NOUVELLES ANNALES DE MATHÉMATIQUES

GUYOU

Sur les approximations numériques

Nouvelles annales de mathématiques 3e série, tome 8 (1889), p. 165-186

<http://www.numdam.org/item?id=NAM_1889_3_8__165_0>

SUR LES APPROXIMATIONS NUMÉRIQUES;

PAR M. GUYOU,
Examinateur d'admission à l'École Navale.

I. — CLASSIFICATION DES ERREURS; OBJET ET DIVISION DU PROBLÈME DES APPROXIMATIONS NUMÉRIQUES; FORMULE GÉNÉRALE DES ERREURS.

L'introduction du problème des approximations numériques dans les programmes d'Arithmétique, très justifiée quand il s'agit d'élèves ne devant pas dépasser les notions les plus élémentaires, me paraît constituer une surcharge inutile pour ceux qui ont acquis les premières notions sur les dérivées; il y aurait, je crois, avantage, pour ces derniers, à restituer à la question sa véritable place, en Algèbre, où elle pourrait être traitée plus simplement et constituerait une sorte d'introduction à la théorie générale des erreurs.

L'ensemble du sujet pourrait être exposé ainsi qu'il suit; je considère le cas où les calculs doivent être effectués sans Tables.

Classification des erreurs du résultat d'une formule numérique. — Le plus généralement, les formules numériques contiennent des nombres obtenus par des mesures directes; ces nombres sont toujours affectés d'erreurs dues à l'imperfection des instruments et des procédés de mesure; de sorte que ces formules, au lieu de représenter la valeur exacte du nombre cherché, n'en représentent qu'une valeur affectée d'une certaine erreur que nous appellerons *erreur provenant des données*.

En outre, lorsque l'on effectuera les opérations indiquées dans la formule donnée, on commettra en général une seconde erreur, que nous appellerons *erreur de calcul.*

L'*erreur totale du résultat* sera égale à la somme algébrique de ces deux erreurs partielles; si l'on désigne en effet par N la valeur exacte du nombre cherché, par N' la valeur exacte du nombre représenté par l'expression numérique dans laquelle on a introduit les nombres obtenus par des mesures directes, et enfin par N'' la valeur approchée de N' obtenue par le calcul, on a, en grandeur et en signe :

Erreur provenant des données.....	$N' - N$
Erreur de calcul..................	$N'' - N'$
Erreur totale....................	$N'' - N$

et enfin

$$N'' - N = (N'' - N') + (N' - N).$$

Les erreurs commises sur les données sont toujours inconnues en grandeur et en sens, mais on peut toujours assigner à la valeur absolue de chacune d'elles une limite supérieure, et, de ces limites supérieures, on peut déduire une limite supérieure de l'erreur provenant des données ($N' - N$); c'est-à-dire une limite que $N' - N$ peut atteindre, mais qu'elle ne peut dépasser.

L'erreur de calcul $N'' - N'$, qui s'ajoute à celle-ci, peut être rendue aussi petite que l'on voudra; mais, lors même qu'elle serait nulle, on ne pourra pas compter au résultat final sur une approximation plus grande que la limite supérieure de l'autre erreur.

Par conséquent, la limite supérieure de l'erreur provenant des données représente l'approximation la plus grande sur laquelle on puisse compter au résultat final.

Objet et division du problème des approximations numériques. — Désignons par $f(a, b, c, \ldots)$ une formule numérique dans laquelle a, b, c représentent les valeurs exactes de certains nombres que l'on obtient par des mesures directes, et soit N la valeur exacte de cette formule :

$$N = f(a, b, c, \ldots);$$

soient a', b', c', ... les valeurs approchées obtenues pour les nombres a, b, c, ...

Le problème des approximations numériques consiste à calculer la valeur de N à une approximation donnée, connaissant a', b', c', ... et les limites supérieures des erreurs commises sur ces nombres.

Il est clair que le problème ne sera possible que si l'erreur provenant des données ne peut pas atteindre l'approximation demandée : il conviendra donc tout d'abord de déterminer la limite supérieure de cette erreur. Il faudra ensuite calculer l'expression

$$N' = f(a', b', c', \ldots)$$

avec une approximation telle que la somme de l'erreur de calcul et de la limite supérieure précédemment obtenue soit au plus égale à l'approximation demandée.

On est ainsi conduit à diviser le problème en deux parties :

1° Déterminer une limite supérieure de l'erreur provenant des données ;

2° Calculer avec une approximation donnée une formule numérique donnée.

Il est clair que, lorsque les nombres contenus dans la formule seront tous connus exactement, ou, du moins, seront susceptibles d'être exprimés avec une approximation indéfinie, la deuxième partie du problème subsistera seule.

La solution de ces deux problèmes est fournie par l'application de la formule générale que nous allons établir.

