On the Rounding Rules for Logarithmic and Exponential Operations

Wei-Da Chen,[1] Wei Lee,[1,*] and Christopher L. Mulliss[2]

[1] *Department of Physics, Chung Yuan Christian University, Chung-Li, Taiwan 32023, R.O.C.*
[2] *Ball Aerospace and Technologies Corporation, Fairborn, Ohio 45324, U.S.A.*

(Received August 12, 2004)

This work is a detailed investigation of the rounding rules for logarithmic and exponential operations, including mathematical derivations and statistical analyses via Monte-Carlo simulations. H. Lawrence Clever suggested rounding rules for logarithmic and exponential operations in 1979. We have found Clever's rules for logarithmic operations to be quite good, especially for natural logarithms, and found that they never lead to a loss of information. We propose a refinement to Clever's rounding rule for common (i.e. base 10) logarithmic operations that significantly increases its accuracy. Clever's rules for exponential operations were found to be extremely poor, almost never predicting the correct number of significant figures in the result. We suggest two alternate rules for exponential operations, both of which are demonstrated to be far more accurate and completely safe for data (i.e. never leading to a loss of information).

PACS numbers: 01.30.Pp, 01.55.+b, 02.70.Uu

I. INTRODUCTION

The concepts of significant figures and rounding rules are discussed in almost every textbook of introductory physics and chemistry. In these textbooks, students are taught two standard rounding rules, one for multiplication and division and the other for addition and subtraction. Over the years, college physics students and teachers have come to rely on these standard rounding rules. Good used a simple division problem to sound an alarm that the application of the standard rounding rule for multiplication and division could give rise to a loss of precision in the result, causing damage to data [1]. Such losses are attributed to the approximate nature of rounding rules and have been well documented by Schwartz [2]. Although some researchers feel that there is no need for a detailed investigation of the effects of significant figures and rounding rules on error propagation, owing to their approximate nature [3], others recognize the importance of rounding rules as common and convenient (even if approximate) tools in error analysis [1, 4]. Several recent detailed investigations of the rounding rules for multiplication/division and addition/subtraction, including mathematical derivations and statistical analyses, have been presented in the literature [5–7].

In addition to the four basic arithmetic operations, logarithmic and exponential operations are also quite frequently encountered in physics and chemistry. There are currently, however, no widely accepted rounding rules for these important mathematical operations. This has long been an unresolved problem. Clever suggested rounding rules in 1979, without giving a mathematical derivation, to deal with the significant figures in logarithmic and exponential operations [8]. Clever's rounding rule for the common (base 10) logarithmic operation requires that the number of significant decimal places in the answer equals one plus the number of significant digits in the argument. For the natural (base e) logarithmic operation, the number of significant decimal places equals the number of significant digits. Clever's rounding rule for the common (base 10) exponential operation requires that the number of significant digits in the answer equals the number of significant decimal places in the argument minus one. For the natural (base e) exponential operation, the number of significant digits equals the number of significant decimal places. The reason why Clever proposed such a rounding rule for the exponential operation is quite clear since the exponential operation is the inverse operation of the logarithmic operation. Similar rounding rules can also be found in Graham's work [9]. In this work, we study the error propagation of both logarithmic and exponential operations via simple mathematical theory and give a quantitative check of Clever's rules through a Monte-Carlo simulation.

As was categorized in our previous works [5–7], the statistical results may fall into several cases: those where the true uncertainty is as large or larger than that predicted by the rule but of the same order of magnitude, those where the true uncertainty is an order of magnitude less than predicted, and those where the true uncertainty is an order of magnitude larger than predicted. The rule is said to "work" in the first case because it predicts the minimum number of significant digits that can be written down without losing precision and hence valuable information in the result. In the second and third cases, the rule apparently "fails", predicting more or fewer significant figures than are needed and, accordingly, overstating or losing precision. This study indicates that Clever's rules are quite good for the logarithmic operations but very poor for the exponential operations. As a result, we propose two alternate rules for exponential operations. In addition, we also propose a refinement for the common logarithmic operation that significantly increases the accuracy of the rounding rule.

II. SIMPLE MATHEMATICAL DERIVATION FOR ROUNDING RULES

II-1. Relations between precision and significant figures

Rounding rules are used to determine the proper number of significant digits or significant decimal places that should be kept after a certain mathematical operation. When dealing with the problems of significant figures, we follow a very simple assumption, which states that the precision (namely, percentage error) of a number is approximately related to the number of significant figures in that number [10, 11]. Written in mathematical form, this assumption expresses the precision in a number x with N_x significant figures in the form [5]:

$$\text{Precision}(x) = C_x \cdot 10^{(2-N_x)}\% \,, \tag{1}$$

where C_x is a positive constant that can range from approximately 0.5 to exactly 5 depending on the actual value of the number x. Moreover, if we write a number x in scientific

TABLE I: Examples illustrating Eqs. (1)–(4).

