CAPITULO I

ARITMÉTICA DE "t" DIGITOS

$$\xi_{c(a)} = 10^{1-t}$$

EXORDIO

Es habitual que los alumnos de ingeniería tomen contacto con los distintos conjuntos numéricos en alguna de las asignaturas iniciales de la carrera.

Así, partiendo de la necesidad de contar objetos, se abstrae el conjunto de los números naturales N; simbolizando la nada se pasa al de los naturales ampliado con el cero, N_0, luego al de los enteros Z incluyendo para ello a los números negativos, luego debido al problema de la medida, se introducen los racionales Q que pueden ser expresados como el cociente entre dos enteros y más adelante se completa el cuadro incluyendo a los irracionales, es decir aquellos que no pueden ser expresados como el cociente, o razón entre dos números enteros.

Racionales e irracionales constituyen el conjunto de los números reales R, fundamento y base para el análisis matemático y problemas relacionados con su fundamentación. R completa la llamada con justicia, recta real, siendo biunívoca la relación entre ambos conjuntos: los puntos de la recta y los elementos de R.

Los grandes problemas de fundamentación fueron tratados, entre otros por Weierstrass, Dedekind, Cantor quienes consolidaron el edificio iniciado por Fermat y llevado a su máximo esplendor por Leibniz y Newton en el siglo XVIII.

La necesidad de resolver problemas donde aparecen cuadrados negativos dio lugar a la consideración, desde la época de Cardano, de una categoría especial de números llamados primero "imposibles", "irreales" y, por fin imaginarios conformando con ellos el conjunto de los números complejos C. Los trabajos de Riemann y Cauchy le dan carta de legitimidad a estos números que hoy se denominan complejos.

De esta forma, un alumno de ingeniería conoce de qué se habla cuando se habla de N, N_0, Z, Q, R y C.

Sin embargo y sin desmerecer un ápice la enorme contribución de cada uno de estos conjuntos al álgebra y al análisis matemático, cotidianamente se trabaja sencillamente con "números". Se los escribe, se los registra, se los guarda en la memoria de una máquina, se opera con ellos, con ellos se obtienen resultados que se interpretan y que se aplican en la solución de problemas concretos.

Estos "números" tienen una característica distintiva: sólo contienen "t" dígitos y llevan implícito, por ese motivo, un error.

Con estos "números" de "t" dígitos se hacen operaciones y, al hacerlas, sus inseparables errores pueden dar desagradables sorpresas.

Estudiar estos "números", sus errores y su forma de propagación es requisito indispensable para iniciar cualquier curso de cálculo numérico.

I INTRODUCCIÓN

1 Sea “a” un número real. Se expresa este número en la base de numeración b escribiendo:

$$a)_b = d_1\ d_2\ d_3\ ,\ d_4\ d_5\ d_6\ d_7\ d_8\ \ldots$$

siendo d_i dígitos del sistema de numeración elegido

d_i 0, 1, 2, 3, 4, 5, 6, 7, 8, 9 en el sistema decimal (b=10)

d_i 0, 1 en el sistema binario (b=2)

d_i 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, A, B, C, D, E, F en el sistema hexadecimal (b=16)

d_i 0, 1, 2, 3, 4, 5, 6, 7 en el sistema octal (b=8)

Por ejemplo

$$a = (375{,}4834\ldots)_{10} \text{ ó } 375{,}4834\ldots$$

$$a = (101110111{,}01111011110000000001101\ldots)_2$$

$$a = (177{,}7BC01A36E2EB1C432)_{16}$$

$$a = (567{,}36740032155\ldots)_8$$

2 Notar que solamente no se consigna explícitamente la base b cuando no hay lugar a confusión, es decir cuando se usa el sistema aritmético decimal (el de todos los días, el que la humanidad utiliza por la simple razón de tener diez dedos (dígitos) en sus manos).

3 Notar también que los dígitos d_1, d_2 y d_3 forman lo que se denomina parte entera del número en cuestión mientras que los restantes forman la parte no entera, para no decir decimal, término que sólo corresponde para la

base b = 10 y se desconoce la existencia de algún término que cumpla la misma función para otras bases. EL carácter, (coma) o. (punto) según la convención que se utilice es el separador entre la parte entera y la no entera

4 La notación $a)_b = d_1\, d_2\, d_3\, ,d_4\, d_5\, d_6\, d_7\, d_8$ ... significa:

$$=d_1 . b^2 + d_2 . b^1 + d_3 . b^0 + d_4 . b^{-1} + d_5 . b^{-2} + d_6 . b^{-3} + d_7 . b^{-4} . d_8 . b^{-5} + ...$$

Esta descomposición del número a según potencias consecutivas de la base b es la clave de la notación posicional y significó un avance trascendente en la historia del pensamiento matemático

5 Como todo avance, no fue sencillo. Durante dos siglos convivieron y combatieron (académicamente) entre sí numeristas y abacistas. Los primeros, defendiendo esta notación, los otros, el iábaco! Corresponde a la renacentista Italia, el mérito de haber sido la puerta de entrada de esta notación a Europa, notación que el lector seguramente le escuchó a su maestra de la escuela primaria cuando ésta le enseñaba

unidades	10^0 (b^0)
decenas	10^1 (b^1)
centenas	10^2 (b^2)
unidades de mil	10^3 (b^3)
décimos	10^{-1} (b^{-1})
centésimos	10^{-2} (b^{-2})
milésimos	10^{-3} (b^{-3})
etc, etc,...	

$$375.4834 = 3 . 10^2 + 7 . 10^1 + 5 . 10^0 + 4 . 10^{-1} + 8 . 10^{-2} + 3 . 10^{-3} + 4 . 10^{-4} + ...$$

3	7	5	.	4	8	3	4
Centenas	Decenas	Unidades	Símbolo	Décimos	Centésimos	Milésimos	Diez milésimos

6 Como breve reseña histórica se menciona que un tal Bonaccio, residente en Pisa Italia, celebra el nacimiento de su hijo Leonardo. Como era vástago de

Bonaccio, casi nunca nadie conoció al niño como Leonardo de Pisa, sino como "el hijo de Bonaccio", esto es, **Fibonacci**. (filius bonnaccio)

7 Bonaccio, por entonces director de una aduana italiana en Argelia, necesita que su hijo sepa de números, por lo que obliga al chiquillo a estudiar aritmética posicional hindú. Milagrosamente, Fibonacci descubrió en las matemáticas el amor de su vida. Nunca más las abandonó.

