LES DEMONSTRATIONS DE LA FORMULE DU BINOME AU XVIIIe SIECLE

Michel Pensivy

La longue histoire de la formule du binôme $(a+b)^n$, pour n entier positif,s'est terminée en 1654, lorsque Blaise Pascal l'a démontrée dans son **Traité du triangle arithmétique** et en a , dans l' **Usage du triangle arithmétique**,donné de nombreux exemples d'applications.

La série du binôme $(1+x)^\alpha = 1 + \alpha x + \frac{\alpha(\alpha-1)}{2} x^2 + \ldots$, pour α non entier , a une histoire beaucoup mieux délimitée dans le temps . Et de 1665 , date où le jeune Newton en note une première version dans la marge de son exemplaire de l'**Arithmetica infinitorum** de Wallis , à 1826, où Niels Henrik Abel en donne pour la première fois une démonstration rigoureuse , elle joue un rôle central dans l'évolution de l'Analyse . Ce n'est pas en effet , pendant toute cette période , un développement en série parmi d'autres , mais le premier énoncé de la quasi-totalité des traités d'Analyse du XVIIIe siècle , à partir duquel on obtient tous les développements en série nécessaires . Elle conserve ce rôle jusqu'au début du XIXe , le perdant au profit de l'exponentielle , après que les notions de convergence aient été précisées .

Nous allons entreprendre de donner les jalons de cette histoire . Il est difficile d'adopter pour ce faire un plan purement chronologique . Les dates qui suivent , si elles ne donnent qu'une idée très schématique de cette chronologie , pourront tout de même servir de points de repère :

1665 : découverte par Newton
1668 : la série chez James Gregory
1685 : le premier énoncé édité (dans l' Algebra de Wallis)
1736 : la traduction par John Colson de la **Méthode des fluxions et des suites infinies** d'Isaac Newton . Démonstration par des méthodes différentielles .
1758 : John Landen : **A Discourse concerning the Residual Analysis** ; refus des méthodes différentielles ; démonstration algébrique .
1761 : la première utilisation d'une équation fonctionnelle par Aepinus .
1774 : une des démonstrations de L . Euler ; multiplication de séries .
1821 : Augustin - Louis Cauchy : **Analyse algébrique** .
1826 : le mémoire d'Abel dans le Journal de Crelle ; première démonstration rigoureuse .

Nous allons choisir un plan différent : partant des deux démonstrations pratiquement définitives (celles de Cauchy et d'Abel) , il sera plus facile , après avoir examiné également les premiers énoncés de la formule à la fin du XVIIe siècle , de suivre son évolution tout au long du XVIIIe , d'observer le rôle fondamental qu'elle y a joué , et , en examinant de près les méthodes mises en oeuvre pour en faire la démonstration , de constater comment on aboutit aux idées de Cauchy et quel fut l'apport original de ce dernier .

1. L' "Analyse algébrique" de Cauchy (1821)

Dans l'introduction de ce livre , rédaction d'un cours donné à l'Ecole Royale Polytechnique , l'auteur indique quel but il s'est fixé : essentiellement de donner plus de rigueur à la présentation de l'Analyse , en renonçant à quelques pratiques courantes au siècle précédent . Il indique par exemple (page II) : " *quant aux méthodes , j'ai cherché à leur donner toute la rigueur qu'on exige en géométrie, de manière à ne jamais recourir aux raisons tirées de la généralité de l'algèbre* " . Et précisant davantage (page IV) : " *il est vrai que , pour rester constamment fidèle à ces principes , je me suis vu forcé d'admettre plusieurs propositions qui paraîtront peut-être un peu dures au premier abord . Par exemple , j'énonce dans le chapitre VI , qu'une série divergente n'a pas de somme ; dans le chapitre VII , qu'une équation imaginaire est seulement la représentation symbolique de deux équations entre quantités réelles , ...* " (rejetant ainsi dès le départ , entre autres , le célèbre "principe de permanence" invoqué au XVIIIe siècle , grâce auquel on affirmait par exemple que toute formule vérifiée pour des nombres réels l'était automatiquement pour des nombres complexes) . Les intentions de Cauchy sont donc claires , et le livre ouvre une période nouvelle dans l'enseignement de l'Analyse .

Or ce livre attire incontestablement l'attention sur l'importance de la série du binôme . Il est clair en effet que l'organisation de ses six premiers chapitres est conçue dans le but principal d'obtenir le développement de $(1+x)^{\mu}$, pour x et μ réels , et celle des trois chapitres suivants pour développer $(1+x)^{\mu}$, où μ est réel et x complexe . Toute l'Analyse est reconstruite à partir de là , dans les deux livres de 1823 (**Leçons sur le calcul infinitésimal**) et de 1829 (**Leçons sur le calcul différentiel**) .

