The error of a difference

If a number a of an approximation to the exact number A, then we call the difference $A-a$ the error in a. The relative error in 'a' as an approximation to A, is defined to be the number $\frac{A-a}{A}$. Note that this number is close to the number $\frac{A-a}{a}$ if it is at all small (precisely, if $\alpha = \frac{A-a}{A}$, then $\frac{A-a}{a} = \frac{\alpha}{1-\alpha}$).

Theorem The absolute error of an algebraic sum of several approximate numbers does not exceed the sum of the absolute errors of the numbers.

Proof. Let $x_1, x_2, \ldots, x_n$ be the given approximate numbers. Consider the algebraic sum

$$u = \pm x_1 \pm x_2 \pm \cdots \pm x_n.$$

Obviously,

$$\Delta u = \pm \Delta x_1 \pm \Delta x_2 \pm \cdots \pm \Delta x_n$$

and, hence,

$$|\Delta u| \le |\Delta x_1| + |\Delta x_2| + \cdots + |\Delta x_n| \quad \text{--- (1)}$$

Corollary. For the limiting absolute error of an algebraic sum we take the sum of the limiting absolute errors of the terms:

$$|\Delta_u| = |\Delta_{x_1}| + |\Delta_{x_2}| + \cdots + |\Delta_{x_n}| \quad \text{--- (2).}$$

Here we consider the difference of two approximate numbers: $u = x_1 - x_2$.

$$\therefore \quad |\Delta_u| = |\Delta_{x_1}| + |\Delta_{x_2}|$$

Whence the limiting relative error of the difference is

$$\delta_u = \frac{|\Delta_{x_1}| + |\Delta_{x_2}|}{A} \quad \text{--- (3)}$$

where A is the exact value of the absolute magnitude of the difference between the numbers x_1 and x_2.

If the approximate numbers x_1 and x_2 are nearly equal numbers and have small absolute errors, the number A is small. From formula (3), it follows that the limiting relative error in this case can be very large whereas the relative errors of the diminuend and subtrahend remain small. This amounts to loss of accuracy.

To illustrate let us compute the difference between two numbers: $x_1 = 47.132$ and $x_2 = 47.111$, each of which is correct to five significant digits. Subtracting, we get $u = 47.132 - 47.111 = 0.021$. The difference u has only two significant digits, of which the last is uncertain since the limiting absolute error of the difference is

$$\Delta u = 0.0005 + 0.0005 = 0.001$$

The limiting relative errors of the diminued, subtrahend and difference are:

$$\delta x_1 = \frac{0.0005}{47.132} \simeq 0.00001,$$

[Note: using $\frac{A-a}{a}$ instead of $\frac{A-a}{A}$].

$$\delta x_2 = \frac{0.0005}{47.111} \simeq 0.00001,$$

$$\delta_u = \frac{0.001}{0.021} \simeq 0.05.$$

The limiting relative error of the difference is roughly 5000 times greater than the limiting relative errors of the original numbers, i.e., x_1 and x_2. It is therefore desirable, in approximate computations, to transform the expressions in which computation of numerical values leads leads to the subtraction of nearly equal numbers.

We give here another example for clear understanding.

Let us find the roots of the quadratic equation

$$ax^2+bx+c=0$$

We know from algebra that the roots are given by the quadratic formula

$$x=\frac{-b\pm\sqrt{b^2-4ac}}{2a} \quad \text{--- (4)}$$

Let us assume that $b^2-4ac>0$, that $b>0$, and we wish to find the root of smaller absolute value using (4); i.e.,

$$x_1=\frac{-b+\sqrt{b^2-4ac}}{2a} \quad \text{--- (5)}.$$

If $4ac$ is small compared with b^2, then $\sqrt{b^2-4ac}$ will agree with to several places. Hence, given that $\sqrt{b^2-4ac}$ will be calculated correctly only to as many places as are used in the calculation, it follows that the numerator of (5), and therefore the calculated root, will be accurate to fewer places than the transformation of the formula by

$$x_1=\frac{(-b+\sqrt{b^2-4ac})(-b-\sqrt{b^2-4ac})}{2a(-b-\sqrt{b^2-4ac})}$$

$$=\frac{-2c}{b+\sqrt{b^2-4ac}} \quad \text{--- (6)}$$

To be specific, take the equation

$$x^2+111.11x+1.2121=0 \quad \text{--- (7)}.$$

Using (5) and five-decimal-digit floating-point chopped arithmetic, we calculate

$$b^2=12{,}345,\ b^2-4ac=12{,}340,\ \sqrt{b^2-4ac}=111.09$$

$$\therefore x_1 = \frac{-b+\sqrt{b^2-4ac}}{2a} = -0.01000$$

while infact, $x_1 = -0.010910$ is the correct root of (7). Here too, the loss of significant digits can be avoided by using (6), viz.,

$$x_1 = \frac{-2c}{b+\sqrt{b^2-4ac}}$$

and using five-decimal-digit arithmetic, we calculate

$$x_1 = -0.010910$$

which is accurate to five digits.