ΑΡΙΘΜΗΤΙΚΗ ΑΝΑΛΥΣΗ

Γ. Παπαευαγγέλου,
Εργαστηριακό Διδακτικό Προσωπικό
Τμήμα Αγρονόμων και Τοπογράφων Μηχανικών, Α.Π.Θ.
Γραφείο: 2ο Υπόγειο πτέρυγας
E-mail geopapaevan@topo.auth.gr

ΣΗΜΕΙΩΣΕΙΣ

Βιβλιογραφία

α. Βιβλιογραφία μαθήματος (Εύδοξος)

«Εισαγωγή στην Αριθμητική Ανάλυση» Ακρίβης, Γεώργιος Δ., Δουγαλής, Βασίλειος Α, ISBN 978-960-524-022-6, 4η έκδοση, έτος τρεχ. έκδοσης 2015, Εκδόσεις: ΙΤΕ-Πανεπιστημιακές Εκδόσεις Κρήτης
ή
«Αριθμητική ανάλυση. Μια αλγοριθμική προσέγγιση», Νικόλαος Μισυρλής, ISBN: 978-960-92031-2-8, ΕΚΔΟΣΕΙΣ Νικόλαος Μισυρλής, Έκδοση 1, 2009.

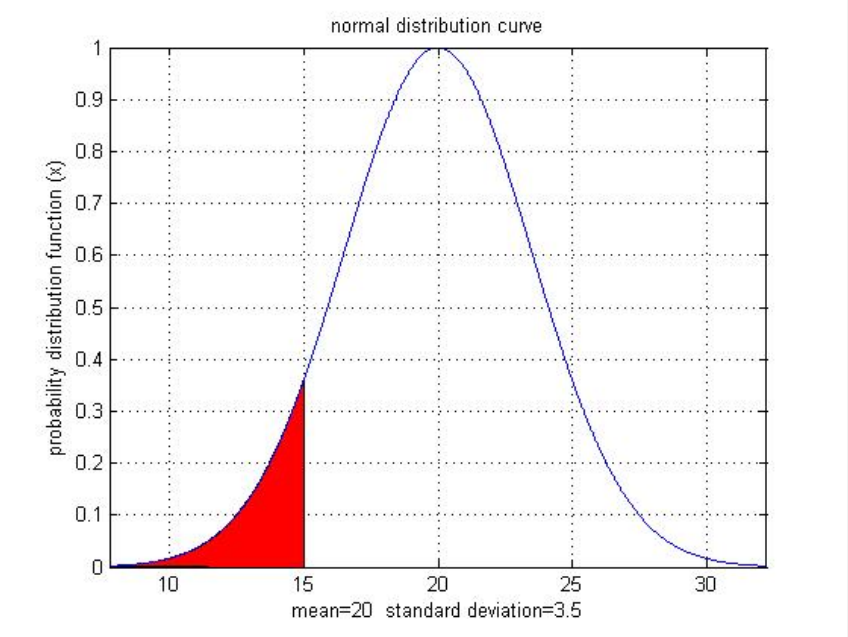
β. Επιπρόσθετη βιβλιογραφία για μελέτη

"Αριθμητική Ανάλυση με χρήση ΗΥ", σημειώσεις Ε. Σιδηρόπουλου, Χ. Φωτιάδη,Α.Π.Θ., 2016.

ΠΟΥ ΕΊΝΑΙ ΑΠΑΡΑΙΤΗΤΗ Η ΑΡΙΘΜΗΤΙΚΗ ΑΝΑΛΥΣΗ ?

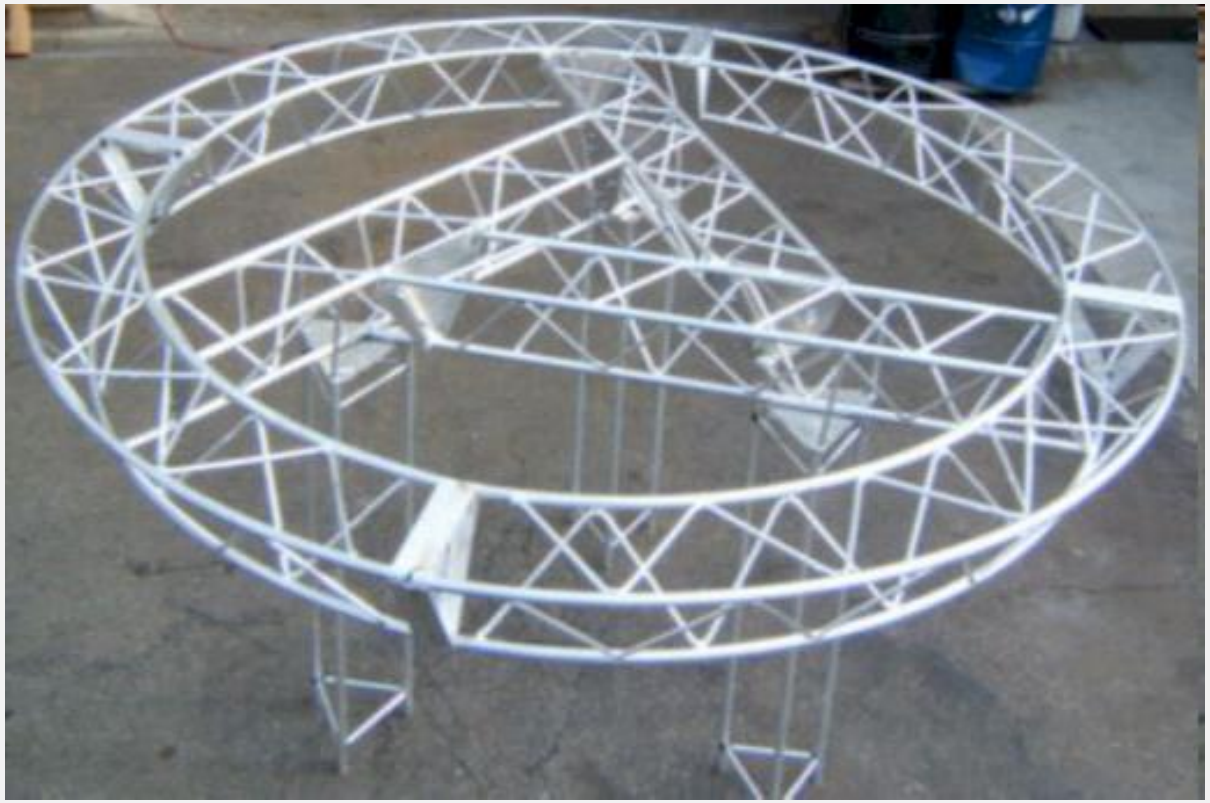
- Για την επίλυση προβλημάτων που δεν λύνονται αναλυτικά

$$\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{-\frac{u^2}{2}}\,du$$



ΠΟΥ Ε΄ΙΝΑΙ ΑΠΑΡΑΙΤΗΤΗ Η ΑΡΙΘΜΗΤΙΚΗ ΑΝΑΛΥΣΗ ?

- Για δισεπίλυτα προβλήματα

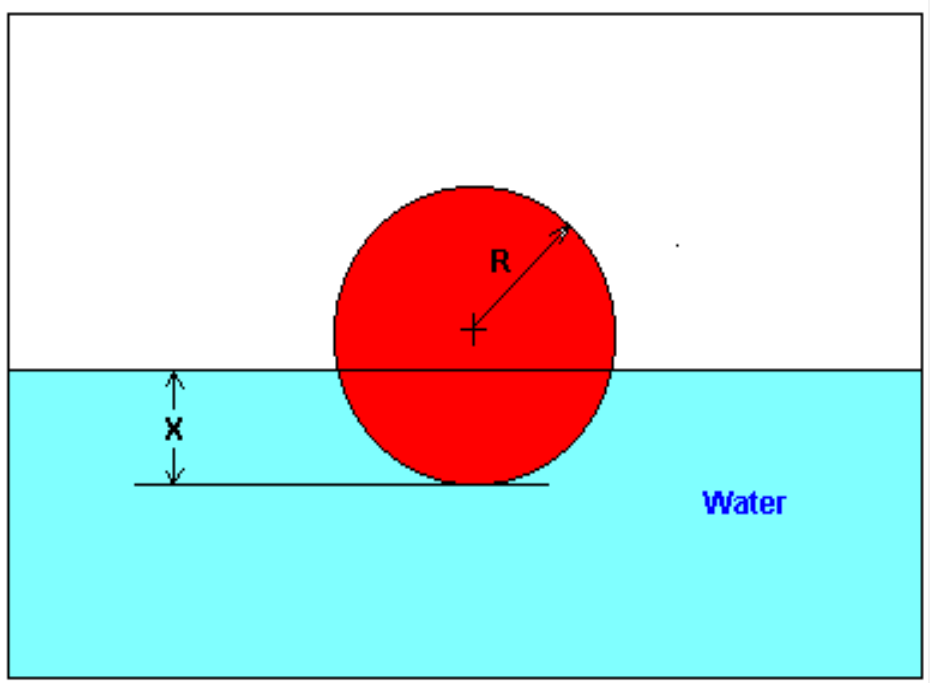


ΜΗ ΓΡΑΜΜΙΚΕΣ ΕΞΙΣΩΣΕΙΣ

Τι ποσοστό της επιπλέουσας σφαίρας είναι έξω από το νερό ?

Διάμετρος=0.11m

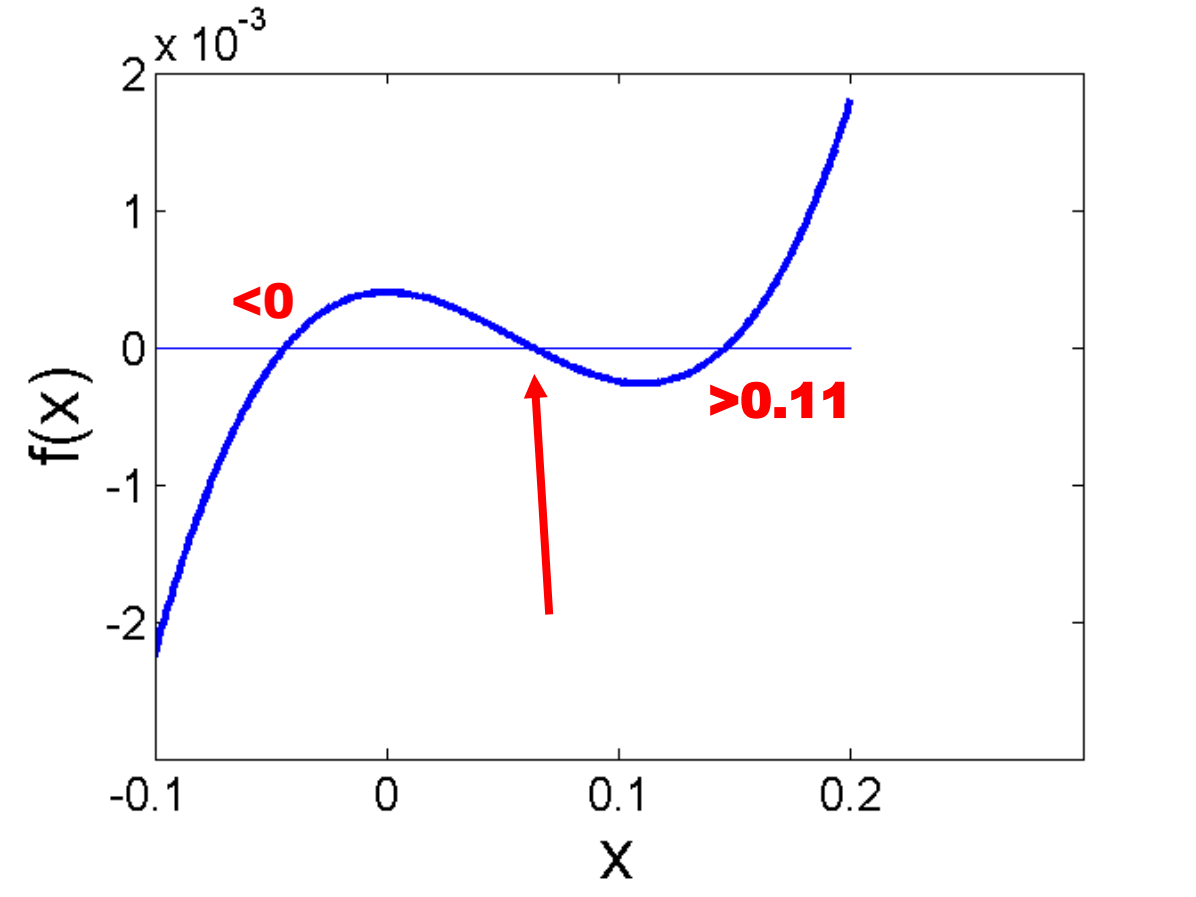
Ειδική βαρύτητα=0.6



$$x^3 - 0.165x^2 + 3.993\times10^{-4} = 0$$

ΜΗ ΓΡΑΜΜΙΚΕΣ ΕΞΙΣΩΣΕΙΣ

Τι ποσοστό της επιπλέουσας σφαίρας είναι έξω από το νερό ?

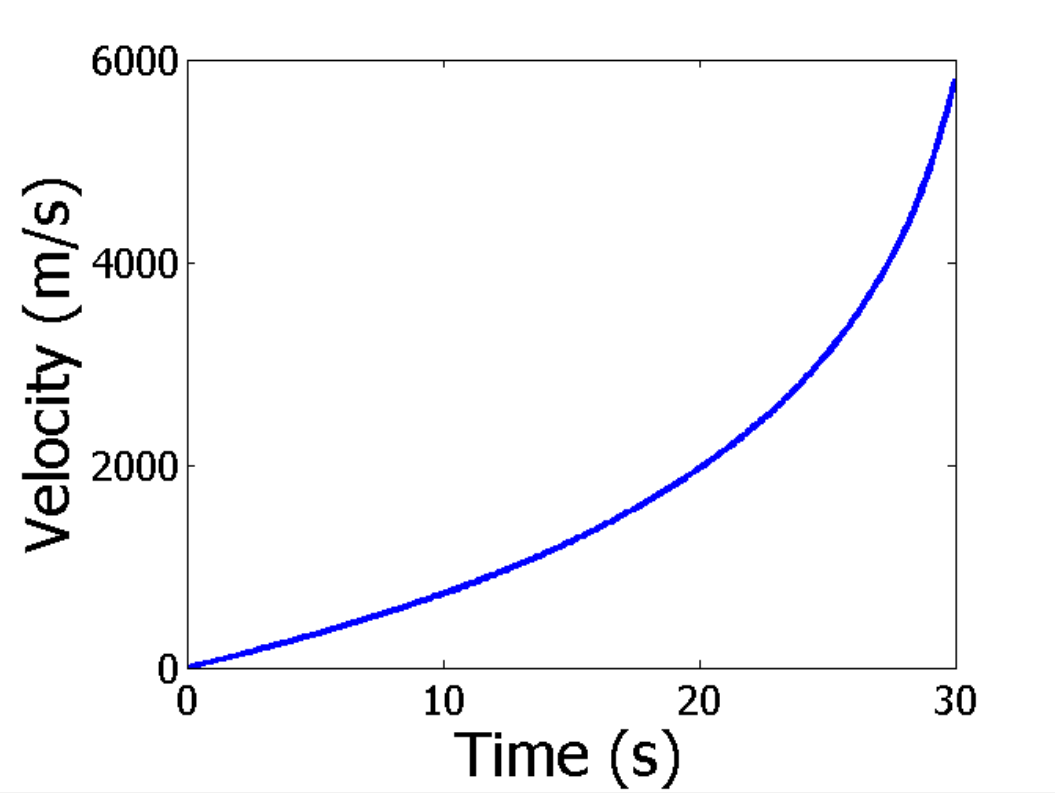


Το x, δεν μπορεί να είναι
Ούτε x>0.11, αλλά ούτε και x<0

$$f(x) = x^3 - 0.165x^2 + 3.993\times10^{-4} = 0$$

ΠΑΡΑΓΩΓΙΣΗ

Πόση είναι η επιτάχυνση σε t=7 s ?



$$v(t) = 2200\ \ln\left(\frac{16\times10^4}{16\times10^4 - 5000t}\right) - 9.8t$$

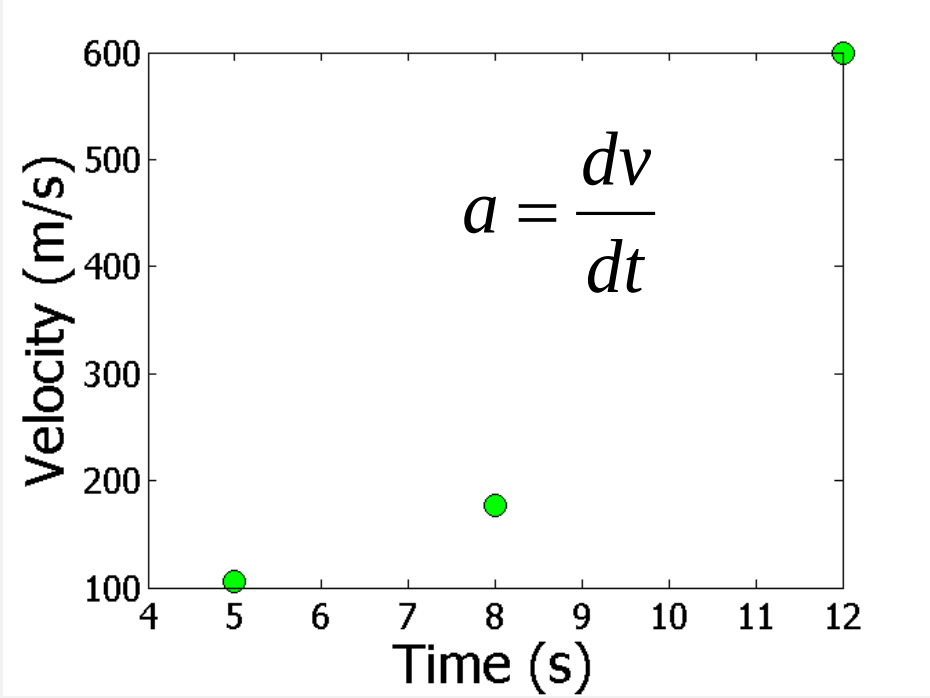
$$a = \frac{dv}{dt}$$

Δυσεπίλυτο αναλυτικά

ΠΑΡΑΓΩΓΙΣΗ

Πόση είναι η επιτάχυνση σε t=7 s ?

