ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
Εργαστήριο Θερμικών Στροβιλομηχανών
Μονάδα Παράλληλης Υπολογιστικής Ρευστοδυναμικής & Βελτιστοποίησης

ΑΡΙΘΜΗΤΙΚΗ ΑΝΑΛΥΣΗ

(4ο Εξάμηνο Σχολής Μηχ.Μηχ. ΕΜΠ)

ΣΦΑΛΜΑΤΑ ΑΡΙΘΜΗΤΙΚΩΝ ΥΠΟΛΟΓΙΣΜΩΝ

Κ.Χ. Γιαννάκογλου, Καθηγητής ΕΜΠ
kgianna@central.ntua.gr
http://velos0.ltt.mech.ntua.gr/research/

Αριθμητική Ανάλυση (ΑΑ) & Σφάλματα Υπολογισμών

- **Οι αριθμητικές μέθοδοι προσεγγίζουν το αποτέλεσμα του προβλήματος και εισάγουν εκ προοιμίου σφάλμα, το οποίο είναι το κόστος για να αποφευχθεί η (ίσως αδύνατη) αναλυτική του λύση.**

- **Τα σφάλματα συνδέονται και με τον τρόπο εκτέλεσης των αριθμητικών πράξεων.**

- **Κατά την εφαρμογή μιας αριθμητικής μεθόδου, πρέπει να (α) εκτιμάται την τάξη του αναμενόμενου σφάλματος & (β) γίνεται προσπάθεια ελαχιστοποίησής του.**

Αρκετά από τα παρακάτω πρέπει ήδη να τα γνωρίζετε είτε από προηγηθέντα μαθήματα ΗΥ και προγραμματισμού τους είτε από τα προηγούμενα κεφάλαια Αριθμητικής Ανάλυσης.

Απόλυτο & Σχετικό Σφάλμα Υπολογισμών

x_t = πραγματική τιμή ενός μεγέθους (true)
x_a = προσεγγιστική τιμή του ίδιου μεγέθους (approximate)

Απόλυτο Σφάλμα: $E_t = x_t - x_a$

Σχετικό Σφάλμα: $\varepsilon_t = \frac{E_t}{x_t} = \frac{x_t - x_a}{x_t} = 1 - \frac{x_a}{x_t}$

Εκτίμηση σχετικού σφάλματος σε επαναληπτικές μεθόδους:

$$\varepsilon_a = \left(1 - \frac{x_a^{k-1}}{x_a^k}\right) \cdot 100\%$$

Κριτήριο τερματισμού: $\left|\varepsilon_a\right| \le \varepsilon_r$

Σημαντικά Ψηφία (ΣΨ) Αριθμών

- Κάθε αριθμός μπορεί να παρασταθεί, ακριβώς ή κατά προσέγγιση, με τη μορφή ενός δεκαδικού με πεπερασμένο πλήθος ψηφίων. Σημαντικά Ψηφία είναι τα ψηφία του που ακολουθούν το πρώτο μη-μηδενικό (συμπ/νου) τα οποία είναι γνωστά με ακρίβεια.
- Η έννοια των ΣΨ σχετίζεται άμεσα με την ακρίβεια των αριθμητικών πράξεων, αφού χρησιμοποιώντας περισσότερα σημαντικά ψηφία αυξάνεται το πλήθος των σωστών ψηφίων στο αποτέλεσμα της πράξης.
- 3 αριθμοί με 4 ΣΨ έκαστος: 4,045 0,4045 0,004045
- Οι 0,17 0,170 0,1700 (με 2, 3, 4 ΣΨ αντίστοιχα) πιθανόν να μην είναι ίσοι μεταξύ τους.
- Μας ενδιαφέρουν τα Σωστά ΣΨ (ΣΣΨ) με τα οποία προσεγγίζεται η πραγματική λύση. Λχ. αν x_t=22,000 τότε δύο αριθμητικές μέθοδοι που καταλήγουν στις λύσεις x_a=22,003 και x_a=21,994 έχουν βρει 4 ΣΨΨ και οι δύο.