8 El aporte de Fibonacci a la matemática es tan grande y tan profundo que prácticamente no puede ser medido. Por la época en la que vivió, el sistema de numeración arábigo era poco menos que una curiosidad: todo el mundo usaba los números romanos. Y ya se sabe lo difícil que es multiplicar por no hablar de dividir con números romanos.

9 Fibonacci, recordando el curso de aritmética hindú aprendido de niño, escribe, en 1202, su tratado *Liber abaci* ("El Libro del Ábaco") que es, ni más ni menos, un tratado sobre el sistema numeral indoarábigo. En él presenta al público y a los científicos europeos los signos hindúes (1, 2, 3...) y el 0 árabe, donde dice que se llama "cero" (*quod arabice zephirum appellatur*).

10 Resultaría injusto olvidar la mención de Alexandre de Villedieu, Franciscano Francés y John de Hallifax, llamado Sacrobosco quienes, junto a Fibonacci merecen el crédito de haber popularizado el “algorism” de la numeración indoarábiga. *Carmen de Algorismo* es un poema de Alexandre de Villedieu donde las operaciones con enteros están descriptas junto al uso del cero como número. Sacrobosco hace lo propio en un tratado de astronomía llamado *Algorismus Vulgaris* utilizado profusamente en la edad media.

11 Cabe también mencionar que hasta alcanzar la fundamentación rigurosa de los números reales debieron pasar varios siglos de arduo trabajo matemático. Los antes nombrados simplemente utilizaban números sin que hayan quedado rastros sobre alguna elaboración conceptual específica sobre la naturaleza de dichos números, tal como se hace en la actualidad.

II NOTACIÓN "t" DÍGITOS

12 El número $a)_b = d_1\, d_2\, d_3\, ,d_4\, d_5\, d_6\, d_7\, d_8 \ldots$ a puede escribirse

$$a)_b = b^q .\ 0,\ d_1\, d_2\, d_3\, d_4\, d_5\, d_6\, d_7\, d_8 \ldots;\ d_1 \neq 0$$

donde b es la base del sistema de numeración elegido q es un exponente que lleva el símbolo que separa la parte entera de la no entera al lugar que le corresponde y la secuencia $d_1\, d_2\, d_3\, d_4\, d_5\, d_6\, d_7\, d_8$ considerada como un todo se denomina **mantisa.(**$d_1 \neq 0$)

Por ejemplo

$$a = 375{,}4834\ldots = 10^3 .\ 0.3754834\ldots$$

$$b = 10$$

$$q = 3$$

mantisa 3754834...

13 Entrando brevemente en criptografía puede hacerse la convención que, si la base b es siempre la misma y que si las mantisas comienzan por izquierda con un dígito distinto de cero, pues si así no fuese se modificaría q hasta lograr un primer dígito por izquierda distinto de cero, no es necesario consignar explícitamente b y los caracteres 0. (cero punto) con lo cual

$$10^3\ 0.\ 3754834\ldots \quad = 33754834\ldots$$

siendo el último número escrito descifrable si se conoce la convención anterior.

14 Como a los números se los usa, se los escribe y se los almacena en la memoria de una computadora las mantisas no pueden tener infinitos dígitos del sistema de numeración elegido, como indica el símbolo matemático ... (tres puntos consecutivos), por la muy sencilla razón que no hay medio físico ni tiempo alguno en aptitud de "soportar infinitos dígitos".

15 En consecuencia, hay que efectuar un corte en alguna parte de la mantisa y este corte se hace luego de t-dígitos (de aquí el nombre del tema) con lo cual, ahora ya en aritmética de t-dígitos.

$$a = 10^q . 0. \delta_1 \ \delta_2 \ \delta_3 \ \delta_4 \ldots \delta_t = q \ \delta_1 \ \delta_2 \ \delta_3 \ \delta_4 \ldots \delta_t$$

donde los dígitos se han escrito con el símbolo δ en lugar del símbolo d para diferenciar y permitir las deducciones que más adelante se incluyen.

16 Cabe señalar que, en la jerga de los especialistas en computadoras la aritmética de t-dígitos se denomina **aritmética de punto flotante** porque el exponente q hace "flotar" al punto decimal hasta colocarlo en su posición correspondiente

III NUMEROS REALES – ARITMÉTICA DE "t" DÍGITOS

17 Como observación de extrema importancia, corresponde señalar que, al trabajar con aritmética de t-dígitos (o en punto flotante) el ***continuo real no existe***. Para un dado **q** luego de la mantisa $d_1 \ d_2 \ d_3 \ldots d_t$ existe la mantisa $d_1 \ d_2 \ d_3 \ldots d_t + 1$ sin posibilidad de escribir algún otro número intermedio.

18 Obsérvese que el proceso de calcular límites, que costó al pensamiento matemático más de dos milenios está basado en el concepto del continuo real, razón por la cual, por lo menos, se debe ser extremadamente prudente para tratar ciertos temas fundamentales de análisis matemático con computadora, sin que esto signifique de manera alguna que no deban aprovecharse y

explotarse sus extraordinarias capacidades de cálculo, graficación y procesamiento de datos.

IV ERROR

19 Naturalmente, al cortar la mantisa, se introduce un error que depende del criterio con que se haga el corte y que, inexorablemente estará presente en todo número de t-dígitos que se utilice y que, además, tendrá la perversa propiedad de propagarse a través de las distintas operaciones en que intervenga dicho número, siendo necesario evaluar en qué medida lo hace y en qué medida se potencia para tener idea sobre la bondad de los resultados a obtener.