Time (s)	5	8	12
Vel (m/s)	106	177	600



Από μετρήσεις με παρεμβολή

ΣΥΣΤΗΜΑ ΓΡΑΜΜΙΚΩΝ ΕΞΙΣΩΣΕΩΝ

Για δεδομένο προφίλ:

Time (s)	5	8	12
Vel (m/s)	106	177	600

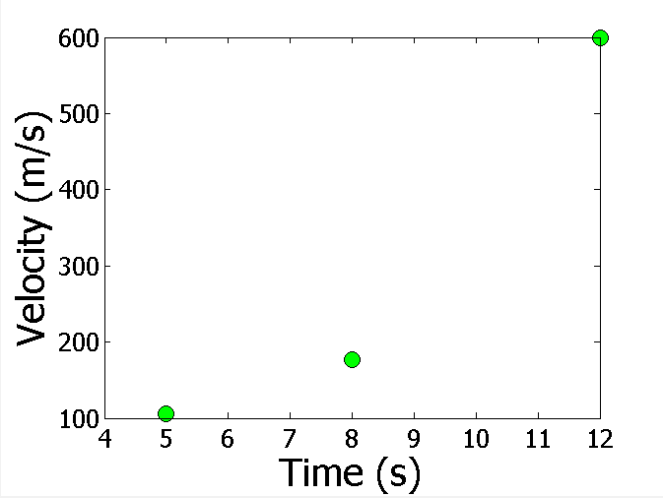
$$v(t) = at^2 + bt + c,\ 5 \le t \le 12$$

Τρείς γραμμικές εξισώσεις:

$$25a + 5b + c = 106$$

$$64a + 8b + c = 177$$

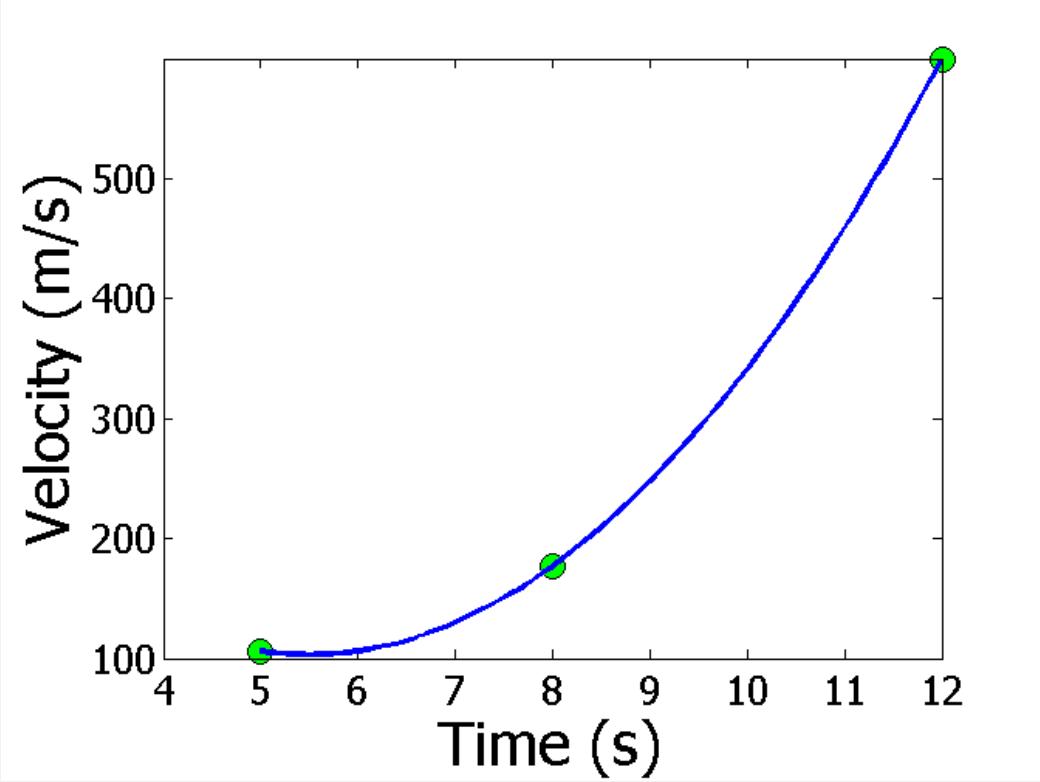
$$144a + 12b + c = 600$$



ΠΑΡΕΜΒΟΛΗ

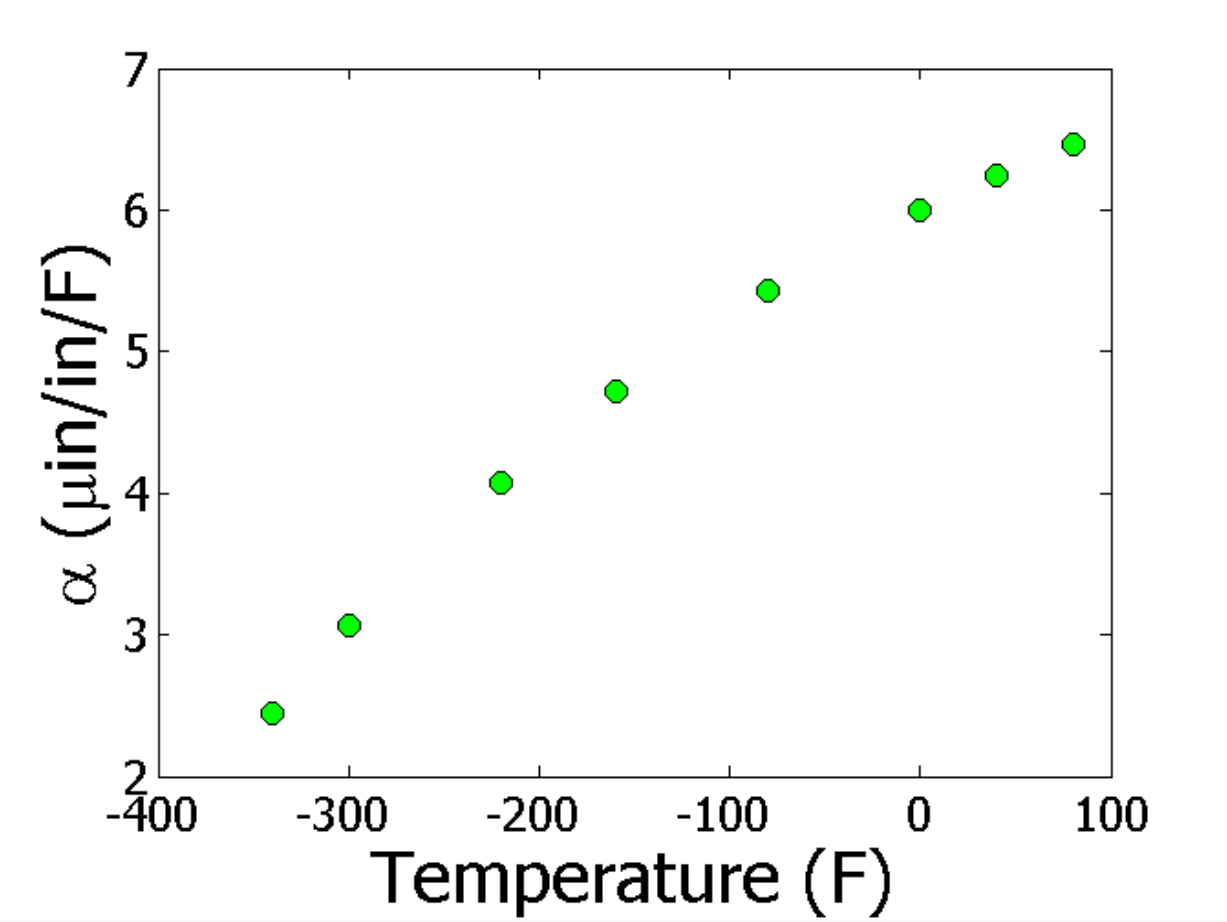
Πόση είναι η ταχύτητα σε t=7 seconds?

Time (s)	5	8	12
Vel (m/s)	106	177	600



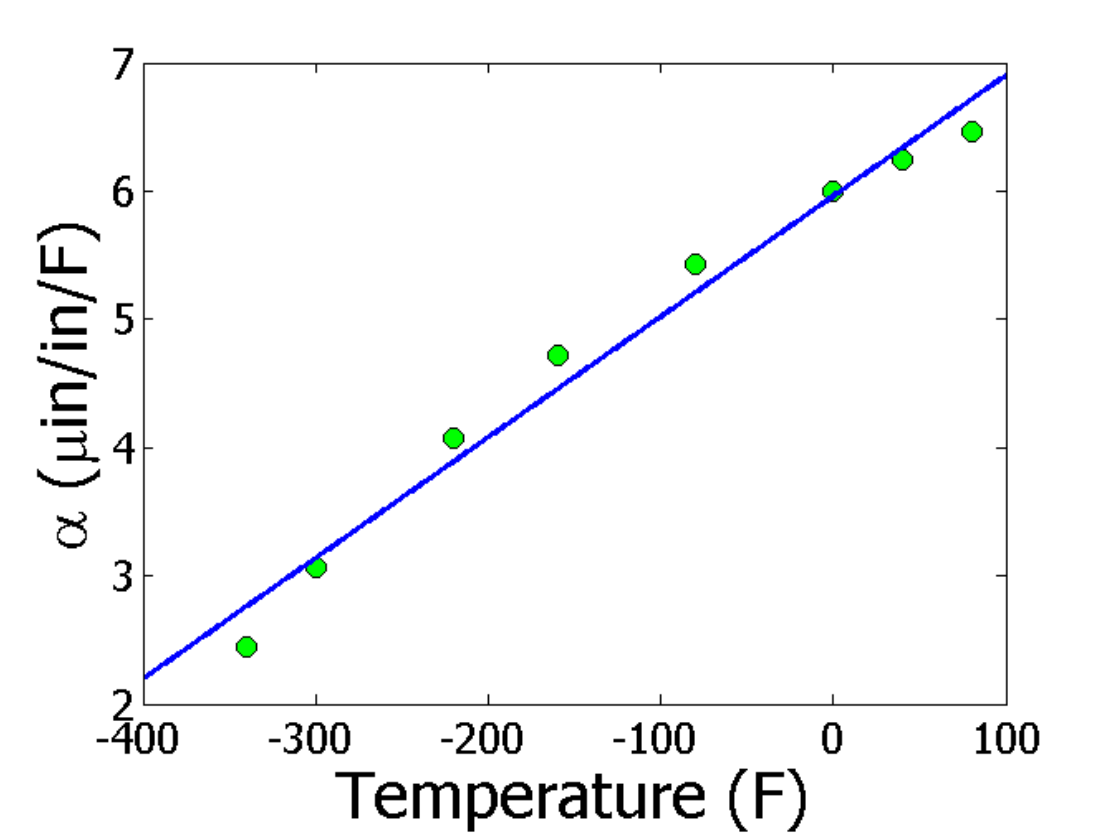
ΠΑΛΙΝΔΡΟΜΗΣΗ

Συντελεστής θερμικής διαστολής για χυτοσίδηρο:



Από τα δεδομένα αυτά, ποια σχέση μπορώ να βρώ για οποιαδήποτε τιμή της θερμοκρασίας?

ΠΑΛΙΝΔΡΟΜΗΣΗ (ΣΥΝΕΧΕΙΑ)



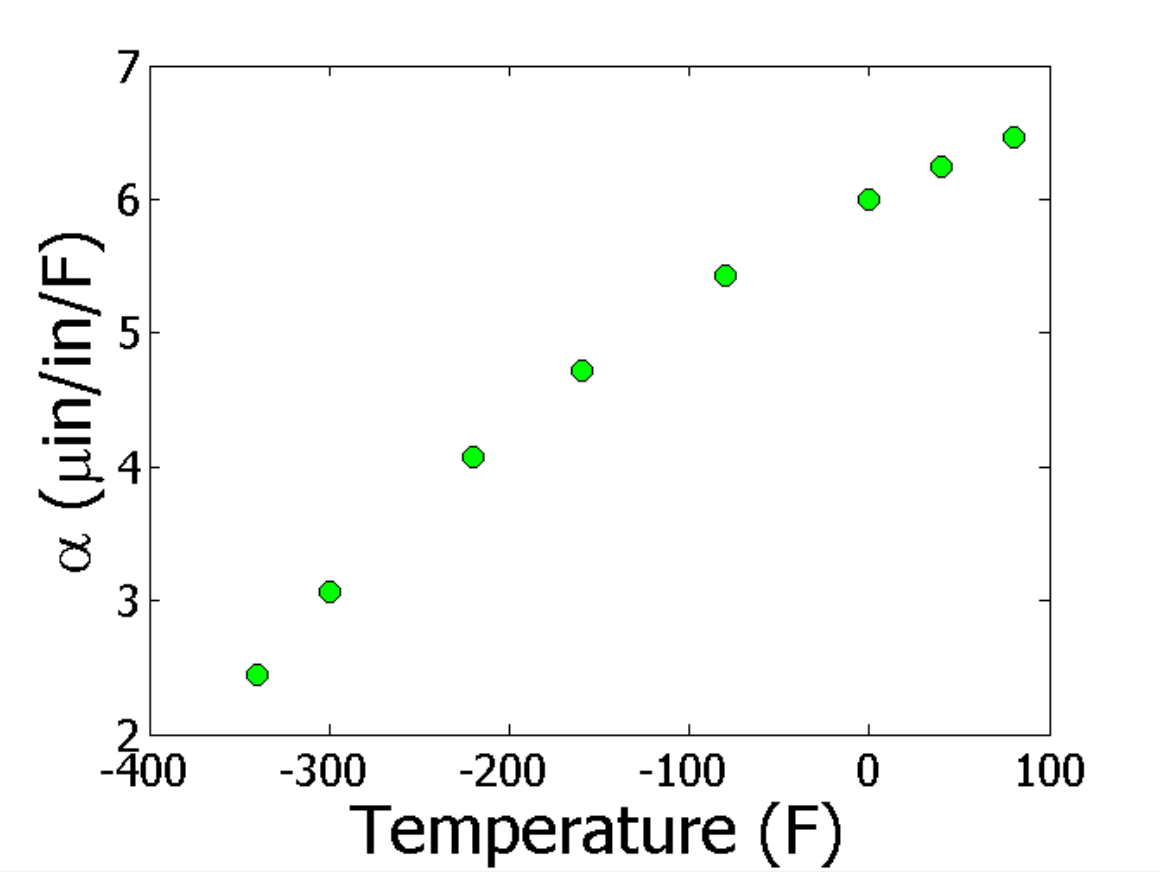
Γραμμική σχέση,
Πρώτη προσέγγιση

ΟΛΟΚΛΗΡΩΣΗ

Συστολή διαμέτρου χαλύβδινου άξονα, όταν βυθιστεί σε υγρό άζωτο

$$\Delta D = D \int_{T_{room}}^{T_{fluid}} \alpha \, dT$$

Η συνολική συστολή, βρίσκεται από το ολοκλήρωμα του συντελεστή για το φάσμα των θερμοκρασιών



ΣΥΝΗΘΕΙΣ ΔΙΑΦΟΡΙΚΕΣ ΕΞΙΣΩΣΕΙΣ

Πόση ώρα χρειάζεται για να μειωθεί η θερμοκρασία ενός στροφέα ρουλεμάν?

$$mc\frac{d\theta}{dt} = -hA(\theta - \theta_a), \quad \theta(0) = \theta_{room}$$

ΕΠΙΛΥΣΗ ΠΡΟΒΛΗΜΑΤΩΝ ΜΗΧΑΝΙΚΟΥ

Περιγραφή προβλήματος

↓

Μαθηματικό μοντέλο

↓

Όταν είναι δυσεπίλυτο αναλυτικά, λύνεται αριθμητικά

Επίλυση μοντέλου

↓

Χρήση της επίλυσης

ΜΑΘΗΜΑΤΙΚΕΣ ΔΙΑΔΙΚΑΣΙΕΣ

- Μη - γραμμικές εξισώσεις
- Διαφόριση
- Σύστημα γραμμικών εξισώσεων
- Προσαρμογή καμπύλης
  - Παρεμβολή
  - Παλινδρόμηση

  Μέθοδοι προσαρμογής μιάς καμπύλης σε σημειακά δεδομένα. Η παρεμβολή βρίσκει την τέλεια καμπύλη (πχ πολυώνυμο). Την παλινδρόμηση, την ενδιαφέρει η διακύμανση (variance) διαφόρων ειδών καμπυλών, και το γενικότερο φυσικό νόημα κάθε είδους
- Ολοκλήρωση
- Συνήθεις διαφορικές εξισώσεις
- Προχωρημένα μαθηματικά προβλήματα:
  - Μερικές διαφορικές εξισώσεις
  - Βελτιστοποίηση
  - Μετασχηματισμοί Fourrier

Τα ορισμένα ολοκληρώματα, οι διαφορικές εξισώσεις, οι μη γραμμικές αλγεβρικές εξισώσεις, τα γραμμικά συστήματα, οι ακολουθίες, οι σειρές είναι θέματα των καθαρών Μαθηματικών που οδηγούν σε υπολογιστικά προβλήματα. Η Αριθμητική Ανάλυση σκοπό έχει την εξεύρεση, διαμόρφωση και ανάλυση αριθμητικών λύσεων σε μαθηματικά προβλήματα.

