perspective

Malpractices in Chemical Calculations

Joe Lee

Department of Chemistry, UMIST, Manchester, M60 1QD, UK
e-mail: joe.lee@umist.ac.uk

Abstract

Chemistry undergraduates are frequently guilty of faulty or inefficient practices in performing physicochemical calculations, possibly leading to incorrect answers, both in the processing of laboratory data and in answers to tutorial/examination questions. The purpose of this article is to draw attention to some of the more common malpractices, but more importantly, to provide a framework for good practice in teaching students how to perform chemical calculations.

Introduction

The purpose of this article is to draw attention to improper or inefficient procedures often encountered in chemical calculations performed by higher education students, as perceived by a lecturer with over 40 years of physical chemistry teaching experience at university level. Sometimes these malpractices are encountered in student answers to isolated problems (perhaps in examinations or tutorials) but more commonly, they are encountered in the processing of laboratory data. The author wishes to identify a number of standard approaches that should help to reduce the confusion of students, and might also provide guidelines that could be followed by instructors and authors of textbooks.

Quantity Calculus vs. Measure Calculus

Before discussing some specific malpractices, it is appropriate to outline the general algebraic system that is at the heart of advanced modern calculation techniques in physical science and engineering. This is based upon the concept that natural relationships exist between *quantities* and not between *measures*, the latter being defined by the general relationship: quantity = measure × unit; (for the special case of dimensionless quantities, the unit is 1). For example, in the statement: mass = 2 kg (× is implicit), mass is a quantity, kg is a unit of that quantity and 2 is the corresponding measure; whereas this mass is invariant, the measure and the unit can be varied in tandem: $2\ \text{kg} = 2 \times 10^3\ \text{g} = 2 \times 10^6\ \text{mg}$ = etc. A simple invariant relationship in quantities is density = mass/volume, but there is no *universal* relationship between measures of these quantities; individual relationships will depend on the choice of units for the three quantities involved. It is this invariance that is responsible for the widespread adoption of the algebraic system known as *Quantity Calculus* (or *quantity algebra*),[1] a system pioneered by people such as E. A. Guggenheim,[2] G. N. Copley[3] and M. L. McGlashan.[4] In this system, symbols represent *quantities*; in the example just given, density, mass and volume might be symbolised ρ, m and V; m will not be (for example) the number of kg in the mass (equivalent to saying that mass = m kg).

The method of quantity calculus contrasts with that traditionally used in pure mathematics, which is normally concerned with pure numbers and their algebraic representatives. However, even when mathematicians deal with non-dimensionless quantities, it has been standard practice to let *measures*, rather than quantities, be represented by algebraic symbols. To illustrate, it was not uncommon to start a problem in the following sort of way: “Let the length of the rod be l cm …” or “Let Mr Smith be aged x years and Mr Jones be aged y years…”. In these cases, the symbols represent *measures*; the algebra resulting from such symbolism could legitimately be called ’measure calculus‘. It *could* be argued that a measure is a *dimensionless quantity* defined as the ratio of a general quantity and a ’standard‘ value of that quantity. For example, we might have a pressure ratio p/p_c or p/p^o (p = pressure, p_c = critical pressure for the specific substance, p^o is a general standard pressure such as 100 kPa or 1 atm = 101.325 kPa). The first ratio, called reduced pressure, is commonly symbolised p_r. One *could* introduce a single symbol for the second, P perhaps, and if p^o is say 1 atm, then $p = P$ atm and P, formally a measure, has become a *special* quantity symbol. If this argument is valid, we are effectively stating that quantity calculus includes measure calculus as a special case, but the reverse is certainly not true, *i.e.*, measure calculus is standalone. There is one important situation where relationships are ultimately required in terms of measures (whether

This journal is © The Royal Society of Chemistry

written as a quantity/unit ratio or as a single symbol) and this is when a computer (capable of processing only numbers) is to be used. The conversion of a quantity formula to the requisite measure formula can cause problems for students and will be considered towards the end of this article. The majority of it will, however, be concerned with faulty or illogical practices in using algebraic quantity formulae.

Examples of Malpractice

Let us now consider these practices. Here, the emphasis will often be made on the correct practice with the implication that such practice is commonly not followed.