Number x	a_x	Number of significant figures, N_x	Place of the right most significant digit, P_x	Uncertainty Δx	Precision(%) $\frac{\Delta x}{x} \times 100\%$	Value of C_x
60	6	1	1	0.5×10^1	0.83×10^1	0.83
72	7.2	2	0	0.5×10^0	0.69×10^0	0.69
64.4	6.44	3	-1	0.5×10^{-1}	0.78×10^{-1}	0.78
92.37	9.237	4	-2	0.5×10^{-2}	0.54×10^{-2}	0.54

notation as

$$x = a_x \times 10^{n(x)} , \tag{2}$$

where $1 \leq |a_x| \leq 10$ and $n(x)$ is an integer, we have the following relation [7]:

$$C_x = 5/|a_x| .$$

Thus, Eq. (1) can be rewritten as

$$\text{Precision}(x) = \frac{5}{|a_x|} \cdot 10^{(2-N_x)}\% . \tag{3}$$

Following Bevington and Robinson [12], the absolute error (or uncertainty) in this number is taken to $\pm 1/2$ in the least significant place. For the sake of convenience, the uncertainty can be expressed mathematically as [6]

$$\text{Uncertainty}(x) = 0.5 \times 10^{P_x} , \tag{4}$$

where P_x is an integer and denotes the place of the right most significant digit in a number x. If $P_x = 1$, the place of the right most significant digit is the tens place. If $P_x = 0$, the place of the right most significant digit is the units place. If $P_x = -1$, the place of the right most significant digit is the tenth decimal place, and so on [6]. All these relations are illustrated in Table I. Let $dp(x)$ denote the number of significant decimal places; i.e., the number of significant digits to the right of the decimal point of a number x. If a number x has $dp(x)$ significant decimal places, clearly we have the relation

$$P_x = -dp(x) . \tag{5}$$

Furthermore, if we divide Eq. (4) by Eq. (2) and compare the ratio with Eq. (3), then we have

$$N_x = n(x) - P_x + 1 \tag{6}$$

or

$$N_x = n(x) + dp(x) + 1 . \tag{7}$$

Eq. (7) has been presented and discussed previously in the literature, first by Graham, who used the relationship to develop significant-figure rules for generic arithmetic functions [9], and then independently by Lee *et al.*, who used the relationship to relate the rounding rule for addition to the rule for multiplication [6].

Throughout this entire paper, N_x denotes the number of significant digits of a number x, P_x the least significant place, and $dp(x)$ the number of significant decimal places. Moreover, we always use x, which is restricted to be a positive number, as the input of an operation and y the output.

Finally, with the above notation, we can summarize Clever's rules in the following mathematical forms:

$$\begin{array}{llll} dp(y) = N_x + 1 & \text{or} & P_y = -(N_x + 1) & \text{for common logarithmic operations}, \\ dp(y) = N_x & \text{or} & P_y = -N_x & \text{for natural logarithmic operations}, \\ N_y = dp(x) - 1 & \text{or} & N_y = -P_x - 1 & \text{for common exponential operations, and} \\ N_y = dp(x) & \text{or} & N_y = -P_x & \text{for natural exponential operations}. \end{array}$$

II-2. Rounding rules for logarithmic operations

Assume that the input x (> 0) has an uncertainty Δx. We would like to calculate the logarithm of x to the base b (> 0), where $b \neq 0$ or 1, and to know how this uncertainty propagates into the output y under the logarithmic operation:

$$y = \log_b x \,. \tag{8}$$

To achieve this goal, let us substitute $x \pm \Delta x$ into Eq. (8) to get

$$\begin{aligned} y \pm \Delta y &= \log_b(x \pm \Delta x) \\ &= \frac{1}{\ln b} \ln(x \pm \Delta x) \\ &= \frac{1}{\ln b} \ln \left[x \left(1 \pm \frac{\Delta x}{x} \right) \right] \\ &= \frac{1}{\ln b} \left[\ln x + \ln \left(1 \pm \frac{\Delta x}{x} \right) \right] \\ &= \frac{\ln x}{\ln b} + \frac{1}{\ln b} \left[\left(\pm \frac{\Delta x}{x} \right) - \frac{1}{2} \left(\pm \frac{\Delta x}{x} \right)^2 + \frac{1}{3} \left(\pm \frac{\Delta x}{x} \right)^3 - \ldots \right], \end{aligned}$$

where we use the series expansion $\ln(1 + x) = x - x^2/2 + x^3/3 - x^4/4 + \ldots$. Since the uncertainty of a number is smaller than that number, $\Delta x/x$ is smaller than 1, and we can therefore neglect the terms of $\Delta x/x$ of higher orders, keeping just the first-order term; i.e.,

$$y \pm \Delta y \approx \frac{\ln x}{\ln b} \pm \frac{1}{\ln b} \left(\frac{\Delta x}{x} \right) = \log_b x \pm \frac{1}{\ln b} \left(\frac{\Delta x}{x} \right).$$

Thus, we can easily recognize that

$$\Delta y \approx \frac{1}{\ln b} \left(\frac{\Delta x}{x} \right),$$

which is equivalent to

$$\text{Uncertainty}(y) \approx \text{Precision}(x)/\ln b\,. \tag{9}$$

This is the error-propagation equation for logarithmic operations. Substitution of the relations (i.e., Eqs. (3)–(5)) described above into Eq. (9) leads to the following formula:

$$dp(y) \approx N_x + [\log(a_x) + \log(\ln b) - 1]\ . \tag{10}$$

This is the formula we use to determine the number of significant decimal places in the output of logarithmic operations.

