On the Standard Rounding Rule for Addition and Subtraction

Wei Lee[1], Christopher L. Mulliss[2], and Hung-Chih Chiu[1]

[1]*Department of Physics, Chung Yuan Christian University, Chung-Li, Taiwan 320, R.O.C.*

[2]*Department of Physics and Astronomy, University of Toledo, Toledo, Ohio 43606, U.S.A.*

(Received June 27, 1999)

An investigation of the commonly suggested rounding rule for addition and subtraction is presented including its derivation from a basic assumption. Through theoretical study and Monte-Carlo simulations, it is shown that this rule predicts the minimum number of significant digits needed to preserve precision 100% of the time. Because the standard rounding rule for these two fundamental operations is accurate and completely safe for data, there is no need to extend this rule to keeping an additional significant figure as has suggested for multiplication and division.

PACS. 01.30.Pp – Textbooks for undergraduates.
PACS. 01.55.+b – Genreal physics.

I. Introduction

Rules for propagation of significant figures are described in virtually every introductory physics textbook. Over the years, college students of physics have come to rely on these "standard" rounding rules for the fundamental mathematical operations; i.e., addition, subtraction, multiplication, and division. It is obvious that rounding, as part of the analysis, must never jeopardize the real data whose experimental error must begin and end with the measuring apparatus. In a recent note by Good [1], a simple division problem is discussed as an example to show that the application of the standard rounding rule leads to a loss of precision in the result. It points out, in a straightforward and stark manner, why the commonly suggested rounding rule is wrong in itself and how its application can be potentially dangerous to data. The fact that rounding rules can lead to such a loss of valuable information is due to their approximate nature and has been well documented by Schwartz [2]. Due to the importance of rounding rules as common and convenient (even if approximate) tools in data manipulation and error analysis [1,3], Mulliss and Lee investigated the rounding rules for multiplication and division in great detail [4]. Through Monte-Carlo simulations, it was shown for the first time that the standard rounding rule for multiplication and division fails to predict the minimum number of significant figures needed to preserve precision approximately 50% of the time. Mulliss and Lee also explain, using their formalism, why the standard rule fails and how it can fail. It was proven that there is no *a priori* rule that always works for multiplication and division. An alternative rule suggested to be adopted as the new standard is a rule which advocates the use of an additional figure over that required by the standard rule [4].

Since the publication of the above-mentioned paper in this *Journal*, the authors have received many positive comments from the concerned educators and textbook authors in the physics community. Indeed, it is because its great impact on science education in general that the American Association of Physics Teachers lists under the on-line PSRC Resource Center: General Physics: Significant Figures [5] the web site maintained by Mulliss [6].

In this paper, the theoretical basis for the standard rounding rule for the even more fundamental operations -addition and subtraction- is presented. The standard rounding rule, as applied to the simple addition and subtraction problems $x = y + z$ and $x = y - z$, is considered. Following Mulliss and Lee [4], a statistical test is applied to quantify the accuracy of the standard rule. Three categories are considered: those where the "true" uncertainty is as large or larger than that predicted by the standard rule but of the same order of magnitude, those where the true uncertainty is less than predicted, and those where the true uncertainty is an order of magnitude larger than predicted. In the first case, the standard rule is said to "work" because it predicts the minimum number of significant digits that can be written down without losing precision, and therefore valuable information, in the result. In the second and third cases the standard rule clearly "fails", predicting fewer or more significant figures than are needed and, therefore, losing or overstating precision.

