

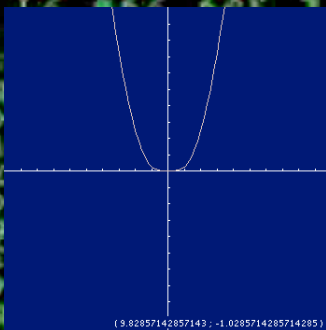
## Introduction

One of the greatest challenges in Scientific Computing is overcoming error in representation. Since the computer can only compute a finite number of operations, one must be aware of data that is lost due to round-off error. In this presentation, we illustrate an example, and a method to reconstruct our problem.

## Method

Let's show where subtraction of near close numbers can go wrong by finding a way of computing  $\sqrt{x^4 + 4} - 2$  without loss of undue significance. If we compute by hand we have the following:

$$\begin{aligned}\sqrt{x^4 + 4} - 2 &= 0 \\ \sqrt{x^4 + 4} &= 2 \\ x^4 + 4 &= 4 \\ x^4 &= 0 \\ x &= 0\end{aligned}$$



(3.82857142857143; -1.0285714285714285)

# Error

## Alexander Graham Department of Mathematics

### Results

Using Numerical Software, such as MatLab, we find that as  $x$  gets closer to 0, our approximation gets worse because we are subtracting close numbers, and the computer can only store a finite number of digits. If we evaluate at  $10^{-10}$ , we find that we lose 9 significant bits by the following theorem:

*Let  $x$  and  $y$  be positive normalized floating point numbers. In the subtraction  $x - y$ ,  $r$  significant bits are lost where  $q \leq r \leq p$  and  $2^q - p \leq 1 - y/x \leq 2^q - q$  for some positive integers  $p$  and  $q$ .*

Our numerical program shows us that  $10^{-10}$  yields .0000000003, while  $10^{-11}$  yields .00000000075, but this is false simply because the function is strictly increasing for  $x > 0$ . Our function becomes unstable for  $x$  close to 0, so we need to make a slight adjustment to avoid subtraction by like terms.

## Summary

We can rewrite the function in the following method to avoid this trouble:

$$\frac{(\sqrt{x^4 + 4} - 2)(\sqrt{x^4 + 4} + 2)}{(\sqrt{x^4 + 4} + 2)} \cdot \frac{x^4}{x^4}$$

These two functions are equivalent because we simply multiplied by a special value of one (the conjugate) to eliminate the loss of significance. Our answer also goes to 0 as  $x$  goes to 0.

## Conclusion

Controlling error is key for many applied mathematicians. It is used frequently in image processing, radio signals, and digital filters to maintain consistency within programs. Similar techniques can be used with transcendental functions, by expanding them into their respective Taylor polynomials. This is one of many tools we can use to increase our understanding of numerical data in Scientific Computing

## Acknowledgements

- Computer Arithmetic Algorithms  
Israel Koren, Prentice-Hall, 2002
- Numerical Analysis  
Kincaid & Cheney, AMS, 2002