Αυτό σημαίνει ότι η Αριθμητική Ανάλυση ανάγει τελικά τα προβλήματα που πραγματεύεται σε υπολογιστικές διαδικασίες. Αυτές αποτελούνται από σαφώς ορισμένα διαδοχικά βήματα και κάθε βήμα συνίσταται στην εκτέλεση μιας ή περισσοτέρων από τις τέσσερεις αριθμητικές πράξεις. Το

ΠΕΔΙΑ ΕΦΑΡΜΟΓΗΣ ΤΗΣ ΑΡΙΘΜΗΤΙΚΗΣ ΑΝΑΛΥΣΗΣ

- Αριθμητικές λύσεις – αποτελέσματα σε υπολογιστικά προβλήματα
- Όπου οι εξισώσεις δεν λύνονται αναλυτικά
- Σχετικά προβλήματα μηχανικού:
- Μηχανική των Ρευστών (meteorology, plastic flow, rheology…)
- Μηχανική των Στερεών (elasticity, fracturing…)
- Προβλήματα διάχυσης και συγκέντρωσης
- Εξισώσεις κυμάτων και άλλα προβλήματα μεταφοράς
- Κβαντομηχανική (Schrödinger's equation)
- Περίπου 1 δισ. ακόμα εφαρμογές

ΠΕΡΙΕΧΟΜΕΝΑ

- Εισαγωγή: σφάλματα, απώλεια σημαντικών ψηφίων, αστάθεια
- Επίλυση μη γραμμικών εξισώσεων (μέθοδος συναρτησιακών επαναλήψεων, μέθοδος Newton – Raphson, επιτάχυνση επαναλήψεων, μέθοδος διχοτόμησης)
- Υπολογισμός ριζών πολυωνύμων (τιμή πολυωνύμου και παραγώγου του, μέθοδος Newton – Raphson, μέθοδος Bernoulli, μιγαδικές ρίζες πολυωνύμων, εντοπισμός διαστημάτων ριζών)
- Επίλυση γραμμικών συστημάτων (απαλοιφή Gauss, σύστημα με τριδιαγωνικό πίνακα, επαναληπτικές μέθοδοι, νόρμες διανυσμάτων και πινάκων, ιδιοτιμές και ιδιοδιανύσματα, παθολογικές καταστάσεις στα γραμμικά συστήματα)
- Παρεμβολή (πολυώνυμο παρεμβολής, Lagrange, σφάλμα πολυωνύμου, επαναλαμβανόμενη γραμμική παρεμβολή, παρεμβολή κατά Newton, παρεμβολή splines)
- Αριθμητική ολοκλήρωση (στοιχειώδης, τραπεζοειδής, κατά Simpson)
- Αριθμητική επίλυση διαφορικών εξισώσεων (Euler, Runge – Kutta, πεπερασμένες διαφορές, Crank – Nicolson)

1ο ΜΑΘΗΜΑ

ΕΙΣΑΓΩΓΗ

1.1 ΑΠΟΛΥΤΟ ΚΑΙ ΣΧΕΤΙΚΟ ΣΦΑΛΜΑ. ΠΑΡΑΣΤΑΣΗ ΑΡΙΘΜΟΥ ΜΕ ΠΕΠΕΡΑΣΜΕΝΟ ΑΡΙΘΜΟ ΨΗΦΙΩΝ

1.2 ΑΠΩΛΕΙΑ ΣΗΜΑΝΤΙΚΩΝ ΨΗΦΙΩΝ. ΑΣΤΑΘΕΙΑ

Προσεγγιστικό αποτέλεσμα:
Πηγές σφαλμάτων

(α) Η παράσταση αριθμών, είτε γίνεται από τον άνθρωπο, είτε από μια μηχανή, περιλαμβάνει πάντα πεπερασμένο αριθμό ψηφίων. Η αναγκαστική παράλειψη ψηφίων εισάγει ένα σφάλμα, που είναι γνωστό ως **σφάλμα αποκοπής ή στρογγύλευσης.**

(β) Το σφάλμα αυτό μεταδίδεται στις αριθμητικές πράξεις που ακολουθούν. Επιπλέον, το αποτέλεσμα κάθε μιας από αυτές υπόκειται σε περιορισμό των ψηφίων του.

(γ) Πολλά προβλήματα παρουσιάζουν μια θεωρητικά ατέρμονη αλληλουχία επί μέρους βημάτων. Τέτοιο πρόβλημα είναι η εύρεση του αθροίσματος μιας συγκλίνουσας σειράς. Οι περιορισμένες δυνατότητες του ανθρώπου ή της μηχανής επιβάλλουν τη διακοπή της διαδικασίας έπειτα από έναν αριθμό βημάτων και την απόκτηση κάποιου αποτελέσματος προσεγγιστικού. Το σφάλμα που εισάγεται με αυτό τον τρόπο μπορεί να ονομαστεί **σφάλμα διακοπής**.

(δ) Οι πραγματικοί αριθμοί αποτελούν ένα συνεχές. Το συνεχές όμως αυτό δεν μπορεί να εισαχθεί με την ιδεατή μορφή του μέσα σε αριθμητικές διαδικασίες. Για παράδειγμα, το διάστημα [0, 1], ενώ περιλαμβάνει έναν άπειρο αριθμό σημείων, αναγκαστικά αντιπροσωπεύεται σε μία αριθμητική διαδικασία με ένα πεπερασμένο αριθμό εσωτερικών του σημείων. Μια τέτοια διαδικασία είναι η αριθμητική ολοκλήρωση, που θα εκτεθεί σε κατοπινό κεφάλαιο. Συνεπώς, για τις αριθμητικές ανάγκες, το συνεχές διακριτοποιείται και το σφάλμα που εισάγεται με αυτό τον τρόπο στους υπολογισμούς ονομάζεται **σφάλμα διακριτοποίησης**.

ΣΦΑΛΜΑΤΑ - ΚΑΤΗΓΟΡΙΕΣ

- Από χρήση πεπερασμένου αριθμού ψηφίων (**σφάλμα αποκοπής ή στρογγύλευσης)**
- Από **μετάδοση σφάλματος** μέσω αριθμητικών πράξεων
- Από τη διακοπή επαναληπτικής / προσεγγιστικής διαδικασίας υπολογισμών έπειτα από ένα αριθμό βημάτων (**σφάλμα διακοπής**).
- Από την διακριτοποίηση των πραγματικών αριθμών στις αριθμητικές διαδικασίες, ενώ είναι συνεχείς (**σφάλμα διακριτοποίησης**)

x: πραγματικός αριθμός, πχ $\sqrt{2}$
X: μία προσέγγισή του, πχ 1.41

ΠΡΑΓΜΑΤΙΚΟ ΣΦΑΛΜΑ ΚΑΙ ΣΧΕΤΙΚΟ ΣΦΑΛΜΑ

Παράδειγμα 1: Μέτρηση πραγματικής απόστασης 100m, εύρεση 99.9m :

Σφάλμα: $\Delta_x = 0.1m$ Σχετικό σφάλμα: $\delta_x = \frac{x-X}{x}$ = 0.001

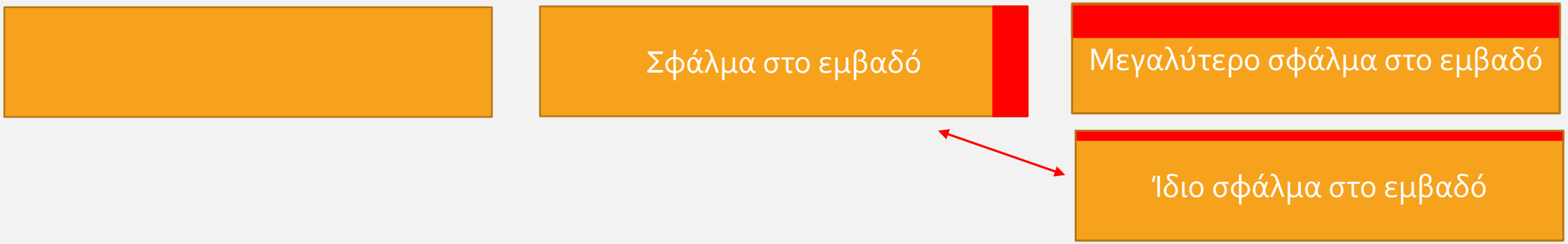
Μέτρηση πραγματικής απόστασης 10m, εύρεση 9.9m :

Σφάλμα: $\Delta_x = 0.1m$ Σχετικό σφάλμα: $\delta_x = \frac{x-X}{x}$ = 0.01

Για το ίδιο σχετικό σφάλμα, χρειαζόμαστε μεγαλύτερη ακρίβεια μέτρησης στα μικρότερα μεγέθη

Παράδειγμα 2: Μέτρηση δύο πλευρών ορθογωνίου:

Για το ίδιο πραγματικό σφάλμα στο εμβαδό, χρειαζόμαστε μεγαλύτερη ακρίβεια στη μέτρηση της μικρής διάστασης



ΓΕΝΙΚΗ ΠΑΡΑΣΤΑΣΗ ΠΡΑΓΜΑΤΙΚΟΥ ΑΡΙΘΜΟΥ

- Γενική παράσταση πραγματικού αριθμού στο δεκαδικό σύστημα:
- Παραδείγματα παράστασης

$$x = a_m.10^m + a_{m-1}.10^{m-1} + \ldots + a_0.10^0 + b_1.10^{-1} + b_2.10^{-2} + \ldots$$

όπου,

$a_m = 1, 2, \ldots, 9$

$a_i = 0, 1, 2, \ldots, 9$ για $i = 0, 1, 2, \ldots, m-1$

$b_i = 0, 1, 2, \ldots, 9$ για $i = 1, 2, 3, \ldots$

Παραδείγματα :

$\pi = 3.14159\ldots = \overset{a_0}{3}*10^0 + \overset{b_1}{1}*10^{-1} + \overset{b_2}{4}*10^{-2} + \ldots$

m=0, η μεγαλύτερη δύναμη του 10 είναι το 0

$\frac{100}{3} = 33.333\ldots = \overset{a_1}{3}*10^1 + \overset{a_0}{3}*10^0 + \overset{b_1}{3}*10^{-1} + \overset{b_2}{3}*10^{-2} + \ldots$

m=1, η μεγαλύτερη δύναμη του 10 είναι το 1

$\frac{401}{4} = 100.25 = \overset{a_2}{1}*10^2 + \overset{a_1}{0}*10^1 + \overset{a_0}{0}*10^0 + \overset{b_1}{2}*10^{-1} + \overset{b_2}{5}*10^{-2}.$

m=2, η μεγαλύτερη δύναμη του 10 είναι το 2

ΑΡΙΘΜΟΣ ΣΩΣΤΩΝ ΨΗΦΙΩΝ

Εάν x=32.26 και X=32.24, τα 3 πρώτα ψηφία είναι σωστά.
Γενικά, τα n πρώτα ψηφία είναι σωστά, εάν:

$$|x - X| \leq 0.5 * 10^{m-n+1}$$

Όπου m = ο μεγαλύτερος εκθέτης του 10 που υπάρχει στον αριθμό.

Τα σωστά ψηφία είναι τα: $a_m, a_{m-1}, \ldots, a_0, b_1, \ldots, b_{m-n-1}$ (το πρώτο είναι πάντα το a_m)

Επομένως, από τον m, και το m-n+1, μπορεί να βρεθεί ο αριθμός των σωστών ψηφίων n.

Παραδείγματα:

Πράγματι, για x = 32.36 και X = 32.34,

m = 1 και |x - X| = 0.02 < 0.05 = 0.05*10^{-1}.

Άρα m - n + 1 = -1 και n = 3.

m=1, και x-X=0.02<0.05=0.5 10^{-1}, άρα m-n+1=-1, και n=3, δηλαδή τα 3 πρώτα ψηφία είναι σωστά

Αν τώρα, x = 32.96 και X = 33.00,

|x - X| = 0.04 < 0.05 = 0.5*10^{-1} και m - n + 1 = -1, όπου m = 1.

Άρα, και πάλι n = 3.

m=1, και x-X=0.05=0.5 10^{-1}, άρα m-n+1=-1, και n=3, δηλαδή τα 3 πρώτα ψηφία είναι σωστά

Στις περιπτώσεις 2 και 3, δεν φαίνονται τα τρία πρώτα ψηφία σωστά αρχικά. Όμως με την στρογγύλευση, είναι!!

Για x = 10 και X = 9.995,

|x - X| = 0.005 = 0.5*10^{-2} και m - n + 1 = -2 με m = 0.

Άρα n = 3.

m=0, και x-X=0.005=0.5 10^{-2}, άρα m-n+1=-2, και n=3, δηλαδή τα 3 πρώτα ψηφία είναι σωστά

<u>ΣΤΗ ΣΥΝΕΧΕΙΑ:</u>

ΣΦΑΛΜΑΤΑ

ΣΗΜΑΝΤΙΚΑ ΨΗΦΙΑ

ΚΙΝΗΤΗ ΥΠΟΔΙΑΣΤΟΛΗ

ΚΑΝΟΝΙΚΟΠΟΙΗΣΗ ΑΡΙΘΜΟΥ

ΣΥΝΟΛΟ ΑΡΙΘΜΩΝ ΕΝΟΣ ΥΠΟΛΟΓΙΣΤΗ

ΑΣΤΑΘΕΙΑ

ΒΑΘΜΟΣ ΚΑΤΑΣΤΑΣΗΣ

ΣΗΜΑΝΤΙΚΟ ΨΗΦΙΟ – ΚΙΝΗΤΗ ΥΠΟΔΙΑΣΤΟΛΗ - ΚΑΝΟΝΙΚΟΠΟΙΗΣΗ

- Σημαντικά ψηφία
- Κινητή υποδιαστολή (floating point)
- Κανονικοποίηση (normalization)

Σε μία δεκαδική παράσταση, **σημαντικό** ψηφίο είναι κάθε ψηφίο διάφορο του 0. Αν κάποιο ψηφίο είναι 0, τότε θεωρείται σημαντικό, όταν βρίσκεται ανάμεσα σε ψηφία διάφορα από το 0. Αν βρίσκεται στο τέλος μιας πεπερασμένης δεκαδικής παράστασης, τότε θεωρείται σημαντικό, εάν ως δεκαδικό ψηφίο είναι σωστό με την έννοια του παραπάνω ορισμού. Για παράδειγμα, στη δεκαδική παράσταση 0.0020080 τα δύο πρώτα μηδενικά μετά την υποδιαστολή δεν είναι σημαντικά, το 2 και το 8 είναι σημαντικά όπως επίσης σημαντικά είναι τα δύο μηδενικά ανάμεσα στο 2 και στο 8. Αν το τελευταίο μηδενικό δεν είναι σημαντικό, τότε δεν πρέπει να αναγράφεται : 0.002008.

Αν οι αριθμητικές πράξεις εκτελούνται ομοιόμορφα με αριθμούς **κινητής υποδιαστολής**, τότε ο παραπάνω αριθμός γράφεται $0.20080*10^{-2}$, έτσι ώστε το πρώτο σημαντικό ψηφίο να βρίσκεται αμέσως μετά την υποδιαστολή. Ας σημειωθεί ακόμη ότι η παραμονή στην παραπάνω παράσταση του μηδενικού μετά το 8 υποδηλώνει πως αυτό θεωρείται σημαντικό. Στις περισσότερες υπολογιστικές μηχανές ακόμα και αριθμοί μεγαλύτεροι από τη μονάδα μετατρέπονται στην παραπάνω μορφή, π.χ. $625.122 = 0.625122*10^{3}$. Με τον τρόπο αυτό κάθε αριθμός μπορεί να θεωρηθεί ότι αναλύεται σε δύο παράγοντες, ένα κλάσμα και μία δύναμη του 10. Η αναγωγή αυτή των αριθμών ονομάζεται **κανονικοποίηση**.

Ερώτηση: γιατί τα 2 πρώτα «0» μετά την υποδιαστολή δεν είναι σημαντικά??
(Διότι μπορούν να παραληφθούν εάν παραστήσουμε τον αριθμό με μορφή αριθμού κινητής υποδιαστολής, ή καλύτερα, κανονικοποιημένο)

ΠΡΩΤΑ ΣΗΜΑΝΤΙΚΑ ΨΗΦΙΑ – ΑΠΟΚΟΠΗ ΚΑΙ ΣΤΡΟΓΓΥΛΕΥΣΗ

(α) Αν το n+1 ψηφίο είναι μικρότερο από το 5, τότε απλώς παραλείπονται όλα τα ψηφία από n+1 και πέρα.