Avoidance of unnecessary repetition

Firstly, one ought to say that in simple calculations, such as those encountered in secondary education, one can completely avoid algebra. To take a simple case of a volumetric analysis calculation, one could proceed as follows (or something like this), using a specific example.[i]

$NaOH + HCl \rightarrow NaCl + H_2O$; known [HCl] = 0.1234 mol dm^{-3}; HCl in burette

Pipetted volume of NaOH = 25.00 cm^3; experimental (mean) titre = 23.25 cm^3

<u>Calculation</u>: Amount of substance of HCl = 0.1234 mol dm^{-3} × 23.25 cm^3 = 0.1234 mol dm^{-3} × (23.25/1000) dm^3 = 0.002869 mol

So amount of substance of NaOH = 0.002869 mol

[NaOH] = 0.002869 mol/25.00 cm^3 = 0.002869 × 1000 mol /25.00 dm^3 = 0.1148 mol dm^{-3}

In pre-HE education, there is admittedly substantial advantage in this kind of procedure. Students often say something like "*We are encouraged not to remember formulas that we do not understand, but to proceed stepwise through the calculation without introducing symbols*". There are common student faults with this kind of procedure, in connection with (i) incorrect names for quantities, (ii) non-inclusion of units, and (iii) insignificant figures; these points will be covered later.

Objections arise to this procedure when an identical analysis has to be applied to several NaOH solutions, using the same pipette and the same HCl solution—the extreme situation would arise if a large number of titrations were required over a time range when NaOH is consumed in a kinetic study. Sensibly, in cases like this, one *must* exploit the constancy of both the pipette volume and the HCl concentration; as a consequence, one can deduce the generic relationship: [NaOH] = (0.1234 mol dm^{-3}/5.00 cm^3) × titre, *i.e.*, [NaOH]/mol dm^{-3} = 0.00494 × titre/cm^3. This formula has but one independent variable, the titre. One would then tabulate titre/cm^3 and [NaOH]/mol dm^{-3} alongside the NaOH sample number/label or time/min, as appropriate. In the case of following the titre (or indeed any other quantity proportional or linearly related to concentration) as a function of reaction time, it may be wasteful of effort to convert each individual titre to a concentration when graphical processing of data is to be used. This point will be discussed again in the section concerned with plotting graphs, below.

Definition of quantity symbols

A statement of a quantity formula should always be accompanied by the definition of its symbols, even if these have a traditional meaning, like V commonly signifying volume; quantity symbols are *recommended*, but not fixed, by bodies such as the International Organization for Standardization (ISO),[5] the International Union of Pure and Applied Chemistry (IUPAC)[6] and the International Union of Pure and Applied Physics (IUPAP).[7] An important point is that one should not include units in such definitions, *e.g.*, 'V = volume', NOT 'V = volume (m^3)'. The choice of unit for the quantity is completely irrelevant—see below. Needless to say, quantity names should be modern ones, *e.g.*, 'amount of substance' or 'chemical amount', NOT 'number of moles' or just 'moles' (equivalent to miscalling 'mass' 'number of grams' or 'grams'), Avogadro constant (for 6.022×10^{23} mol^{-1}) and NOT Avogadro number—incidentally, the addition of the possessive: 's' after a person's name is not current scientific practice.

The need to substitute quantity symbols with both measures and units

Substitution of values into a quantity formula requires that both the measure *and* the corresponding unit be inserted. To take a simple physical chemistry example, consider the perfect (or ideal) gas relationship $V = nRT/p$ to be used to calculate gaseous volume V from pressure p, amount of substance n and thermodynamic (absolute) temperature T; R is the gas constant. It is essential to replace each of the right side symbols by both the measure AND corresponding unit; furthermore, the units should be there at *all* stages of the calculation—it is an option whether to keep each unit alongside its measure or to isolate the units in a composite unit term. Following this, one should handle the unit manipulation with the same care as the measure simplification; (if the relationship is wrong, unit manipulation will often reveal this). If the units are so-called *coherent*—SI units for example—they will 'cancel' nicely and, using this same example, give a simple, recognisable unit of volume. But, even if they are

---

[i] In what follows, there is no specific objection to using M as a synonym of mol dm^{-3}, or litre—safest symbol L — and mL instead of dm^3 and cm^3.

