COSASIN.TXT 28 Jan. 2008

Computing $1 - s^2$ Accurately in Binary Floating-Point

Suppose $0 < s = \sin(\theta) < 1$. We wish to compute $c := \cos(\theta)$ from s as accurately as we can from formulas $c^2 = 1 - s^2 = (1-s)\cdot(1+s)$. Which formula should we use to minimize rounding error? Let formulas

$F0(s) := 1 - s^2$ compute $1 - (s^2 \pm \delta s^2) \pm \delta c^2 = c^2 \pm E0\cdot\epsilon$ and

$F1(s) := (1-s)\cdot(1+s)$ compute $(1-s \pm \delta s)\cdot(1+s \pm \frac{1}{2}\epsilon) \pm \delta c^2 = c^2 \pm E1\cdot\epsilon$

where

$$E0(s) = (\delta s^2 + \delta c^2)/\epsilon \quad \text{and} \quad E1(s) \approx \tfrac{1}{2}(1-s) + (1+s)\delta s/\epsilon + \delta c^2/\epsilon \ .$$

Assume $\epsilon := \text{NextAfter}(1, +\infty) - 1$, as is the case for MATLAB whose eps = 1/2^52 , and that arithmetic is rounded correctly "to nearest".

Evidently E0 is smaller than E1 if s is small enough, and E1 is the smaller if s is close enough to 1 . We seek a threshold σ optimized to minimize our rounding error-bound if we use F0(s) when $0 \le s < \sigma$ and F1(s) otherwise. $\sigma = 3/4$ because ...

If $1/\sqrt{8} < s < \frac{1}{2}$: $1/8 < s^2 < \frac{1}{4}$ and $3/4 < c^2 < 7/8$.
- In E0, $\delta s^2 = \epsilon/16$ and $\delta c^2 = \frac{1}{4}\epsilon$ so $E0(s) = 5/16$.
- In E1, $\delta s = \frac{1}{4}\epsilon$ and $\delta c^2 = \frac{1}{4}\epsilon$ so $E1(s) \approx \frac{1}{2}(1-s) + \frac{1}{4}(1+s) + \frac{1}{4}$.
- Evidently $E1(s) > E0(s)$, so use F0(s) to compute c^2 .

If $\frac{1}{2} < s < \sqrt{\frac{1}{2}}$: $\frac{1}{4} < s^2 < \frac{1}{2}$ and $\frac{1}{2} < c^2 < 3/4$.
- In E0, $\delta s^2 = \epsilon/8$ and $\delta c^2 = \frac{1}{4}\epsilon$ so $E0(s) = 3/8$.
- In E1, $\delta s = 0$ and $\delta c^2 = \frac{1}{4}\epsilon$ so $E1(s) \approx \frac{1}{2}(1-s) + \frac{1}{4}$.
- Evidently $E1(s) > E0(s)$, so use F0(s) to compute c^2 .

If $\sqrt{\frac{1}{2}} < s < \sqrt{3}/2$: $\frac{1}{2} < s^2 < 3/4$ and $1/4 < c^2 < \frac{1}{2}$.
- In E0, $\delta s^2 = \frac{1}{4}\epsilon$ and $\delta c^2 = 0$ so $E0(s) = \frac{1}{4}$.
- In E1, $\delta s = 0$ and $\delta c^2 = \epsilon/8$ so $E1(s) \approx \frac{1}{2}(1-s) + 1/8$.
- Evidently $E1(s) > E0(s)$ just when $s < 3/4$ so ...
  - if $s < 3/4$ use F0(s) to compute c^2 , else use F1(s) .

If $\sqrt{3}/2 < s < \sqrt{(7/8)}$: $3/4 < s^2 < 7/8$ and $1/8 < c^2 < \frac{1}{4}$.
- In E0, $\delta s^2 = \frac{1}{4}\epsilon$ and $\delta c^2 = 0$ so $E0(s) = \frac{1}{4}$.
- In E1, $\delta s = 0$ and $\delta c^2 = \epsilon/16$ so $E1(s) \approx \frac{1}{2}(1-s) + 1/16$.
- Evidently $E1(s) < E0(s)$, so use F1(s) to compute c^2 .