Σημαντικά Ψηφία Αριθμών – Παραδείγματα

- **Αν το σχετικό σφάλμα ενός αριθμού είναι μικρότερο μιας τιμής ε_r (%), τότε ο αριθμός θα έχει τουλάχιστον n ΣΣΨ, σύμφωνα με:**

$$\varepsilon_r = 0{,}5 \times 10^{-n}$$

ή

$$\varepsilon_r \leq (0{,}5 \times 10^{2-n})\ \%$$

<u>Παράδειγμα:</u>

Ακριβής/αναλυτική λύση = 4,80
Αριθμητική λύση Μεθόδου Α = 4,807
Αριθμητική λύση Μεθόδου Β = 4,794

Σφάλμα Στρογγυλοποίησης

Η καταχώριση αριθμού στη μνήμη ΗΥ γίνεται με περιορισμένο αριθμό ΣΨ. Επίσης, υπάρχουν αριθμοί που, ενώ στο δεκαδικό σύστημα παριστάνονται ακριβώς, δεν έχουν πεπερασμένη παράσταση στο δυαδικό σύστημα. Το σφάλμα λόγω παράλειψης ΣΨ ενός αριθμού ονομάζεται **σφάλμα στρογγυλοποίησης**.

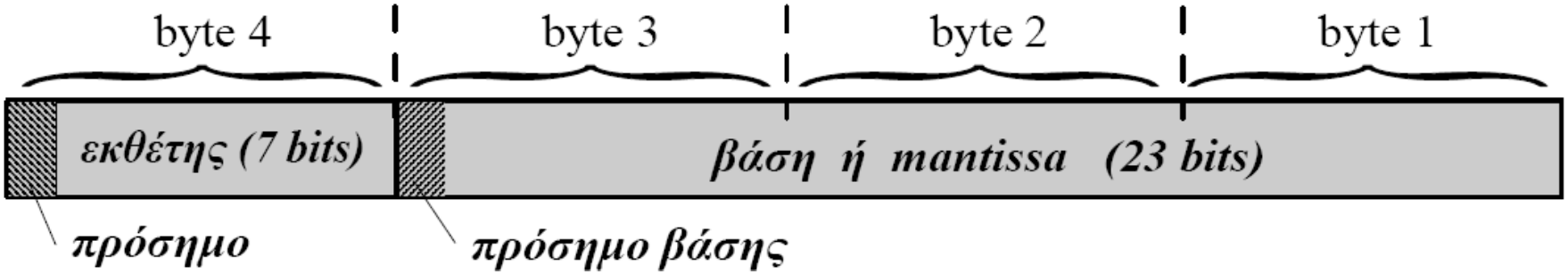
Τρόποι στρογγυλοποίησης:

- **Απλή Στρογγυλοποίηση:** αγνοούνται όλα τα σημαντικά ψηφία που δεν μπορούν να καταχωρηθούν.

- **Συμμετρική Στρογγυλοποίηση:** το τελευταίο ψηφίο που διατηρείται αυξάνεται κατά μία μονάδα αν το πρώτο ψηφίο που παραλείπεται είναι ίσο ή μεγαλύτερο του 5.

Καταχώρηση Αριθμών στη Μνήμη ΗΥ

Καταχώρηση πραγματικού αριθμού 32 bits



Στην αριθμητική κινητής υποδιαστολής (floating point), κάθε αριθμός εκφράζεται σε εκθετική μορφή

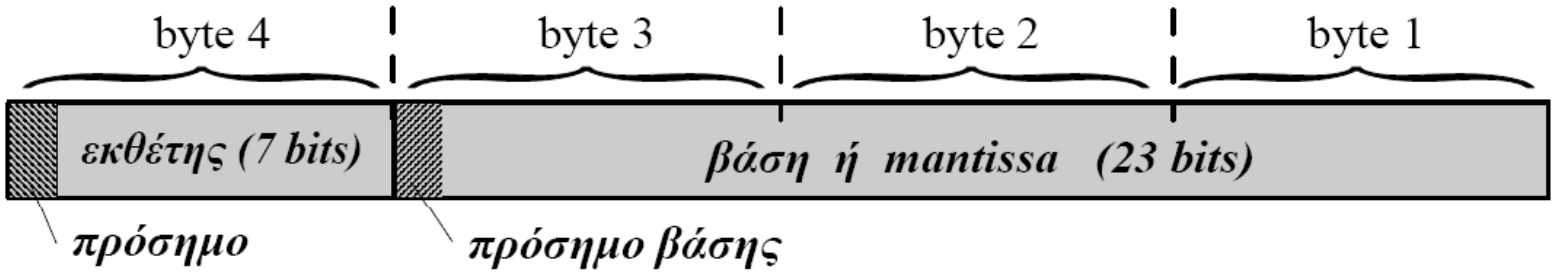
$$x = f \cdot b^m$$

εκθέτης (exponent)

κλασματικό μέρος ή βάση (mantissa). Κανονικοποιημένο, δηλαδή το πρώτο ψηφίο αμέσως μετά την υποδιαστολή όχι 0.