Rappelons la démarche suivie par Cauchy dans l' **Analyse algébrique** : après avoir évoqué de façon assez intuitive la notion de limite , il donne ses définitions d'un infiniment petit (c'est une suite convergeant vers zéro) et d'une fonction continue (" *la fonction f(x) restera continue par rapport à x entre les limites données si , entre ces limites , un accroissement infiniment petit de la variable produit toujours un accroissement infiniment petit de la fonction elle-même* ") ; ce n'est donc pas la définition d'une continuité en un point , et les notions d'uniformité ne sont pas encore perçues . Il énonce ensuite (chapitre 2) quelques théorèmes généraux sur les fonctions continues et les limites , en particulier celui-ci : pour $f > 0$, si $\lim\limits_{+\infty} \frac{f(x+1)}{f(x)} = k$, alors $\lim\limits_{+\infty} [f(x)]^{\frac{1}{x}} = k$ (le critère de Cauchy découle du critère de d'Alembert) . Suivent deux chapitres (3 et 4) sur les fonctions alternées et l'interpolation des fonctions , dont on se demande en première lecture pour quelle raisons Cauchy les place à cet endroit du livre (la clé sera fournie par le traitement de la série du binôme , deux chapitres plus loin) . Au chapitre 5 , il résout ensuite quatre équations fonctionnelles et démontre en particulier, grâce à sa notion de fonction continue : si $f(\mu + \mu') = f(\mu)\, f(\mu')$, alors $f(x) = A^x = [f(1)]^x$. Enfin viennent au chapitre 6 les définitions sur les séries (convergence , divergence , somme) , des critères de convergence (dont le critère "de Cauchy" pour lequel la réciproque est seulement affirmée , le statut alors très flou des nombres réels ne lui permettant pas d'en démontrer davantage) , un résultat faux où il "démontre" que la somme d'une série de fonctions continues est continue (l'erreur venant bien sûr de sa définition de la continuité , mais aussi des notations utilisées) , puis des théorèmes sur la multiplication des séries et sur les séries entières .

Cauchy applique alors la totalité des résultats précédents pour étudier la série du binôme , ce qu'il fait de la manière suivante :

il pose :

$$g(\mu) = 1 + \mu x + \frac{\mu(\mu-1)}{2} x^2 + \frac{\mu(\mu-1)(\mu-2)}{2.3} x^3 + \ldots\ldots ,$$

montre que $g(\mu)$ est définie pour $x \in]-1, 1[$, que g est continue (en utilisant le théorème faux du chapitre 6), que g vérifie l'équation fonctionnelle $g(\mu)\, g(\mu') = g(\mu + \mu')$ (d'après les résultats des chapitres 3 et 4, et c'est le seul endroit du livre où ils sont utilisés), d'où $g(x) = A^x$, d'après le chapitre 5.

Ce qui lui permet enfin d'écrire :

$g(\mu) = A^\mu = [g(1)]^\mu = (1 + x)^\mu$.

Les développements de l'exponentielle et du logarithme sont ensuite obtenus selon le principe suivant :

$$e^x = \lim_{\alpha \to 0} (1 + \alpha x)^{\frac{1}{\alpha}}$$

$$\ell(1 + x) = \lim_{\mu \to 0} \frac{(1 + x)^\mu - 1}{\mu}$$

(les notations sont celles qu'a utilisées Cauchy, mais le principe du passage du binôme aux logarithmes avait été indiqué pour la première fois par Halley dans un article de 1695, et effectué de façon voisine pendant tout le XVIIIe siècle).

Dans les chapitres 7 à 9, tous les calculs et définitions sont repris, en les adaptant aux fonctions à valeurs complexes.

Dans ce livre, la place et l'importance (une trentaine de pages) des deux chapitres sur l'interpolation des fonctions, utilisés uniquement pour démontrer que la somme $g(\mu)$ de la série du binôme vérifie $g(\mu)\, g(\mu') = g(\mu + \mu')$, permettent sans nul doute d'affirmer qu'un des buts, peut-être même le seul but, de Cauchy dans l' Analyse algébrique était d'établir ce développement en série.

2. Le traitement définitif par Abel (1826)

C'est Abel qui signale le premier l'erreur de Cauchy sur la somme d'une série de fonctions continues, dans une lettre de Janvier 1826 adressée à Holmbœ, un de ses anciens professeurs resté en Norvège. Il donne le contre-exemple de la série convergente: $\sin x - \frac{\sin 2x}{2} + \ldots + (-1)^n \frac{\sin nx}{n} + \ldots$, égale à $\frac{x}{2}$ si $x \in [0, \pi[$ et à 0 si $x = \pi$.

Dans cette même lettre, il écrit : " *je crois que tu ne pourras me proposer qu'un très petit nombre de théorèmes contenant des séries infinies, à la démonstration desquels je ne puisse faire des objections bien fondées. La formule du binôme elle-même n'est pas encore rigoureusement démontrée* ".

Dans un mémoire de 1826, paru dans le Journal de Crelle, il indique qu'" *on n'a même pas examiné tous les cas où cette série est convergente*" et que" *le but de ce mémoire est d'essayer de remplir une lacune par la solution du problème suivant : trouver la somme de la série :* $1 + \frac{m}{1} x + \frac{m(m-1)}{2} x^2 + \ldots$, *pour toutes les valeurs réelles ou imaginaires de* x *et de* m *pour* *lesquelles la série est convergente"* .

Se plaçant donc dans un cadre plus général que celui de l'Analyse algébrique (l'exposant m peut prendre ici des valeurs complexes),il va suivre une démarche très voisine (équations fonctionnelles),utilisant toutefois des théorèmes plus précis sur les séries , sur les produits de séries , et surtout une définition de la continuité d'une fonction beaucoup mieux adaptée : " *une fonction* fx *sera dite continue de* x *entre les limites* $x = a$ *et* $x = b$ *, si pour une valeur quelconque de* x *comprise entre ces limites , la quantité* $f(x-\beta)$ *, pour des valeurs toujours décroissantes de* β *, s'approche indéfiniment de la limite* fx " ; (il s'agit bien ici de la définition de la continuité de f au point x , ce qui n'était pas le cas de la définition de Cauchy) .