Formule générale des erreurs. — Soit $f(x, y, z)$ une fonction de trois variables indépendantes; désignons par δf l'accroissement que prend cette fonction lorsque, après avoir donné aux variables les valeurs a, b, c, on leur donne les valeurs $a + \delta\alpha$, $b + \delta b$, $c + \delta c$; on aura

$$\delta f = f(a + \delta a,\ b + \delta b,\ c + \delta c) - f(a,\ b,\ c).$$

que l'on peut écrire identiquement

$$\begin{aligned}\delta f = f(a + \delta a,\ b + \delta b,\ c + \delta c) &- f(a,\ b - \delta b,\ c - \delta c)\\ + f(a,\ b + \delta b,\ c + \delta c) &- f(a,\ b,\ c + \delta c)\\ + f(a,\ b,\ c + \delta c) &- f(a,\ b,\ c).\end{aligned}$$

Appliquant à chacune de ces différences la formule des accroissements finis, on obtient

$$\begin{aligned}\delta f = \delta a f'_a(a &+ 0\,\delta a,\ b - \delta b,\ c + \delta c)\\ + \delta b f'_b(a,\ b &+ 0'\,\delta b,\ c + \delta c)\\ + \delta c f'_c(a,\ b,\ c &+ 0''\,\delta c).\end{aligned}$$

Désignons d'une manière générale par a', a'', a''', b', b'', b''', c', c'', c''' des nombres compris entre a et $a + \delta a$, b et $b + \delta b$, c et $c + \delta c$, ou égaux à ces limites elles-mêmes, on pourra écrire plus simplement

$$\delta f = \delta a f'_a(a', b', c') + \delta b f'_b(a'', b'', c'') + \delta c f'_c(a''', b''', c''').$$

Cette formule donne une expression de l'erreur que l'on commet lorsque, dans une formule numérique

$$f(a, b, c),$$

on remplace les nombres exacts a, b, c par les valeurs approchées $a + \delta a$, $b + \delta b$, $c + \delta c$; et il est clair qu'elle est applicable à un nombre quelconque de valeurs approchées a, b, c, ...

Les nombres a', b', c', a'', b'', ... qu'il faut introduire dans les dérivées partielles f'_a, f''_b, f'_c, ..., pour obtenir les facteurs des erreurs δa, δb, δc, sont inconnus, mais, en prenant pour ces nombres leurs valeurs maxima ou leurs valeurs minima, de manière à donner à ces facteurs les plus grandes valeurs qu'ils puissent prendre, on pourra déduire de cette formule une limite supérieure de l'erreur δf.

Remarque. — Il n'est pas inutile de faire remarquer ici que les erreurs δa, δb, δc, ... étant en général très petites, les valeurs des dérivées f'_a, f'_b, ... varient très peu lorsque l'on fait varier les nombres a', b', c', ... entre leurs limites extrêmes ; par conséquent, cette formule permet d'obtenir non seulement une limite supérieure de l'erreur δf, mais encore une limite très voisine de la valeur de l'erreur elle-même.

II. — Première partie du problème des approximations numériques.

Soit proposé de calculer l'erreur provenant des données a, b, c, ..., sur le résultat d'une formule numérique

$$f(a, b, c, \ldots)$$

dans laquelle les nombres a, b, c sont connus à $\pm \Delta a$, $\pm \Delta b$, $\pm \Delta c$ près.

On a, en désignant par δf, δa, δb, ... les valeurs exactes des erreurs en grandeur et en signe

$$\begin{aligned}\delta f = {} & \delta a f'_a(a', b', c', \ldots) \\ & + \delta b f'_b(a'', b'', c'', \ldots) + \delta c f'_c(a''', b''', c''', \ldots) + \ldots.\end{aligned}$$

Désignons par a_1, b_1, c_1 les valeurs approchées connues de a, b, c.

Les sens et les grandeurs des erreurs δa, δb, δc, ... sont inconnus, mais leurs valeurs absolues sont plus petites que Δa, Δb, Δc, ...; les nombres a', b', c' compris entre a et a_1, b et b_1, c et c_1 sont *a fortiori* compris entre

$$\begin{array}{lcl} a_1 + \Delta a & \text{et} & a_1 - \Delta a, \\ b_1 + \Delta b & \text{et} & b_1 - \Delta b, \\ c_1 - \Delta c & \text{et} & c_1 - \Delta c; \end{array}$$

si, par conséquent, dans les dérivées partielles, on remplace le nombre a partout où il se trouve, soit par $a_1 + \Delta a$, soit par $a_1 - \Delta a$, de manière à forcer la valeur de cette dérivée, et que l'on agisse de même pour les nombres b, c, ..., on obtiendra des valeurs supérieures à celles que pourront prendre les facteurs des erreurs des données.

On obtiendra donc une limite supérieure de l'erreur δf en multipliant ces limites supérieures des dérivées partielles respectivement par Δa, Δb, Δc, ..., et en faisant la somme des valeurs absolues des produits.