II-3. Rounding rules for exponential operations

As far as exponential operations are concerned, we consider the following exponential operation; i.e., the antilogarithm of x to the base b:

$$y = \text{antilog}_b\, x = b^x\,. \tag{11}$$

In a similar way, we carry out the same procedure as we did for logarithmic operations to get

$$\begin{aligned} y \pm \Delta y &= y \cdot \left(1 \pm \frac{\Delta y}{y}\right) \\ &= b^{x \pm \Delta x} \\ &= \exp\left[(x \pm \Delta x) \cdot \ln b\right] \\ &= \exp\left[x \cdot \ln b \pm \Delta x \cdot \ln b\right] \\ &= \exp(x \cdot \ln b) \times \exp(\pm\Delta x \cdot \ln b)\,. \end{aligned}$$

Thus, with the recognition that $y = b^x = \exp(x \cdot \ln b)$, we have

$$\begin{aligned} 1 \pm \frac{\Delta y}{y} &= \exp(\pm\Delta x \cdot \ln b) \\ &= 1 + (\pm\Delta x \cdot \ln b) + \frac{1}{2!}(\pm\Delta x \cdot \ln b)^2 + \frac{1}{3!}(\pm\Delta x \cdot \ln b)^3 + \dots\ , \end{aligned}$$

where we have expanded the exponential function in a Taylor series. We restrict our investigation to a range where $\Delta x \cdot \ln b$ is less than 1 for both bases considered ($b = 10$ and $b = e$). This restriction requires that $dp(x) \geq 0$. This allows us to neglect the terms of orders higher than 2 so that

$$1 \pm \frac{\Delta y}{y} \approx 1 \pm (\Delta x \cdot \ln b)\,.$$

Hence, we have

$$\frac{\Delta y}{y} \approx \Delta x \cdot \ln b\,,$$

or

$$\text{Precision}(y) \approx \text{Uncertainty}(x) \times (\ln b)\,. \tag{12}$$

This is the error-propagation equation for exponential operations. Similarly, substituting all the relations above into Eq. (12), we come to the following formula that determines the significant digits of the output of exponential operations:

$$N_y \approx dp(x) - [\log(a_y) + \log(\ln b) - 1]\ . \tag{13}$$

There is one thing worth noting. If we exchange x and y in Eq. (10) and rearrange the terms, we will recover Eq. (13). This is expected, considering that the logarithmic and exponential functions are inverse functions of each other. The error-propagation equations can be easily verified through differentiation, since we have the following relation for any function $y = f(x)$ [10]:

$$\Delta y \approx \left(\frac{df}{dx}\right) \cdot \Delta x\,.$$

III. A STATISTICAL STUDY OF THE ROUNDING RULES

III-1. The method

To investigate the statistical properties of the rounding rules, a Monte-Carlo procedure was used. A computer code was written in Fortran 90 and run on a Pentium(R) 4 PC equipped with a *Compaq Visual Fortran complier* (Version 6.5.0). The code uses a random number generator based on Ran2 [13] to create numbers. Each number 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. We restrict the random numbers to the range from 1 to 9.9999×10^5 for logarithmic operations and to the range from 0.01 to 9.9999 for exponential operations. The range of numbers investigated ensures that the assumptions used to derive the error-propagation equations, given in Eqs. (9) and (12), are valid. The program calculates the common and natural logarithms or exponentials of the generated numbers, and determines the number of significant figures that should be kept according to each of the rounding rules. It uses Eq. (9) to directly compute the "true" absolute error for logarithmic operations. It uses Eq. (12) to compute the "true" precision for exponential operations, which is then converted to the "true" absolute error. The absolute error in the result predicted by each rule is taken to $\pm1/2$ in the least significant decimal place. The true absolute error and the value predicted by each rule are compared, and the operation is counted as one of the cases described previously. The program repeats the calculation for one million logarithmic and exponential problems and computes statistics.

TABLE II: The statistical results of the application of Clever's rounding rules for logarithmic operations.

Category	$y = \log_b x$	
	Base 10	Base e
1 more digit needed	0.00%	0.00%
Worked	57.18%	97.00%
1 digit too many	42.82%	3.00%

III-2. Clever's rounding rules for logarithmic operations

Table II shows the statistical results of Clever's rounding rules for common logarithmic (base 10) and natural logarithmic (base e) operations. As is shown in the table, Clever's rounding rules work 57.18% of the time for common logarithmic operations and 97.0% of the time for natural logarithmic operations. It is apparent that Clever's rounding rules are excellent for natural logarithmic operations and acceptable for common logarithmic operations. Although it does not work as well for common logarithmic operations as it does for natural logarithmic operations, it is fortunate that when it fails it always predicts one more digit than is actually needed. It will never predict fewer significant digits than are actually needed. Thus it will never lead to a loss of information in the result. Because of its high accuracy for natural logarithmic operations and its acceptable accuracy plus safety against information loss for common logarithmic operations, Clever's rules are deemed to be satisfactory overall.

We must mention that our statistical results do not agree with Graham's prediction that Clever's rule for common logarithmic operations would fail about 64% of the time [9]. Graham's prediction was based upon an incorrect derivation that lead him to the following incorrect version of our Eq. (10): $dp(y) = N_x + [\log(\ln b)]$. Note that Graham's equation has been converted to our notation for the reader's convenience. He approximated $\log(x)$ as $n(x)+1$, but actually $\log(x)$ should be expressed as $n(x)+\log(a_x)$. With this modification and Eq. (7), Graham's formula would be the same as our Eq. (10). Therefore, we see that Graham's formula is only an approximation, which neglects the effect of, to first order, the leading digit of the number.