Il commence par étudier en détail tous les cas où la série converge (pour $|x| < 1$, quel que soit m ; pour $|x| = 1$, $a = \mathrm{Re}(x) \neq -1$, convergence si $k = \mathrm{Re}(m) > -1$; pour $|x| = 1$ et $a = -1$, convergence si $k > 0$; divergence dans tous les autres cas) , puis notant $\varphi(m)$ la somme de la série , vérifie la relation $\varphi(m+n) = \varphi(m)\,\varphi(n)$, pose $x = a + i\,b = \alpha(\cos\varphi + i\sin\varphi)$, $m = k + i\,k'$,

$\varphi(m) = p + q\,i = r(\cos s + \sin s)$, puis $r = f(k, k')$, $s = \psi(k, k')$, démontre que p et q sont continues . Il établit ensuite des équations fonctionnelles vérifiées par les fonctions continues f et ψ, ce qui lui permet d'expliciter ces deux fonctions puis , après un calcul très détaillé , d'obtenir la formule :

$$1 + m\,x + \frac{m(m-1)}{2}x^2 + \ldots$$
$$= \left[\cos\left(k \operatorname{Arctg}\frac{b}{1+a} + \frac{1}{2}k' \operatorname{Log}((1+a)^2+b^2)\right) + i\sin\left(k \operatorname{Arctg}\frac{b}{1+a} + \frac{1}{2}k' \operatorname{Log}((1+a)^2+b^2)\right)\right]$$
$$\times\left[(1+a)^2 + b^2\right]^{\frac{k}{2}} e^{-k' \operatorname{Arctg}\frac{b}{1+a}} .$$

La formule du binôme pour m et x réels s'obtient donc en prenant $b = 0$ et $k' = 0$, ce qui donne bien :

$$\varphi(m) = (1+a)^k = (1+x)^m .$$

Abel écrira , dans une autre lettre à Holmbœ (Décembre 1826) : " *j'ose dire que c'est la première démonstration rigoureuse de la formule binôme , dans tous les cas possibles* " .

3 . Recherche de paternité : les premiers énoncés de la formule , fin XVIIe .

a) Le premier énoncé publié du développement en série d'un binôme avec exposant rationnel se trouve dans l'Algebra de Wallis (1685) . Il l'attribue à Newton et , effectivement , les deux paragraphes du livre qui sont consacrés à la formule sont une reproduction abrégée des deux lettres adressées neuf années auparavant (les 13 Juin et 24 Octobre 1676) par Isaac Newton à Oldenburg , Secrétaire de la Royal Society , en réponse à Leibniz qui demandait ce que savaient faire les mathématiciens anglais sur les séries infinies .

Dans la première lettre (dite *epistola prior*) , Newton énonce le théorème du binôme pour un exposant rationnel $\frac{m}{n}$, sous la forme :

$$\overline{P+PQ}^{\frac{m}{n}} = P^{\frac{m}{n}} + \frac{m}{n} A\,Q + \frac{m-n}{2n} B\,Q + \frac{m-2n}{3n} C\,Q + \frac{m-3n}{4n} D\,Q + \ldots$$

(où $A = P^{\frac{m}{n}}$, $B = \frac{m}{n} A\,Q$, $C = \frac{m-n}{2n} B\,Q$, ...) .

Suivent de nombreux exemples , le choix de P et Q étant fait dans chaque cas de façon à ce que les termes de la série deviennent de plus en plus petits .

Leibniz répond en demandant l'origine de la formule .

La seconde lettre de Newton (dite *epistola posterior*) contient ses explications . Il rappelle les travaux de Wallis sur les courbes d'équations :

$$y = (1-x^2)^{\frac{0}{2}} \ , \ y = (1-x^2)^{\frac{1}{2}} \ , \ y = (1-x^2)^{\frac{2}{2}} \ , \ y = (1-x^2)^{\frac{3}{2}} \ , \ y = (1-x^2)^{\frac{4}{2}} \ , \ldots\ldots$$

et les quadratures correspondantes :

$$x \qquad , \qquad\qquad , \ x - \frac{1}{3}x^3 \qquad , \qquad\qquad , \ x - \frac{1}{3}x^3 + \frac{1}{5}x^5 \ , \ldots\ldots\ldots$$

puis remarque qu'on peut , en interpolant , en déduire les aires sous les autres courbes : par exemple pour $y = (1-x^2)^{\frac{1}{2}}$ et pour $y = (1-x^2)^{\frac{3}{2}}$; ce qu'il fait à partir de constatations sur les premiers termes, sur les dénominateurs , sur la forme des seconds termes des quadratures déjà écrites , et sur les développements des termes : $(1+1)^0$, $(1+1)^1$, $(1+1)^2$, $(1+1)^3$, $(1+1)^4$, ...

D'où la forme pressentie pour les coefficients .Il écrit les séries correspondant aux exposants $\frac{1}{2}$, $\frac{3}{2}$, $\frac{1}{3}$, et fait une vérification en élevant au carré le résultat obtenu pour $(1-x^2)^{\frac{1}{2}}$.