Técnica de Corte

20 Sea a un número real (En base b = 10)

$$a = 10^q \times 0.d_1\, d_2\, d_3\, d_4 \ldots d_t\, d_{t+1}\, d_{t+2} \ldots \; ; \; d_1 \neq 0$$

y C (a) el "cortado" de a en t-dígitos

$$C\,(a) = 10^q\; 0.\, \delta_1\; \delta_2\; \delta_3\; \delta_4 \ldots \delta_t \qquad ; \; \delta_1 \neq 0$$

donde $\delta_i = d_i$ para i = 1, hasta t

Es decir que, en esta técnica, la mantisa del número **a** se corta impiadosamente luego de t dígitos sin importar cual sea el valor de d_{t+1}

Por ejemplo, siendo $a = 10^3 \,.\, 0{,}29871999\ldots$

el $C(a)$ en aritmética de t = 5 dígitos es

$$C(a) = 10^3 \,.\, 0{,}29871$$

sin importar la secuencia de números siguientes (999 que indudablemente tientan a transformar en dos al uno final).

Error en la técnica de corte

21 El error absoluto en la técnica de corte es

$$e_{c(a)} = | a - C(a)| = 10^q \ . \ 0.\ 0000 \ . \ . \ . \ 0 \ d_{t+1} \ d_{t+2} \ d_{t+3} \ . \ . \ .$$

suponiendo que $d_{t+1} \neq 0$ resulta

$$e_{c(a)} = 10^{q-t} \ 0.\ d_{t+1} \ d_{t+2} \ d_{t+3} \ . \ .$$

y, el correspondiente error relativo

$$\varepsilon_{c(a)} = \frac{e_{c(a)}}{a} = \frac{10^{q-t} 0.d_{t+1}d_{t+2}d_{t+3}...}{10^q 0.d_1 d_2 d_3 d_4 ...}$$

Si bien no hay problema con el cociente de exponentes, el cociente de mantisas, por tener un divisor menor que la unidad puede dar un cociente mayor que uno lo que dificulta la acotación de este error relativo.

Sin embargo, puede hacerse

$$0.\ d_1 d_2 d_3 \ d_4 \ . \ . \ . = 10^{-1} \quad d_1 .d_2 \ d_3 \ d_4 \ . \ . \ .$$

y siendo $d_1 \neq 0$, el número $d_1 \ . \ d_2 \ d_3 \ d_4$ será mayor que uno lo que garantiza un cociente menor que la unidad.

De esto resulta

$$\varepsilon_{c(A)} = \frac{10^{q-t}}{10^{q-1}} \cdot \frac{0.d_{t+1}d_{t+2}d_{t+3}...}{d_1.d_2 d_3 d_4 ...} < 10^{1-t}$$

$$\varepsilon_{c(a)} < 10^{1-t}$$

Desigualdad que permite acotar el error relativo de cualquier número expresado en aritmética de t-dígitos "cortado", error que depende exclusivamente de la cantidad de dígitos de la mantisa (t) y **no** de la magnitud del número a.

Cualquier número expresado en aritmética de t-dígitos "cortado" tiene inexorablemente asociado un error relativo cuya cota es $10^{(1-t)}$ que afecta a todas las operaciones en que dicho número interviene.

Técnica de Redondeo

22 Sea a un número real

$$a = 10^q \,.\, 0.\, d_1 d_2 d_3 d_4 \ldots d_t\, d_{t+1}\, d_{t+2} \ldots \;;\; d_1 \neq 0$$

R (a) el "redondeado" de a en t dígitos

$$R(a) = 10^q\, 0.\, \delta_1\, \delta_2\, \delta_3\, \delta_4 \ldots \delta_t$$

donde $\delta_i = d_i$ para i = 1 hasta t-1

$\delta_t = d_t + 1$ si $d_{t+1} \geq 5$

$\delta_t = d_t$ si $d_{t+1} < 5$

Es decir que, en esta técnica, todos los dígitos de la mantisa se hacen iguales al del número a dado, excepto el último de ellos cuyo valor se establece según sea el valor del primer dígito de a ubicado en la posición t+1.

Siendo $a = 10^3 \,.\, 0.\, 29871999 \ldots$

el R (a) en aritmética de t = 5 dígitos es

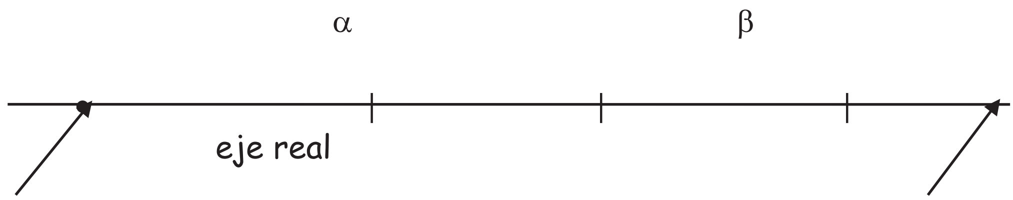
$$R(a) = 10^3 \,.\, 0.\, 29872$$

porque $d_{t+1} = d_6 = 9 > 5 \Rightarrow \delta_5 = d_5 + 1 = 1 + 1 = 2$

Error en la técnica de redondeo:

23 Como ya ha sido dicho, en la aritmética de t dígitos el continuo real no puede ser expresado, teniéndose una situación como la siguiente:

•



$r = 10^q\, 0.\, \delta_1 \delta_2 \delta_3 \,...\, \delta_t$ $\qquad$ $s = 10^q\, 0.\, \delta_1 \delta_2 \delta_3 \,...\, \delta_t + 1$

donde r y s son dos números consecutivos en dicha aritmética.

Suponiendo que el real a esté representado en el eje real por el punto α, su aproximación en t dígitos redondeado estará dado por r; mientras que si está representado por β, su aproximación será s.