This journal is © The Royal Society of Chemistry

not coherent, one will still get the correct answer (because the rules of substitution have been followed), *e.g.*, one might obtain an answer 1234 J atm^{-1} for the volume and if (as likely) that unit is not liked, it can be converted into m^3 by use of atm = 101 325 Pa and J $\text{Pa}^{-1} = \text{m}^3$. Unit prefixes can be manipulated in the same sort of way, *e.g.*, the composite unit $\text{m}^3\ \text{cm}^{-3}$ (produced as such perhaps in some calculation) can be reduced to a number as follows: $\text{m}^3\ \text{cm}^{-3} = \text{m}^3\ (10^{-2}\ \text{m})^{-3} = \text{m}^3 \times 10^{+6}\ \text{m}^{-3} = 10^6$.

Use of proper functional arguments

If a function (f), by its definition, has a non-dimensionless argument [x in $f(x)$], then this requirement *must* be satisfied. Perhaps the commonest example of such a function in chemistry is the logarithm; one can take logarithms only of *positive numbers*. Physical scientists (textbook writers and, sadly, mathematicians who teach physical science students) should not pretend that terms, such as pH = $-\log_{10}([\text{H}^+])$, $\ln(p)$ or $\ln(k)$ (where k is a rate constant) are meaningful. The following are all right: pH = $-\log_{10}([\text{H}^+]/\text{mol dm}^{-3})$ (/ means ’divided by‘), $\ln(p/\text{atm})$, $\ln(p/p^{\text{o}})$ (p^{o} = a chosen standard pressure), $\ln(p_1/p_2)$, $\ln(p_{\text{r}})$ (p_{r} = a relative pressure, which could be p/p^{o} or even p/atm). The indefinite integral of $1/p$ is $\ln(kp)$ with k a positive integration constant with dimension of reciprocal pressure. The (strictly meaningless) differential coefficient $\text{d}\ln(p)/\text{d}T$ is correctly (but cumbersomely) written $\text{d}\ln(kp)/\text{d}T$; however, it is also correctly (*and* simply) written $(1/p)\ \text{d}p/\text{d}T$.

Dimensional homogeneity in expressions and equations

Additive/subtractive expressions, and equations generally, should be ’dimensionally homogeneous‘. For example, in connection with the thermal expansion of a solid, the following are all examples of correct forms:

$V/\text{cm}^3 = 1.234 + 2.345 \times 10^{-4}\ t/^{\text{o}}\text{C}$ (t = Celsius temperature)

$V = 1.234\ \text{cm}^3 + 2.345 \times 10^{-4}\ \text{cm}^3\ {}^{\text{o}}\text{C}^{-1}\ t$

$V = (1.234 + 2.345 \times 10^{-4}\ t/^{\text{o}}\text{C})\ \text{cm}^3$

but the following are improper:

$V = 1.234 + 2.345 \times 10^{-4}\ t$

$V/\text{cm}^3 = 1.234 + 2.345 \times 10^{-4}\ t$

Modern tabulation

Tabulation, for the sake of conciseness, is of numbers and the names of columns and rows should reflect this. For example, we might give a set of volumes as

V/cm^3 1.234 2.456 3.789…

not $V(\text{in cm}^3)$ … or $V(\text{cm}^3)$ …

Formally, V 1.234 cm^3 2.456 cm^3 3.789 cm^3…

is correct but clearly cumbersome; transfer of the common unit once and for all to the row designator makes obvious sense. The same point applies to a repetitive power-of-ten multiplier;

V/m^3 1.234×10^{-3} 2.456×10^{-3} 3.789×10^{-3}…

is (deliberately avoiding the prefixes milli and centi to make the point) more concisely written as

$V/10^{-3}\ \text{m}^3$ 1.234 2.456 3.789…

or as $10^3\ V/\text{m}^3$ 1.234 2.456 3.789…

In this latter case, the first table entry 1.234 is exactly what the row name says it is, *viz.*, $10^{+3}\ V/\text{m}^3$ (author’s + sign emphasis) so that $V = 1.234 \times 10^{-3}\ \text{m}^3$. Incidentally, V/cm^3 can alternatively be written $V \times \text{cm}^{-3}$ or just $V\ \text{cm}^{-3}$, but the quotient format is usually regarded as clearer.