βάση συστήματος αρίθμησης

Καταχώρηση Αριθμών στη Μνήμη ΗΥ



Μέγιστη τιμή εκθέτη στα 7 διαθέσιμα bits:

$m = 1111111_2 = (2^7 - 1)_{10} = 127$

Απλή Ακρίβεια - Single Precision

Μέγιστος (κατά απόλυτη τιμή) πραγματικός αριθμός που μπορεί να καταχωρηθεί: $x = 1 \cdot 2^{127} \approx 2. \times 10^{38}$ → Σφάλμα Υπερχείλισης (overflow).

Ελάχιστος (κατά απόλυτη τιμή) πραγματικός αριθμός που μπορεί να καταχωρηθεί: $x = 0{,}1 \cdot 2^{-127} \approx 0{,}6 \cdot 10^{-39}$, → Σφάλμα «Υπερ»χείλισης (underflow).

Διπλή Ακρίβεια – Double Precision

Διατίθενται 64 bits, για έναν αριθμό. Τα bits του εκθέτη γίνονται τουλάχιστον 10, οπότε καταχωρούνται αριθμοί μέχρι και $10^{\pm 308}$, ενώ παρέχεται ακρίβεια της τάξης των 15 ÷ 16 σημαντικών ψηφίων.

Σήμερα, δεν έχει κανένα νόημα η υλοποίηση μεθόδων ΑΑ με χρήση απλής ακρίβειας. *Εξαιρέσεις??*

Σφάλμα Στρογγυλοποίησης κατά την Καταχώρηση Αριθμών

Η πραγματική λύση του προβλήματος:

$$x_t = f \cdot b^m + g \cdot b^{m-k}$$

k πρώτα ΣΨ που μπορούν να καταχωρηθούν

f, g κανονικοποιημένα:

$$\frac{1}{b} \le |f| < 1 \qquad 0 \le |g| < 1$$

Απλή στρογγυλοποίηση:

$$x_a = f \cdot b^m \quad \Rightarrow \quad |\varepsilon_t| = \left|\frac{x_t - x_a}{x_t}\right| \le \left|\frac{g \cdot b^{m-k}}{x_t}\right| \le \left|\frac{1 \cdot b^{m-k}}{x_t}\right| \le \frac{b^{m-k}}{b^{m-1}} = b^{1-k}$$

Συμμετρική στρογγυλοποίηση:

$$x_a = \begin{cases} f \cdot b^m, & |g| < 0{,}5 \\ f \cdot b^m \pm b^{m-k}, & |g| > 0{,}5 \end{cases}$$

Δείξτε ότι το σχετικό σφάλμα της συμμετρικής είναι το μισό αυτού της απλής στρογγυλοποίησης.

Άσκηση

Άσκηση 1: **Έστω Η/Υ (που χρησιμοποιεί απλή στρογγυλοποίηση) με 3 διαθέσιμες θέσεις στη βάση για την αποθήκευση πραγματικών αριθμών.**

(α) Ποιοι αριθμοί μπορούν να αποθηκευθούν με ακρίβεια στο διάστημα [0,5 , 1];

(β) Πώς καταχωρούνται στην μνήμη του Η/Υ οι αριθμοί

0,63 0,68 0,62499 0,62501 ;

Τι πρόβλημα μπορεί να υπάρξει στη σύγκριση των δυο ζευγών αριθμών; Πώς αντιμετωπίζεται το πρόβλημα;

Σφάλμα Στρογγυλοποίησης σε Αριθμητικές Πράξεις

Πρόσθεση δύο αριθμών:

- Αρχικά γράφονται με ίδιο εκθέτη (τον μεγαλύτερο).
- Η βάση του αριθμού με το μικρότερο εκθέτη παύει να είναι κανονικοποιημένη και κάποια ΣΨ μπορεί να χαθούν.
- Παράδειγμα:

$0{,}4658 \cdot 10^{1} + 0{,}3765 \cdot 10^{-1}$

στο δεκαδικό σύστημα, σε βάση 4 θέσεων (καταχωρημένοι με ακρίβεια).

$0{,}4658 \cdot 10^{1} + 0{,}0037 \cdot 10^{1} = (0{,}4658 + 0{,}0037) \cdot 10^{1} = 0{,}4695 \cdot 10^{1}$

αντί του σωστού $0{,}469565 \cdot 10^{1}$.

