On the maximum relative error when computing x^n in floating-point arithmetic

Stef Graillat[1] and Vincent Lefèvre[2] and Jean-Michel Muller[2]

[1] Université Pierre et Marie Curie, Paris France
[2] Laboratoire LIP, CNRS, ENS de Lyon, Inria and Université Claude Bernard, Lyon, France

jean-michel.muller@ens-lyon.fr

Keywords: Floating-Point Arithmetic, Rounding errors

We improve the usual relative error bound for the computation of x^n through iterated multiplications by x in binary, precision-p, floating-point arithmetic. More precisely, we analyze the following algorithm.

$y \leftarrow x$
for $k = 2$ to n do
  $y \leftarrow \mathrm{RN}(x \cdot y)$
end for
return(y)

where RN means "round to nearest" (that is, $\mathrm{RN}(x \cdot y)$ is the result of the floating point multiplication `x * y` in round-to-nearest mode).

We show the following result:

Theorem 1. *Assume $p \geq 5$ (which holds in all practical cases). If*

$$n \leq \sqrt{2^{1/3} - 1} \cdot 2^{p/2},$$

then

$$|\widehat{x}_n - x^n| \leq (n-1) \cdot u \cdot x^n.$$

The obtained relative error bound $(n-1) \cdot u$ is only slightly better than the usual one γ_{n-1}, but it is simpler. We also discuss the more general problem of computing the product of n terms.

References:

[1] Muller et al., *Handbook of Floating-Point Arithmetic*, Birkhauser, 2010.

[2] Graillat, Lefevre and Muller, *On the maximum relative error when computing x^n in floating-point arithmetic*, Technical report, 2013. Available at http://hal-ens-lyon.archives-ouvertes.fr/ensl-00945033