Plotting graphs

Just as one tabulates numbers, so one also plots *numbers*[8] (just as mathematicians do when symbols such as x and y represent numbers). However, this is only when we have discrete plot points, corresponding to measured data, say, a set of temperatures and corresponding volumes where we might plot V/cm^3 against T/K (or perhaps against $T/\text{kK} \equiv T/1000\ \text{K}$) with axis tick marks properly labelled as numbers.[ii] It must necessarily follow that graphical intercepts, slopes and areas must also be *numbers* and, furthermore, the algebraic equivalents of these parameters must reflect this. To exemplify (also bringing in the point about proper logarithmic arguments): in the field of reaction kinetics, one encounters the Arrhenius (quantity) relationship $k = A \exp(-E/RT)$ [k = a rate constant (here exemplified by a first order one with dimension [TIME^{-1}]), A = pre-exponential factor, E = activation energy, T = thermodynamic (or absolute) temperature]. Suppose, one has a set of experimental values for k and corresponding T and one wishes to determine A and E by graphical means. The proper logical procedure, starting with division by a sensible unit of k, is as follows:

$k/\text{min}^{-1} = (A/\text{min}^{-1}) \exp(-E/RT)$

$\ln(k/\text{min}^{-1}) = \ln(A/\text{min}^{-1}) - E/RT = \ln(A/\text{min}^{-1}) - E/(R \times 1000\ \text{K}) \times 1000\ \text{K}/T$

Then plot the *number* $\ln(k/\text{min}^{-1})$ against the (conveniently sized) *number* $1000\ \text{K}/T$ to give a straight line of slope $-E/(R \times 1000\ \text{K})$ and intercept $\ln(A/\text{min}^{-1})$ from which E can be obtained by multiplying the slope by $-R \times 1000$ K and from which A can be obtained by taking the natural antilogarithm of the intercept and then multiplying this by min^{-1}. If axes are labelled as numbers, it would be wrong to say that $\ln(k)$— meaningless anyway—is plotted against $1/T$ and the slope is $-E/R$.

[ii] Formally, there is nothing improper in plotting in a quantity space, V against T with tick marks tediously labelled as *quantities* (1 cm^3, 2 cm^3, 3 cm^3, …, 290 K, 300 K, 310 K, …).

This journal is © The Royal Society of Chemistry

However, one must distinguish here such number plots from so-called ’sketch graphs‘ where one merely wants to show *with no experimental points*, for example, the general trend of *V* (for a solid, say) against *T*. Here one plots quantities *V* and *T*, with common implicit assumptions that the axes are linear and cross at the origin.

It is now appropriate to return to the point deferred earlier, concerning the optimum processing of titre vs time data in a kinetics experiment, avoiding the conversion of each individual HCl titre to a NaOH concentration. Suppose that one wishes to test whether the process is first order in NaOH (other reactants being in excess and essentially remaining at constant concentration). The equation to be tested is then $\ln([\text{NaOH}]/[\text{NaOH}]_o) = -kt$ {subscript o indicates a value at zero time, k = first order rate constant, t = time}. The first point to note is that $[\text{NaOH}]_o$ is a constant and division of each [NaOH] by that constant can easily be avoided, but one can go further. Since $[\text{NaOH}] \propto$ titre, *T*, the equation to be tested can be transformed as follows: $\ln(T/T_o) = -kt$; $\ln(T/\text{cm}^3) = \ln(T_o/\text{cm}^3) - k \times \text{min} \times t/\text{min}$

Plot $\ln(T/\text{cm}^3)$ against t/min and *if* the order is 1 (and the reaction has been followed to high degree of completion), the graph will be a straight line of slope $-k \times$ min from which $k = -\text{slope} \times \text{min}^{-1}$ can be obtained. Similar arguments will apply to the testing of other kinds of rate equation, *i.e.*, not first order, except that there will then be need for a final *single* use of the proportionality constant in $[\text{NaOH}] \propto$ titre when converting a graphical slope into a rate constant. If a *product* rather than a reactant is being monitored, again similar arguments may be used, but there will be a further need to obtain and involve a titre at ’infinite‘ time.