Μικρή απώλεια ακρίβειας κατά την πρόσθεση, που ίσως γίνει σημαντική αν προστίθεται τεράστιο πλήθος πραγματικών αριθμών.

Σφάλμα Στρογγυλοποίησης σε Αριθμητικές Πράξεις

Πρόσθεση πολλών αριθμών:

- **Δεν** ισχύει η προσεταιριστική ιδιότητα:

Παράδειγμα: άθροισμα των $x_1 = 5675$, $x_2 = -5673$ και $x_3 = 3{,}457$, σε βάση τεσσάρων θέσεων:

$$(x_1 + x_2) + x_3 = (0{,}5675 \cdot 10^4 - 0{,}5673 \cdot 10^4) + 0{,}3457 \cdot 10^1 = 0{,}0002 \cdot 10^4 + 0{,}3457 \cdot 10^1 =$$
$$= 0{,}2000 \cdot 10^1 + 0{,}3457 \cdot 10^1 = 0{,}5457 \cdot 10^1 = 5{,}457, \quad \text{ενώ:}$$
$$x_1 + (x_2 + x_3) = 0{,}5675 \cdot 10^4 + (-0{,}5673 \cdot 10^4 + 0{,}0003 \cdot 10^4) = 0{,}5675 \cdot 10^4 - 0{,}5670 \cdot 10^4 =$$
$$= 0{,}0005 \cdot 10^4 = 5{,}000$$

$$(x_1 + x_2) + x_3 \neq x_1 + (x_2 + x_3)$$

Το τελικό σφάλμα στρογγυλοποίησης μειώνεται αν προστεθούν πρώτα οι αριθμοί που δεν διαφέρουν σημαντικά μεταξύ τους.

Διαβάστε μια σειρά από επιλεγμένα αριθμητικά παραδείγματα στο βιβλίο του μαθήματος.

Σφάλμα Στρογγυλοποίησης σε Αριθμητικές Πράξεις

Αφαίρεση δύο (πολύ κοντινών) αριθμών: **Παράδειγμα:**

$0{,}4658\cdot10^{1} - 0{,}4592\cdot10^{1} = 0{,}0066\cdot10^{1}$ → **Καταχώριση:** $0{,}6600\cdot10^{1}$

και το σφάλμα μεταφέρεται σε επόμενες πράξεις…

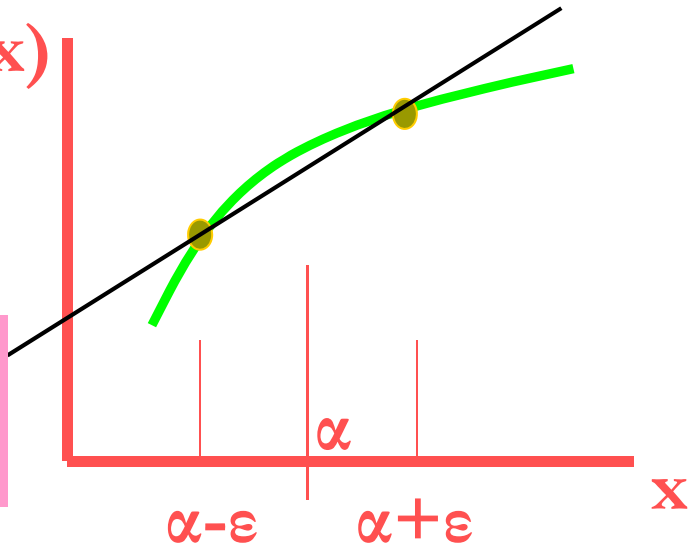
Ένα «πιο ακραίο» παράδειγμα (βάση 4 θέσεων) **:**

$4{,}65855 - 4{,}65845 = 0{,}4659\cdot10^{1} - 0{,}4658\cdot10^{1} = 0{,}0001\cdot10^{1} = 0{,}1000\cdot10^{-2}$

Σφάλμα 900%.

Το πρόβλημα ακρίβειας
των <u>Πεπερασμένων Διαφορών</u>.

$$\frac{f(\alpha+\varepsilon)-f(\alpha-\varepsilon)}{2\varepsilon} = f'(\alpha)+\mathrm{O}(\varepsilon^{2})$$



Σφάλμα Αποκοπής

Στις αριθμητικές μεθόδους, η παράλειψη όρων της πλήρους μαθ. έκφρασης μιας παράστασης. προκαλεί σφάλμα αποκοπής.

$$f(x) = f(a) + f'(a)(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \ \ldots + \frac{f^{(n)}(a)}{n!}(x-a)^n + R_n$$

Σφάλμα Αποκοπής:

$$R_n = \frac{f^{(n+1)}(\xi)}{(n+1)!}(x-a)^{n+1}$$

όπου $\xi \in [a, x]$ ένας συγκεκριμένος αριθμός, που εξαρτάται από το n.