To see the mathematical background of Clever's rounding rules, let us specify the value of b in Eq. (10) to get

$$dp(y) \approx N_x + [\log(a_x) - 0.638] \quad \text{for common logarithmic operations}\,,$$

and

$$dp(y) \approx N_x + [\log(a_x) - 1] \quad \text{for natural logarithmic operations.}$$

To examine the contribution of the bracketed terms to the number of significant figures, we calculate the values of the bracketed terms for all possible values of $\mathrm{INT}(a_x)$. The results are displayed in Table III. It is worth mentioning that $\mathrm{INT}(w)$ here, as a convention, represents an integer converted from the positive number w by truncating. Consequently, $\mathrm{INT}(a_x)$

TABLE III: Contribution of the leading digit of the input $x = a_x \times 10^{n(x)}$ to the number of significant figures in the output of logarithmic operations.

INT(a_x)	Base 10		Base e	
	Value of $[\log(\mathrm{INT}(a_x)) - 0.638]$	Proper choice of $dp(y)$	Value of $[\log(\mathrm{INT}(a_x)) - 1]$	Proper choice of $dp(y)$
1	-0.638	N_x	-1.000	$N_x - 1$
2	-0.337	N_x	-0.699	N_x
3	-0.161	N_x	-0.523	N_x
4	-0.036	N_x	-0.398	N_x
5	0.061	$N_x + 1$	-0.301	N_x
6	0.140	$N_x + 1$	-0.222	N_x
7	0.207	$N_x + 1$	-0.155	N_x
8	0.265	$N_x + 1$	-0.097	N_x
9	0.316	$N_x + 1$	-0.046	N_x

stands exactly for the leading digit (namely, the left-most significant digit) of the input x. Note that $dp(y)$ listed in Table III must be a positive integer. The proper choice of $dp(y)$ should be made in accordance with

$$dp(y) = \mathrm{RINT}\left\{N_x + [\log(a_x) + \log(\ln b) - 1]\right\} ,$$

where RINT(w) denotes the smallest integer that is greater than or equal to w.

For common logarithmic operations, Table III shows that there are two possible cases: $dp(y) = N_x$ ($P_y = -N_x$) or $dp(y) = N_x + 1$ ($P_y = -(N_x + 1)$). In order to avoid losing valuable information carried by the digits, we would rather slightly overstate the precision than predict fewer digits than actually needed. That is to say, the most appropriate rule which is independent of a_x is $dp(y) = N_x{+}1$ ($P_y = -(N_x{+}1)$). This is just the mathematical statement of Clever's rule for common logarithmic operations. The two possible cases in Table III (N_x and $N_x + 1$) ensure that Clever's rule for common logarithmic operations ($dp(y) = N_x + 1$) can only predict one significant decimal place too many when it does fail.

As Table III shows, the transition from $dp(y) = N_x$ ($P_y = -N_x$) to $dp(y) = N_x + 1$ ($P_y = -(N_x + 1)$) happens somewhere between $a_x = 4$ and $a_x = 5$. Since a_x can be represented, to zeroth order, by the leading (significant) digit of the input x, we can refine Clever's rounding rule in order to attain a significantly higher probability of success. We provide the following rounding rule for common logarithmic operations, called the CLM rule for common logarithmic operations:

> If the leading digit of the input x is greater than or equal to 4, report one more significant decimal place in the answer than there are significant digits in the number x; otherwise, keep the same number of significant decimal places in the answer as there are significant digits in the number x.

The CLM rule for common logarithmic operations works very well, having a probability of

TABLE IV: The statistical results of the application of Clever's rules for exponential operations.

Category	$y = b^x$	
	Base 10	Base e
2 more digits needed	61.46%	0.00%
1 more digit needed	38.54%	97.89%
Worked	0.00%	2.11%
1 digit too many	0.00%	0.00%

success of 90.21%. It only predicts one digit too many 9.79% of the time and no other cases arise.

For natural logarithmic operations, Table III shows that there are two possible cases: $dp(y) = N_x$ ($P_y = -N_x$) or $dp(y) = N_x - 1$ ($P_y = 1 - N_x$). Due to the same reason described above, we adopt the rule $dp(y) = N_x$, which is exactly Clever's rule for natural logarithmic operations. Table III clearly illustrates why this rule works so well for natural logarithms.

III-3. Clever's rounding rules for exponential operations

Table IV shows the statistical results of Clever's rounding rules for exponential operations. We see that Clever's rules for exponential operations are extremely poor. They *never* work for common exponential operations and work only 2.11% of the time for natural exponential operations. Even worse is the fact that when it does fail, it predicts *fewer* digits than actually needed, causing damage to data. Therefore, Clever's rules for exponential operations cannot be adopted as the standard.

To see why Clever's rules fail, we substitute $b = 10$ and $b = e$ into Eq. (13) to obtain, respectively,

$$N_y \approx dp(x) - [\log(a_y) - 0.638] \quad \text{for common exponential operations,}$$

and

$$N_y \approx dp(x) - [\log(a_y) - 1] \quad \text{for natural exponential operations.}$$

Table V shows the contribution of the bracketed terms in the above two equations for integer values of a_y from 1 to 9.