Notar que según la técnica de corte los números reales representados por α y β se aproximan ambos mediante el número r.

Por eso, heuristicamente, puede decirse que

$$\varepsilon_R(a) < \frac{1}{2}\, 10^{1-t}$$

V PROPAGACIÓN DEL ERROR:

24 Habiendo quedado establecido que a cada número expresado (o almacenado en la memoria de una computadora) en aritmética de t-dígitos se le asocia un error relativo, corresponde ahora analizar cómo se propaga dicho error en las operaciones aritméticas.

Para ello, solo se considerarán las cuatro operaciones fundamentales: adición, sustracción, multiplicación y división; con la siguiente notación: letra sin subíndice, verdadero valor; letra con subíndice t, número representado por la letra, expresado en aritmética de t-dígitos y e con letra de subíndice, error absoluto correspondiente.

$$a = a_t + e_a$$

25 **Adición:**

Sea calcular

$$z = x + y$$

con $x = x_t + e_x$

e $y = y_t + e_y$

$$z = x_t + e_x + y_t + e_y = z_t + e_z$$

donde $z_t = x_t + y_t$

y $e_z = e_x + e_y$

En términos de error relativo

$$\varepsilon_z = \frac{e_z}{z} = \frac{e_x + e_y}{x+y} = \frac{e_x}{x+y} + \frac{e_y}{x+y} = \frac{e_x}{x+y} \cdot \frac{x}{x} + \frac{e_y}{x+y} \cdot \frac{y}{y} = \frac{x}{x+y}\varepsilon_x + \frac{y}{x+y}\varepsilon_y$$

de donde, el error relativo de una suma es igual a la suma del error relativo de las sumandos afectados por coeficientes de ponderación iguales al valor del sumando dividido por la suma.

$$\varepsilon_z = \frac{x}{x+y}\varepsilon_x + \frac{y}{x+y}\varepsilon_y$$

26 **Sustracción:**

Sea ahora calcular

$$z = x - y$$

con la misma notación anterior, resulta

$$z = x_t + e_x - (y_t + e_y) = x_t - y_t + (e_x - e_y)$$

donde $z_t = x_t - y_t$

$e_z = e_x - e_y$

En términos de error relativo

$$\varepsilon_z = \frac{e_z}{z} = \frac{e_x - e_y}{x-y} = \frac{e_x}{x-y} - \frac{e_y}{x-y} = \frac{e_x}{x-y} \cdot \frac{x}{x} - \frac{e_y}{x-y} \cdot \frac{y}{y} = \frac{x}{x-y}\varepsilon_x - \frac{y}{x-y}\varepsilon_y$$

de donde, el error relativo de una resta es igual a la diferencia de los errores relativos de minuendo y sustraendo afectados por coeficientes de ponderación iguales a cada uno de ellos dividido por la diferencia.

$$\varepsilon_z = \frac{x}{x-y}\varepsilon_x - \frac{y}{x-y}\varepsilon_y$$

27 **Cuidado:** se han respetado los signos matemáticos resultantes de la deducción pero debe tenerse en cuenta que, el caso más desfavorable se presenta cuando los errores relativos de minuendo y sustraendo son de distinto signo, razón por la cual el error relativo de la diferencia será la suma de los términos de la expresión anterior.

28 Lo dicho en el párrafo anterior no es nada comparado con esta verdadera ***catástrofe:***

Sea $x_t = 10^3 \,.\, 0.9876578$

e $y_t = 10^3 \,.\, 0.9876577$

ambos en aritmética de t = 7 dígitos.

$z_t = x_t - y_t = 10^3 \,.\, 0.0000001 = 10^{-4} .\, 0.1000000$

Dado que tanto x como y llevan asociado un error relativo de corte 10^{-6}, por la aritmética utilizada, resulta:

$$\varepsilon_z = \frac{10^3 0.9876578}{10^{-4} 0.1000000} 10^{-6} + \frac{10^3 0.9876577}{10^{-4} 0.1000000} 10^{-6}$$

de donde

$$\varepsilon_z = 197.53155$$

y, en %

$$\varepsilon_z = 19753,3155\%$$

Alerta roja. Jamás reste números parecidos o muy parecidos porque pasan este tipo de cosas y, de allí en más, ningún resultado tiene sentido.

29 **Multiplicación:**

Sea ahora calcular

$$z = x \,.\, y$$

con la misma rotación anterior, resulta

$$z = (x_t + e_x)\,(y_t + e_y) = x_t\, y_t + x_t\, e_y + y_t\, e_x + e_x\, e_y$$

donde $z_t = x_t\, y_t$

$$e_z = x_t\, e_y + y_t\, e_x + e_x\, e_y$$

Se puede suponer que, si todo está razonablemente bien, e_x y e_y son pequeños, y que, en consecuencia su producto $e_x \,.\, e_y$ es mucho más pequeño y no será catastrófico olvidarse del último término de la expresión anterior.

Cuando se haya decolado y se adquiera nivel de vuelo de crucero se dirá simplemente que $e_x\, e_y$ es infinitésimo de orden superior al de los otros términos y se logrará el mismo resultado.