En fait , en lisant les seules explications contenues dans cette lettre , on ne peut pas vraiment savoir si Newton a réellement suivi cette voie pour découvrir la formule , ou si , après avoir généralisé la formule de Pascal qu'il savait valable pour tout exposant entier positif , il a ensuite explicité quelques cas particuliers et donné après coup ces justifications heuristiques .

b) A la même époque , on trouve aussi , sous une forme plus cachée , un développement en série d'un binôme dans deux lettres de James Gregory adressées à Collins (les 23 novembre et 19 décembre 1670) . Gregory y est conduit par des calculs de logarithmes : partant de deux nombres b et d tels que $\log b = e$ et $\log(b+d) = e + c$, il cherche un nombre dont le logarithme est $e + a$.

Or ,

$$e + a = \log b + \frac{a}{c}\,[\log(b+d) - \log b] = \log b\left(1 + \frac{d}{b}\right)^{\frac{a}{c}} .$$

Pour obtenir le développement du binôme (dont l'exposant est d'ailleurs quelconque) , il utilise la formule d'interpolation dite aujourd'hui " de Newton-Gregory" , de la manière suivante (mais avec des notations bien entendu tout à fait différentes) :

$\log b = e$ et $\log(b+d) = e + c$ donnent $\log b\left(1 + \frac{d}{b}\right)^n = e + n\,c$.

Prenant alors : $f(0) = b$, $f(c) = b + d$, ... , $f(n\,c) = b\left(1 + \frac{d}{b}\right)^n$, on a :

$$\Delta f(0) = d \quad , \quad \Delta^2 f(0) = \frac{d^2}{b} \quad , \quad \Delta^3 f(0) = \frac{d^3}{b^2} \quad , \ \ldots\ldots$$

et , en reportant dans la formule :

$$f(a) = f(0) + \frac{a}{c}\,\Delta f(0) + \frac{\frac{a}{c}\left(\frac{a}{c} - 1\right)}{2}\,\Delta^2 f(0) + \ldots\ldots \ ,$$

on obtient :

$$b\left(1 + \frac{d}{b}\right)^{\frac{a}{c}} = b + \frac{a}{c}\,d + \frac{\frac{a}{c}\left(\frac{a}{c} - 1\right)}{2}\frac{d^2}{b} + \frac{a}{c}.\frac{a-c}{2c}.\frac{a-2c}{3c}.\frac{d^3}{b^2} + \ldots\ldots ,$$

ce qui est bien le développement en série d'un binôme .

c) Il semble donc que Newton ait été devancé sur ce point par James Gregory , d'autant que ce dernier connaissait la formule dès 1668 . En effet , une lettre de Collins à Moray , destinée à arbitrer une querelle de priorité entre Gregory et Huygens au sujet de la quadrature du cercle et de l'hyperbole , fait le point sur leurs connaissances respectives à cette date et cite en particulier des calculs analogues aux précédents .

Mais on est certain malgré tout , d'après Whiteside , que Newton avait découvert la formule dès l'hiver 1664-65 : jeune étudiant , il annotait ses livres , et on trouve la série du binôme , sous la forme d'une généralisation compliquée du triangle de Pascal , dans la marge de son exemplaire de l'**Arithmetica infinitorum** de Wallis . On peut donc sans doute possible lui attribuer la formule , et c'est ce que l'Histoire a retenu .

d) Dès 1620 , on peut trouver la trace d'une utilisation de $(1 + \alpha)^{1/2}$ dans l'**Arithmetica logarithmica** de Briggs (qui ne paraîtra qu'en 1624) . Pour l'interpolation de logarithmes , il calcule en effet les racines successives de nombres proches de 1 , mais il ne s'agit que des quatre premiers termes du développement , utilisé dans quelques cas particuliers , et non d'une formule générale .

4 . Les méthodes de démonstrations au XVIIIe siècle .

a) On ne trouve pas d'essai de démonstration avant 1736 , date de la première édition de **The method of Fluxions** , traduction anglaise par John Colson du manuscrit latin de Newton . Le livre sera traduit en français quatre années plus tard par Buffon , puis retraduit en latin en 1742 , à partir de la version anglaise de Colson, par Castillon (dans les **Opuscula Neutoni** édités à Lausanne) . La première édition d'**Oeuvres complètes** de Newton ne paraîtra qu'en 1779 , comprenant cinq volumes en latin , accompagnés de notes du Révérend Horsley respectant très fidèlement les notations de l'auteur .

b) Jusqu'en 1736 , on trouve tout de même quelques articles intéressants sur le sujet , notamment dans les **Philosophical Transactions** :

- un article (déjà cité) de l'astronome et mathématicien Halley , en 1695 , appliquant le théorème du binôme à des calculs de logarithmes .
- deux articles dus à De Moivre , en 1697 et 1698 , où il donne une généralisation du théorème , élevant une série entière à une puissance entière ou rationnelle . Ces calculs seront à l'origine de quelques démonstrations de la formule du binôme au XVIIIe siècle , mais serviront aussi de point de départ à l'Ecole Combinatoire allemande , groupée autour de Hindenburg vers 1800 .

- un échange de lettres a lieu sur le sujet entre Leibniz et Nicolas Bernoulli (en 1713-14) où ils évoquent, sur quelques cas particuliers seulement, la convergence de la série.

- le théorème figure sans démonstration dans les **Oeuvres complètes** de Jean Bernoulli (quatre volumes parus de 1727 à 1742).

c) En revanche, de 1736 (Colson) à 1821 (Cauchy), on peut trouver de nombreuses démonstrations, au moins une cinquantaine par des mathématiciens ayant quelque importance, dans toutes les revues scientifiques européennes et les comptes rendus des diverses Académies.