For common exponential operations, Table V shows that there two possible cases: $N_y = dp(x)$ ($N_y = -P_x$) or $N_y = dp(x) + 1$ ($N_y = 1 - P_x$). The two possible cases in Table V ensure that Clever's rounding rule for common exponential operations ($N_y = dp(x) - 1$) cannot work at all and will predict one or two fewer significant digits than actually needed.

For natural exponential operations, Table V also shows that there are two possible cases: $N_y = dp(x)$ ($N_y = -P_x$) and $N_y = dp(x) + 1$ ($N_y = 1 - P_x$). The two possible cases in Table V ensure that Clever's rounding rule for natural exponential operations ($N_y = dp(x)$) will predict one less significant digit than is needed when it does fail.

TABLE V: Contribution of the leading digit of the output $y = a_y \times 10^{n(y)}$ of exponential operations to the number of significant figures in the output.

	Base 10		Base e	
INT(a_y)	Value of $-[\log(\mathrm{INT}(a_y)) - 0.638]$	Proper choice of N_y	Value of $-[\log(\mathrm{INT}(a_y)) - 1]$	Proper choice of N_y
1	0.638	$dp(x) + 1$	1	$dp(x) + 1$
2	0.337	$dp(x) + 1$	0.699	$dp(x) + 1$
3	0.161	$dp(x) + 1$	0.523	$dp(x) + 1$
4	0.036	$dp(x) + 1$	0.398	$dp(x) + 1$
5	-0.061	$dp(x)$	0.301	$dp(x) + 1$
6	-0.140	$dp(x)$	0.222	$dp(x) + 1$
7	-0.207	$dp(x)$	0.155	$dp(x) + 1$
8	-0.265	$dp(x)$	0.097	$dp(x) + 1$
9	-0.316	$dp(x)$	0.046	$dp(x) + 1$

III-4. Alternatives to Clever's rounding rules for exponential operations

In addition to understanding why Clever's rules for exponential operations fail, we can propose an alternate rounding rule based upon the mathematical analysis. Because the possible cases are $N_y = dp(x)$ ($N_y = -P_x$) and $N_y = dp(x) + 1$ ($N_y = 1 - P_x$) for both common and natural operations, we adopt the rule $N_y = dp(x) + 1$ ($N_y = 1 - P_x$) for the sake of saving the important information carried by the digits. To put it into words, we state the rounding rule in the following way:

> For both common and natural exponential operations, report one more significant digit in the output y than there are significant digits decimal places in the input x.

For convenience, we call this rule the CLM rule for exponential operations. We present the statistical results of this rule in Table VI. It turns out that this CLM rule for exponential operations is excellent for natural exponential operations and fairly good for common exponential operations. It works 64.46% and 98.13% of the time for common and natural exponential operations, respectively. Equally important, it never brings about a loss in valuable information carried by the digits. Although it will slightly overstate the precision, this is a minor shortcoming when compared with the risk of causing damage to data.

Note that the results in Tables V and VI may cause some readers concern. Seeing that in Table V, 4 out of 10 rows have $dp(x)+1$ as the right answer for common exponential operations, some readers may expect the success rate of the CLM rule to be about 40%. This logic holds true for the results in Table II and Table III. This logic is, however, invalid for the results in Table V and Table VI. Table III shows the contribution of the leading digit of the input which are generated as uniformly random numbers; i.e., the leading digit has a uniform frequency distribution. It is natural to expect the success rate in Table II to be about 60% for common logarithmic operations, because there are 6 out 10 rows in Table

TABLE VI: The statistical results of the application of the CLM rule for exponential operations.

Category	$y = b^x$	
	Base 10	Base e
1 more digit needed	0.00%	0.00%
Worked	61.46%	98.13%
1 digit too many	38.54%	1.87%

TABLE VII: The frequency distribution of the leading digit of the output y for common exponential operations.

Leading digit of $y = 10^x$	Digit frequency (%)
1	46.55%
2	12.40%
3	11.47%
4	5.61%
5	6.74%
6	6.07%
7	5.49%
8	3.04%
9	2.63%

III having $N_x + 1$ as the right answer. In Table V, however, we present the contribution of the leading digit of the output. Although we do generate uniformly random numbers for the input x, the leading digit of the answer will generally not have a uniform frequency distribution after a certain mathematical operation. We show the frequency distribution of a_y in Table VII to illustrate this idea.

Similarly, from Table V, we can refine the CLM rule for common exponential operations, so as to significantly increase the probability of success. We propose the following refined CLM rule for common exponential operations:

> If the leading digit of the output y is smaller than or equal to 4, report one more significant digit in the output y of common exponential operations than there are significant decimal places in the input x; otherwise, keep the same number of significant digits in the output y as there are significant decimal places in the input x.

It turns out that this refined CLM rule for common exponential operations works extremely well (90.16% of the time) and never leads to a loss of information (predicting one significant digit too many 9.84% of the time).

IV. SUMMARY

In most situations the problems encountered by physics students, including those in textbooks, do not deal with quantities where the uncertainties are explicitly stated. When this is the case significant figures play an important role in indicating the uncertainty of a number, and a rounding rule becomes meaningful. How to round a number properly after certain mathematical operations is a seemingly trivial, but fundamentally important, question. Rounding rules give us a simple and quick way to determine the proper number of significant figures to retain after simple mathematical operations. This is why an appropriate rounding rule is desired.