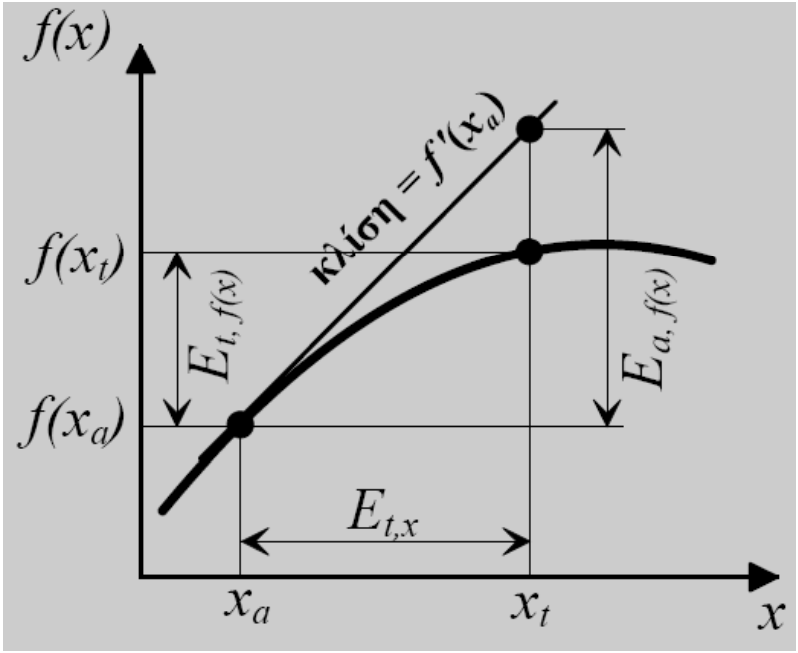
►Σε μια αριθμητική μέθοδο, το σφάλμα κάθε ενδιάμεσου αποτελέσματος μεταφέρεται στην επόμενη σχέση που θα χρησιμοποιηθεί, όπου μπορεί να μικρύνει ή να μεγαλώσει. Να ελέγχεται η επίδραση που έχει ένα σφάλμα σε επόμενες πράξεις και να αποφεύγεται μια μεγάλη αύξηση κατά τη μετάδοσή του !!!

Μετάδοση Σφάλματος σε Συναρτήσεις Μιας Μεταβλητής

Απόλυτο Σφάλμα: $E_{t,f(x)} = f(x_t) - f(x_a)$

$$f(x_t) \cong f(x_a) + f'(x_a)(x_t - x_a) \Rightarrow f(x_t) - f(x_a) \cong f'(x_a)(x_t - x_a)$$

$$E_{t,f(x)} \cong f'(x_a) E_{t,x} = E_{a,f(x)}$$



Ακριβής εκτίμηση σφάλματος για γραμμική συνάρτηση…

Μετάδοση Σφάλματος σε Συναρτήσεις Μιας Μεταβλητής

Εκτίμηση του Σχετικού Σφάλματος:

$$e_{a,f(x)} \cong \frac{E_{a,f(x)}}{f(x_a)} \cong \frac{f'(x_a)(x_t - x_a)}{f(x_a)} = \frac{f'(x_a)\,x_a}{f(x_a)}\,\frac{(x_t - x_a)}{x_a} = \boxed{\frac{f'(x_a)\,x_a}{f(x_a)}}\,e_{a,x}$$

Δείκτης Κατάστασης **της f(x) στο σημείο x_a. Εάν |δείκτης|=~1 ή <1, τότε το σφάλμα παραμένει περίπου σταθερό ή μειώνεται κατά τη μετάδοση Αλλιώς, το σφάλμα αυξάνεται.**

Άσκηση 2: Η θερμότητα που εκπέμπεται μέσω ακτινοβολίας από ένα σώμα δίνεται από τον τύπο: $q=e\sigma AT^4$, όπου e, σ είναι σταθερές, A η επιφάνειά του και T η θερμοκρασία του. Η T μετρήθηκε εργαστηριακά και βρέθηκε ίση με 1200±45K. Να βρεθεί το σχετικό σφάλμα υπολογισμού της θερμότητας.