On peut, en les classant suivant la méthode de démonstration utilisée, les répartir en cinq groupes :

- par interpolation.
- élévation d'une série entière à une puissance entière (De Moivre, Castillon, livres d'enseignement).
- méthodes différentielles.
- méthodes algébriques.
- utilisation d'équations fonctionnelles.

On peut d'ailleurs passer très rapidement sur les démonstrations par interpolation car elles concernent, soit les premiers énoncés de la formule (Newton, J. Gregory), soit des mathématiciens d'importance secondaire (par exemple : Walz, dans les **Actes de Leipzig**, en 1745. ou Gourieff, dans les **Comptes rendus de Saint-Pétersbourg**, en 1799).

d) Pour illustrer le deuxième type de démonstration, on peut donner le schéma suivi par Castillon en 1742, dans l'article des **Philosophical Transactions** où il annonce la parution de ses **Opuscula Neutoni**.

Dans le cas d'un exposant n entier positif, il se sert d'un dénombrement :

$$(p+q)^n = (p+q_1)(p+q_2)\ldots\ldots(p+q_n) \quad , \text{ où } \quad q_1 = q_2 = \ldots = q_n = q$$

puis, pour un exposant fractionnaire positif, pose a priori le développement sous la forme :

$$(p+q)^{\frac{r}{n}} = A\,p^{\frac{r}{n}} + B\,p^{\frac{r}{n}-1}\,q + C\,p^{\frac{r}{n}-2}\,q^2 + \ldots\ldots$$

Castillon élève alors les deux membres de cette égalité à la puissance n, en se référant à l'article de De Moivre de 1697 :

$$(p+q)^r = p^r\,(A + B\,p^{-1}\,q + C\,p^{-2}\,q^2 + \ldots\ldots)^n$$

il développe alors, puis identifie, ce qui conduit à :

$1 = A^n$, donc $A = 1$

$r = n\,A^{n-1}\,B$, donc $B = \frac{r}{n}$

$\frac{r\,(r-1)}{2} = n\,A^{n-1}\,C + \frac{n\,(n-1)}{2}\,A^{n-2}\,B^2$, donc $C = \frac{\frac{r}{n}\left(\frac{r}{n}-1\right)}{2}$, etc ..

dans le cas d'un exposant fractionnaire négatif, il écrit $\overline{p+q}^{\,-s} = \frac{1}{\overline{p+q}^{\,s}}$, et effectue une division suivant les puissances croissantes.

Cette démonstration se retrouve (à cause de sa brièveté ?) dans la plupart des traités destinés à l'enseignement : dans les **Eléments d'Algèbre** de Clairaut (1746), le **Traité de calcul différentiel et intégral** de Cousin (1796), les **Eléments raisonnés d'Algèbre** de L'huilier (1804), etc .....

Elle trouve un prolongement dans les travaux de l'Ecole Combinatoire allemande d'Hindenburg (suivant De Moivre, lorsqu'on calcule $(a + bx + cx^2 + dx^3 + \ldots)^n$, on obtient par exemple des exposants de la forme $s + 2t + 3v + 4w + \ldots$, et il s'agit donc de rechercher les entiers $s, t, v, w, \ldots$ tels que $s + 2t + 3v + 4w + \ldots$ soit un entier donné) . Si elle a abouti à une impasse, cette " *Kombinatorischeschule*" a au moins eu une influence sur les notations (Kramp introduit le terme "*factorielle*" , au début "*faculté* " , et la notation C_p^n) .

e) Les trois autres types de démonstrations (méthodes différentielles , algébriques et équations fonctionnelles) sont plus importants et demandent à être détaillés , car ils mettent en jeu des questions plus fondamentales .

Il est intéressant de constater déjà qu'on ne les rencontre pas simultanément , mais qu'ils se succèdent dans le temps . De plus , le passage d'une méthode à l'autre est le résultat soit d'un refus (on passe aux méthodes algébriques quand on n'accepte plus les méthodes différentielles) soit d'une insuffisance (les méthodes algébriques ne permettent pas d'aborder le cas d'un binôme où l'exposant est irrationnel) .

Enfin , jusqu'en 1760 environ , on envisage uniquement les exposants rationnels , pour considérer après cette date des binômes dont les exposants sont réels quelconques . A cet égard , le changement de point de vue est dû sans aucun doute à l'**Introduction à l'Analyse des infinis** , publié par Euler en 1748 , qui exerça une grande influence sur les mathématiciens de l'époque , livre où pour la première fois l'exponentielle figure avant les logarithmes , enfin détachés du support géométrique constitué par l'hyperbole , et où Euler explique très en détail le maniement des fonctions exponentielles .

5 . Utilisation de méthodes différentielles .

Le principe en est très simple . Citons par exemple la démonstration faite par Colson , en 1736 , dans les commentaires qu'il ajoute à sa traduction du livre de Newton .

Il pose a priori le développement sous la forme :

$(1 + x)^n = 1 + A x + B x^2 + C x^3 + \ldots\ldots\ldots$ (où n est rationnel)
puis , utilisant les notations et le vocabulaire dus à Newton , il prend la "fluxion" des deux membres :

$$n \dot{x} (1 + x)^{n-1} = A \dot{x} + 2 B x \dot{x} + 3 C x^2 \dot{x} + \ldots\ldots$$

d'où , en simplifiant par $\dot{x}$ puis en prenant $x = 0$, $A = n$. Les autres coefficients sont obtenus de la même façon , par d'autres dérivations .