The rounding rules for multiplication/division and addition/subtraction have been discussed in the literature [5–7]. Besides these four basic arithmetic operations, physics students also often have to deal with logarithmic and exponential operations. Clever proposed rounding rules for logarithmic and exponential operations in 1979. After our detailed study and statistical analysis, Clever's rules for logarithmic operations were found to be fairly good. They work 57.18% and 97.0% of the time for common (base 10) and natural (base e) logarithmic operations, respectively. They will overstate the precision 42.82% of the time for common (base 10) operation and 3.0% of the time for natural (base e) logarithmic operations. Fortunately, Clever's rules for logarithmic operations never lead to a loss in precision. As far as Clever's rounding rules for exponential operations are concerned, they never work properly for common (base 10) exponential operations and work only 2.11% of the time for natural (base e) exponential operations.

Because Clever's rule for common logarithmic operations is so much less accurate (57.18%) than that for natural logarithmic operations (97%), we propose a refined rule for common logarithmic operations that is more accurate. The CLM rule for common logarithmic operations predicts the proper number of significant digits 90.21% of the time and one more digit than needed 9.79% of the time. However, the accuracy of the rounding rule is increased at the cost of simplicity.

Because Clever's rules for exponential operations are extremely poor, two alternative rules for exponential operations are proposed — the CLM rule for exponential operations and the refined CLM rule for common exponential operations. The CLM rule for exponential operations is shown to be accurate 61.46% and 98.13% of the time, respectively, for common and natural exponential operations. It will slightly overstate the precision 38.54% and 1.87% of the time for common and natural exponential operations, respectively. It will, however, never give rise to a loss in valuable information carried by the digits. Since the CLM rule for common exponential operations has an accuracy of only 61.46%, we further refine the rule so as to increase the accuracy. The refined CLM rule for common exponential operations works 90.16% of the time and predicts one digit too many 9.84% of the time. Again, the accuracy is increased in the expense of simplicity.

In conclusion, for the sake of convenience, we suggest adopting Clever's rounding rules for logarithmic operations and the CLM rule for exponential operations as the standard rounding rules for general purposes. However, if higher accuracy is required, one may well use the CLM rule for common logarithmic operations and the refined CLM rule for common

TABLE VIII: Summary of the suggested rounding rules.

Operation	Name of the rounding rule	Rounding rule	
Common logarithm	Clever's rule	$dp(y) = N_x + 1$	$P_y = -(N_x + 1)$
Natural logarithm	Clever's rule	$dp(y) = N_x$	$P_y = -N_x$
Common exponential	CLM rule	$N_y = dp(x) + 1$	$N_y = 1 - P_x$
Natural exponential	CLM rule	$N_y = dp(x) + 1$	$N_y = 1 - P_x$

TABLE IX: Examples illustrating the usage of the rounding rules for logarithmic operations.

Number x	Value of $\log(x)$	Rounded result by Clever's rule	Rounded result by the CLM rule
34.5	1.53782	1.5378	1.538
358	2.55388	2.5539	2.554
853	2.93094	2.9309	2.9309

Number x	Value of $\ln(x)$	Rounded result by Clever's rule
34.5	3.5409	3.541
358	5.8805	5.881
853	6.7488	6.749

exponential operations. One should choose the appropriate rounding rule to use depending on one's own purpose and situation. To familiarize the readers with the rounding rules, we summarize the suggested rounding rules in Table VIII and give some examples to illustrate the usage of the rounding rules in Tables IX and X.

TABLE X: Examples illustrating the usage of the rounding rules for exponential operations.

Number x	Value of 10^x	Rounded result by the CLM rule	Rounded result by the refined CLM rule
2.856	717.79	717.8	717.8
3.258	1811.3	1811	1811
5.05	112200	112000	110000

Number x	Value of e^x	Rounded result by the CLM rule
2.856	17.392	17.39
3.258	25.997	26.00
5.05	156.0	156

Acknowledgments

Wei-Da Chen is currently pursuing his M.S. degree in the Graduate Institute of Astronomy, National Central University. He acknowledges Chung Yuan Christian University for supporting this study via an undergraduate research grant.

APPENDIX: TOWARDS A GENERAL PRINCIPLE OF ROUNDING RULES

Seeing that there are many other mathematical functions of one variable, we would like to propose a general principle based on which one can determine the specific form of the rounding rule — which we consider the most proper one and should be adopted as a standard rule — for each function of one variable. To do so, we investigate in detail the error propagation of uncertainty according to a more general statistical approach and discuss the validity of the general principle.

We note first that the result of any measurement of a quantity x is, in general, stated as

$$x = x_{\text{best}} + \Delta x\,,$$

where x_{best} is the best estimate value of x in the measurement and Δx the uncertainty [10]. We then consider the Taylor series expansion about x_{best} of an operation of the form of $y = f(x)$

$$y = f(x) = f(x_{\text{best}}) + f'(x_{\text{best}})(x - x_{\text{best}}) + \frac{f''(x_{\text{best}})}{2!}(x - x_{\text{best}})^2 + \dots\,.$$

Moreover, the best estimate value of x can be proved to be the average of many measurements of x [10]; i.e.,

$$x_{\text{best}} = \langle x \rangle\,,$$

and for any symmetric error distribution function (such as the normal distribution) it can also be shown that

$$\sum_i (x_i - x_{\text{best}}) = 0 \qquad \text{or} \qquad \langle \Delta x \rangle = 0\,.$$

where x_i denotes the result of the ith measurement. Hence, we have

$$y = f(x) = f(\langle x \rangle) + f'(\langle x \rangle)(\Delta x) + \frac{f''(\langle x \rangle)}{2!}(\Delta x)^2 + \dots\,.$$