(β) Αν το n+1 ψηφίο είναι μεγαλύτερο από 5, τότε προστίθεται μία μονάδα στο n ψηφίο και παραλείπονται όλα τα ψηφία από το n+1 και πέρα.

(γ) Αν το n+1 ψηφίο είναι το 5 και υπάρχουν μη μηδενικά ψηφία μετά από αυτό, τότε εφαρμόζεται η παραπάνω περίπτωση (β).

(δ) Αν το n+1 ψηφίο είναι το 5 και ακολουθείται από μηδενικά, τότε εξετάζεται το n-οστό ψηφίο. Αν αυτό είναι άρτιο εφαρμόζεται η περίπτωση (β).

Για παράδειγμα, ο π = 3.1415926535. . . μπορεί να αποκοπεί ή να στρογγυλευτεί, όπως παρακάτω :

αποκοπή :	3.14	3.141	3.1415	3.14159
στρογγύλευση :	3.14	3.142	3.1416	3.14159

Διαφορά αποκοπής από στρογγύλευση

- Ο αριθμός των ψηφίων είναι πεπερασμένος
- Εάν n=τα πρώτα σημαντικά ψηφία, παραλείπονται αυτά από το n+1 και πέρα (αποκοπή)
- Εφαρμόζεται στρογγύλευση
- Παράδειγμα στρογγύλευσης

Είναι εύκολο να αποδειχτεί ότι, στην περίπτωση της αποκοπής, $|\Delta| \leq 10^{-t}$, όπου t η τελευταία παραμένουσα δεκαδική θέση. Για τη στρογγύλευση, $|\Delta| \leq 0.5*10^{-t}$.

ΚΑΝΟΝΙΚΟΠΟΙΗΜΕΝΗ (NORMALIZED) ΠΑΡΑΣΤΑΣΗ ΑΡΙΘΜΟΥ

Το 10 είναι η βάση των υπολογισμών για τον άνθρωπο.
Στις υπολογιστικές μηχανές είναι το 2, 8, 16, … γιατί ταιριάζει καλύτερα στη δομή τους.
Χρησιμοποιείται η κανονικοποιημένη μορφή με σύστημα κινητής υποδιαστολής. Ο πραγματικός x αναπαρίσταται σαν:

$$fl(x) = \pm(d_1.\beta^{-1} + d_2.\beta^{-2} + \ldots + d_n.\beta^{-n}).\beta^{e},$$

όπου β είναι η βάση και d_i $(i = 1, 2, \ldots, n)$ τα ψηφία. Για τα d_i ισχύει :

$$1 \le d_1 \le \beta-1, \qquad 0 \le d_i \le \beta-1 \text{ για } i = 2, 3, \ldots, n.$$

Ο πεπερασμένος n των πρώτων σημαντικών ψηφίων είναι χαρακτηριστικό της μηχανής. Ο εκθέτης e είναι ακέραιος και κυμαίνεται ανάμεσα στην κατώτερη τιμή –L και στην ανώτερη U:

$$-L \le e \le U$$

Τα L και U εκφράζουν χαρακτηριστικά της μηχανής

ΠΑΡΑΔΕΙΓΜΑΤΑ ΚΑΝΟΝΙΚΟΠΟΙΗΜΕΝΟΥ ΑΡΙΘΜΟΥ

$$fl(x) = \pm(d_1.\beta^{-1} + d_2.\beta^{-2} + \ldots + d_n.\beta^{-n}).\beta^{e},$$

όπου β είναι η βάση και d_i $(i = 1, 2, \ldots, n)$ τα ψηφία. Για τα d_i ισχύει :

$$1 \le d_1 \le \beta-1, \qquad 0 \le d_i \le \beta-1 \text{ για } i = 2, 3, \ldots, n.$$

Στο δεκαδικό σύστημα:

$$fl(x)= 53.241 = .53241*10^2 = (5*10^{-1}+3*10^{-2}+2*10^{-3}+4*10^{-4}+1*10^{-5})*10^2$$

Στο δυαδικό σύστημα:

$$fl(x)= (11010)_2 = (.11010)_2*2^5 = (1*2^{-1}+1*2^{-2}+0*2^{-3}+1*2^{-4}+0*10^{-5})*2^5 = (0.5+0.25+0+0.0625+0)*32 = 26$$

πράγματι, όπως θα δούμε: $(11010)_2 = 0*2^0+1*2^1+0*2^2+1*2^3+1*2^4 = 0+2+0+8+16 = 26$

ΑΡΙΘΜΟΣ ΜΕΛΩΝ ΤΟΥ F (ΣΥΝΟΛΟΥ ΑΡΙΘΜΩΝ ΠΟΥ ΜΠΟΡΟΥΝ ΝΑ ΑΝΑΠΑΡΑΣΤΑΘΟΥΝ)

$$fl(x) = \pm(d_1.\beta^{-1} + d_2.\beta^{-2} + \ldots + d_n.\beta^{-n}).\beta^e ,$$

όπου β είναι η βάση και d_i (i = 1, 2, . . . , n) τα ψηφία. Για τα d_i ισχύει :

$$1 \le d_1 \le \beta-1, \qquad 0 \le d_i \le \beta-1 \text{ για } i = 2, 3, \ldots, n.$$

Ασκηση. Να αποδειχτεί ότι ο αριθμός των μελών του συνόλου **F** είναι ίσος με $2.(\beta-1).\beta^{n-1}.(U-L+1)+1$.

Λόγω των δύο προσήμων (όσοι θετικοί, άλλοι τόσοι αρνητικοί)

Οι συνδυασμοί του πρώτου ψηφίου (πχ στο δυαδικό, τα ψηφία είναι 0 και 1, άρα το πρώτο μπορεί να είναι μόνο 1)

Οι συνδυασμοί των υπόλοιπων ψηφίων, τα οποία είναι (n-1), και το καθένα παίρνει β τιμές

Οι συνδυασμοί των τιμών του e

- Ο αριθμός των μελών του F, είναι ο αριθμός όλων των αριθμών που μπορούν να αναπαρασταθούν από έναν υπολογιστή, με βάση το β, με αριθμό σημαντικών ψηφίων n, και με εκθέτη e που παίρνει τιμές από –L έως U.

ΑΣΚΗΣΗ: Να βρεθεί ο αριθμός των μελών του F αν είναι β=2, n=52, L=2048 και U=-2048

β=	2
n=	52
U=	2048
L=	-2048
F=	1.85E+19

ΜΕΓΙΣΤΗ ΤΙΜΗ ΣΦΑΛΜΑΤΟΣ ΑΠΟΚΟΠΗΣ

Για κάθε πραγματικό αριθμό x βρίσκεται ένας αριθμός $fl(x) \in F$, με αποκοπή ή στρογγύλευση. Προκύπτει συνεπώς η ανάγκη να εκτιμηθεί το σφάλμα που υπεισέρχεται στην προσέγγιση αυτή.

Ο πραγματικός αριθμός x μπορεί να γραφτεί ως εξής, με βάση το β :

$$x = \pm(d_1.\beta^{-1} + d_2.\beta^{-2} + \ldots + d_n.\beta^{-n} + d_{n+1}.\beta^{-n-1} + d_{n+2}.\beta^{-n-2} + \ldots).\beta^e.$$

Τότε για την περίπτωση της αποκοπής,

Τα φηφία που παραλείπονται και δημιουργείται το σφάλμα Δx

$$|\Delta_x| = |x - fl(x)| = (d_{n+1}.\beta^{-n-1} + d_{n+2}.\beta^{-n-2} + \ldots).\beta^e =$$

$$= \beta^{-n}.\underbrace{(d_{n+1}.\beta^{-1} + d_{n+2}.\beta^{-2} + \ldots)}_{<1}.\beta^e \leq \beta^{-n}.\beta^e .$$

Αλλά $|\Delta_x| = |x|.|\delta_x|$.

Σχετικό σφάλμα: $\left|\delta_x\right| \leq \dfrac{\beta^{-n}\beta^e}{|x|}$

Άρα,

Από την παράσταση του x,

$$|x| = \beta^{-1}.(d_1 + d_2.\beta^{-1} + \ldots + d_n.\beta^{1-n} + d_{n+1}.\beta^{-n} + d_{n+2}.\beta^{-1-n} + \ldots).\beta^e$$

$$> d_1.\beta^{-1}.\beta^e$$

Άρα, $\dfrac{1}{|x|} < \dfrac{\beta}{d_1.\beta^e}$ και $\boxed{|\delta_x| < \dfrac{1}{d_1}\beta^{1-n}}$

Συνεπώς,

Το σχετικό σφάλμα στην αναπαράσταση κινητής υποδιαστολής πραγματικού αριθμού x είναι μικρότερο σε απόλυτη τιμή από $|\delta_x| \leq \frac{\beta^{1-n}}{d_1}$ και επειδή $d_1 \geq 1$ είναι $|\delta_x| \leq \beta^{1-n}$ πχ 10^{-n}

ΜΕΓΙΣΤΗ ΤΙΜΗ ΣΦΑΛΜΑΤΟΣ ΑΠΟΚΟΠΗΣ - ΠΑΡΑΔΕΙΓΜΑ

$$x = \pm(d_1.\beta^{-1} + d_2.\beta^{-2} + \ldots + d_n.\beta^{-n} + \underbrace{d_{n+1}.\beta^{-n-1} + d_{n+2}.\beta^{-n-2} + \ldots}).\beta^{e}.$$

Τα ψηφία που παραλείπονται και δημιουργείται το σφάλμα Δx

Σχετικό σφάλμα: $\boxed{|\delta_x| < \frac{1}{d_1}\beta^{1-n}}$

Άρα, το σχετικό σφάλμα είναι τόσο μικρότερο, όσο μεγαλύτερο είναι το πρώτο ψηφίο του αριθμού στην αποκοπή

Ποιό είναι το σχετικό σφάλμα του 432.56847 εάν αποκόψουμε (παραλείψουμε) τα 3 τελευταία του ψηφία?

Ποιο είναι εάν επρόκειτο για τον 432.56999?

ΥΠΟΛΟΓΙΣΤΕ ΤΟ ΣΑΝ ΑΣΚΗΣΗ!

Το σχετικό σφάλμα στην αναπαράσταση κινητής υποδιαστολής πραγματικού αριθμού **x** είναι μικρότερο σε απόλυτη τιμή από $|\delta_x| \leq \frac{\beta^{1-n}}{d_1}$ και επειδή $d_1 \geq 1$ είναι $|\delta_x| \leq \beta^{1-n}$ πχ 10^{-n}

ΜΕΓΙΣΤΗ ΤΙΜΗ ΣΦΑΛΜΑΤΟΣ ΑΠΟΚΟΠΗΣ - ΠΑΡΑΔΕΙΓΜΑ

$$x = \pm(d_1.\beta^{-1} + d_2.\beta^{-2} + \ldots + d_n.\beta^{-n} + \underbrace{d_{n+1}.\beta^{-n-1} + d_{n+2}.\beta^{-n-2} + \ldots}_{\text{Τα ψηφία που παραλείπονται και δημιουργείται το σφάλμα } \Delta x}).\beta^e.$$

Σχετικό σφάλμα: $\boxed{|\delta_x| < \frac{1}{d_1}\beta^{1-n}}$

Ποιό είναι το σχετικό σφάλμα του 432.56847 εάν αποκόψουμε (παραλείψουμε) τα 3 τελευταία του ψηφία?

Γίνεται 432.56, που σημαίνει σφάλμα 0.00847, και σχετικό σφάλμα 0.00847/432.56847 = 1.96 10^{-5}

Είναι d_1=4, β=10, n=5, άρα το σχετικό σφάλμα θα είναι μικρότερο από $10^{-4}/4$ = 0.000025 = 2.5 10^{-5}

Εάν επρόκειτο για τον 432.56999, το σχετικό σφάλμα θα ήταν: 2.31 10^{-5}

Το σχετικό σφάλμα στην αναπαράσταση κινητής υποδιαστολής πραγματικού αριθμού **x** είναι μικρότερο σε απόλυτη τιμή από $|\delta_x| \leq \frac{\beta^{1-n}}{d_1}$ και επειδή $d_1 \geq 1$ είναι $|\delta_x| \leq \beta^{1-n}$ πχ 10^{-n}

ΑΠΩΛΕΙΑ ΣΗΜΑΝΤΙΚΩΝ ΨΗΦΙΩΝ. ΑΣΤΑΘΕΙΑ: ΔΙΑΦΟΡΕΣ ΑΡΙΘΜΩΝ

Τα σφάλματα που οφείλονται στην πεπερασμένη αναπαράσταση των πραγματικών αριθμών μεταδίδονται στις αριθμητικές πράξεις. Ιδιαίτερη προσοχή απαιτεί η πράξη της αφαίρεσης, όταν αφαιρούνται αριθμοί που δεν διαφέρουν πολύ μεταξύ τους. Έστω ότι ζητείται η διαφορά μεταξύ των αριθμών x = 48.128 και y = 48.107, που υποτίθενται γνωστοί με πέντε σωστά σημαντικά ψηφία. Αυτό σημαίνει ότι $|\Delta_x| \le 0.5 * 10^{+1-5+1} = 0.5 * 10^{-3}$ και $|\Delta_y| \le 0.5 * 10^{-3}$.

Αν z = x - y, τότε προφανώς $|\Delta_z| \le |\Delta_x| + |\Delta_y| \le 10^{-3}$ ή $|\Delta_z| \le 0.001$.

υπενθυμίζεται ότι, εάν έχουμε αριθμό με n σωστά ψηφία και m την μεγαλύτερη δύναμη του 10 σ'αυτόν, το μέγιστο σφάλμα του είναι: $|x - X| \le 0.5 * 10^{m-n+1}$

Γράφουμε τους x και y με την κανονικοποιημένη μορφή και αφαιρούμε :

$$z = 0.48128 * 10^2 - 0.48107 * 10^2 = 0.00021 * 10^2 = 0.21 * 10^{-1}.$$

Συνεπώς το z έχει μόνο δύο σημαντικά ψηφία από τα οποία το δεύτερο είναι αμφίβολης ορθότητας, επειδή $|\Delta_z|$ είναι της τάξης του $0.01*10^{-1}$.

Αν τώρα παρατηρήσουμε τα επί μέρους σχετικά σφάλματα,

$$|\delta_x| \le \frac{0.5*10^{-3}}{0.48128*10^2} \approx 10^{-5}$$

$$|\delta_y| \le \frac{0.5*10^{-3}}{0.48107*10^2} \approx 10^{-5}$$

$$|\delta_z| \le \frac{10^{-3}}{0.21*10^{-1}} \approx 0.05,$$

το $|\delta_z|$ φαίνεται αμέσως ότι είναι αρκετές τάξεις μεγέθους μεγαλύτερο.

Επειδή η διαφορά τους είναι μικρή, και βρίσκεται στον παρονομαστή του μέγιστου πιθανού σφάλματος, η τελευταία είναι μεγάλη

ΑΚΡΙΒΕΙΑ ΣΤΙΣ ΔΙΑΦΟΡΕΣ ΑΡΙΘΜΩΝ

Ανάλογο φαινόμενο στον υπολογισμό της παρακάτω παράστασης, όπου e μικρός αριθμός:

$$\sqrt{x+\varepsilon}-\sqrt{x}$$

Στην περίπτωση αυτή, το μειονέκτημα της αφαίρεσης μπορεί να διορθωθεί αν η παράσταση μετασχηματιστεί:

$$\sqrt{x+\varepsilon}-\sqrt{x}=\frac{\varepsilon}{\sqrt{x+\varepsilon}+\sqrt{x}}$$

Στη μετασχηματισμένη παράσταση δεν συμβαίνει απώλεια σημαντικών ψηφίων. Το ίδιο ισχύει και στην f(x) = 1-cosx για μικρές τιμές του x. Η συνάρτηση αυτή μετασχηματίζεται ως εξής:

$$f(x)=1-\cos x=\frac{1-cos^2x}{1+\cos x}=\frac{sin^2x}{\cos x+1}$$

Δηλαδή, και στις δύο περιπτώσεις, μετασχηματίζουμε τη συνάρτηση σε λόγο, αντί διαφορά, ώστε να αποφευχθεί η απώλεια των σημαντικών ψηφίων

Τυπογραφικά λάθη στο βιβλίο: + αντί – και 1+cosx αντί cosx

ΑΣΚΗΣΕΙΣ

- Παρακαλώ να υπολογίσετε την τιμή της συνάρτησης

$$\sqrt{x+\varepsilon}-\sqrt{x}=\frac{\varepsilon}{\sqrt{x+\varepsilon}+\sqrt{x}}$$

για x = 2 και ε = 0.01 και με τις δύο μορφές της

με τρία δεκαδικά ψηφία, και να συγκρίνετε τα σχετικά σφάλματα των δύο μορφών

(διαφορά και λόγος). Ως πραγματική τιμή να ληφθεί η τιμή με όλα τα διαθέσιμα ψηφία.