Remarque. — Pour montrer par un exemple comment on peut obtenir des limites supérieures des valeurs des dérivées, supposons que l'on ait

$$f'_c = \frac{ac}{\sqrt{a-b}+b};$$

les valeurs approchées connues a_1, b_1, c_1 et les valeurs exactes a, b, c sont comprises entre

$$\begin{array}{lcl} a_1 - \Delta a & \text{et} & a_1 + \Delta a, \\ b_1 - \Delta b & \text{et} & b_1 + \Delta b, \\ c_1 - \Delta c & \text{et} & c_1 + \Delta c. \end{array}$$

Pour obtenir la valeur exacte du coefficient de δc, il faudrait remplacer a, b, c par des valeurs a', b', c' comprises entre a et a_1, b et b_1, c et c_1 ou égales à ces limites; en remplaçant a au numérateur par $a_1 + \Delta a_1$ et au dénominateur par $a_1 - \Delta a_1$, nous augmenterons le numérateur et nous diminuerons le dénominateur; nous forcerons donc la valeur de l'expression. Nous obtiendrons le même résultat en remplaçant au numérateur c par $c_1 + \Delta c$, et enfin en remplaçant au dénominateur b par $b_1 + \Delta b$ sous le radical et par $b_1 - \Delta b$ en dehors; on aura ainsi évidemment

$$f'_c(a', b', c') < \frac{(a_1 + \Delta a)(c_1 + \Delta c)}{\sqrt{(a_1 - \Delta a) - (b_1 + \Delta b)} + (b_1 - \Delta b)}.$$

On peut donc formuler la règle suivante :

Pour obtenir une limite supérieure de l'erreur provenant de données incertaines, on remplacera dans la formule numérique ces données par des lettres a, b, c, ...; on prendra les dérivées partielles de l'expression ainsi obtenue par rapport à chacune d'elles; on remplacera dans ces dérivées les lettres a, b, c, ... par des valeurs limites par excès ou par défaut des nombres qu'elles représentent, de manière à forcer leurs valeurs; on multipliera les résultats obtenus respectivement par les limites des approximations Δa, Δb, Δc, ..., *et l'on fera la somme des valeurs absolues des produits.*

Exemple I. — Dans la formule numérique

$$N = \frac{(0,117\pi + 1,43)^2}{\sqrt{2\sqrt{3} + 5}},$$

les nombres 0,117 et 1,43, obtenus par mesures directes, sont connus à une unité près de l'ordre de leur dernier chiffre; on demande une limite supérieure de l'erreur qui peut affecter le résultat N.

En suivant textuellement la règle, nous avons

$$f(a, b) = \frac{(a\pi + b)^2}{\sqrt{2\sqrt{3} + 5}},$$

$$f'_a = \frac{2\pi(a\pi + b)}{\sqrt{2\sqrt{3} + 5}}, \qquad f'_b = \frac{2(a\pi + b)}{\sqrt{2\sqrt{3} + 5}}.$$

Remplaçons aux numérateurs a et b par des valeurs par excès 0,118 et 1,44, il vient

$$f'_a < \frac{2\pi(0,118\pi + 1,44)}{\sqrt{2\sqrt{3} + 5}}, \qquad f'_b < \frac{2(0,118\pi + 1,44)}{\sqrt{2\sqrt{3} + 5}}.$$

En multipliant ces nombres par 0,001 et par 0,01 et en faisant la somme, nous obtiendrons non seulement une limite supérieure de l'erreur, mais encore une limite très voisine de celle qui peut être atteinte, puisque Δa et Δb peuvent atteindre 0,001 et 0,01; mais dans l'application on n'a pas intérêt à connaître une valeur aussi précise, et l'on évalue en nombres ronds les valeurs des facteurs des erreurs, en ayant soin toutefois de toujours arrondir les nombres de manière à forcer les valeurs de ces facteurs; ainsi au numérateur nous remplacerons π par 4, au dénominateur $\sqrt{3}$ par 1 et $\sqrt{7}$ par 2; on aura ainsi

$$f'_a < \frac{8(0,472 + 1,44)}{2} < 8.$$

$$f'_b < \frac{2 \times 1,912}{2} < 2.$$

On a donc finalement

$$\delta N < 8 \times 0,001 + 2 \times 0,01 \qquad \text{ou} \qquad \delta N < 0,028.$$

Remarque. — Il ne faut pas perdre de vue que, lors même que les dérivées eussent été négatives, il aurait fallu faire la somme des produits.

Exemple II. — Dans un triangle ABC rectangle en A, on a obtenu par mesures directes

$$c = 12^{m},5 \text{ à } \pm 0,1 \text{ près},$$
$$C = 22^{\circ} \quad \text{à } \pm 30' \text{ près}.$$

On demande une limite supérieure de l'erreur qui peut affecter le côté b calculé avec ces éléments.