Entonces, en términos de error relativo

$$\varepsilon_z = \frac{e_z}{z} = \frac{x_t e_y + y_t e_x}{x.y} = \frac{x_t e_y}{x.y} + \frac{y_t e_x}{x.y}$$

de donde, aproximadamente

$$\varepsilon_z = \varepsilon_x + \varepsilon_y$$

30 **División:**

Sea, por fin, calcular

$$z = \frac{x}{y}$$

en este caso resulta

$$z = \frac{x_t + e_x}{y_t + e_y}$$

efectuando la división y olvidando términos relativamente muy pequeños (en vuelo de crucero, depreciando infinitésimos de orden superior), resulta:

$$z = \frac{x_t}{y_t} + \frac{e_x}{y_t} - \frac{x_t}{y_t^2} e_y$$

donde

$$z_t = \frac{x_t}{y_t}$$

y $$e_z = \frac{e_x}{y_t} - \frac{x_t}{y_t^2} e_y$$

y, en términos de error relativo

$$\varepsilon_z = \varepsilon_x - \varepsilon_y$$

Un detalle operativo

31 Sean x e y números en aritmética de t = 5 dígitos con los siguientes valores:

$$x = 10^2\ 0.\ 37425$$

$$y = 10^{-1}\ 0.\ 43258$$

Si se tiene que calcular $z_t = x_t + y_t$ ambos números deben estar expresados con el mismo exponente. Para ello hay que desplazar a derecha la mantisa de y hasta que su exponente sea 2 pero, teniendo en cuenta que desplazar a la derecha equivale a dividir, habría que multiplicar por la potencia de diez que corresponde para que su valor no cambie, y se obtenga un exponente igual a dos, entonces:

$$10^{-1}\ 0.\ 43258 = 10^2\ 0.\ 00043$$

2, 5 y 8
se caen

Al hacer esto, en aritmética de t = 5 dígitos, hay dígitos que se pierden. Literalmente se caen del registro en el que estaban grabados por la muy sencilla razón que el registro no tiene lugar donde grabarlos. Solamente pueden registrarse en él t = 5 dígitos.

Por este motivo cuando se hacen adiciones y sustracciones en aritmética de t dígitos se hace lo mismo que al trabajar manualmente: se habilita más lugar para minimizar el problema.

Usualmente se toman para estas operaciones una aritmética de 2t dígitos con la cual

$$x_{2t} = 10^2\ 0.\ 3742500000$$

$$y_{2t} = 10^2\ 0.\ 0004325800$$

$$z_{2t} = 10^2\ 0.\ 3746825800$$

Una vez hecha la operación, se vuelve a la aritmética de t dígitos, introduciéndose un error por esta reducción de mantisa.

Algo similar ocurre al calcular el producto $x_t\ y_t$. El exponente es la suma algebraica de los exponentes dados 2 + (-1) y la mantisa es el producto de las mantisas de los números dados. Esta operación tiene, en general, 2t dígitos que deben ser considerados antes de proceder a su reducción a t dígitos, con el consecuente error.

Lo mismo ocurre en las sustracciones y divisiones, razón por la cual, en general, puede decirse que cada una de las operaciones fundamentales lleva asociado, también, un error, que en lo sucesivo, se notarán:

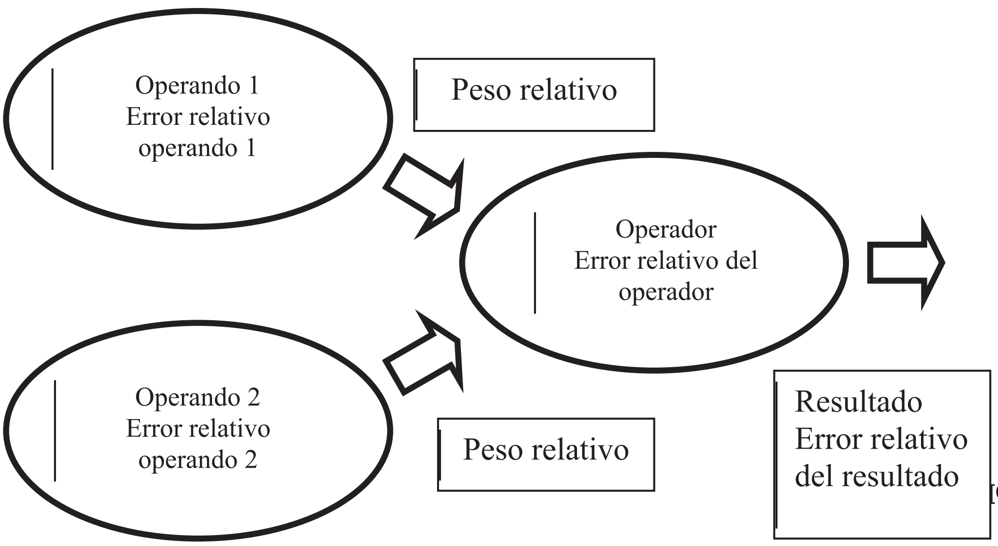
$$\varepsilon_{+}\ ;\ \varepsilon_{-}\ ;\ \varepsilon_{x}\ \text{ y }\ \varepsilon_{\%}$$

Errores que pueden leerse "error en la máquina de sumar"; "error en la máquina de restar"; "error en la máquina de multiplicar" y "error en la máquina de dividir" respectivamente.

Con estos elementos se está en condiciones de establecer cómo se propaga el error.

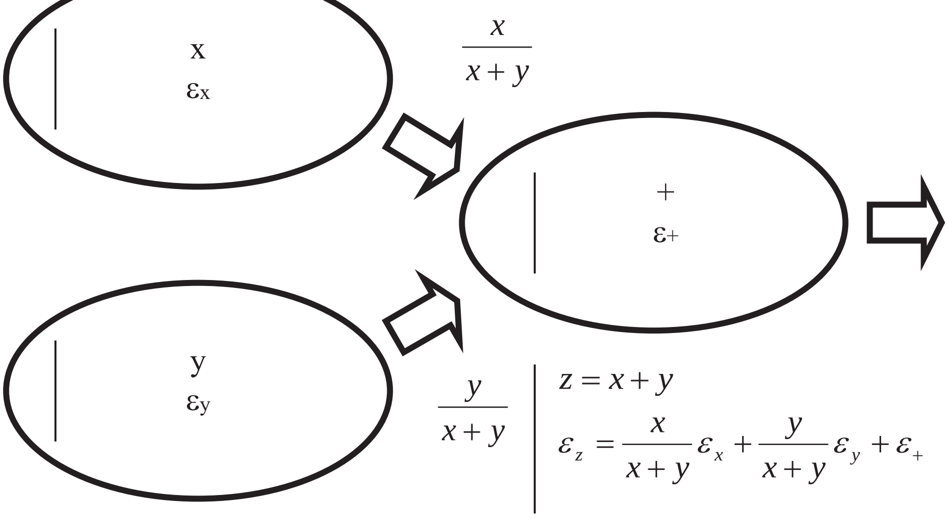
Gráfico de procesos:

32 Cada una de las cuatro operaciones fundamentales se representa mediante el siguiente esquema



En el lugar de los operandos se coloca el nombre de los números (o los números) que intervienen en la operación; en el lugar del operador se coloca el símbolo correspondiente + ; -; x o %; en peso relativo los coeficientes de ponderación mencionados al analizar la propagación del error relativo y un error ε_+ ; ε_- ; ε_x o $\varepsilon_\%$.