(η λύση στην επόμενη διαφάνεια)

$$f(x)=1-\cos x=\frac{1-cos^2x}{1+\cos x}=\frac{sin^2x}{\cos x+1}$$

- Παρακαλώ να υπολογίσετε την τιμή της συνάρτησης

για x = 0.05 και 0.02 και με τις δύο μορφές της

με τρία δεκαδικά ψηφία, και να συγκρίνετε τα σχετικά σφάλματα των δύο μορφών

(διαφορά και λόγος). Ως πραγματική τιμή να ληφθεί η τιμή με όλα τα διαθέσιμα ψηφία.

ΣΦΆΛΜΑΤΑ ΜΕΤΑΣ

στο excel, καμμία διαφορά μεταξύ των δύο, λόγω χρήσης πολλών δεκαδικών		όταν όμως χρησιμοποιήσουμε μόνο τρία δεκαδικ					
Διαφορά	Λόγος	x	ε	sqrt(x+ε)	sqrt(x)	Διαφορά	Λόγος
0.03492411	0.034924112	2	0.1	1.449	1.414	0.035	0.0349284
0.00353113	0.003531126	2	0.01	1.418	1.414	0.004	0.0035311
0.00035351	0.000353509	2	0.001	1.415	1.414	0.001	0.0003535
0.02224998	0.022249981	5	0.1	2.258	2.236	0.022	0.0222519
0.00223495	0.002234951	5	0.01	2.238	2.236	0.002	0.0022351
0.0002236	0.000223596	5	0.001	2.236	2.236	0	0.0002236
		x (rad)	cosx (3 ψηφ)	sinx (3ψηφ)		Διαφορά	Λόγος
5E-05	4.99996E-05	0.01	1.000	0.010		0	0.00005
0.00019999	0.000199993	0.02	1.000	0.020		0	0.0002
0.00044997	0.000449966	0.03	1.000	0.030		0	0.00045
0.00079989	0.000799893	0.04	0.999	0.040		0.001	0.0008004
0.00124974	0.00124974	0.05	0.999	0.050		0.001	0.0012506
0.00179946	0.00179946	0.06	0.998	0.060		0.002	0.0018018

$$\sqrt{x+\varepsilon}-\sqrt{x}=\frac{\varepsilon}{\sqrt{x+\varepsilon}-\sqrt{x}}.$$

$$f(x)=1-\cos x=\frac{1-\cos^2 x}{1+\cos x}=\frac{\sin^2 x}{\cos x}$$

0.01	0.009999833	0.99995
0.04	0.039989334	0.999200107
0.05	0.0499792	0.9987503
0.06	0.059964	0.9982005

Σφάλματα στις δύο συναρτήσεις: μικρότερα όταν μετασχηματίζονται από διαφορά σε λόγο

$$\sqrt{x+\varepsilon}-\sqrt{x} = \frac{\varepsilon}{\sqrt{x+\varepsilon}+\sqrt{x}}$$

Διαφορά	Λόγος	x	ε	sqrt(x+ε)	sqrt(x)	Διαφορά	Λόγος
0.03492411	0.034924112	2	0.1	1.449	1.414	0.035	0.0349284
0.00353113	0.003531126	2	0.01	1.418	1.414	0.004	0.0035311
0.00035351	0.000353509	2	0.001	1.415	1.414	0.001	0.0003535
0.02224998	0.022249981	5	0.1	2.258	2.236	0.022	0.0222519
0.00223495	0.002234951	5	0.01	2.238	2.236	0.002	0.0022351
0.0002236	0.000223596	5	0.001	2.236	2.236	0	0.0002236

$$f(x) = 1-\cos x = \frac{1-cos^2 x}{1+\cos x} = \frac{sin^2 x}{\cos x+1}$$

		x (rad)	cosx (3 ψηφ)	sinx (3ψηφ)		Διαφορά	Λόγος
5E-05	4.99996E-05	0.01	1.000	0.010		0	0.00005
0.00019999	0.000199993	0.02	1.000	0.020		0	0.0002
0.00044997	0.000449966	0.03	1.000	0.030		0	0.00045
0.00079989	0.000799893	0.04	0.999	0.040		0.001	0.0008004
0.00124974	0.00124974	0.05	0.999	0.050		0.001	0.0012506
0.00179946	0.00179946	0.06	0.998	0.060		0.002	0.0018018

ΑΚΡΙΒΕΙΑ ΣΤΙΣ ΔΙΑΦΟΡΕΣ ΑΡΙΘΜΩΝ

Μια γενικότερη μέθοδος για τον χειρισμό της f(x) στηρίζεται στο ανάπτυγμα Taylor του cos x :

$$\cos x = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \ldots$$

Χαρακτηριστικό παράδειγμα για την χρήση των σειρών Taylor προσφέρει η συνάρτηση $\dfrac{\tan x - \sin x}{x^3}$

Το θεώρημα της μέσης τιμής μπορεί επίσης να χρησιμοποιηθεί σε πολλές περιπτώσεις, π.χ. αν το x είναι μικρό,

$$\ln(a+x) - \ln a = \frac{x}{a+\theta} \approx \frac{x}{a+\frac{x}{2}} \quad \text{όπου } 0 < \theta < x.$$

Ο μετασχηματισμός των παραπάνω παραστάσεων προτείνεται σαν άσκηση :

$\dfrac{1}{\sqrt{x+1}} - \dfrac{1}{x}$, $\cos(x+\varepsilon) - \cos x$, $\sqrt[3]{x+1} - \sqrt[3]{x}$,

$\dfrac{1}{\sqrt{x+1}} - \dfrac{1}{\sqrt{x}}$, $(x+\varepsilon)^{2/3} - x^{2/3}$, $\tan(x+\varepsilon) - \tan x$,

$1 - e^{-x}$, $\ln(1 - x)$.

Οι παραπάνω περιπτώσεις χαρακτηρίζονται από το γεγονός ότι μικρό σφάλμα στην ανεξάρτητη μεταβλητή επιφέρει πολύ μεγαλύτερο σφάλμα στον υπολογισμό της συνάρτησης. Εμφανίζεται δηλαδή αυξημένη "ευαισθησία" μιας συνάρτησης f(x) ως προς μεταβολές του x.

ΒΑΘΜΟΣ ΚΑΤΑΣΤΑΣΗΣ ΣΥΝΑΡΤΗΣΗΣ

Η απόλυτη τιμή του λόγου του σχετικού σφάλματος συνάρτησης προς το σχετικό σφάλμα της μεταβλητής της, δίνει ένα ποσοτικό κριτήριο της ευαισθησίας της συνάρτησης σε μια περιοχή τιμών της μεταβλητής. Ονομάζεται βαθμός κατάστασης, και ορίζεται (με **x** την τιμή της μεταβλητής και $\bar{x}$ μια προσεγγιστική της τιμή):

$$\boxed{\beta\kappa = \frac{\left|\frac{f(x) - f(\bar{x})}{f(x)}\right|}{\left|\frac{x - \bar{x}}{x}\right|} := \left|\frac{f'(x) * x}{f(x)}\right|}$$

Για παράδειγμα, έστω η συνάρτηση: f(x)=1-cosx για πολύ μικρές τιμές του x. Στην περίπτωση αυτή:

$$\left|\frac{f'(x) * x}{f(x)}\right| = \left|\frac{x * \sin x}{1 - \cos x}\right| = \left|\frac{(1 + \cos x)\sin x}{1 - cos^2 x} * x\right| = \left|\frac{(1 + \cos x) * x}{sin \quad x}\right| := \frac{1 + \cos x}{x}$$

Για μικρές τιμές του **x** ο β.κ. μπορεί να πάρει πολύ μεγάλες. Σαν άσκηση να βρεθεί ο β.κ. που αντιστοιχεί στην μετασχηματισμένη παράσταση:

$$f(x) = \frac{sin^2 x}{1 + \cos x}$$

Ένα άλλο παράδειγμα, είναι η: $f(x) = \frac{1}{1 - x^2}$ όπου: $\left|\frac{f'(x) * x}{f(x)}\right| = \left|\frac{2x^2}{1 - x^2}\right|$

που πράγματι αυξάνει απεριόριστα όταν το x βρίσκεται στην περιοχή του 1.

ΑΣΚΗΣΗ

- Βρείτε τον βαθμό κατάστασης της συνάρτησης **y = sqrt(x)**.

(η λύση στην επόμενη διαφάνεια)

- Ποια σημασία έχει το αποτέλεσμα?

ΑΣΤΑΘΕΙΑ

Απεναντίας, η $f(x)=\sqrt{x}$ παρουσιάζει αρκετά μικρή τιμή του β.κ. :

$$\beta\kappa = \frac{\left|\frac{f(x)-f(\bar{x})}{f(x)}\right|}{\left|\frac{x-\bar{x}}{x}\right|} := \left|\frac{f'(x)*x}{f(x)}\right| = \left|-\frac{x}{2\sqrt{x}\sqrt{x}}\right| = \frac{1}{2}$$

β.κ. σταθερός, δεν μπορεί να πάρει μεγάλες τιμές για οποιοδήποτε **x**, και άρα δεν υπάρχει αστάθεια στη συνάρτηση αυτή!

Μια αριθμητική διαδικασία, ένας αλγόριθμος, αποτελείται από επί μέρους υπολογιστικά βήματα. Αν ένα ή περισσότερα από αυτά τα βήματα συνίστανται στον υπολογισμό μιας συνάρτησης με μεγάλο β.κ., τότε τα αριθμητικά σφάλματα, που υπεισέρχονται στον υπολογισμό της συνάρτησης αυτής, διαδίδονται και μεγεθύνονται ανεξέλεγκτα στην πορεία του αλγορίθμου.

Το φαινόμενο αυτό ονομάζεται **αστάθεια** και αναφέρεται στον αλγόριθμο συνολικά και όχι σε έναν περιορισμένο αριθμητικό υπολογισμό.

ΠΑΡΑΔΕΙΓΜΑΤΑ ΣΦΑΛΜΑΤΩΝ - ΑΣΚΗΣΕΙΣ

ΠΡΑΓΜΑΤΙΚΟ ΣΦΑΛΜΑ

- Η διαφορά μεταξύ της πραγματικής τιμής ενός μεγέθους και της προσεγγιστικής τιμής που έχει βρεθεί από μία αριθμητική μέθοδο

Πραγματικό σφάλμα = πραγματική τιμή – προσεγγιστική τιμή

ΠΑΡΑΔΕΙΓΜΑ – ΠΡΑΓΜΑΤΙΚΟ ΣΦΑΛΜΑ

Η παράγωγος, $f'(x)$ συνάρτησης $f(x)$ μπορεί να προσεγγιστεί από την εξίσωση,

$$f'(x) \approx \frac{f(x+h) - f(x)}{h}$$

για $f(x) = 7e^{0.5x}$ και $h = 0.3$

a) βρείτε την προσεγγιστική τιμή $f'(2)$

b) Την πραγματική τιμή $f'(2)$

c) Το πραγματικό σφάλμα της (a)

ΛΥΣΗ

$$f'(x) \approx \frac{f(x+h)-f(x)}{h}$$ για $f(x) = 7e^{0.5x}$ και $h = 0.3$

a) για $x = 2$ και $h = 0.3$

$$f'(2) \approx \frac{f(2+0.3)-f(2)}{0.3}$$

$$= \frac{f(2.3)-f(2)}{0.3}$$

$$= \frac{7e^{0.5(2.3)} - 7e^{0.5(2)}}{0.3}$$

$$= \frac{22.107 - 19.028}{0.3} = 10.263$$

b) Η ακριβής τιμή της $f'(2)$ μπορεί να βρεθεί ως εξής:

$$f(x) = 7e^{0.5x}$$

$$f'(x) = 7 \times 0.5 \times e^{0.5x}$$

$$= 3.5e^{0.5x}$$

Ετσι:

$$f'(2) = 3.5e^{0.5(2)} = 9.5140$$

Πραγματικό σφάλμα = πραγματική τιμή – προσεγγιστική τιμή:

$$E_t = = 9.5140 - 10.263 = -0.722$$

ΣΧΕΤΙΚΟ ΠΡΑΓΜΑΤΙΚΟ ΣΦΑΛΜΑ

- Λόγος πραγματικού σφάλματος ως προς την πραγματική τιμή:

$$\text{Σχετικό πραγματικό σφάλμα } (\in_t) = \frac{\text{Πραγματικό σφάλμα}}{\text{Πραγματική τιμή}}$$

ΠΑΡΑΔΕΙΓΜΑ ΣΧΕΤΙΚΟΥ ΠΡΑΓΜΑΤΙΚΟΥ ΣΦΑΛΜΑΤΟΣ

Ακολουθώντας το προηγούμενο παράδειγμα:

Να βρεθεί το σχετικό σφάλμα $f(x) = 7e^{0.5x}$ στη θέση $f'(2)$ με $h = 0.3$

Από το προηγούμενο παράδειγμα,

$$E_t = -0.722$$

Το σχετικό πραγματικό σφάλμα είναι $\in_t = \dfrac{\text{True Error}}{\text{True Value}}$

$$= \frac{-0.722}{9.5140} = -0.075888$$

η ποσοστιαία:

$$\in_t = -0.075888 \times 100\% = -7.5888\%$$

ΣΦΑΛΜΑ ΠΡΟΣΕΓΓΙΣΤΙΚΗΣ ΔΙΑΔΙΚΑΣΙΑΣ (ΔΙΑΡΚΩΣ ΜΕΤΑΒΑΛΛΟΜΕΝΟ)

- Σε προσεγγιστική διαδικασία διαδοχικών επαναλήψεων

Σφάλμα (E_a) = τρέχουσα προσέγγιση – προηγούμενη προσέγγιση

ΠΑΡΑΔΕΙΓΜΑ—ΣΦΑΛΜΑ ΠΡΟΣΕΓΓΙΣΗΣ

Για $f(x) = 7e^{0.5x}$ και $x = 2$ βρείτε

a) $f'(2)$ με $h = 0.3$

b) $f'(2)$ με $h = 0.15$

c) το σφάλμα προσέγγισης της τιμής $f'(2)$ στο (b)

λύση: a) για $x = 2$ και $h = 0.3$

$$f'(2) \approx \frac{f(2+0.3) - f(2)}{0.3} = \frac{f(2.3) - f(2)}{0.3} = 10.263$$

b) για $x = 2$ και $h = 0.15$

(Μικραίνοντας το h, προσεγγίζουμε καλύτερα την τιμή)

$$f'(2) \approx \frac{f(2+0.15) - f(2)}{0.15} = \frac{f(2.15) - f(2)}{0.15} = \frac{7e^{0.5(2.15)} - 7e^{0.5(2)}}{0.15}$$

$$= \frac{20.50 - 19.028}{0.15} = 9.8800$$

c) Το σφάλμα προσέγγισης E_a είναι

$E_a =$ Τρέχουσα προσέγγιση – προηγούμενη προσέγγ. $= 9.8800 - 10.263 = -0.38300$

Σε τί μπορεί να είναι ιδιαίτερα χρήσιμο το σφάλμα προσέγγισης?

(όταν η διαφορά διαδοχικών προσεγγίσεων μπεί κάτω από κάποιο όριο, μπορεί να σταματήσει η διαδικασία. Ιδιαίτερα χρήσιμο σε αυτόματες διαδικασίες)

ΣΧΕΤΙΚΟ ΠΡΟΣΕΓΓΙΣΤΙΚΟ ΣΦΑΛΜΑ

- Ο λόγος του σφάλματος προσέγγισης προς την τρέχουσα προσέγγιση

$$\text{Σχετικό προσεγγιστικό σφαλμα } (\in_a) = \frac{\text{Προσεγγιστικό σφάλμα}}{\text{Τρέχουσα προσεγγιστική τιμή}}$$

Από το προηγούμενο παράδειγμα, να βρεθεί το σχετικό σφάλμα προσέγγισης

$$f(x) = 7e^{0.5x} \qquad x = 2 \qquad h = 0.3 \qquad h = 0.15$$

Η προσεγγιστική τιμή είναι: $f'(2) = 10.263$ με $h = 0.3$ και $f'(2) = 9.8800$ με $h = 0.15$

$$E_a \equiv 9.8800 - 10.263 = -0.38300$$

$$\in_a = \frac{-0.38300}{9.8800} = -0.038765$$ και ποσοστιαία: $\in_a = -0.038765 \times 100\% = -3.8765\%$

Η απόλυτη τιμή : $|\in_a| = |-0.038765| = 0.038765 \text{ or } 3.8765\%$

Σε τί μπορεί να είναι ιδιαίτερα χρήσιμο το σχετικό σφάλμα προσέγγισης?

ΠΩΣ ΧΡΗΣΙΜΟΠΟΙΕΙΤΑΙ ΤΟ ΣΦΑΛΜΑ ΣΑΝ ΚΡΙΤΗΡΙΟ ΓΙΑ ΝΑ ΣΤΑΜΑΤΗΣΕΙ ΜΙΑ ΠΡΟΣΕΓΓΙΣΤΙΚΗ ΔΙΑΔΙΚΑΣΙΑ ΥΠΟΛΟΓΙΣΜΟΥ?