Μετάδοση Σφάλματος σε Συναρτήσεις Πολλών Μεταβλητών

Εκτίμηση του Απόλυτου Σφάλματος:

$$E_{t,f(x_1,x_2,\ldots x_n)} \cong \frac{\partial f_a}{\partial x_1} E_{t,x_1} + \frac{\partial f_a}{\partial x_2} E_{t,x_2} + \ldots + \frac{\partial f_a}{\partial x_n} E_{t,x_n}$$

όπου $f_a = f(x_{1,a}, x_{2,a}, \ldots x_{n,a})$

Εκτίμηση του Σχετικού Σφάλματος:

$$\varepsilon_{a,f(x_1,x_2,\ldots x_n)} \cong \frac{\partial f_a}{\partial x_1}\frac{x_{1,a}}{f_a} \varepsilon_{a,x_1} + \frac{\partial f_a}{\partial x_2}\frac{x_{2,a}}{f_a} \varepsilon_{a,x_2} + \ldots + \frac{\partial f_a}{\partial x_n}\frac{x_{n,a}}{f_a} \varepsilon_{a,x_n}$$

Άσκηση 3 (Σεπτ. 2015): **Η ροπή αδράνειας (ως προς τον άξονα συμμετρίας) ομοιόμορφου και συμπαγούς κυλινδρικού σώματος μάζας M και ακτίνας R δίνεται από τη σχέση $I=MR^2/2$. Τα M, R μετρήθηκαν με σχετικό σφάλμα ±2% και ±1% αντίστοιχα. Ποιό το αντίστοιχο σχετικό σφάλμα της ροπής αδράνειας I.**

Μια ματιά στις 4 Βασικές Αριθμητικές Πράξεις

Πρόσθεση: $f(x,y) = x + y$

$$E_{t,(x+y)} \cong E_{t,x} + E_{t,y} \qquad \text{και} \qquad \varepsilon_{a,(x+y)} \cong \frac{x}{x+y}\varepsilon_{a,x} + \frac{y}{x+y}\varepsilon_{a,y}$$

Αφαίρεση: $f(x,y) = x - y$

$$E_{t,(x-y)} \cong E_{t,x} - E_{t,y} \qquad \text{και} \qquad \varepsilon_{a,(x-y)} \cong \frac{x}{x-y}\varepsilon_{a,x} - \frac{y}{x-y}\varepsilon_{a,y}$$

Πολλαπλασιασμός: $f(x,y) = x \cdot y$

$$E_{t,(x\cdot y)} \cong y \cdot E_{t,x} + x \cdot E_{t,y} \qquad \text{και} \qquad \varepsilon_{a,(x\cdot y)} \cong \varepsilon_{a,x} + \varepsilon_{a,y}$$

Διαίρεση: $f(x,y) = x/y$

$$E_{t,(x/y)} \cong \frac{1}{y}E_{t,x} - \frac{x}{y^2}E_{t,y} \qquad \text{και} \qquad \varepsilon_{a,(x/y)} \cong \varepsilon_{a,x} - \varepsilon_{a,y}$$

Σε όλες τις πράξεις, πλην αφαίρεσης, το μέγιστο σχετικό σφάλμα του αποτελέσματος θα είναι το πολύ ίσο με το άθροισμα των σφαλμάτων των δύο μεταβλητών:

$$\left|\varepsilon_{a,f(x,y)}\right| \leq \left|\varepsilon_{a,x}\right| + \left|\varepsilon_{a,y}\right|$$

Κανόνες Προστασίας από Αριθμητικά Σφάλματα

Το συνολικό αριθμητικό σφάλμα προκύπτει ως το άθροισμα του σφάλματος στρογγυλοποίησης και αποκοπής. Η προσπάθεια μείωσης του ενός εκ των δύο συνήθως προκαλεί αύξηση του άλλου. Στους σύγχρονους 32/64-bit υπολογιστές, το σφάλμα στρογγυλοποίησης είναι πολύ μικρότερο από το σφάλμα αποκοπής.

Κανόνες: αναδιάταξη των αριθμητικών παραστάσεων, ώστε:

- Να αποφεύγεται η αφαίρεση δύο παραπλήσιων αριθμών ή η πρόσθεση δύο αριθμών πολύ διαφορετικού μεγέθους.
- Να αθροίζονται πρώτα οι μικρότεροι και μετά οι μεγαλύτεροι όροι μιας παράστασης.
- Να μειώνεται κατά το δυνατόν ο αριθμός των πράξεων.