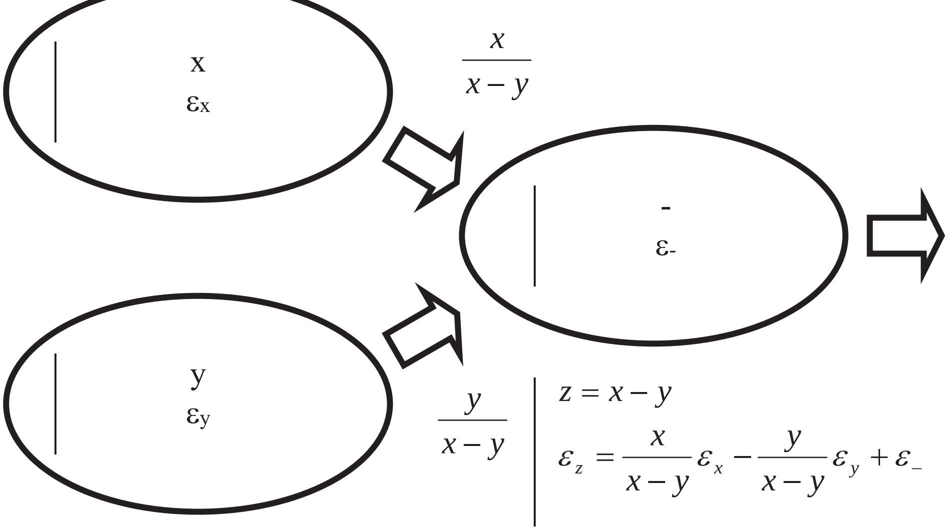
Así, para la adición



$$z = x + y$$

$$\varepsilon_z = \frac{x}{x+y}\varepsilon_x + \frac{y}{x+y}\varepsilon_y + \varepsilon_+$$

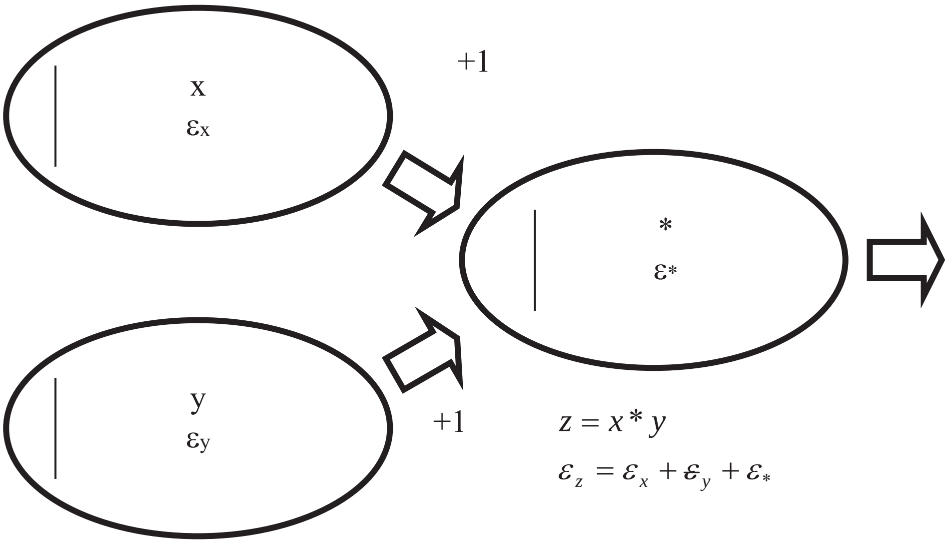
Para la sustracción



$$z = x - y$$

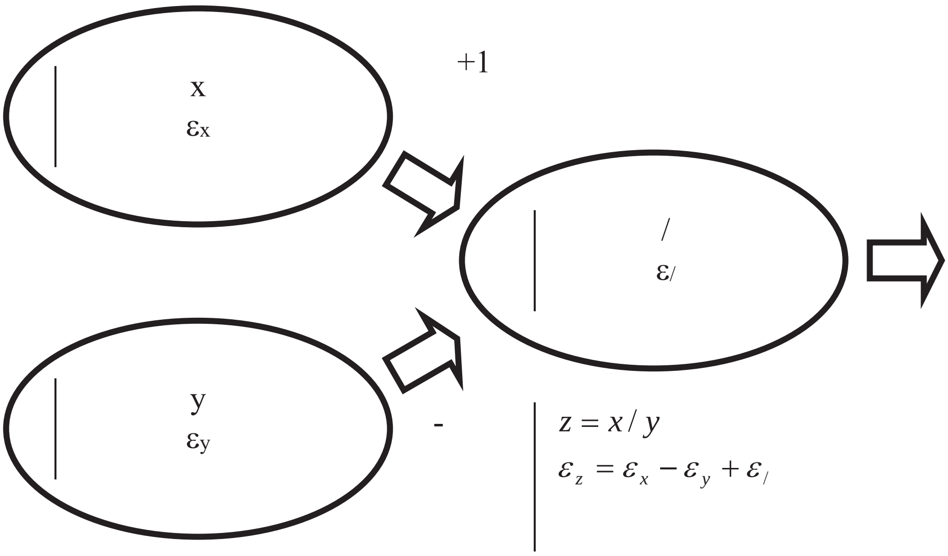
$$\varepsilon_z = \frac{x}{x-y}\varepsilon_x - \frac{y}{x-y}\varepsilon_y + \varepsilon_-$$