On a

$$b = c \cot C = f(c, C),$$

$$f'_c = \cot C, \qquad f'_C = \frac{-c}{\sin^2 C}$$

et, par suite,

$$f'_c < \cot 21^{\circ}30', \qquad f'_C < \frac{12,6}{\sin^2 21^{\circ}30'}.$$

On trouve, à l'aide d'une Table des valeurs naturelles des lignes trigonométriques,

$$f'_c < 2,6, \qquad f'_C < 12,6 \times (2,7)^2 < 117;$$

il reste à multiplier par Δc et ΔC. On ne doit pas perdre de vue ici que les expressions usuelles des dérivées des lignes trigonométriques ont été obtenues en supposant l'accroissement de l'arc exprimé en fonction du rayon; on devra donc exprimer ΔC de cette manière : on aura ainsi

$$\Delta c = 0,1, \qquad \Delta C = \frac{\pi}{360};$$

il viendra donc enfin, en faisant les produits et ajoutant les valeurs absolues,

$$\delta b < 2,6 \times 0,1 + \frac{117\pi}{360},$$

$$\delta b < 0,26 + \frac{120}{360}\pi,$$

$$\delta b < 0,26 + \frac{\pi}{3} \qquad \text{ou} \qquad < 1^{m},31.$$

Remarque. — La formule que nous avons considérée ici ne peut pas être calculée arithmétiquement, mais, cette première partie du problème des approximations étant indépendante du procédé de calcul, la même méthode convient à tous les cas.

III. — Deuxième partie du problème des approximations.

Actuellement, sans nous préoccuper de l'origine ni de l'exactitude des nombres donnés, nous avons à calculer à une approximation donnée le résultat d'une formule numérique donnée.

Résultat complet d'une formule ou d'une opération arithmétique. — Nous appellerons résultat complet d'une formule numérique complexe ou d'une opération arithmétique simple la valeur exacte de cette formule ou du résultat de l'opération; ce résultat ne peut être exprimé le plus souvent qu'à l'aide d'un nombre indéfini de décimales ou du moins à l'aide d'un nombre de décimales supérieur à celui dont on a besoin.

Ordre du dernier chiffre à conserver dans le résultat approché d'une formule numérique. — Il ne suffit pas, pour résoudre le problème que nous avons en vue, de calculer un nombre différant du résultat complet de la formule d'une erreur moindre qu'une quantité donnée, il faut encore que ce résultat soit exprimé avec le plus petit nombre possible de chiffres; or, quel que soit le degré de précision avec lequel on effectuera les calculs, le résultat complet de la dernière opération différera toujours du résultat complet de la formule elle-même d'une petite quantité, et, si l'on supprime les chiffres à partir d'un certain ordre, même en forçant le dernier chiffre conservé, on ajoutera à cette erreur une nouvelle erreur qui pourra atteindre 5 unités de l'ordre du premier chiffre supprimé; la nouvelle erreur pourra, il est vrai, être moindre, mais, comme le nombre cherché est inconnu, on ne peut pas affirmer *a priori* qu'elle sera moindre que 5 unités de l'ordre du premier chiffre supprimé.

Supposons que l'on demande le résultat à n unités de l'ordre α, n étant au moins égal à un; si n est plus grand que 5, soit 7 par exemple, on pourra exprimer le résultat en unités de l'ordre qui précède α : il suffira en effet de conduire les opérations de manière que le résultat complet de la dernière, celle qui donne le résultat final, soit erroné de moins de deux unités en plus ou en moins de l'ordre α, car, en supprimant les chiffres de l'ordre $\alpha+1$ et en forçant au besoin le dernier chiffre conservé, on ajoutera à cette erreur une erreur moindre que 5 : le résultat conservé sera donc erroné de moins de 7 unités en plus ou en moins. Mais si n est plus petit que 5, on ne pourra pas faire cette suppression; pour n'avoir qu'une seule règle convenant à tous les cas, nous conviendrons que, lorsque l'on demandera le résultat d'une formule numérique à moins de n unités ($n > 1$) de l'ordre α, nous devrons exprimer ce résultat en unités de cet ordre.

Ainsi un résultat demandé à moins de $\pm$ 0,00352, c'est-à-dire de $\pm$ 3^{mm},52, devra être exprimé en millièmes; en d'autres termes, *l'ordre du dernier chiffre conservé devra être celui du premier chiffre significatif à droite du nombre qui exprime l'approximation.*

Des formules numériques simples. — Nous appellerons *formule numérique simple* une formule numérique dont le résultat peut s'obtenir par une seule opération arithmétique; ainsi les formules

$$\sqrt{3,58932}, \quad \frac{42,85741}{27,412}, \quad 0,854 + 0,4305 + 7,325$$

sont des formules numériques simples.

Mais une formule telle que $(6,45)^3$ n'est pas une formule simple, car le résultat ne peut être obtenu que par deux multiplications.