II. Simple derivation of the standard rounding rule

A commonly suggested rounding rule for addition and subtraction [7] states: "*When numbers are added or subtracted, the number of decimal places in the result should equal the smallest number of decimal places of any term in the sum.*" The standard rounding rule for the addition and subtraction of two numbers can be inferred from one very simple assumption. According to the concept of uncertainty, the absolute error in a number, written without any indication of its real uncertainty, is taken to be $\pm 1/2$ in the least significant decimal place [8]. As an illustration of this relationship, Table I presents the number 815.24 written to 1, 2, 3, 4, and 5 significant figures; also included are the corresponding uncertainty and "place" of the least significant digit. Let the integer Px denote the place of the right most significant digit in a number x, such as $P(3.01) = -2$, $P(48) = 0$, or $P(620) = 1$. One obtains the uncertainty in a number:

$$\text{Uncertainty}\,(x) = 0.5 \times 10^{P_x}. \tag{1}$$

For the derivation of the standard rounding rule for addition and subtraction, consider the fundamental mathematical operations of two numbers, $y \pm \Delta y$ and $z \pm \Delta z$. In the simplest approximation for $x = y + z$ or $x = y - z$, it is customary to take the maximum uncertainty to be the one quoted, leading to

$$\text{Uncertainty}\,(x) = \text{Uncertainty}\,(y) + \text{Uncertainty}\,(z). \tag{2}$$

Substituting the relationship Eq. (1) into Eq. (2) gives

$$10^{Px} = 10^{Py} + 10^{Pz}, \tag{3}$$

or

$$Px = P + \log(1 + 10^{P\theta - P}), \tag{4}$$

TABLE I. An illustration of the relationship between the place of the right most significant digit and the uncertainty of a quantity.

Number	Place of the least significant digit	Uncertainty
800	2	$\pm$ 50
820	1	$\pm$ 5
815	0	$\pm$ 0.5
815.2	-1	$\pm$ 0.05
815.24	-2	$\pm$ 0.005

where $P = \max(Py, Pz)$ and $P' = \min(Py, Pz)$.

At this point in the derivation, two separate cases must be considered: the case where y and z do not have the same place of the least significant digit and the case where they do.

Case 1 $(Py \neq Pz)$ When Py does not equal Pz, they must be different by at least 1. Even when they differ by this minimum amount, the term in the logarithm in Eq. (4) involving both P' and P is 1/10 of unity. Hence one is justified in replacing Eq. (4) by

$$Px = P + \log(1) = P, \tag{5}$$

implying that $Px = \max(Py, Pz)$. This is, clearly, a statement of the standard rounding rule.

Case 2 $(Py = Pz)$ When Py equals Pz, both terms in the logarithm in Eq. (4) become equally dominant. Under this condition, Eq. (4) reduces to

$$Px = P + \log(2), \tag{6}$$

where $P = Py = Pz$. One can see that the integer $Px = \text{Nint}\,[P + \log(2)]$, where Nint (w) is defined to be the closest integer to the number w. Because $\log(2)$ is smaller than 0.5, it can never cause Px to be rounded up to $P + 1$. Thus $Px = P = \max(Py, Pz)$, which is, again, consistent with the standard rounding rule.

When measured quantities are added or subtracted, the uncertainties add. It is worth mentioning that if the original uncertainties are *independent* and *random*, a more realistic estimate of the resulting uncertainty is given by the quadratic sum, which is never larger than their ordinary sum [9]. In other words, Eq. (2) is now rewritten as

$$(\text{Uncertainty}\,(x))^2 = (\text{Uncertainty}\,(y))^2 + (\text{Uncertainty}\,(z))^2. \tag{7}$$

Under these conditions, Eq. (4) becomes

$$Px = P + (1/2)\log[1 + 10^{2(P^{0} - P)}], \tag{8}$$

and it can easily be proven that the standard rounding rule for the simple addition or subtraction problem of two numbers is still well justified.