Εαν $|\in_a| \leq \in_s$ όπου $\in_s$ η απαιτούμενη ακρίβεια δεν χρειάζονται επιπλέον προσεγγίσεις

Εάν τουλάχιστον m σημαντικά ψηφία απαιτείται να είναι σωστά στην τελική τιμή, τότε, το ποσοστιαίο σχετικό προσεγγιστικό σφάλμα πρέπει να είναι μικρότερο από:

$$|\in_a| \leq 0.5 \times 10^{2-m}\%$$

Παράδειγμα: απαιτούνται 3 σημαντικά ψηφία, άρα το ποσοστιαίο αυτό σφάλμα πρέπει να είναι μικρότερο από 0.05%, ή 0.0005 (και άρα το τρίτο ψηφίο είναι εξασφαλισμένα ακριβές)

ΠΙΝΑΚΑΣ ΤΙΜΩΝ ΑΠΌ ΤΟ ΠΡΟΗΓΟΥΜΕΝΟ ΠΑΡΑΔΕΙΓΜΑ

για $f(x) = 7e^{0.5x}$ και $x = 2$ με διάφορα βήματα h

h	$f'(2)$	$\lvert\in_a\rvert$	m
0.3	10.263	---	0
0.15	9.8800	3.877%	1
0.10	9.7558	1.273%	1
0.01	9.5378	2.285%	1
0.001	9.5164	0.2249%	2

Σχετικό προσεγγιστικό σφάλμα

Αριθμός σωστών σημαντικών ψηφίων

ΣΦΆΛΜΑ ΣΤΡΟΓΓΎΛΕΥΣΗΣ

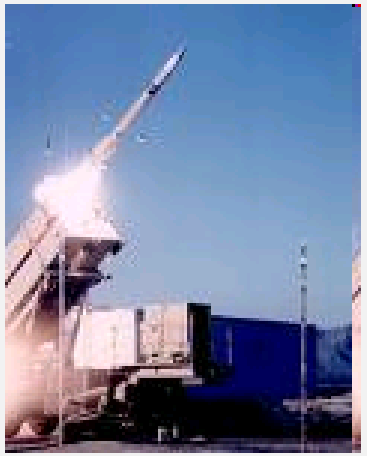
- Δημιουργείται από προσεγγιστική αναπαράσταση αριθμού

$$\frac{1}{3} \cong 0.333333$$

$$\sqrt{2} \cong 1.4142...$$

ΠΑΡΑΔΕΙΓΜΑ ΠΡΟΒΛΗΜΑΤΟΣ ΑΠΌ ΣΦΑΛΜΑ ΣΤΡΟΓΓΥΛΕΥΣΗΣ

- Αδυναμία αναχαίτισης πυραύλου λόγω σφάλματος στρογγύλευσης του κύκλου του ωρολογίου του αυστήματος αναχαίτισης:



- Ο κύκλος ωρολογίου 1/10 seconds αναπαρίσταται από 24-bit register δημιουργώντας σφάλμα 9.5 x 10^{-8} sec ανά sec.
- Η μπαταρία λειτουργούσε επί 100 συνεχόμενες ώρες, δημιουργώντας σφάλμα:

$$= 9.5 \times 10^{-8} \frac{\text{s}}{\text{s}} \times 100\text{hr} \times \frac{3600\text{s}}{1\text{hr}} = 0.342s$$

- Αποτέλεσμα: σφάλμα στη θέση του πυραύλου: 687 m.
- Ο στόχος θεωρείται εκτός βολής σε απόσταση μεγαλύτερη από 137 m.

ΜΕΤΑΦΟΡΑ ΣΗΜΑΝΤΙΚΩΝ ΨΗΦΙΩΝ ΣΤΟΥΣ ΥΠΟΛΟΓΙΣΜΟΥΣ

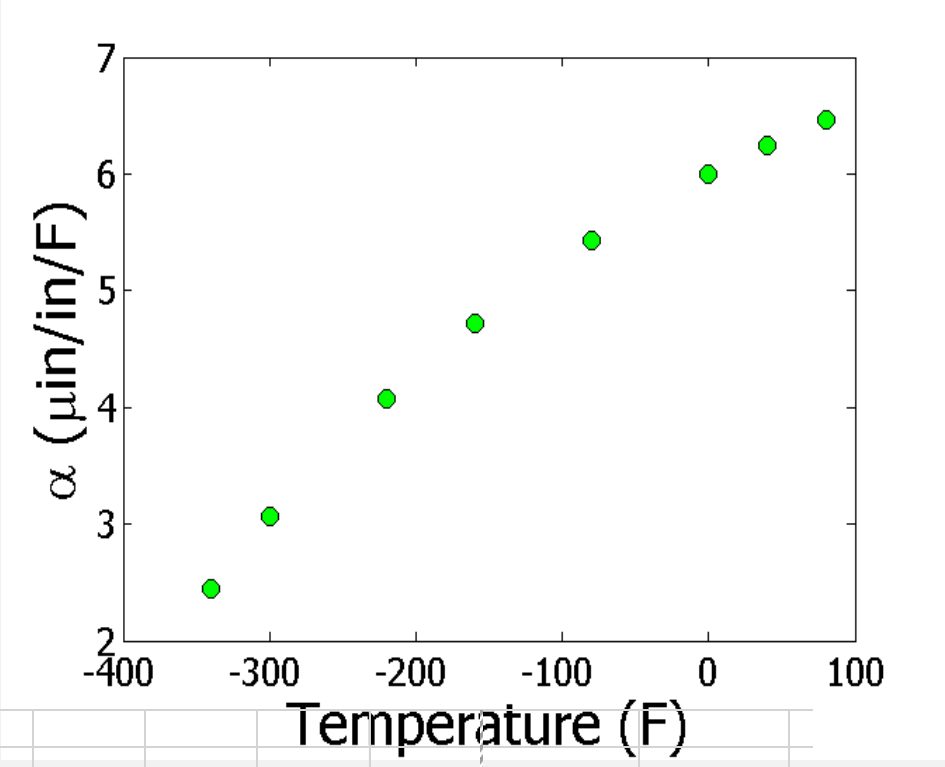
ΕΥΡΕΣΗ ΤΗΣ ΣΥΣΤΟΛΗΣ ΤΗΣ ΔΙΑΜΕΤΡΟΥ

$$\Delta D = D\int_{T_a}^{T_c} \alpha(T)\,dT$$

T_a=80°F; T_c=-108°F; D=12.363”

$$\alpha = a_0 + a_1 T + a_2 T^2$$

ΣΥΝΤΕΛΕΣΤΗΣ ΘΕΡΜΙΚΗΣ ΔΙΑΣΤΟΛΗΣ ΣΕ ΣΥΝΑΡΤΗΣΗ ΜΕ ΤΗ ΘΕΡΜΟΚΡΑΣΙΑ

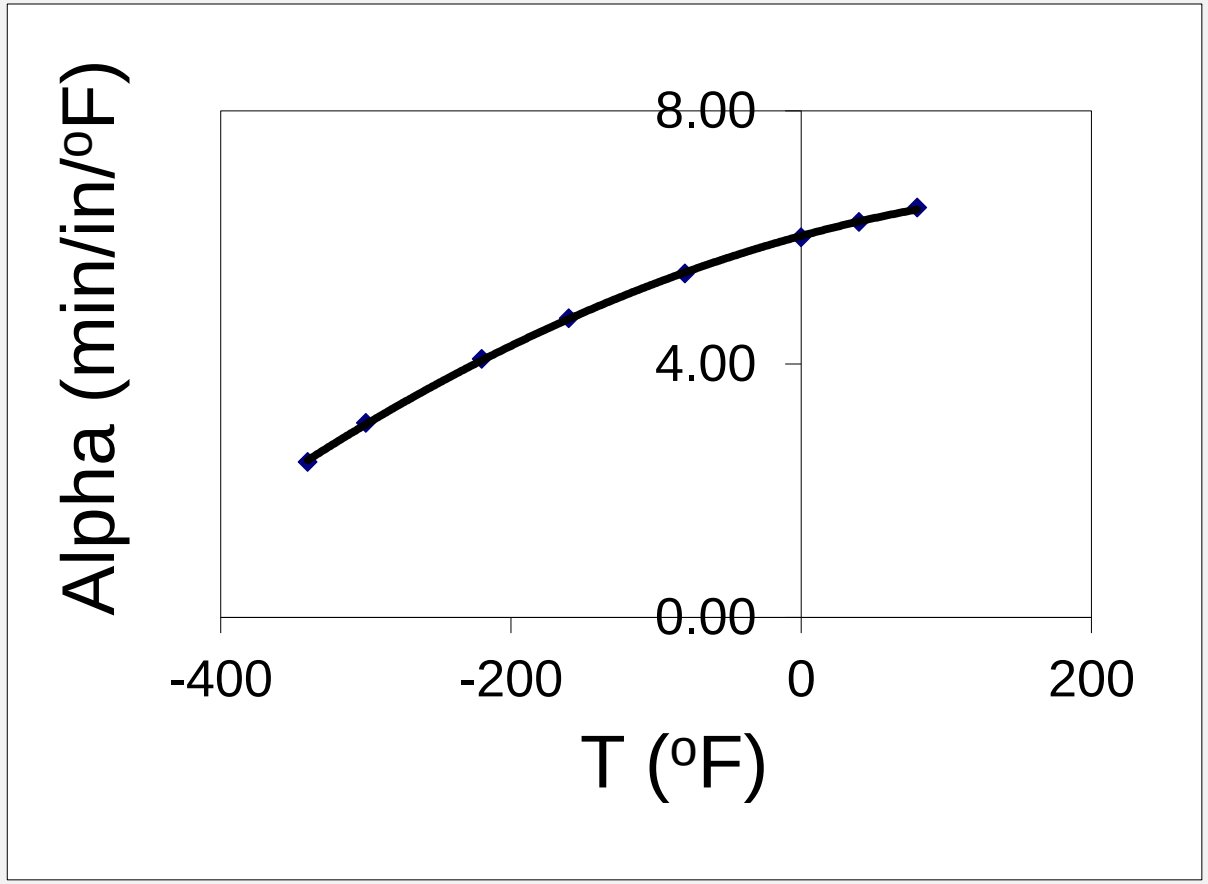


T(°F)	α (μin/in/°F)
-340	2.45
-300	3.07
-220	4.08
-160	4.72
-80	5.43
0	6.00
40	6.24
80	6.47

Exit and click:

-340	2.45
-300	3.07
-220	4.08
-160	4.72
-80	5.43
0	6
40	6.24
80	6.47

ΠΑΛΙΝΔΡΟΜΗΣΗ (GENERAL FORMAT)



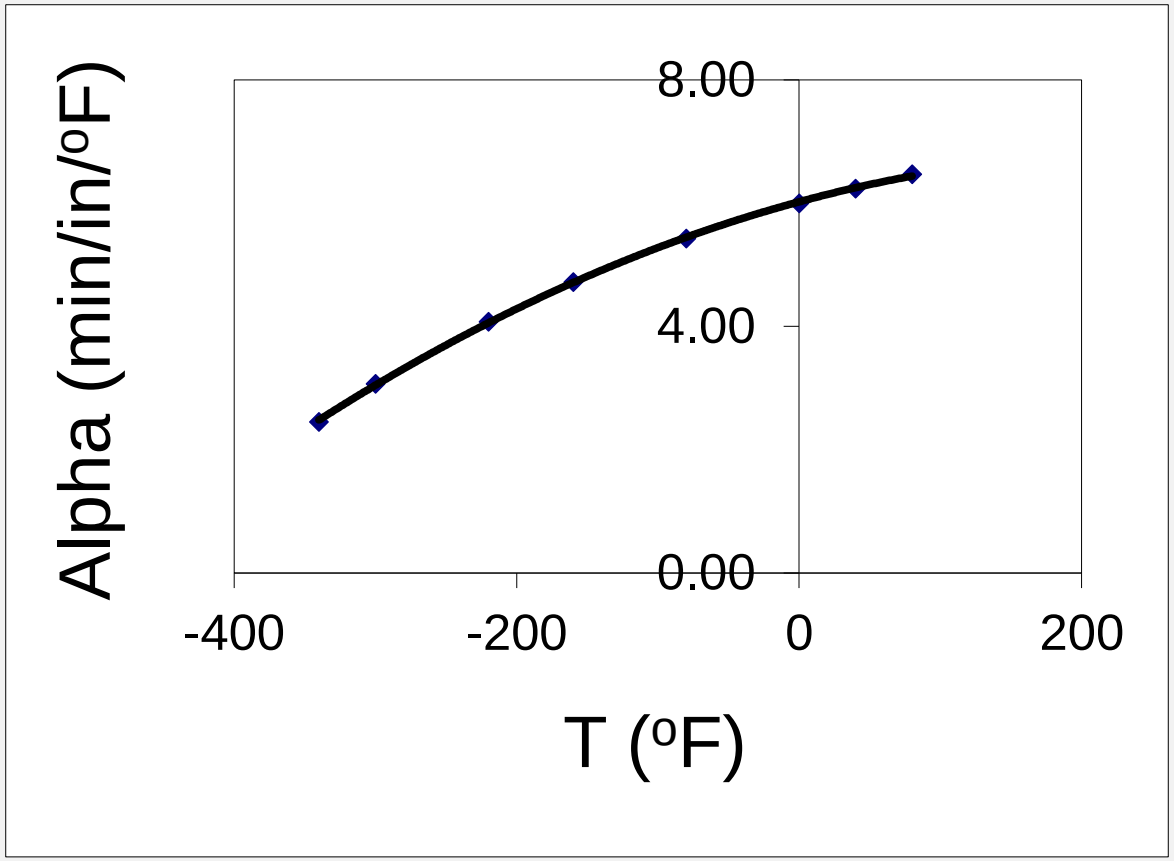
$$\alpha = -1E\text{-}05T^2 + 0.0062T + 6.0234$$

ΠΑΡΑΤΗΡΟΥΜΕΝΕΣ ΚΑΙ ΥΠΟΛΟΓΙΣΜΕΝΕΣ ΤΙΜΕΣ

$$\alpha = -1E\text{-}05T^2 + 0.0062T + 6.0234$$

T(ºF)	α (μin/in/ºF) Given	α (μin/in/ºF) Predicted
-340	2.45	2.76
-300	3.07	3.26
-220	4.08	4.18
-160	4.72	4.78
-80	5.43	5.46
0	6.00	6.02
40	6.24	6.26
80	6.47	6.46

ΠΑΛΙΝΔΡΟΜΗΣΗ (SCIENTIFIC FORMAT)



$$\alpha = -1.2360E\text{-}05T^2 + 6.2714E\text{-}03T + 6.0234$$

ΠΑΡΑΤΗΡΟΥΜΕΝΕΣ ΚΑΙ ΥΠΟΛΟΓΙΣΜΕΝΕΣ ΤΙΜΕΣ

$$\alpha = -1.2360E\text{-}05T^2 + 6.2714E\text{-}03T + 6.0234$$

T(°F)	α (μin/in/°F) Given	α (μin/in/°F) Predicted
-340	2.45	2.46
-300	3.07	3.03
-220	4.08	4.05
-160	4.72	4.70
-80	5.43	5.44
0	6.00	6.02
40	6.24	6.25
80	6.47	6.45

ΠΑΡΑΤΗΡΟΥΜΕΝΕΣ ΚΑΙ ΥΠΟΛΟΓΙΣΜΕΝΕΣ ΤΙΜΕΣ

$$\alpha = -1.2360E\text{-}05T^2 + 6.2714E\text{-}03T + 6.0234$$

$$\alpha = -1E\text{-}05T^2 + 0.0062T + 6.0234$$

T(°F)	α (μin/in/°F) Given	α (μin/in/°F) Predicted	α (μin/in/°F) Predicted
-340	2.45	2.46	2.76
-300	3.07	3.03	3.26
-220	4.08	4.05	4.18
-160	4.72	4.70	4.78
-80	5.43	5.44	5.46
0	6.00	6.02	6.02
40	6.24	6.25	6.26
80	6.47	6.45	6.46

Scientific

general

ΣΦΑΛΜΑ ΑΠΟΚΟΠΗΣ - ΠΑΡΑΔΕΙΓΜΑ

Χρήση μερικών μόνο όρων μιάς σειράς Maclaurin για τον υπολογισμό του e^x

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \ldots\ldots\ldots\ldots\ldots\ldots\ldots$$

Εάν μόνο 3 όροι χρησιμοποιηθούν:

$$\text{Σφάλμα αποκοπής (truncation)} = e^x - \left(1 + x + \frac{x^2}{2!}\right)$$