Mais à la fin de son calcul , Colson remarque lui-même le cercle vicieux : on obtient la formule $\dot{\overline{x^n}} = n x^{n-1} \dot{x}$ (où n est rationnel) en utilisant le théorème du binôme !

On trouve , de 1736 à 1760 pour la majorité d'entres elles , une douzaine de démonstrations du même type , utilisant soit les notations des fluxions (Maclaurin , Thomas Simpson , jusqu'à Horsley dans la première édition des **Oeuvres complètes de Newton**) , soit les notations du calcul différentiel (certaines démonstrations de Kæstner , ou deux autres dues à Euler , l'une faite oralement à l'astronome Clemm , qui la rédige en 1752 , l'autre dans les **Institutions du Calcul Différentiel** , en 1755) . Tous remarquent le cercle vicieux et le disent : l'article de Clemm s'intitule: **Lettre à Monsieur Euler sur quelques paradoxes du calcul analytique** ; L'huilier , qui fait cette démonstration dans son **Exposition élémentaire des principes des calculs supérieurs**, primé par l'Académie de Berlin , écrit :" *comme la détermination précédente dépend de la généralisation du théorème binomial pour un exposant quelconque , qui n'est pas un nombre entier positif ; et que plusieurs mathématiciens ont cru que la démonstration de ce théorème , prise dans ce sens général , dépend du calcul différentiel : on pourrait élever des doutes sur la légitimité de son application dans un écrit où l'on tâche de démontrer les principes de ces calculs* " . Et il ajoute , non sans humour : " *mais comme je suis sûr qu'on peut démontrer ce théorème général d'une manière rigoureuse (plus longue à la vérité que celle qui est déduite du calcul différentiel) par les éléments du* *calcul algébrique , je ne me laisse pas arrêter par cette difficulté . Je développerais cette démonstration élémentaire , si je ne craignais pas de me trop écarter par ce hors-d'œuvre du but principal que je me dois proposer "* !

Or il est clair qu'on peut très bien contourner cette difficulté , en démontrant d'abord la formule $\dot{\overline{x^n}} = n\,x^{n-1}\,\dot{x}$ pour n entier positif (soit par récurrence , soit en utilisant la formule du binôme avec exposant entier) , puis en obtenant $\dot{\overline{x^{p/q}}}$ par les seules règles de la dérivation ($y = x^{\frac{p}{q}}$ donc $y^q = x^p$ et $q\,y^{q-1}\,\dot{y} = p\,x^{p-1}\,\dot{x}$, ce qui donne $\dot{y} = \frac{p}{q}\,x^{\frac{p}{q}-1}\,\dot{x}$), utilisant ensuite la méthode de Colson pour obtenir les coefficients de la série du binôme .

Il est très curieux de constater que ces mathématiciens savent parfaitement qu'on pourrait procéder ainsi . La phrase exacte de Colson à la fin de la démonstration est la suivante : *" en vérité on peut difficilement dire que ceci , ou tout autre calcul dérivé de la méthode des fluxions , est une preuve rigoureuse du théorème . Parce que cette méthode provient elle-même à l'origine de la méthode d'élévation en puissances , au moins en puissances entières , et suppose donc la connaissance préalable des coefficients "*.
D'ailleurs Maclaurin , dans son célèbre Traité des fluxions (1742) , démontre d'abord la formule $\dot{\overline{x^n}} = n\,x^{n-1}\,\dot{x}$ pour n entier positif (par un procédé d'encadrement assez compliqué se référant à la méthode d'exhaustion des Anciens) puis termine comme on l'a indiqué dans le paragraphe précédent .

Or le livre de Maclaurin est très connu à l'époque , et largement diffusé (il est traduit en français dès 1749) . Si les démonstrations du théorème du binôme utilisant le calcul différentiel ne suivent pas ce plan , qui nous semble aujourd'hui plus logique , et qui est connu des mathématiciens concernés , c'est donc qu'ils éprouvent une certaine réticence à le faire , et qu'ils souhaiteraient respecter un certain ordre dans la rédaction des Traités , à savoir :

- algèbre
- série du binôme
- autres développements en série
- dérivée de x^n (pour n fractionnaire)
- et enfin tout le calcul différentiel .

Ceci permet aussi de constater quel est à cette époque le statut exact des développements en série : ils ne font pas partie de l'Analyse , on n'y voit aucune notion de limite ; au XVIIIe siècle , développer en série est une opération purement algébrique .

6 . John Landen et les démonstrations algébriques .

Dans un petit livre (une quarantaine de pages) paru en 1758 , **A Discourse concerning the Residual Analysis** , John Landen formule les mêmes réticences et , dans un long commentaire, explique qu'à son avis la méthode des fluxions n'est pas le moyen le plus naturel pour résoudre de nombreux problèmes auxquels on l'applique d'habitude . Et que , d'après lui , les principes empruntés par Newton à la *" doctrine du mouvement "* (notamment la notion de vitesse) pour établir les bases de son calcul sont étrangers à l'Algèbre . Il affirme donc qu'il faut utiliser les *"anciennes méthodes"* algébriques à chaque fois qu'il est possible de le faire ; il aborde de cette façon le théorème du binôme et en donne une démonstration où les seules opérations utilisées sur les séries sont l'addition et la multiplication par un nombre quelconque , démonstration où il utilise à plusieurs reprises la formule suivante (appelée "formule de Landen" par les commentateurs de la fin du siècle) :

$$\frac{x^{\frac{m}{n}} - v^{\frac{m}{n}}}{x - v} = x^{\frac{m}{n} - 1} \frac{1 + \left(\frac{v}{x}\right) + \left(\frac{v}{x}\right)^2 + \left(\frac{v}{x}\right)^3 + \ldots\ldots + \left(\frac{v}{x}\right)^{m-1}}{1 + \left(\frac{v}{x}\right)^{\frac{m}{n}} + \left(\frac{v}{x}\right)^{\frac{2m}{n}} + \left(\frac{v}{x}\right)^{\frac{3m}{n}} + \ldots\ldots + \left(\frac{v}{x}\right)^{(n-1)\frac{m}{n}}}$$