III. A statistical study of the standard rule

To investigate the statistical properties of the standard rounding rule, a Monte-Carlo procedure was used. A computer code was written in FORTRAN77 and run on a PentiumII350 PC equipped with a Microsoft Fortran PowerStation compiler (Version 4.0). The code uses a random number generator based upon Ran2 [10] to create two numbers. Each of these numbers has a randomly determined number of significant figures ranging from 1 to 5 and a randomly determined number of places to the left of the decimal point ranging from 0 to 5. Each digit in these numbers is randomly assigned a value from 0 to 9, except for the leading digit that is randomly assigned a value from 1 to 9. Each resulting number can range from the smallest and least precise value of 0.1 to the largest and most precise value 99999. The program calculates the sum or difference of the two generated numbers and determines the number of significant figures, which should be kept according to the standard rule. It then uses Eqs. (1) and (2) (or Eqs. (1) and (7)) to compute the "true" uncertainty from the two original uncertainties. The uncertainty in the sum or difference, as predicted by the standard rule, is taken to be $\pm 1/2$ in the least significant decimal place [8]. The true uncertainty and the value predicted by the standard rule are compared and the addition or subtraction problem is counted as one of the three cases described previously. The program repeats the calculation for one million additions or subtractions and computes statistics.

IV. Results and discussion

Unlike for the case of multiplication and division, where the application of the standard rule works only 46.4% of the time [4], the statistics in this study shows that the standard rounding rule for addition and subtraction will neither overstate precision nor cause a loss of valuable information. The rule simply works perfectly. A careful check of the formalism given by Eq. (4) or (8) conveys that these results are not unexpected at all. Notice that the integer $(P' - P)$ is either zero or negative, indicating that the sum at the right hand side of the equation is greater than P but not beyond $P + \log\,(2)$. The fact that Px is "substantially" greater than P (in a minute amount) suggests that the standard rule preserves precision well. Likewise, the fact that it can never cause Px to be rounded up to $P + 1$ implies that the standard rule never predicts more significant figures than are needed, indicating that the standard rule never overstates precision. Consequently, the standard rule for addition and subtraction of two numbers is never "wrong".

Now consider a series of additions and subtractions. One can show mathematically that the standard rule always works as long as there are nine numbers or less in the series and it may not fail for a series consisting more than ten numbers. For ten numbers in a series (or 100 numbers in a series if all of the associated uncertainties are independent and random), the extreme case takes place as they all have the same place of the least significant digit, causing the standard rule to overstate precision. Of course, in a series of additions and ubtractions, the longer the series, the more likely one is to need less significant figures to adequately represent the results because the errors of $\pm 1/2$ in the last decimal place add up. It is true that the standard rule may fail for a long series of additions and subtractions, however, it is a large enough number to ensure the validity of the rule in most practical situations. Because the standard rounding rule for addition and subtraction never leads to a loss of precision no matter how long a series is, it is always safe for data.

The fact that the standard rounding rule for addition must preserve precision should be able to imply the major finding from the Mulliss and Lee paper, which reveals the fact that the alternate rounding rule for multiplication must preserve precision [4]. To relate an addition problem to a multiplication problem, consider an addition problem involving the sum of n identical numbers y, where n is a precise positive integer. Assume that y and thus the sum x are positive. One can see that $Px = \text{Rint}(\log x) - Nx$, where Rint (w) denotes the smallest integer that is greater than the number w and the notation Nx stands for the number of significant figures in x. With this key relationship, some algebraic procedures similar to those appearing in Section II lead to the result $Nx = [\text{Rint}\,(\log x) - \text{Rint}\,(\log y)] + Ny - \log\, n$. Consider the special case where $n = 2$; i.e., $x = y + y$. The result becomes $Ny - \log\,(2) \leq Nx \leq Ny + 1 - \log\,(2)$, implying that the positive integer $Nx = Ny + 1$ must preserve precision for multiplication, which is the major point of the alternate rounding rule for multiplication and division. For the case where 10 identical terms are summed up, one can obtain $Nx = Ny$. Recall that it is a case where the standard rule for addition predicts an extra significant digit in the final result. The equality represents that the application to this case of the alternate rule for multiplication falls into the category of "1 digit too many". Therefore, the standard rounding rule for addition and the alternate rounding rule for multiplication form a self-consistent set of rounding rules that serve to preserve precision.

IV. Conclusions

Many problems encountered by physics students in daily life, including those in textbooks, do not deal with quantities where the uncertainties are explicitly stated. In these cases, the number of significant figures is the only available information upon which to base an error estimate. One could of course calculate percentage error at every step. The reason one cannot just abandon all rounding procedures is to avoid grossly overstating the precision every time an elementary calculation is painstakingly done. A satisfactory rounding rule provides the easiest way to protect against the loss of valuable information, pending the more careful determination of experimental error that should accompany the final result.