We will approximate the uncertainty (Δx) as the standard deviation σ_x [12]; i.e.,

$$(\Delta x)^2 \approx \sigma_x^2 = \left\langle (x - \langle x \rangle)^2 \right\rangle = \langle x^2 \rangle - \langle x \rangle^2\,. \tag{A.1}$$

If the uncertainty is small, the first-order approximation leads to

$$y = f(x) \approx f(\langle x \rangle) + f'(\langle x \rangle)(\Delta x)\,.$$

Thus the average of the outcome of the mathematical operation is

$$\langle y \rangle \approx f(\langle x \rangle) + \langle \Delta x \rangle f'(\langle x \rangle) \approx f(\langle x \rangle)\,.$$

We are now in a position to relate the uncertainty Δy, which we will approximate as the standard deviation σ_y, to the uncertainty Δx:

$$(\Delta y)^2 \approx \sigma_y^2 = \left\langle (y - \langle y \rangle)^2 \right\rangle = \left\langle \left[f'(\langle x \rangle)(\Delta x) \right]^2 \right\rangle = \left[f'(\langle x \rangle) \right]^2 \sigma_x^2 \approx \left[f'(\langle x \rangle) \right]^2 (\Delta x)^2\,.$$

Substituting the relation between uncertainty and the place of least significance (Eq. (4)) into the above equation, we arrive at, after some algebra,

$$P_y \approx P_x + \log f'(\langle x \rangle)\,. \tag{A.2}$$

This is the equation on which the rounding rule for any arbitrary function of one variable is based; it is equivalent to Graham's equation (Eq. (11)) [9].

To work out a specific form of the rounding rule for a specific function, one must investigate the contribution of the logarithmic term in Eq. (A.2). In general, P_y and P_x are required to be integers whereas the logarithmic term, which depends on the value of $\langle x \rangle$, may not be. This implies that there is no perfect rounding rule in most mathematical operations, as predicted by Mulliss and Lee [5] in an investigation of rounding rules for multiplication and division. Nevertheless, we can propose a proper rounding rule following the principle of least damage. Significant figures can be used as a rough indication of the precision of a number. Therefore, the significant figures cannot be overstated and rounding rules are used to provide a quick and convenient way to determine the proper number of digits that should be kept after a certain mathematical operation. Since no perfect rule exists, the application of rounding rules can give rise to an overestimate or underestimate of the number of significant figures that the result of an operation should have. An overestimate means overstating the precision of the result, while an underestimate indicates a loss of important information carried by the digits that are rounded off, implying damage to data. Accordingly, to strike a balance between overstating the precision and protecting the data, we must take into account the relative importance of overstating one or two more digits and of losing important information. In our point of view, overstating only one or two more digits is a minor defect compared with losing important information. This is why we always choose the rounding rules in such a way that the rules never result in cases of underestimate, thus causing damage to data. In order not to lose the information carried by the digits, obviously we should adopt the rule

$$P_y = \mathrm{LINT}\left[P_x + \log f'(\langle x \rangle) \right]\,, \tag{A.3a}$$

where LINT(w) denotes the largest integer that is smaller than or equal to w (L stands for the *left*-hand side and INT represents integers). By substituting Eq. (6) into Eq. (A.3a), a similar relationship can be obtained for N_y:

$$N_y = \mathrm{RINT}\left[1 + n(f(\langle x\rangle)) - P_x - \log f'(\langle x\rangle)\right] , \tag{A.3b}$$

where RINT(w) denotes the smallest integer that is larger than or equal to w (R stands for the *right*-hand side and INT represents integers). Note that $n(f(\langle x\rangle)) = \log f(\langle x\rangle) - \log a_y$, from taking the common logarithm of Eq. (2). In a similar manner, a_x may enter into Eq. (A.3a) or Eq. (A.3b) depending on the function $f(\langle x\rangle)$.

Rounding rules provide us with a quick and convenient tool to determine, *in advance*, the proper number of significant figures in the result of a mathematical operation. However, the logarithmic term in the general principle, Eq. (A.3a) or (A.3b), depends on the value of the average of x, or the value of the best estimate of x. Consequently, we cannot know beforehand the proper value of P_y since we have to compute the logarithmic term first. To overcome this problem, we choose the general principle as

$$P_y = \mathrm{LINT}\left[\min\left(P_x + \log f'(\langle x\rangle)\right)\right] \tag{A.4a}$$

or

$$N_y = \mathrm{RINT}\left[\max\left(1 + n(f(\langle x\rangle)) - P_x - \log f'(\langle x\rangle)\right)\right] , \tag{A.4b}$$

where the minimum or maximum function, $\min(w)$ or $\max(w)$, is an extreme value taken over the range of possible w, depending effectively on a_x and a_y. For many functions, Eq. (A.4a) or (A.4b) will reduce to cancelling out all dependence on $\langle x\rangle$ entirely, although function parameters such as bases and exponents may remain. For other functions, it may be necessary to evaluate the minimum or maximum function over a reasonable range of expected $\langle x\rangle$ values. In conclusion, as a consequence of all the above arguments, it is obvious that the rounding rules, so chosen, can give us a quick and convenient way to determine in advance the number of significant figures of the result of a mathematical operation, without losing important information carried by the digits.

