

Accuracy of Floating Point Binary Numbers



<http://nm.MathForCollege.com>

Transforming Numerical Methods Education for STEM Undergraduates



For more details on this topic

- Go to <http://nm.MathForCollege.com>
- Click on Floating Point Binary Representation



You are free

- to **Share** – to copy, distribute, display and perform the work
- to **Remix** – to make derivative works

under the following conditions

- **Attribution** — You must attribute the work in the manner specified by the author or licensor (but not in any way that suggests that they endorse you or your use of the work).
- **Noncommercial** — You may not use this work for commercial purposes.
- **Share Alike** — If you alter, transform, or build upon this work, you may distribute the resulting work only under the same or similar license to this one.



How do you determine the accuracy of a floating-point binary number?





Q: What is the significance of machine epsilon for a student in an introductory course in numerical methods?

A: The machine epsilon is an upper bound on the absolute relative true error in representing a number. Let y be the machine number representation of a number x , then we can show that the absolute relative true error in the representation is

$$\left| \frac{x-y}{x} \right| \leq \epsilon_{mach}$$



A machine stores floating-point numbers in a hypothetical 10-bit binary word. It employs the first bit for the sign of the number, the second one for the sign of the exponent, the next four bits for the magnitude of the exponent, and the last four bits for the magnitude of the mantissa. Confirm that the magnitude of the relative true error that results from approximate representation of 0.02832 in the 10-bit format is less than the machine epsilon.





Acknowledgement

This instructional power point brought to you by
Numerical Methods for STEM undergraduate

<http://nm.MathForCollege.com>

Committed to bringing numerical methods to the
undergraduate

This material is based upon work supported by the National
Science Foundation under Grant #2013271. Any opinions,
findings, and conclusions or recommendations expressed in this
material are those of the author(s) and do not necessarily
reflect the views of the National Science Foundation.



THE END

<http://nm.MathForCollege.com>

USF UNIVERSITY OF
SOUTH FLORIDA