Il tentera par la suite de donner à l'Analyse une nouvelle présentation , basée entièrement sur la formule précédente et des procédés algébriques , dans un second livre paru en 1764 : **The Residual Analysis** . C'est une tentative analogue , toutes proportions gardées , à celle que fera Lagrange à la fin du siècle dans sa **Théorie des fonctions analytiques** , où il citera d'ailleurs dans l'Introduction les livres de Landen , lui reprochant seulement ses notations compliquées .

On trouve d'autres démonstrations basées sur le même principe dans la seconde moitié du XVIIIe siècle : dans les **Philosophical Transactions** (par exemple : Sewell , en 1796 : **the binomial theorem legally demonstrated by Algebra**) , ou dans les Comptes rendus de Saint-Pétersbourg (par exemple deux autres démonstrations d'Euler , en 1776 et 1779) . Toutes concernent un exposant rationnel , et aucune n'utilise de multiplications de séries .

Enfin , la **Théorie des fonctions analytiques** de Lagrange (1796) entre évidemment dans cette catégorie , même si , concernant la série du binôme , il ne fait que poser la forme du développement:
$(x + i)^m = x^m + A\,i + B\,i^2 + C\,i^3 + \ldots$, montrer qu'il suffit de connaître le terme A , et écrire :" *or, il est facile de démontrer , soit par les simples règles de l'Arithmétique , soit par les premières opérations de l'Algèbre , que les deux premiers termes de la puissance* m *du binôme sont* $x^m + m\,x^{m-1}\,i$ *, que* m *soit un nombre entier ou fractionnaire , positif ou négatif* " !

7 . La première utilisation d'une équation fonctionnelle : Aepinus (1761) .

Les méthodes algébriques ne permettant pas d'aborder le cas d'un binôme où l'exposant est réel , on voit apparaître , après 1760 , pour obtenir dans ce cas le développement en série , un certain nombre de démonstrations utilisant des équations fonctionnelles et qu'on peut répartir en deux catégories : celles qui donnent une équation fonctionnelle vérifiée par le second terme de la série , et celles qui indiquent une équation fonctionnelle vérifiée par la somme de la série . Ces dernières contiennent donc les premiers énoncés sur la multiplication des séries , et posent des problèmes qui mèneront aux énoncés définitifs de Cauchy et Abel .

On considère généralement que ces méthodes ont pour point de départ un article d'Euler publié à Saint-Pétersbourg et écrit en 1774 . Or , Euler y cite un autre article , écrit par Aepinus en 1761 , et publié dans les Comptes rendus de la même Académie . Aepinus , connu pour l'invention du condensateur et par ses travaux sur l'électricité , s'était d'abord intéressé aux mathématiques , et la démarche qu'il suit dans cet article retient l'attention par un certain nombre d'idées intéressantes , et de remarques peu formulées à ce moment du XVIIIe siècle .

Il pose : $(x + 1)^m = A\,x^m + B\,x^{m-1} + C\,x^{m-2} + D\,x^{m-3} + \ldots\ldots$
et tente , ce qui est rare à l'époque , de justifier cette écriture . Il énonce un argument qui ne prouve que l'unicité (si on peut calculer les coefficients , c'est que l'écriture est possible!) , mais il est tout de même intéressant qu'il se soit posé la question .

Il indique ensuite que les coefficients ne dépendent que de m, et pose par conséquent :

$$(x+1)^m = A^m x^m + B^m x^{m-1} + C^m x^{m-2} + D^m x^{m-3} + \ldots\ldots$$

(les notations sont exactement celles d'Aepinus , qui précise que m désigne dans certains cas un indice , dans d'autres un exposant) .

Puis , développant les deux membres de l'égalité $(2(x+1))^m = ((2x+1)+1)^m$, et identifiant les coefficients , il obtient :

$$A^m = 1 = A^{m-1} = A^{m-2} = \ldots\ldots$$

$$2^m B^m = 2^m B^m$$

$$C^m = \frac{B^m B^{m-1}}{1.2} \quad ; \qquad D^m = \frac{B^m B^{m-1} B^{m-2}}{1.2.3} \quad ; \quad \ldots\ldots$$

Aepinus fait alors une récurrence très détaillée (ce qui est rare également) pour démontrer la formule générale permettant d'exprimer un coefficient quelconque en fonction des coefficients B .

Il reste donc à déterminer B^m .

Or $(x+1)^{r+s} = (x+1)^r (x+1)^s$ conduit à $B^{r+s} = B^r + B^s$, équation fonctionnelle vérifiée par B . Pour la résoudre , il remarque que :

si $m = \ldots\ldots, -3s, -2s, -s, 0, s, 2s, 3s, \ldots\ldots$

alors $B^m = \ldots\ldots, -3B^s, -2B^s, -B^s, 0, B^s, 2B^s, 3B^s, \ldots\ldots$,

deux progressions arithmétiques où B^m et m sont dans un rapport constant .