para la multiplicación



$$z = x * y$$
$$\varepsilon_z = \varepsilon_x + \varepsilon_y + \varepsilon_*$$

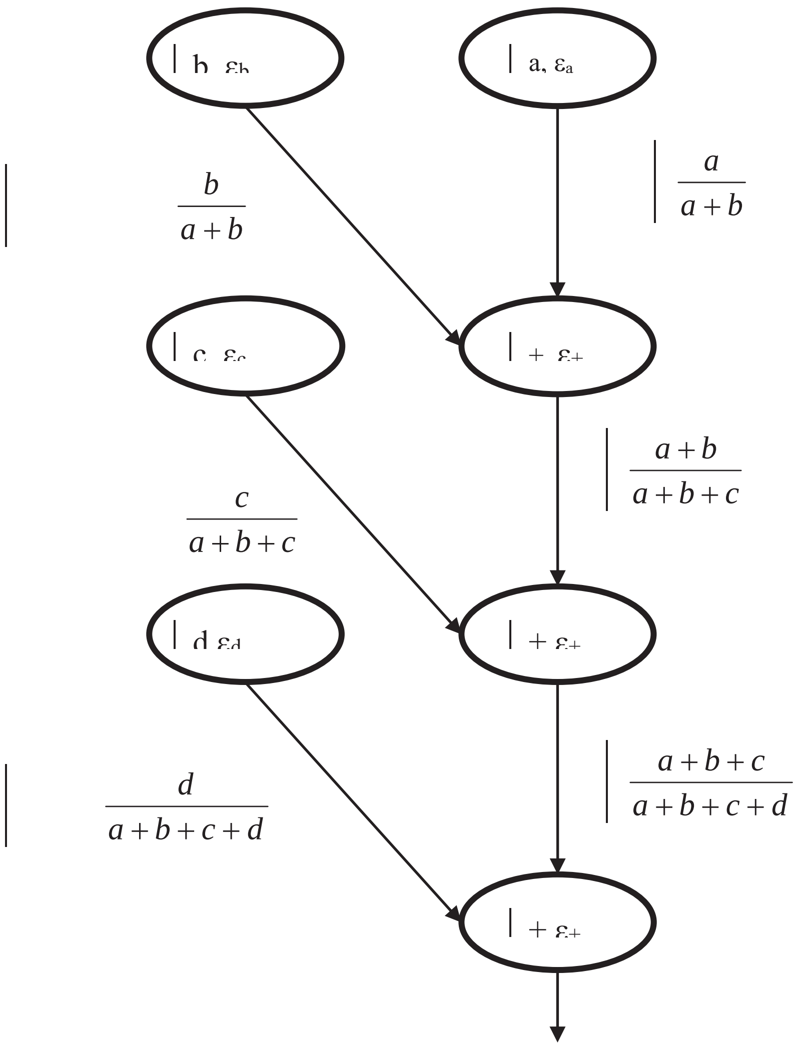
y para la división



$$z = x / y$$
$$\varepsilon_z = \varepsilon_x - \varepsilon_y + \varepsilon_/$$

Combinando entre sí estos gráficos surge la secuencia de operaciones a realizar y se obtiene un medio apto para analizar el error relativo en el resultado.

33 Para estudiar, por ejemplo el error resultante de la suma de cuatro números, a, b, c d puede utilizarse el siguiente gráfico:



Obsérvese que este gráfico representa el algoritmo de la suma en la manera habitual de hacerlo. Se suman los dos primeros números, en este caso a y b, luego a la suma a+b se le suma el número c y, por último a la suma a+b+c se le suma el número c. Esta última suma es el número z buscado.

34 Para estudiar los errores relativo y absoluto en el resultado z se escribe

$$\left[\left(\frac{a}{a+b}\varepsilon_a+\frac{b}{a+b}\varepsilon_b+\varepsilon_+\right)\frac{a+b}{a+b+c}+\frac{c}{a+b+c}\varepsilon_c+\varepsilon_+\right]\frac{a+b+c}{a+b+c+d}+\frac{d}{a+b+c+d}\varepsilon_d+\varepsilon_+=\varepsilon_z$$

donde se ha seguido formalmente lo consignado en el gráfico de procesos correspondiente a la suma, afectando a las sucesivas sumas por los factores de peso correspondientes.

35 A efectos de poner en evidencia un hecho sorpresivo, se toma como hipótesis que cada uno de los números a, b, c y d son almacenados sin errores relativos, con lo cual sólo afecta al resultado el error relativo correspondiente a la "máquina de sumar" ε_+. La expresión del párrafo anterior se transforma de la siguiente manera

$$\frac{a+b}{a+b+c+d}\varepsilon_+ + \frac{a+b+c}{a+b+c+d}\varepsilon_+ + \varepsilon_+ = \varepsilon_z$$

$$\frac{(a+b)\varepsilon_+ + (a+b+c)\varepsilon_+ + (a+b+c+d)\varepsilon_+}{a+b+c+d} = \varepsilon_z$$

$$\frac{(3a+3b+2c+d)\varepsilon_+}{a+b+c+d} = \varepsilon_z$$

36 Siendo el denominador de la última expresión escrita el valor de la suma z buscada, el error absoluto será

$$e_z = \varepsilon_z z = \varepsilon_z(a+b+c+d) = (3a+3b+2c+d)\varepsilon_z$$

Lo que implica que el error absoluto con que se obtiene la suma z depende del orden en que se consideran los sumandos. En efecto, si en los primeros pasos se toman los sumandos más grandes, el error absoluto será mayor al que se obtiene cuando se consideran en primera instancia los sumandos más chicos.