Lorsque l'on aura à calculer la valeur d'une expression de ce genre à une approximation donnée, on sera en général conduit à remplacer les valeurs exactes des nombres qui entrent dans l'expression donnée par des valeurs approchées plus simples; si l'approximation du résultat complet de la nouvelle expression est de $\pm p$ unités de l'ordre α et que l'on supprime les chiffres de l'ordre $\alpha + 1$ en forçant au besoin le dernier chiffre conservé, on commettra une nouvelle erreur plus petite que ± 5 unités de l'ordre $\alpha + 1$, ou $\pm \frac{1}{2}$ unité de l'ordre α: le résultat conservé sera donc approché à moins de $\pm (p + \frac{1}{2})$ unités de l'ordre α. On est donc conduit à la règle suivante, applicable à tous les cas :

Si l'on veut obtenir le résultat d'une opération simple à moins de n unités de l'ordre $\alpha (n > 1)$, on pourra substituer aux nombres à opérer des nombres tels que la somme des influences des erreurs sur le résultat complet soit plus petite que $(n - \frac{1}{2})$ unités de l'ordre α; on effectuera ensuite l'opération sur ces nombres plus simples et l'on supprimera au résultat tous les chiffres qui suivent celui qui exprime des unités de l'ordre α, en forçant ce dernier d'une unité si le premier chiffre supprimé est plus grand que 5 ou égal à 5.

Remarque. — Lorsque les nombres à opérer sont donnés directement, au lieu d'être les résultats d'opérations antérieures, on disposera du sens des erreurs qui seront commises sur ces nombres lorsqu'on les remplacera par des nombres plus simples : on pourra ainsi prendre ces erreurs dans un sens tel que l'erreur sur le résultat complet soit dans un sens connu ; supposons que ce soit par excès : alors le résultat complet sera approché *par excès* à moins de n unités de l'ordre α ; si l'on supprime à ce résultat tous les chiffres qui suivent celui d'ordre α, on commettra une erreur *par défaut* ε moindre qu'une unité de cet ordre ; par conséquent, le résultat simplifié sera approché à moins de n unités de l'ordre α par excès, et à moins de ε par défaut ; ε étant plus petit que 1 et par suite que n, le résultat sera encore approché à moins de n unités d'ordre α, mais on ne connaîtra plus le sens de l'approximation.

On pourrait donc, dans ce cas, se borner à prendre les nombres à opérer avec une approximation telle que l'erreur sur le résultat complet soit plus petite que n unités de l'ordre α au lieu de $n - \frac{1}{2}$, comme dans le cas précédent ; mais l'avantage obtenu ainsi est si faible, lorsqu'il existe, qu'il est préférable de s'en tenir à la règle précédente, qui convient à tous les cas.

Calcul des formules numériques simples. — Une formule numérique simple peut contenir un nombre seulement, c'est le cas des carrés, des racines carrées et des racines cubiques, ou deux nombres, comme la multiplication, la division, la soustraction, ou enfin plusieurs nombres, comme l'addition.

Pour calculer à moins de n unités de l'ordre α, $(n > 1)$, la valeur d'une formule numérique simple, on commencera par remplacer par des lettres A, B, C, ... les nombres que l'on supposera susceptibles d'être simplifiés, on appliquera ensuite, à l'expression obtenue, la formule générale des erreurs.

On obtiendra ainsi une expression de la forme

$$\delta f = \mathrm{M}\,\delta\mathrm{A} + \mathrm{N}\,\delta\mathrm{B} + \mathrm{P}\,\delta\mathrm{C} + \ldots.$$

On déterminera, comme on va le voir, les valeurs limites à donner à δA, δB, δC, ... pour que δf soit plus petit que $n - \frac{1}{2}$ unités de l'ordre α; on prendra ensuite les nombres A_1, B_1, C_1, ... avec les approximations ainsi déterminées; on effectuera l'opération avec ces nombres, et l'on supprimera les chiffres qui suivent celui qui exprime des unités de l'ordre α, en forçant ce chiffre ou en le conservant, suivant la valeur du premier chiffre supprimé.

Dans le cas de l'addition et de la soustraction, la formule des erreurs est de la forme

$$\mathrm{S} = \mathrm{A} + \mathrm{B} + \mathrm{C}, \quad \ldots, \quad \delta\mathrm{S} = \delta\mathrm{A} + \delta\mathrm{B} + \delta\mathrm{C},$$
$$\mathrm{D} = \mathrm{A} - \mathrm{B}, \quad \ldots, \quad \delta\mathrm{D} = \delta\mathrm{A} - \delta\mathrm{B};$$

dans les autres cas, les facteurs M, N, P dépendent en général des nombres à opérer, et ces nombres doivent y être remplacés par des valeurs comprises entre les valeurs exactes et les valeurs approchées; ces dernières ne sont pas connues; il en est de même des premières lorsque les nombres doivent être déterminés par des opérations antérieures; mais, d'une manière générale, on peut toujours prendre deux valeurs assez écartées l'une de l'autre pour être certain *a priori* qu'elles comprennent la valeur exacte et la valeur approchée. On peut donc ainsi déterminer des limites M', N', P' supérieures des valeurs de ces facteurs.