Unit conversion

Students often find unit conversion difficult. To take a simple example, suppose one had a concentration of 1.234 mol dm^{-3} and wanted to convert this into so many mol m^{-3}. Students would often argue (correctly) as follows: 1 m^3 is 1000 times as big as 1 dm^3, so this would require a 1000 times as much amount of substance, *i.e.*, 1234 mol, so 1.234 mol dm^{-3} = 1234 mol m^{-3}. But the following (taken in more steps than is conventional for emphasis) is simpler and more generally applicable: $1.234 \text{ mol dm}^{-3} = 1.234 \text{ mol } (10^{-1}\text{ m})^{-3} = 1.234 \times (10^{-1})^{-3} \text{ mol m}^{-3} = 1.234 \times 10^{+3} \text{ mol m}^{-3} = 1234 \text{ mol m}^{-3}$.

Significant figures

It is very common for students to create far too many figures in a calculation. A simple rule in a multiplicative/divisive calculation is to give the answer to the same *relative* error (roughly the same number of *significant figures*, but this can be wrong by ±1) as the least accurate datum in the calculation. Additive/subtractive calculations should give an answer to the same number of *decimal places* as the least precise datum, so that (1.15 – 0.1234) cm^3 = 1.03 cm^3, but 1(exactly) – 0.1234 = 0.8766. As a rough guide, taking a logarithm of a number with *n* significant figures would produce *n* decimal places. One can do better than this by applying statistical considerations, of course, (‘cumulation of errors’), and including ‘± *e*’ after the number (preferably, stating whether *e* is a standard error of the mean or a 95% confidence deviation or whatever). In this connection, a simple rule of thumb is to give *e* to one figure (or possibly two if the first figure is a 1) and then round the mean to the same number of *decimal places*. If the last decimal place is a zero, it should be present; the same point applies to the primary data, *i.e.*, significant terminal zeros should be included. A related matter arises when the number of significant figures is less than the number of figures preceding the decimal point (or decimal symbol); 1.234×10^5 is clear in significant figures but 12340 is ambiguous. On a minor issue, it is the author’s opinion, that for measures below 1, an initial zero before the decimal point is generally clearer than when that zero is omitted. Two other small points agreed internationally: (1) a comma should not be used to space sets of three digits (either side of a decimal symbol)—when there are more than four digits, one can use a space instead, *e.g.*, 12 345.123 45; (2) the comma can be used instead of a stop for a decimal symbol—this is standard European (but not British) practice

Use of computers

Although the previous emphasis has been on the use of algebraic symbols as quantity representatives, one may have to consider an effective deviation from this if data are to be processed by computers, which cannot (generally) handle units. In this context, it may be that, for example, instead of *c* standing for concentration, it may have to stand for concentration/mol dm^{-3}; strictly speaking *c* would be the name of a register containing the number concentration/mol dm^{-3}, a point more appreciated when one had to write a computer programme to perform the task and less appreciated in the present age of general computer packages. As a precursor to computer processing, some manipulation of the quantity formula to involve units is necessary.

This journal is © The Royal Society of Chemistry

This will be illustrated by the calculation of a gas volume from the perfect gas equation. Suppose that one wanted to (i) calculate the number of cm^3 in the volume, (ii) input the constant R as the number of $\mathrm{J\ K^{-1}\ mol^{-1}}$ in this quantity, and (ii) input information on quantities n, T and p as the number of moles (symbol mol), the number of K and the number of atm, respectively. Then, we would have to manipulate the original equation $pV = nRT$ as follows:

$$V = \frac{nRT}{p}$$

$$\frac{V}{\mathrm{cm}^3} = \frac{(n/\mathrm{mol})(R/\mathrm{J\,K^{-1}\,mol^{-1}})(T/\mathrm{K})}{(p/\mathrm{atm})} \times \frac{\mathrm{mol} \times \mathrm{J\,K^{-1}\,mol^{-1}} \times \mathrm{K}}{\mathrm{atm} \times \mathrm{cm}^3}$$

$$= \frac{(n/\mathrm{mol})(R/\mathrm{J\,K^{-1}\,mol^{-1}})(T/\mathrm{K})}{(p/\mathrm{atm})} \times \frac{\mathrm{mol} \times \mathrm{J\,K^{-1}\,mol^{-1}} \times \mathrm{K}}{101\,325\ \mathrm{Pa} \times 10^{-6}\mathrm{m}^3}$$