2Ο ΠΑΡΑΔΕΙΓΜΑ ΣΦΑΛΜΑΤΟΣ ΑΠΟΚΟΠΗΣ

Χρήση πεπερασμένου Δx για την τιμή της $f'(x)$

$$f'(x) \approx \frac{f(x+\Delta x) - f(x)}{\Delta x}$$

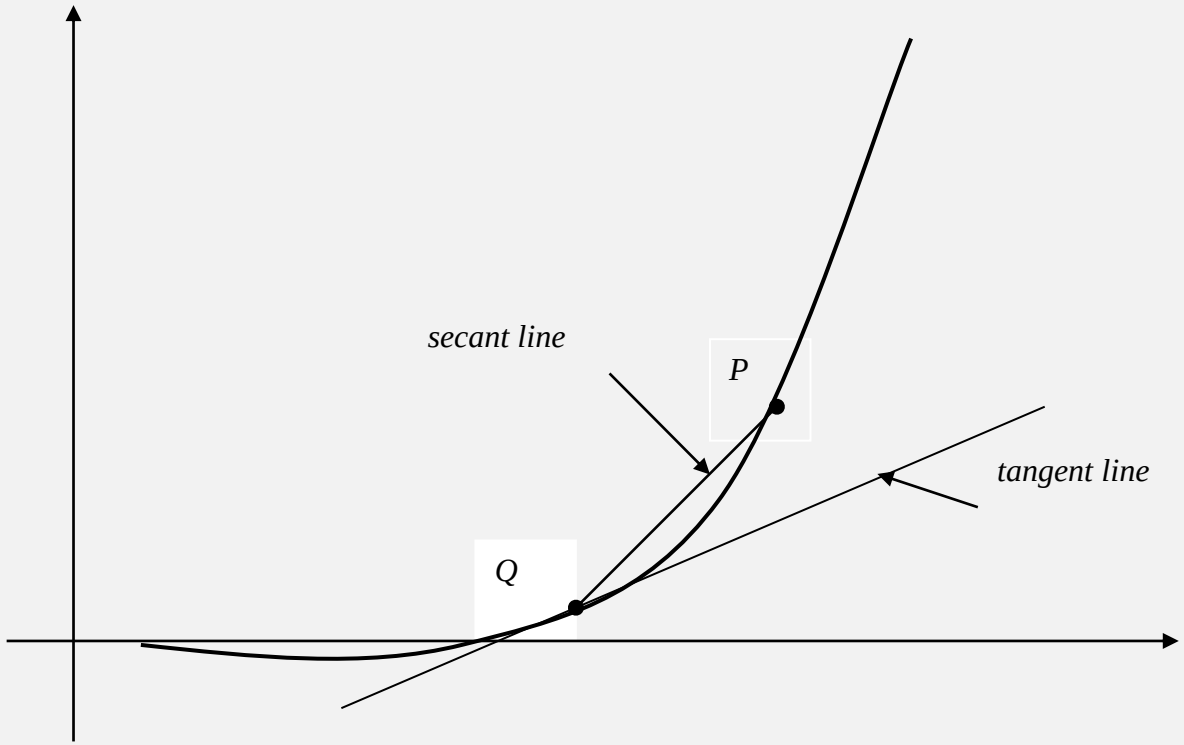
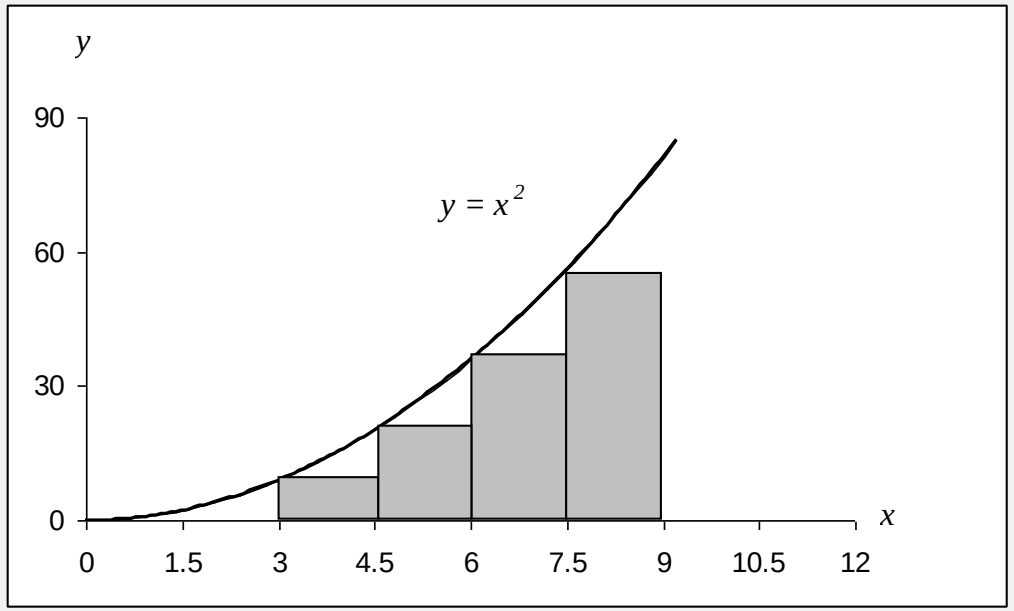


Figure 1. Approximate derivative using finite Δx

3Ο ΠΑΡΑΔΕΙΓΜΑ ΣΦΑΛΜΑΤΟΣ ΑΠΟΚΟΠΗΣ

Χρήση πεπερασμένων ορθογωνίων για τον προσδιορισμό ολοκληρώματος



ΠΑΡΑΔΕΙΓΜΑ—ΣΕΙΡΕΣ MACLAURIN

Υπολογίστε το: $e^{1.2}$ με σχετικό προσεγγιστικό σφάλμα μικρότερο από 1%.

$$e^{1.2} = 1 + 1.2 + \frac{1.2^2}{2!} + \frac{1.2^3}{3!} + \ldots\ldots\ldots\ldots\ldots\ldots$$

n	$e^{1.2}$	E_a	$\lvert \in_a \rvert \%$
1	1	__	___
2	2.2	1.2	54.545
3	2.92	0.72	24.658
4	3.208	0.288	8.9776
5	3.2944	0.0864	2.6226
6	3.3151	0.020736	0.62550

6 όροι είναι απαραίτητοι. Πόσοι όροι χρειάζονται για τουλάχιστον 1 σημαντικό ψηφίο σωστό (γιατί)?

Γιατί ισχύει $|x - X| \leq 0.5 * 10^{m-n+1}$, με m=0 και n=1, το σφάλμα θα πρέπει να είναι μικρότερο από 0.5 (ώστε να έχω 1 τουλάχιστον σημαντικό ψηφίο σίγουρα σωστό)

ΠΑΡΑΔΕΙΓΜΑ 2 —ΠΑΡΑΓΩΓΙΣΗ

Να βρεθεί η $f'(3)$ για $f(x) = x^2$ με $f'(x) \approx \dfrac{f(x+\Delta x) - f(x)}{\Delta x}$ και $\Delta x = 0.2$

$$f'(3) = \frac{f(3+0.2) - f(3)}{0.2} = \frac{f(3.2) - f(3)}{0.2} = \frac{3.2^2 - 3^2}{0.2} = \frac{10.24 - 9}{0.2} = \frac{1.24}{0.2} = 6.2$$

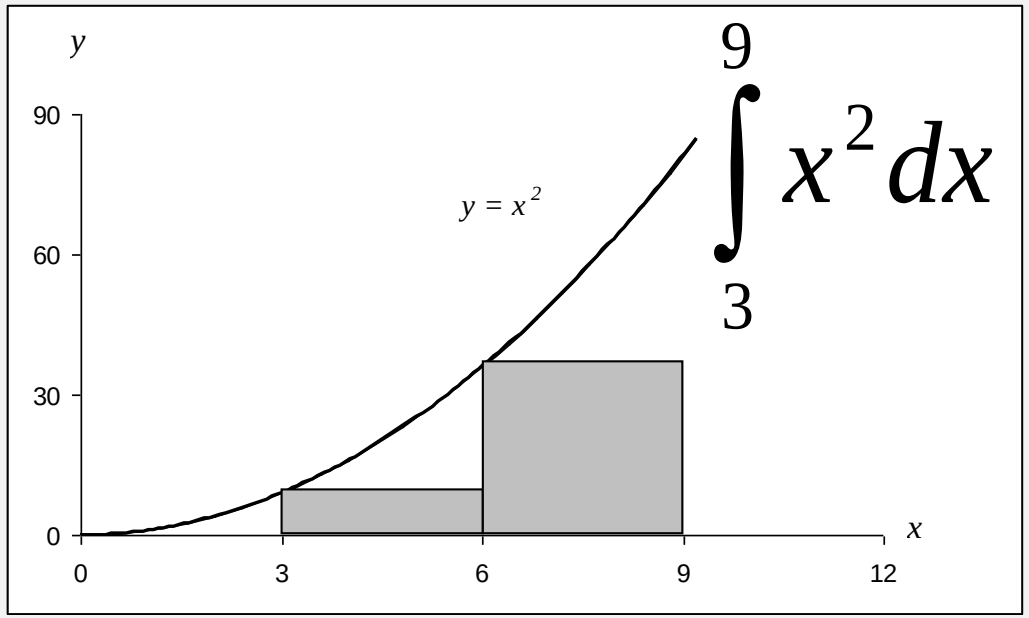
Η πραγματική τιμή: $f'(x) = 2x, \quad f'(3) = 2 \times 3 = 6$

Σφάλμα αποκοπής: $6 - 6.2 = -0.2$

Να βρεθεί το σφάλμα αποκοπής με: $\Delta x = 0.1$

ΠΑΡΑΔΕΙΓΜΑ 3 — ΟΛΟΚΛΗΡΩΣΗ

Με τη χρήση δύο ορθογωνιων ίσου πλάτους να προσεγγιστεί το εμβαδό κάτω από την καμπύλη $f(x) = x^2$ στο διάστημα $[3,9]$



$$\int_3^9 x^2 dx$$

Με πλάτος 3: $\int_3^9 x^2 dx = (x^2)\Big|_{x=3}(6-3) + (x^2)\Big|_{x=6}(9-6)$

(Δύο διαστήματα)

$$= (3^2)3 + (6^2)3 = 27 + 108 = 135$$

Πραγματική τιμή $\int_3^9 x^2 dx = \left[\frac{x^3}{3}\right]_3^9 = \left[\frac{9^3 - 3^3}{3}\right] = 234$

Σφάλμα αποκοπής: $234 - 135 = 99$

Να βρεθεί το σφάλμα αποκοπής για 4 ορθογώνια (διαστήματα)

… ΛΥΣΗ ΜΕ 4 ΔΙΑΣΤΗΜΑΤΑ …:

	x2	πλάτος	εμβαδό
3	9	1.5	13.5
4.5	20.25	1.5	30.375
6	36	1.5	54
7.5	56.25	1.5	84.375
9	81	1.5	
για 4 διαστήματα (πλάτος 1.5) =			182.25
σφάλμα αποκοπής =			60.75
μικρότερο από 99 που ήταν με 2 διαστήματα			

ΔΥΑΔΙΚΗ ΑΝΑΠΑΡΑΣΤΑΣΗ

Αναπαράσταση δεκαδικου αριθμου:

$$257.76 = 2\times10^{2} + 5\times10^{1} + 7\times10^{0} + 7\times10^{-1} + 6\times10^{-2}$$

Αναπαράσταση αριθμού στο δυαδικό σύστημα:

$$(1011.0011)_2 = \begin{pmatrix} (1\times2^{3} + 0\times2^{2} + 1\times2^{1} + 1\times2^{0}) \\ +(0\times2^{-1} + 0\times2^{-2} + 1\times2^{-3} + 1\times2^{-4}) \end{pmatrix}_{10}$$

$$= 11.1875$$

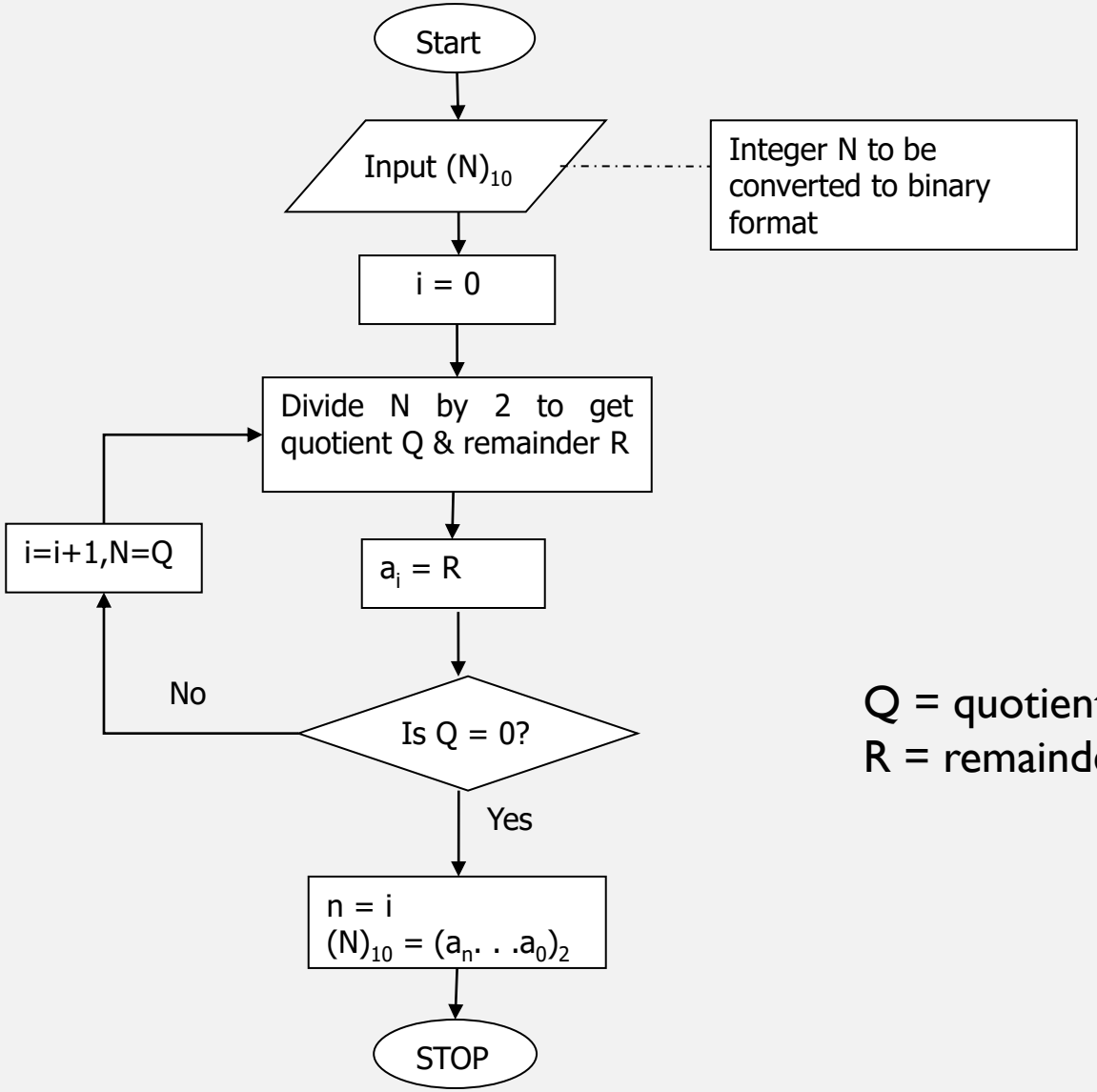
ΜΕΤΑΤΡΟΠΗ ΑΚΕΡΑΙΟΥ ΣΤΟ ΔΥΑΔΙΚΟ ΣΥΣΤΗΜΑ

Πίνακας μετατροπής του αριθμού 11.

	Ψηφίο πηλίκου	υπόλοιπο
11/2	5	$1 = a_0$
5/2	2	$1 = a_1$
2/2	1	$0 = a_2$
1/2	0	$1 = a_3$

Επομένως:

$$(11)_{10} = (a_3 a_2 a_1 a_0)_2$$
$$= (1011)_2$$



Q = quotient = πηλίκο
R = remainder = υπόλοιπο

Μετατροπή δεκαδικού στο δυαδικό σύστημα

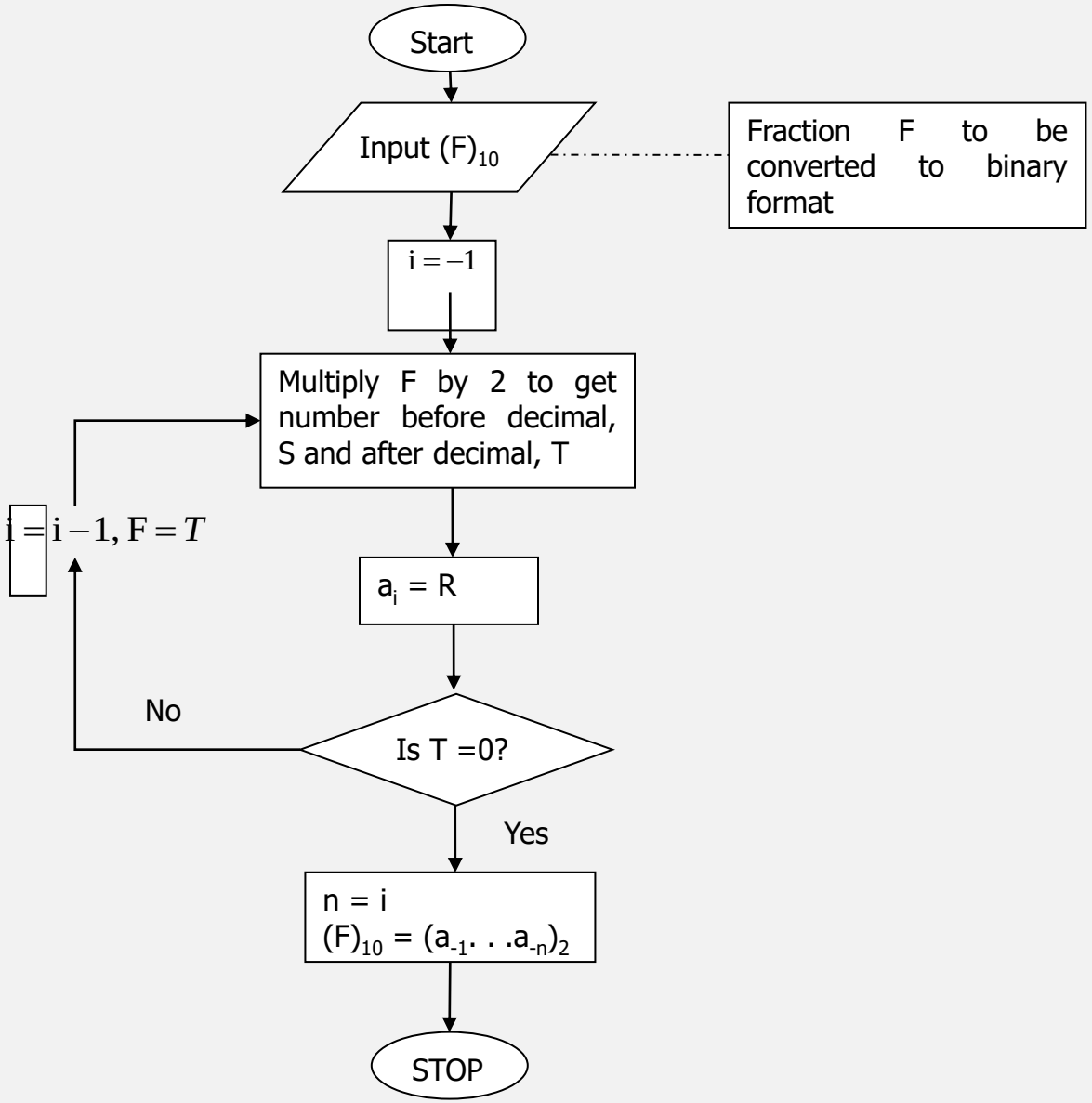
Πίνακας. Μετατροπή δεκαδικού στο δυαδικό σύστημα.