On prendra alors pour δA, δB, δC des valeurs telles que la somme des valeurs absolues des produits $M'\delta A$,

$N' \delta B$, $P' \delta C$ soit plus petite que $n - \frac{1}{2}$ unités de l'ordre α; si l'on dispose du sens des erreurs δA, δB, δC, on pourra encore prendre ces erreurs de manière qu'une partie des produits ait le signe $+$ et l'autre partie le signe $-$, et l'on prendra ensuite les erreurs telles que chacune des sommes des produits de même signe soit plus petite que $n - \frac{1}{2}$ unités de l'ordre α, car la somme des erreurs sera plus petite que le plus grand des deux groupes.

Exemple. — Soit à calculer à $\pm 0,00475$ la valeur de la formule simple

$$Q = \frac{3,5684}{3,1415}.$$

Nous poserons

$$Q = \frac{A}{\pi};$$

on en déduira

$$\delta Q = \frac{1}{\pi'} \delta A - \frac{A'}{\pi'^2} \delta\pi:$$

l'approximation demandée est $0,00475$, soit $4^{\text{mm}},75 = n$; on devra prendre δA et $\delta\pi$ de manière que δQ soit plus petit que $(n - \frac{1}{2})$ millièmes ou $4^{\text{mm}},25$.

Si l'on prend δA et $\delta\pi$ par défaut l'un et l'autre, il suffira que chacun des deux termes soit plus petit que $(n - \frac{1}{2})$ millièmes; il faudra donc que

$$\frac{\delta A}{\pi'} < 4^{\text{mm}},25, \qquad \frac{A' \delta\pi}{\pi'^2} < 4^{\text{mm}},25.$$

Si l'on ne veut pas s'astreindre à fixer le sens des approximations δA et $\delta\pi$, on prendra chacun des termes plus petit que la moitié de $(n - \frac{1}{2})$ millièmes, savoir

$$\frac{\delta A}{\pi'} < 2^{\text{mm}},125, \qquad \frac{A\, \delta\pi}{\pi'^2} < 2^{\text{mm}},125.$$

Résolution des inégalités. — On sera donc ainsi conduit en général à résoudre les inégalités de la forme

$$f(A', B')\,\partial A < m.$$

Ces inégalités se résolvent à vue en nombres ronds, mais il importe de ne pas oublier qu'en arrondissant les nombres il faudra toujours agir de manière à forcer les premiers membres et à réduire les seconds, afin que les inégalités que l'on a à résoudre soient vérifiées *a fortiori;* de plus, il ne faudra pas hésiter à prendre pour A′ et B′ des valeurs maxima et minima assez écartées l'une de l'autre pour être certain qu'elles comprendront entre elles la valeur exacte et la valeur approchée, car si l'on prenait des limites trop rapprochées, on pourrait constater, une fois le calcul terminé, qu'elles ne comprennent pas ces deux valeurs, et il faudrait recommencer.

Exemple. — Considérons l'exemple cité plus haut, et résolvons les deux inégalités

$$\frac{\partial A}{\pi} < 2^{\text{mm}},125, \qquad \frac{A'\,\partial\pi}{\pi'^2} < 2^{\text{mm}},125.$$

Pour *forcer* les premiers membres, nous remplacerons π par la valeur par défaut 3, et A par la valeur par excès 4; on aura

$$\frac{\partial A}{3} < 2^{\text{mm}},125, \qquad \frac{4}{9}\,\partial\pi < 2^{\text{mm}},125;$$

chassant les dénominateurs et arrondissant les seconds membres en les *réduisant*, il vient

$$\partial A < 6^{\text{mm}}, \qquad 4\,\partial\pi < 18^{\text{mm}},$$
$$\partial\pi < 4^{\text{mm}};$$

on prendra donc

$$A = 3,57, \qquad \pi = 3,14:$$

on fera le quotient 3,57 par 3,14 jusqu'aux dix-millièmes ; on trouve

$$Q = 1,1369 ;$$

on supprime le chiffre 9 et l'on force le 6 d'une unité ; il vient donc enfin

$$Q = 1,137 \text{ à } \pm 4^{\text{mm}},75 \text{ près.}$$

Formules complexes. — Les formules numériques complexes sont celles dont la valeur ne peut être obtenue que par une série d'opérations simples successives, dont la dernière a pour résultat le nombre demandé.

Il est clair que, de l'approximation exigée pour le résultat final, on peut, en appliquant la règle qui précède, déduire l'approximation avec laquelle doivent être connus les éléments de la dernière opération ; de l'approximation de ces éléments, on pourra déduire celle qui est nécessaire pour les nombres à l'aide desquels on les obtient, et ainsi de suite ; on arrivera ainsi à connaître l'approximation des nombres de la première opération, on effectuera ensuite les opérations en s'arrêtant aux unités de l'ordre indiqué par l'approximation qu'on a reconnue être nécessaire pour les résultats, et en forçant au besoin le dernier chiffre.

