Numerical Analysis

Math 370 Fall 2004 — MWF 2:30 - 3:25pm

 — Fowler North 3

Class 2: **Wednesday September 8**

SUMMARY Round-off Error and The Significance Floating Point Arithmetic
CURRENT READING Recktenwald 5.2

k-th digit Chopping
In this case all the digits after d_k are **ignored** ("chopped off")
k-th digit Rounding
In this case if the value of $d_{k+1} \geq 5$ then d_k is replaced by $d_k + 1$
<u>GROUPWORK</u>
Show that the expression involving k which gives you an upper bound for the relative error involved in using chopping arithmetic is $\epsilon_{rel} = 10^{-k+1}$

It can also be shown that a bound for the relative error involved in using **rounding arithmetic** is *half* that for chopping, $\epsilon_{rel} = 0.5 \times 10^{-k+1} = 5 \times 10^{-k}$.
Round-off Errors in the Quadratic Formula
Recall that the common formula for the roots of a quadratic equation $ax^2 + bx + c = 0$ is

$$x_1 = \frac{-b + \sqrt{b^2 - 4ac}}{2a} \qquad \text{and } x_2 = \frac{-b - \sqrt{b^2 - 4ac}}{2a}$$

Round-off error can wreak havoc with the numerical implementation of this formula. Consider

$$x^2 + 62.10x + 1 = 0$$

which has the approximate roots $x_1 = -0.01610723$ and $x_2 = -62.08390$
Because of the size of the parameters in the quadratic equation, b^2 is much bigger than $4ac$, so $\sqrt{b^2 - 4ac}$ is very close to b. $a = 1$, $b = 62.10$, $c = 1$

$b^2 =$ $\qquad\qquad$ $4ac =$ $\qquad\qquad$ $b^2 - 4ac =$

GROUPWORK

Using 4-digit rounding arithmetic compute the first root x_1

What's the relative error in this calculation?

Solution: change the formula for x_1 so that we don't have to subtract b from $\sqrt{b^2 - 4ac}$
Now, a new formula for $x_1 =$

Use a similar new formula to compute x_2 (using 4-digit precision) and compute the relative error in x_2

What's the problem?
Solution: Use the new formula for x_1 when you have to subtract numbers which are similar in size, use the traditional formula for the other root.

The Ultimate Quadratic Formula

$$q \equiv -\frac{1}{2}\left[b + \text{sign}(b)\sqrt{b^2 - 4ac}\right]$$

where

$$\text{sign(b)} = \begin{cases} 1 & b \geq 0 \\ -1 & b < 0 \end{cases}$$

and

$$x_1 = \frac{q}{a} \qquad \text{and } x_2 = \frac{c}{q}$$