To illustrate the proper manner of applying our approach, we will use it to recover the CLM rounding rule for the exponential function. We start with Eq. (A.4b), ignoring the RINT and maximum functions for now:

$$N_y \approx 1 + n(f(\langle x\rangle)) - P_x - \log f'(\langle x\rangle) .$$

We next substitute the derivative of the exponential function into the above expression:

$$N_y \approx 1 + n(f(\langle x\rangle)) - P_x - \log \ln b - x \log b .$$

We next recognize that $n(f(\langle x\rangle)) = \log f(\langle x\rangle) - \log a_y$, and substitute the result, evaluated for the exponential function, into the above expression:

$$N_y \approx 1 + x \log b - \log a_y - P_x - \log \ln b - x \log b .$$

We then simplify the expression and include the RINT and maximum functions to obtain:

$$N_y = \mathrm{RINT}\left[\max\left((1 - P_x) - (\log a_y - \log \ln b)\right)\right] .$$

For both $b = e$ and $b = 10$, and over the entire range of a_y, the above expression reduces to $N_y = 1 - P_x$, which is consistent with the results found by the more straightforward method employed in the main body of this paper (refer to Table VIII). The general statistic approach outlined in this appendix, and described by Eq. (A.4a) and Eq. (A.4b), appears to be a plausible approach that may be applied to a wide class of single-variable functions.

Finally, there is one point to note. In arriving at Eq. (A.2), we made a first-order approximation. How good is this approximation? Or, how large is the domain of validity of our general principle as expressed by Eq. (A.4a) or (A.4b)? To find the domain of validity of our general principle, we examine the condition that will make our approximation (to the first order) a good one. Apparently, the condition imposed is

$$\left|\frac{f^{(n)}(\langle x\rangle)}{n!}(\Delta x)^n\right| > \left|\frac{f^{(n+1)}(\langle x\rangle)}{(n+1)!}(\Delta x)^{n+1}\right|$$

or

$$\left|\frac{f^{(n+1)}(\langle x\rangle)}{f^{(n)}(\langle x\rangle)}\frac{\Delta x}{n+1}\right| < 1 ,$$

where $f^{(n)}$ denotes the nth derivative of the function f. We check the domain of validity for the cases of the logarithmic and exponential operations. First, note that the nth derivative of a logarithmic function of the form of Eq. (8) is

$$f^{(n)} = \frac{1}{\ln b}\frac{1}{x^n} ,$$

which gives the relationship

$$\left|\frac{f^{(n+1)}}{f^n}\frac{\Delta x}{n+1}\right| = \frac{1}{n+1}\left|\frac{\Delta x}{x}\right| < 1 .$$

The above inequality holds since n is a non-negative integer and $(\Delta x/x)$ is smaller than unity. This means that Clever's rounding rules for logarithmic operations can be applied to all real numbers. In this case, there is no restriction on the domain of validity. Next, let us check for the cases of exponential operations. The nth derivative of an exponential function given by Eq. (11) is

$$f^{(n)} = (\ln b)^n f ,$$

which leads to

$$\left|\frac{f^{(n+1)}}{f^n}\frac{\Delta x}{n+1}\right| = \frac{|\ln b \cdot \Delta x|}{n+1} .$$

For the criterion to hold, we must require $|\Delta x \cdot \ln b| < n + 1$, where n is a non-negative integer. Since a number that is smaller than 1 will also be smaller than 2, 3, etc., we demand, to be conservative, the condition to be

$$|\Delta x \cdot \ln b| < 1\,,$$

which requires $P_x \leq 0$, or $dp(x) \geq 0$, the same restriction as we have previously imposed in deriving the rounding rule for exponential operations.

References

* Electronic address: `wlee@phys.cycu.edu.tw`

[1] R. H. Good, Phys. Teach. **34**, 192 (1996).

[2] L. M. Schwartz, J. Chem. Edu. **62**, 693 (1985).

[3] B. L. Earl, J. Chem. Edu. **65**, 186 (1988).

[4] S. Stieg, J. Chem. Edu. **64**, 471 (1987).

[5] C. L. Mulliss and W. Lee, Chin. J. Phys. **36**, 479 (1998).

[6] W. Lee, C. L. Mulliss, and H.-C. Chiu, Chin. J. Phys. **38**, 36 (2000).

[7] W.-D. Chen, W. Lee, and C. L. Mulliss, Chin. J. Phys. **42**, 335 (2004).

[8] H. L. Clever, J. Chem. Edu. **56**, 824 (1979).

[9] D. M. Graham, J. Chem. Edu. **66**, 573 (1989).

[10] J. R. Taylor, *An Introduction to Error Analysis: The Study of Uncertainties in Physical Measurement*, 2nd ed. (University Science Books, Sausalito, C.A., 1997), pp. 30–31 and pp. 117–119.

[11] B. M. Shchigolev, *Mathematical Analysis of Observations* (London, Iliffe Books, London, 1965), p. 22.

[12] P. R. Bevington and D. K. Robinson, *Data Reduction and Error Analysis for the Physical Sciences* (McGraw-Hill, New York, 1992), p. 5 and p. 12.

[13] 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.