Le calcul d'une formule complexe n'offre donc aucune difficulté nouvelle ; mais les calculs successifs doivent être faits avec beaucoup d'ordre, et nous engageons au moins les débutants à se conformer à la règle pratique énoncée ci-après.

Pour mieux faire saisir les différentes parties de cette règle, nous les appliquerons à l'exemple suivant, à mesure que nous les formulerons.

Exemple. — Calculer à $\pm 0,035$ la valeur de l'expression

$$X = \sqrt{\frac{45 \times 50,213642}{3,1415926}}.$$

Règle. — 1° *Remplacer par des lettres* A, B, C, ..., *dans l'expression donnée, le nombre ou les nombres auxquels on espère pouvoir substituer des nombres plus simples.*

On écrira ainsi

$$N = \sqrt{\frac{45 \times A}{\pi}}.$$

2° *Préparer un Tableau en quatre colonnes ayant respectivement pour titres :* (1) Opérations à effectuer, (2) Résultats approchés par excès et par défaut, (3) Approximations nécessaires, (4) Résultats définitifs.

(Les nombres inscrits dans les colonnes du Tableau ci-après seront obtenus comme on va l'indiquer.)

(1).	(2).		(3).	(4).
	Résultats approchés			
Opérations à effectuer.	par défaut.	par excès.	Approximations nécessaires.	Résultats définitifs.
A	50	51	0,01	50,21
$P = 45 \times A$	2200	2300	1	2259
π	3	4	0,001	3,141
$Q = \frac{P}{\pi}$	550	800	1,2	7,9
$N = \sqrt{Q}$	20	30	0,035	26,81

3° *Indiquer dans la colonne* (1) *les opérations à effectuer successivement, en désignant par de nouvelles lettres les résultats de ces opérations, et en ayant soin, lorsque, dans une de ces opérations, on aura à utiliser un des nombres qui ont été désignés au début par des lettres, de placer la lettre qui désigne ce nombre avant l'opération elle-même.*

Ainsi, pour l'exemple considéré, la première opération à effectuer est le produit $45 \times A$: on le désigne par

P et on inscrit avant P la lettre A; on aura ensuite à effectuer le coefficient $Q = \frac{P}{\pi}$; ayant à utiliser le nombre π dans cette opération, on inscrira cette lettre avant l'opération. On indiquera enfin la dernière opération à effectuer $N = \sqrt{Q}$; de cette manière, tous les nombres dont il y aura lieu de déterminer l'approximation sont inscrits dans la colonne (1) et dans l'ordre même dans lequel on aura à les employer.

4° *Dans la colonne* (2), *on inscrira des valeurs grossièrement approchées par défaut et par excès des nombres mentionnés dans la colonne* (1).

Il est possible que dans la suite quelques-uns de ces nombres restent inutiles, mais l'étude préliminaire qui serait nécessaire pour reconnaître *a priori* ceux dont on pourrait n'avoir pas besoin demanderait un travail au moins égal à celui qu'exige leur évaluation; il n'y a donc aucun avantage à faire cette étude.

5° *Inscrire sur la dernière ligne de la colonne* (3), *en regard de N, l'approximation demandée, déterminer ensuite l'approximation nécessaire aux éléments de la dernière opération, l'inscrire en regard de ces éléments, et continuer en remontant jusqu'à ce que la colonne soit remplie.*

On écrira ainsi, en regard de N, l'approximation demandée 0,035; on aura ensuite, conformément à la règle des opérations simples,

$$\delta N = \frac{1}{2N}\,\delta Q < 0{,}035 - \tfrac{1}{2}\text{ cent.}, \quad \frac{\delta Q}{10} < 0{,}03, \quad \delta Q < 1{,}2,$$

$$\delta Q = \frac{1}{\pi}\,\delta P - \frac{P'}{\pi'^2}\,\delta\pi < 1{,}2 - \tfrac{1}{2}\text{ unité} \quad \text{ou} \quad 0{,}7,$$

c'est-à-dire

$$\frac{1}{\pi'}\,\partial P < \frac{0,7}{2}, \quad \frac{\partial P}{3} < \frac{0,7}{2}, \quad \partial P < 1,$$

$$\frac{P'}{\pi'^2}\,\partial\pi < \frac{0,7}{2}, \quad \frac{2300}{9}\,\partial\pi < 0,35, \quad 300\,\partial\pi < 0,35, \quad \partial\pi < 0,001,$$

et enfin

$$\partial P - 45\,\partial A < 1 - \tfrac{1}{2} \text{ ou } 0,5, \qquad \partial A < \frac{0,5}{50}, \qquad \partial A < 0.01.$$

Il ne faut pas perdre de vue qu'en arrondissant il faut forcer les premiers membres et réduire les seconds.