37 Sean, por ejemplo, a = 20; b = 17; c = 8 y d = 4. Si se hace a+b = 20 + 17 =37; se le suma c = 8 resulta 37 + 8 = 45 y, por último 45 + 4 = 49 pero el error absoluto en este caso es

$$(3*20+3*17+2*8+4)\varepsilon_+ = (60+51+16+4)\varepsilon_+ = 131\varepsilon_+$$

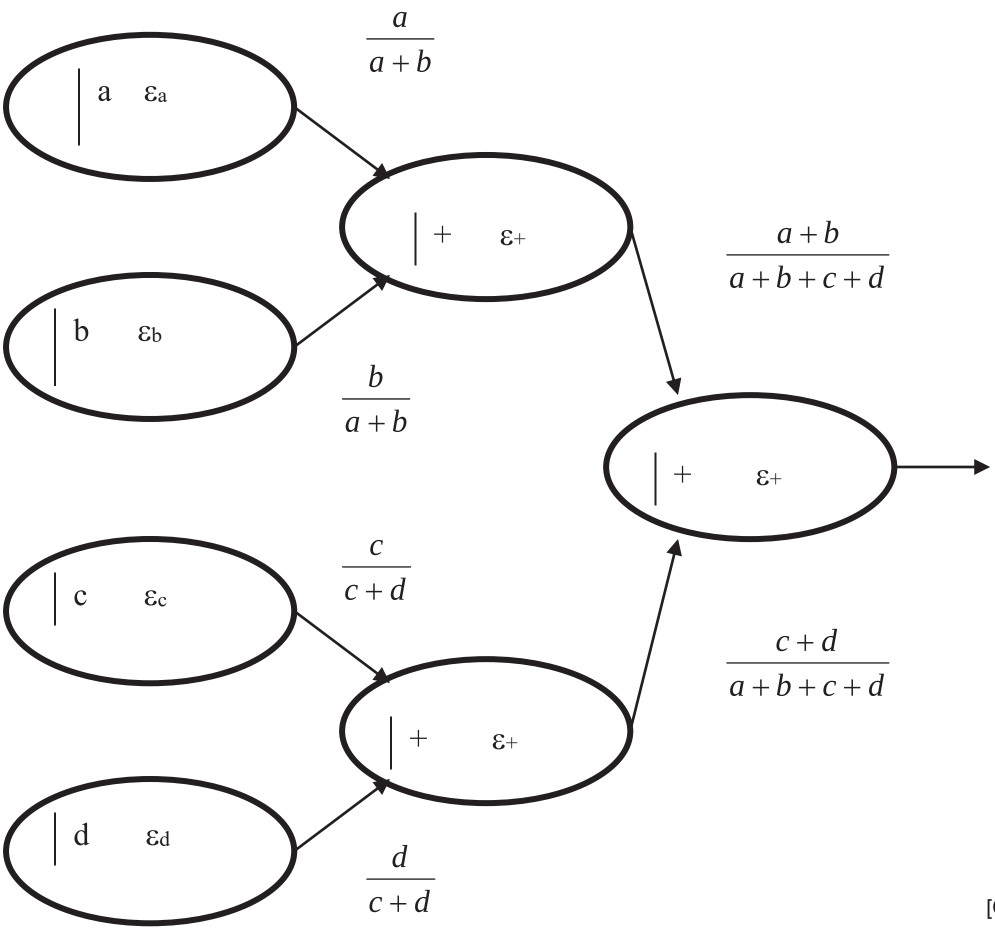
Si, en cambio, se consideran en primera instancia los sumandos más chicos, resulta

$$(3*4+3*8+2*17+20)\varepsilon_z = (12+24+34+20)\varepsilon_z = 90\varepsilon_+$$

68.70%

38 Esta característica debe ser tomada en cuenta cuando, por ejemplo, se efectúa un refinamiento de la solución de un sistema de ecuaciones lineales puesto que las correcciones, mucho más chicas que los valores originales deben ser tomadas como primeros sumandos para evitar fenómenos como el expuesto.

39 Otra forma de efectuar esa suma es calcular, primero a+b; luego c+d y por último
z = a+ b +c + d como indica el siguiente gráfico de proceso.



40 Para estudiar los errores relativo y absoluto en el resultado z se escribe

$$\varepsilon_z = \left(\frac{a}{a+b}\varepsilon_a + \frac{b}{a+b}\varepsilon_b + \varepsilon_+ \right)\frac{a+b}{a+b+c+d} + \left(\frac{c}{c+d}\varepsilon_c + \frac{d}{c+d}\varepsilon_d + \varepsilon_+ \right)\frac{c+d}{a+b+c+d} + \varepsilon_+$$

41 Haciendo la misma hipótesis del caso anterior, resulta

$$\varepsilon_z = \frac{(a+b)\varepsilon_+}{a+b+c+d} + \frac{(c+d)\varepsilon_+}{a+b+c+d} + \varepsilon_+ = \frac{(a+b)\varepsilon_+ + (c+d)\varepsilon_+ + (a+b+c+d)\varepsilon_+}{a+b+c+d} = \frac{2(a+b+c+d)}{a+b+c+d}\varepsilon_+$$

En términos de error absoluto queda

$$e_z = \varepsilon_z z = \varepsilon_z (a+b+c+d) = 2(a+b+c+d)\varepsilon_+$$

Donde resulta evidente que el orden en que se consideran los sumandos no influye en la suma.

42 Con los valores numéricos a = 20; b = 17; c = 8 y d = 4 anteriores resulta

$$e_z = 2(20+17+8+4)\varepsilon_+ = 2*49\varepsilon_+ = 98\varepsilon_+$$

Lo que indica que el procedimiento, si bien no requiere una especial consideración de orden sobre la magnitud de los sumandos, arroja un resultado mejor que el primero tratado y un poco peor que el segundo, siempre en términos del error absoluto del resultado.

volver al Índice General