	Number	Number after decimal	Number before decimal
0.1875×2	0.375	0.375	$0=a_{-1}$
0.375×2	0.75	0.75	$0=a_{-2}$
0.75×2	1.5	0.5	$1=a_{-3}$
0.5×2	1.0	0.0	$1=a_{-4}$

Πολλαπλασιασμός Χ2, και συνεχείς πολλαπλασιασμοί του υπολοίπου Χ2, αφού αφαιρεθεί η μονάδα όταν υπάρχει. Όταν δεν υπάρχει, μπαίνει μηδέν.

Επομένως:

$$(0.1875)_{10}=(a_{-1}a_{-2}a_{-3}a_{-4})_2$$
$$=(0.0011)_2$$


Start
Input $(F)_{10}$
Fraction F to be converted to binary format
$i = -1$
Multiply F by 2 to get number before decimal, S and after decimal, T
$i = i - 1, F = T$
$a_i = R$
No
Is T =0?
Yes
n = i
$(F)_{10} = (a_{-1}. . .a_{-n})_2$
STOP

ΜΕΤΑΤΡΟΠΗ ΠΡΑΓΜΑΤΙΚΟΥ ΣΤΟ ΔΥΑΔΙΚΟ ΣΥΣΤΗΜΑ

$$(11.1875)_{10} = (\ ?.?\ )_2$$

Since

$$(11)_{10} = (1011)_2$$

and

$$(0.1875)_{10} = (0.0011)_2$$

we have

$$(11.1875)_{10} = (1011.0011)_2$$

ΔΕΝ ΑΝΑΠΑΡΙΣΤΩΝΤΑΙ ΟΛΟΙ ΟΙ ΔΕΚΑΔΙΚΟΙ ΑΡΙΘΜΟΙ ΑΚΡΙΒΩΣ ΣΤΟ ΔΥΑΔΙΚΟ ΣΥΣΤΗΜΑ…

Πίνακας: μετατροπή δεκαδικού (0.3) στο δυαδικό σύστημα.

	Number	**Number after decimal**	**Number before Decimal**
0.3×2	0.6	0.6	$0=a_{-1}$
0.6×2	1.2	0.2	$1=a_{-2}$
0.2×2	0.4	0.4	$0=a_{-3}$
0.4×2	0.8	0.8	$0=a_{-4}$
0.8×2	1.6	0.6	$1=a_{-5}$

$$(0.3)_{10}\approx(a_{-1}a_{-2}a_{-3}a_{-4}a_{-5})_2=(0.01001)_2=0.28125$$

Με 5 ψηφία μετά την υποδιαστολή, έχουμε 0.28125 αντί 0.3)

… Η ΧΡΕΙΑΖΟΝΤΑΙ ΠΑΡΑ ΠΟΛΛΑ ΨΗΦΙΑ

			ψηφίο δύναμης του 2	δύναμη του 2			συνολική τιμή
0,3X2	0.6	0.6	0	-1	0,5	0	0
0,6X2	1.2	0.2	1	-2	0,25	0,25	0,25
0,2X2	0.4	0.4	0	-3	0,125	0	0,25
0,4X2	0.8	0.8	0	-4	0,0625	0	0,25
0,8X2	1.6	0.6	1	-5	0,03125	0,03125	0,28125
0,6X2	1,2	0,2	1	-6	0,015625	0,01563	0,296875
0,2X2	0,4	0,4	0	-7	0,0078125	0	0,296875
0,4X2	0,8	0,8	0	-8	0,00390625	0	0,296875
0,8X2	1,6	0,6	1	-9	0,001953125	0,00195	0,298828125
επομένως, $(0.010011001)_2$ = 0,298828125							

Με 9 ψηφία μετά την υποδιαστολή, έχουμε 0.298828125 αντί 0.3)

ΕΝΑΛΛΑΚΤΙΚΟΣ ΤΡΟΠΟΣ ΜΕΤΑΤΡΟΠΗΣ ΣΤΟ ΔΥΑΔΙΚΟ ΣΥΣΤΗΜΑ

Μετατροπή του $(11.1875)_{10}$ στο δυαδικό σύστημα

$$
\begin{aligned}
(11)_{10} &= 2^3 + 3 \\
&= 2^3 + 2^1 + 1 \\
&= 2^3 + 2^1 + 2^0 \\
&= 1\times 2^3 + 0\times 2^2 + 1\times 2^1 + 1\times 2^0 \\
&= (1011)_2
\end{aligned}
$$

$$
\begin{aligned}
(0.1875)_{10} &= 2^{-3} + 0.0625 \\
&= 2^{-3} + 2^{-4} \\
&= 0\times 2^{-1} + 0\times 2^{-2} + 1\times 2^{-3} + 1\times 2^{-4} \\
&= (.0011)_2
\end{aligned}
$$

$$(11.1875)_{10} = (1011.0011)_2$$

Μετάδοση σφαλμάτων

Στις αριθμητικές μεθόδους, οι υπολογισμοί δεν γίνονται με ακριβείς αριθμούς. Πώς μεταδίδονται αυτές οι ανακρίβειες (σφάλματα) με τους υπολογισμούς?

παράδειγμα 1:

Μετάδοση σφάλματος στην πρόσθεση. Πρόσθεση $X + Y$ όπου

$X = 1.5 \pm 0.05$

$Y = 3.4 \pm 0.04$

λύση

Μέγιστες τιμές $X = 1.55$ and $Y = 3.44$

Μέγιστη τιμή αθροίσματος $X + Y = 1.55 + 3.44 = 4.99$

Ελάχιστες τιμές $X = 1.45$ and $Y = 3.36$.

Ελάχιστη τιμή αθροίσματος $X + Y = 1.45 + 3.36 = 4.81$

Επομένως

$$(1.5+3.4)-(0.05+0.04) = 4.81 \leq X + Y \leq 4.99 = (1.5+3.4)+(0.05+0.04)$$

Μετάδοση σφαλμάτων σε τύπους

f συνάρτηση μεταβλητών $X_1, X_2, X_3, \ldots\ldots, X_{n-1}, X_n$

Μέγιστη τιμή σφάλματος στην f

$$\Delta f \approx \left|\frac{\partial f}{\partial X_1}\Delta X_1\right| + \left|\frac{\partial f}{\partial X_2}\Delta X_2\right| + \ldots\ldots + \left|\frac{\partial f}{\partial X_{n-1}}\Delta X_{n-1}\right| + \left|\frac{\partial f}{\partial X_n}\Delta X_n\right|$$

Δηλαδή το μέγιστο σφάλμα είναι ίσο με το άθροισμα για κάθε μεταβλητή, του γινομένου του σφάλματος της μεταβλητής επί την παράγωγο της συνάρτησης ως προς τη μεταβλητή.

παράδειγμα 2:

Η παραμόρφωση σε αξονικό στοιχείο τετραγωνικής διατομής:

$$\in = \frac{F}{h^2 E}$$

Όπου:

$$F = 72 \pm 0.9 \text{ N}$$
$$h = 4 \pm 0.1 \text{ mm}$$
$$E = 70 \pm 1.5 \text{ GPa}$$

Μέγιστο σφάλμα στην μετρημένη παραμόρφωση?

παράδειγμα 2:

$$\in = \frac{72}{(4\times10^{-3})^2(70\times10^9)} = 64.286\times10^{-6} = 64.286\mu$$

Παρακαλώ υπολογίστε το μέγιστο σφάλμα:

$$\Delta\in = \left|\frac{\partial\in}{\partial F}\Delta F\right| + \left|\frac{\partial\in}{\partial h}\Delta h\right| + \left|\frac{\partial\in}{\partial E}\Delta E\right|$$

$$\in = \frac{F}{h^2E}$$

$$F = 72\pm0.9\ \text{N}$$
$$h = 4\pm0.1\ \text{mm}$$
$$E = 70\pm1.5\ \text{GPa}$$

παράδειγμα 2:

$$\in = \frac{F}{h^2 E}$$

$$F = 72 \pm 0.9 \text{ N}$$
$$h = 4 \pm 0.1 \text{ mm}$$
$$E = 70 \pm 1.5 \text{ GPa}$$

$$\frac{\partial \in}{\partial F} = \frac{1}{h^2 E} \qquad \frac{\partial \in}{\partial h} = -\frac{2F}{h^3 E} \qquad \frac{\partial \in}{\partial E} = -\frac{F}{h^2 E^2}$$

Αρα

$$\Delta E = \left|\frac{1}{h^2 E}\Delta F\right| + \left|\frac{2F}{h^3 E}\Delta h\right| + \left|\frac{F}{h^2 E^2}\Delta E\right|$$

$$= \left|\frac{1}{(4\times10^{-3})^2(70\times10^9)}\times 0.9\right| + \left|\frac{2\times 72}{(4\times10^{-3})^3(70\times10^9)}\times 0.0001\right|$$

$$+ \left|\frac{72}{(4\times10^{-3})^2(70\times10^9)^2}\times 1.5\times10^9\right|$$

$$= 5.3955\mu$$

Επομένως

$$\in = (64.286\mu \pm 5.3955\mu)$$

παράδειγμα 3:

Αφαίρεση αριθμών περίπου ίσων μπορεί να δημιουργήσει ανακρίβειες.

$$z = x - y$$

$$|\Delta z| = \left|\frac{\partial z}{\partial x}\Delta x\right| + \left|\frac{\partial z}{\partial y}\Delta y\right| = |(1)\Delta x| + |(-1)\Delta y| = |\Delta x| + |\Delta y|$$

Το σχετικό σφάλμα είναι:

$$\left|\frac{\Delta z}{z}\right| = \frac{|\Delta x| + |\Delta y|}{|x - y|}$$

Για παράδειγμα: $x = 2 \pm 0.001$

$y = 2.003 \pm 0.001$

$$\left|\frac{\Delta z}{z}\right| = \frac{|0.001| + |0.001|}{|\,2 - 2.003\,|} = 0.6667 = 66.67\%$$

ΑΝΑΠΑΡΑΣΤΑΣΗ ΚΙΝΗΤΗΣ ΥΠΟΔΙΑΣΤΟΛΗΣ (FLOATING POINT REPRESENTATION)

ΚΙΝΗΤΗ ΥΠΟΔΙΑΣΤΟΛΗ: ΚΑΝΟΝΙΚΟΠΟΙΗΜΕΝΗ ΜΟΡΦΗ

256.78 γράφεται $+\,2.5678 \times 10^{2}$

0.003678 γράφεται $+\,3.678 \times 10^{-3}$

-256.78 γράφεται $-\,2.5678 \times 10^{2}$

Η μορφή είναι: $\text{sign} \times \text{mantissa} \times 10^{\text{exponent}}$ $\qquad \sigma \times m \times 10^{e}$

Παράδειγμα: -2.5678×10^{2}

$$\sigma = -1$$

$$m = 2.5678$$

$$e = 2$$

ΜΟΡΦΗ ΚΙΝΗΤΗΣ ΥΠΟΔΙΑΣΤΟΛΗΣ ΔΥΑΔΙΚΩΝ ΑΡΙΘΜΩΝ

$y = \sigma \times m \times 2^e$
$\sigma = \text{πρόσημο}$ (**0 για θετικό, 1 για αρνητικό**)
$m =$ **mantissa** $[(1)_2 < m < (10)_2]$ $\delta\eta\lambda\alpha\delta\eta\ \mu\epsilon\tau\alpha\xi\acute{\upsilon}\ 1\ \kappa\alpha\iota\ 2$

Το 1 δεν αποθηκεύεται, καθώς είναι η βάση στην μηδενική δύναμη!

$e =$ **θετικός εκθέτης**

ΠΑΡΑΔΕΙΓΜΑ ΑΠΟΘΗΚΕΥΣΗΣ ΔΥΑΔΙΚΟΥ

9 bit-υποθετική λέξη

- Το πρώτο bit αποθηκεύει το πρόσημο,
- Το δεύτερο bit το πρόσημο του εκθέτη,
- Τα επόμενα 4 bits τα σημαντικά ψηφία (mantissa),
- Τα τελευτάια 3 bits τον εκθέτη

Μετατροπή σε δυαδικο

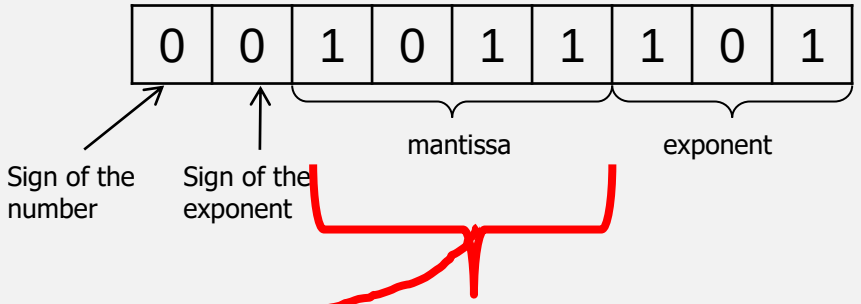
Μετατροπή σε floating point

μετακίνηση υποδιαστολής κατά 5

Αποκοπή mantissa, γιατί υπάρχουν μόνο 4 ψηφία

$$(54.75)_{10} = (110110.11)_2 = (1.1011011)_2 \times 2^5$$
$$\cong (1.1011)_2 \times (101)_2$$

Η αναπαράστασή του είναι:



Λόγω της αποκοπής έχει αποθηκευτεί σαν: (1+0.5+0.125+0.0625)*2^(1+4) =1.6875*32 = 54

Machine Epsilon

Ορίζεται ως το μέτρο της ακρίβειας: η διαφορά μεταξύ του 1 και του αμέσως επόμενου μεγαλύτερου αριθμού που μπορεί να αναπαρασταθεί.

ΠΑΡΑΔΕΙΓΜΑ: EPSILON ΣΕ ΛΕΞΗ 10 BIT ΔΥΑΔΙΚΟΥ ΣΥΣΤΗΜΑΤΟΣ

Λέξη 10 bit

- Πρόσημο αριθμού
- Πρόσημο εκθέτη
- 4 ψηφία για τον εκθέτη
- 4 ψηφία για mantissa

0	0	0	0	0	0	0	0	0	0

$$= (1)_{10}$$

Next number →

0	0	0	0	0	0	0	0	0	1

$$= (1.0001)_2 = (1.0625)_{10}$$

$$\in_{mach} = 1.0625 - 1 = 2^{-4}$$

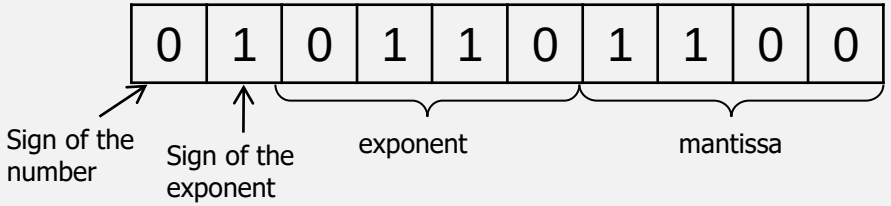
ΣΧΕΤΙΚΟ ΣΦΑΛΜΑ ΚΑΙ MACHINE EPSILON

Το σφάλμα στην αναπαράσταση ενός αριθμού, είναι μικρότερο από το machine epsilon

Example $(0.02832)_{10} \cong (1.1100)_2 \times 2^{-5}$

$$= (1.1100)_2 \times 2^{-(0110)_2}$$

10 bit word (sign, sign of exponent, 4 for exponent, 4 for mantissa)



$$(1.1100)_2 \times 2^{-(0110)_2} = 0.0274375$$

$$\in_a = \left| \frac{0.02832 - 0.0274375}{0.02832} \right|$$

$$= 0.034472 < 2^{-4} = 0.0625$$