6° *On écrira immédiatement, dans la colonne* (4), *les valeurs des nombres qui ont été désignés au début par des lettres, avec l'approximation indiquée en regard, et l'on effectuera enfin les opérations indiquées dans la colonne* (1) *en poussant les calculs jusqu'aux unités de l'ordre du premier chiffre significatif à gauche de l'approximation inscrite dans la colonne* (3), *et en forçant ou en conservant le dernier chiffre, suivant que le premier chiffre supprimé est plus grand que* 5 *ou égal à* 5, *ou qu'il est plus petit que* 5.

Ainsi, dans l'exemple considéré, nous écrirons pour A et π les valeurs 50, 21 et 3,141; nous calculerons P en conservant au résultat le chiffre des unités qu'il n'y a pas lieu de forcer; nous calculerons ensuite le quotient $\frac{P}{\pi}$ en dixièmes; nous extrairons enfin la racine carrée en poussant jusqu'au chiffre des centièmes, l'aspect du reste montrant que le chiffre suivant est plus petit que 5; nous conserverons le résultat 26,81, qui est ainsi la valeur de X à 0,035 près.

Exercice.

Calculer $N = \dfrac{(0,11702\,\pi + 1,43115)^2}{\sqrt{2\sqrt{3} + 5,01264}}$ à $\pm$ 0,02 près,

$$N = \frac{(A\pi + B)^2}{\sqrt{2\sqrt{3} + C}}.$$

	Opérations à effectuer.	Résultats approchés par défaut.	Résultats approchés par excès.	Approximations necessaires.	Résultats définitifs.
Numérateur.	A	0,1	0,12	0,00007	0,117
	π	3	4	0,002	3,14
	$P = A.\pi$	0,3	0,5	0,001	0,367
	B	1,4	1,5	0,001	1,431
	$S = P + B$	1,7	2	0,0025	1,798
	$K = S^2$	2	4	0,015	3,23
Dénominateur.	$R = \sqrt{3}$	1,7	2	0,0025	1,733
	$D = 2R$	3	4	0,01	3,47
	C	5	5,1	0,01	5,01
	$E = C + D$	8	9,1	0,026	8,48
	$F = \sqrt{E}$	2	3	0,007	2,912
	$N = \dfrac{K}{F}$	»	»	0,02	1,11

Détail du calcul des approximations.

$$N = \frac{K}{F}, \quad \delta N = \frac{1}{F'}\delta K - \frac{K'}{F'^2}\delta F < 0,02 - \tfrac{1}{2}\,\text{cent. ou} < 0,015,$$

$$\frac{\delta K}{F'} < \frac{0,015}{2}, \quad \frac{\delta K}{2} < \frac{0,015}{2}, \quad \delta K < 0,015,$$

$$\frac{K'\,\delta F}{F'^2} < \frac{0,015}{2}, \quad \frac{4}{4}\delta F < \frac{0,015}{2}, \quad \delta F < 0,007.$$

Numérateur.

$$K = S^2, \qquad \partial K = 2S' \,\partial S < 0{,}015 - \tfrac{1}{2}\text{ cent.} < 0{,}01,$$

$$4\,\partial S < 0{,}01, \qquad \partial S < 0{,}0025,$$

$$S = P + B, \qquad \partial S = \partial P + \partial B < 0{,}0025 - \tfrac{1}{2}\text{ mill.},$$

$$\partial P + \partial B < 0{,}002,$$

$$\partial P < 0{,}001,$$

$$\partial B < 0{,}001,$$

$$P = A\pi, \qquad \partial P = A' \,\partial\pi + \pi' \,\partial A < 0{,}001 - \tfrac{1}{2}\text{ mill.},$$

$$A' \,\partial\pi + \pi' \,\partial A < 0{,}0005,$$

$$A' \,\partial\pi < \frac{0{,}0005}{2},$$

$$0{,}12\,\partial\pi < \frac{0{,}0005}{2},$$

$$\partial\pi < 0{,}002,$$

$$\pi' \,\partial A < \frac{0{,}0005}{2},$$

$$\partial A < 0{,}00007.$$

Dénominateur.

$$F = \sqrt{E}, \qquad \partial F = \frac{1}{2F'}\,\partial E < 0{,}007 - \tfrac{1}{2}\text{ mill.},$$

$$\frac{\partial E}{4} < 0{,}0065,$$

$$\partial E < 0{,}026,$$

$$E = C + D, \qquad \partial E = \partial C + \partial D < 0{,}026 - \tfrac{1}{2}\text{ cent.},$$

$$\partial C + \partial D < 0{,}020,$$

$$\partial C < 0{,}01,$$

$$\partial D < 0{,}01,$$

$$D = 2R, \qquad \partial D = 2\,\partial R < 0{,}01 - \tfrac{1}{2}\text{ cent.},$$

$$2\,\partial R < 0{,}005,$$

$$\partial R < 0{,}0025.$$