The previous work by Mulliss and Lee [4] proves that there is no *a priori* rounding rule for multiplication and division that can accurately predict the number of significant digits in all cases. It is also shown that the alternate rule, advocating the use of an extra significant figure, is far superior to the standard rounding rule. While the standard rounding rule for multiplication and division leads to a loss of precision over 50% of the time, the current work does indicate that the standard rounding rule for addition and subtraction is completely satisfactory in this regard.

Preserving precision under addition must preserve precision under multiplication because multiplication is directly related to addition. In the case where the tightest constraints on Nx exist, the standard rule for addition implies that, for multiplication, $Nx = Ny + 1$ guaranties that precision. Thus the use of the standard rule for addition points towards the alternate rule for multiplication as the correct rule that preserves precision.

This paper and the previous one together provide important new insights into the standard rounding rules for the fundamental, mathematical operations and, thus, have implications for the way that the rules should be taught to science students. The following steps are suggested:

Step 1) Correct statements that are factually incorrect. Avoid any misleading statements such as: "In multiplication and division, the product or quotient can *not* have more significant digits than the minimum number of significant digits used in the calculation."

Step 2) Emphasize that rounding rules are simply *conventions*. Rounding rules provide a convenient way to handle the propagation of errors when dealing with numbers that do not have explicitly stated uncertainties. It may be obvious to teachers that these rules are only approximate, but many students interpret a "rule" to mean "something that always works". These rounding conventions are called "rules" only because they provide a set of instructions.

Step 3) Decide which rounding rule to teach; the standard or the alternate rule. For multiplication and division, the standard rule is less accurate and allows valuable information to be discarded over half of the time. The alternate rule is more accurate and perfectly safe, but can overstate precision by having one or, on very rare occasions, two unneeded digits. For addition and subtraction, the standard rounding rule works well. The chance is slim for the rule to overstate precision; the best part is that it is always safe for data.

Step 4) Tell students what can happen and what to expect when they use the rounding rule. There is no *a priori* rule for multiplication and division because the proper number of significant figures depends critically on the result of the calculation. The application of the standard rule for multiplication and division can yield only three possible results. On very rare occasions it predicts one digit too many, overstating the precision. Most of the time it predicts one digit too few, causing valuable information to be lost. The application of the alternate rule for multiplication and division can also yield three possible results. It predicts one digit too many less than half of the time; on very rare occasions it can predict two digits too many, overstating the precision. However, it is always safe for data. As for the standard rule for addition and subtraction, it is always safe for data and the application of this rule can overstate precision only when dealing with a very long series of operations.

References

[1] R. H. Good, Phys. Teach. **34**, 192 (1996).
[2] L. M. Schwartz, J. Chem. Educ. **62**, 693 (1985).
[3] S. Stieg, J. Chem. Educ. **64**, 471 (1987).
[4] Christopher L. Mulliss and Wei Lee, Chin. J. Phys. **36**, 479 (1998).
[5] http://www.psrc-online.org/.
[6] http://www.angelfire.com/oh/cmulliss/other.html.
[7] R. A. Serway, *Physics for Scientists and Engineers with Modern Physics*, 4th ed. (Saunders College Publishing, San Francisco, C.A., 1996), p. 16.
[8] J. R. Taylor, *An Introduction to Error Analysis: The Study of Uncertainties in Physical Measurements*, 2nd ed. (University Science Books, Sausalito, C.A., 1997), pp. 30-31.
[9] P. R. Bevington and D. K. Robinson, *Data Reduction and Error Analysis for the Physical Sciences* (McGraw-Hill, New York, 1992), pp. 41-43.
[10] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in FORTRAN: The Art of Scientific Computing*, 2nd ed. (Cambridge University Press, New York, 1992), pp. 272-273.