$$= \frac{(n/\mathrm{mol})(R/\mathrm{J\,K^{-1}\,mol^{-1}})(T/\mathrm{K})}{(p/\mathrm{atm})} \times \frac{\mathrm{J}}{101\,325\ \mathrm{Pa} \times 10^{-6}\mathrm{m}^3}$$

$$= \frac{(n/\mathrm{mol})(R/\mathrm{J\,K^{-1}\,mol^{-1}})(T/\mathrm{K})}{(p/\mathrm{atm})} \times \frac{1}{0.101\,325}$$

$$= 9.869\ \frac{(n/\mathrm{mol})(R/\mathrm{J\,K^{-1}\,mol^{-1}})(T/\mathrm{K})}{(p/\mathrm{atm})}$$

and then assign each measure (= quantity/unit) to a particular computer register (sensibly, given the same name as the quantity, of course). Naturally, if the units involved are coherent, *e.g.*, (unprefixed) SI units, there is no multiplying factor like the 9.869 above.

Unit names and symbols

One should use the correct (mandatory) symbols for units, *e.g.*, s not sec for seconds, h not hr for hour, l or L not lit for litres (L better because of possible confusion of l with the digit 1), atm not atmos for atmospheres. Generally, unit *names* based upon scientists' names have an initial lower case letter, *e.g.* joule, but the corresponding symbols have a capital letter, *e.g.* J; otherwise the symbol is all lower case—litre is the exception to the rule. Unit symbols (which in print are roman, unlike italic symbols for quantities in print) should not be pluralised or followed by an abbreviation point. Of course, there is a severe font problem when writing on paper or a blackboard; italicisation can be simulated by underlining but this is tedious—students could, for example, confuse K (for kelvin) with K (for equilibrium constant). There should be a space between a measure and a unit and between components of a composite unit; compare ms^{-1} = 1/millisecond and $\mathrm{m\ s^{-1}}$ = metre per second. Although a strong supporter of SI units, the author does not believe that students should not use other units.

Use of the solidus

A small point arises in connection with the use of the solidus (/) for division in quantity or unit expressions. The modern convention is that everything that follows the solidus (unless bracketing indicates otherwise) is part of the divisor. Thus a/bc means a divided by the product of b and c; in other words multiplication has priority over division; J/K mol is the same as $\mathrm{J\ K^{-1}\ mol^{-1}}$—a second solidus after K would be redundant. This contrasts with the usage in computer programming coding, which is a different matter. However, it is better to add extra bracketing if there is any danger of confusion.

Conclusion

It is my view that students (like us!) must use completely rigorous and consistent methods for performing chemical calculations. Whilst this might seem to necessitate extensive defining of rules and procedures, it is an essential part of gaining an education in chemistry if our students are to understand their subject properly and are to use their knowledge as professional chemists. Although many of the principles underpinning this article have been expounded by others, my experience, and that of virtually all my colleagues, is that students receive conflicting advice from lecturers, tutors, demonstrators and textbooks. Also, it is widely felt that students come to university with a weaker grasp of physics and mathematics than (say) 20 years ago. University chemical educators must be absolutely consistent in the way they expect students to manipulate data if they are to grasp the essential principles, and I hope that the guidelines provided in this article will help both colleagues and students.

References

1. I. Mills, T. Cvitaš, K. Homann, N. Kallay and K. Kuchitsu, *Quantities, Units and Symbols in Physical Chemistry*, 2nd ed., Blackwells, Oxford, 1988, pp 107-9.
2. E.A. Guggenheim, *J. Chem. Ed.*, 1958, **35**, 606.
3. G.N. Copley; *J. Roy. Inst. Chem.*, 1959, **83**, 477.
4. M.L. McGlashan, *Physicochemical Quantities and Units*, Royal Institute of Chemistry Monograph for Teachers No. 15, 1971.
5. *ISO 31-0* to *ISO 31-13*, International Organization for Standardization, Geneva, 1992.

6. Ref. 1, p 9.
7. E.R. Cohen and P. Giacomo, *Symbols, Units, Nomenclature and Fundamental Constants in Physics* (Document IUPAP-25), International Union of Pure and Applied Physics, 1987.
8. S. Le Vent; *Educ. Chem.*, 1987, **24**, 14.

This journal is © The Royal Society of Chemistry