L'argument utilisé pour poursuivre mérite aussi d'être signalé : Aepinus indique que ceci reste vrai si m est un nombre réel quelconque , car m peut " *parcourir de façon continue*" les intervalles de la progression précédente .

D'où $B^m = \lambda m$

et $\lambda = 1$, puisque $B^1 = 1$.

Ce qui lui permet d'obtenir tous les coefficients de la série du binôme pour un exposant m réel quelconque .

Aepinus n'en reste d'ailleurs pas là ; il est aussi le premier à aborder le cas d'un exposant imaginaire m , avec un raisonnement malheureusement beaucoup plus flou , qui le conduit à "démontrer" qu'on a aussi $B^m = m$ dans ce cas ! En revanche , il ne précise nulle part la nature du nombre x .

On trouve , près de quarante plus tard , une démonstration du même type dans le **Traité du calcul différentiel et intégral** de Lacroix (1797) . Même s'il est certain , par les nombreuses citations qu'il fait , que Lacroix a lu de très nombreux ouvrages , il semble tout de même peu probable qu'il ait trouvé son inspiration dans cet article d'Aepinus , qu'Euler est d'ailleurs le seul à citer . Il est beaucoup plus vraisemblable qu'à l'origine de la démonstration de Lacroix se trouve la remarque que Lagrange avait faite en 1796 sur le deuxième terme de la série du binôme . En revanche , Lacroix aborde aussi le cas d'un exposant imaginaire ; avec Aepinus , et bien sûr Abel , ce sont les seuls à le faire .

Ce premier type d'équation fonctionnelle n'est donc pas celui qui conduira à la démonstration définitive , et on comprend aisément pourquoi : même après que Cauchy , puis Abel , aient donné leur définition de la continuité , il n'était pas évident de démontrer que le second terme du développement de $(1+x)^m$ était une fonction continue de m .

8 . Multiplication des séries : d'Euler (1774) à Cauchy et Abel .

Dans sa démonstration de 1774 , Euler pose :

$$[m] = 1 + m\,x + \frac{m\,(m-1)}{1\,.\,2}\,x^2 + \frac{m\,(m-1)\,(m-2)}{1\,.\,2\,.\,3}\,x^3 + \ldots\ldots$$

et remarque que , pour m et n entiers positifs , on a :

$$[m]\,[n] = [m+n] \ .$$

Mais , dit-il , " *la composition des termes de ce produit doit demeurer la même , soit que les lettres m et n représentent des nombres entiers , soit des nombres quelconques* " .

Ainsi , Euler estime qu'on peut appliquer à des valeurs quelconques des lettres m et n l'équation fonctionnelle trouvée lorsque ces lettres représentent des entiers positifs . Cette variante du "principe de permanence" , alors fréquemment utilisé , ne convaincra pas tous ses contemporains et sera à l'origine de nombreuses autres tentatives de démonstrations .

Euler en déduit :

$$[m]\,[n]\,[p] = [m+n+p] \qquad \text{et} \qquad [n]^i = [i\,n]$$

puis $\quad [n] = [i\,n]^{\frac{1}{i}} = ((1+x)^{i\,n})^{\frac{1}{i}} = (1+x)^n$

(en choisissant i tel que i n soit entier)

enfin , prenant $n = -\,m$: $[n] = \dfrac{1}{[m]} = \dfrac{1}{(1+x)^m} = (1+x)^n$.

Le calcul est donc fait pour un exposant rationnel , mais la remarque faite plus haut semble indiquer qu'Euler pense la propriété vraie pour un exposant quelconque .

La phrase d'Euler sur l'équation fonctionnelle n'ayant pas été admise par tous , elle est à l'origine de plusieurs articles ayant pour but de combler cette lacune en calculant avec soin les termes de la série -produit , d'abord quand l'exposant est rationnel (par exemple Segner dans les Comptes rendus de l'Académie de Berlin en 1777) , puis quand l'exposant est réel (le calcul le plus voisin de celui de Cauchy se trouve dans un article paru aux Annales de Gergonne en 1818 , trois ans avant l'Analyse algébrique , article écrit par De Stainville , Répétiteur du Cours de Cauchy à l'Ecole Polytechnique) .

Il faut citer aussi les articles de Pfaff (1788) et Rothe (1796) , où ils résolvent plusieurs équations fonctionnelles par dérivation , notamment $\varphi(x)\,\varphi(y) = \varphi(x+y)$, ce qu'ils appliquent ensuite à la série du binôme :

$$\varphi(x) = C^x = [\varphi(1)]^x = (1+a)^x \ .$$

Dans cette évolution qui conduit à Cauchy et Abel , les étapes significatives sont celles où apparaissent progressivement les résultats précis sur la convergence des séries . Sur ce point , l'article de 1812 où Gauss fait une étude rigoureuse et détaillée de la série hypergéométrique

$$\frac{\alpha(\alpha+1)\ldots\ldots(\alpha+n-1)\,\beta(\beta+1)\ldots\ldots(\beta+n-1)}{\gamma(\gamma+1)\ldots\ldots(\gamma+n-1)}\,\frac{x^n}{n!}$$

constitue un moment décisif, par l'influence qu'il exerce sur les mathématiciens de l'époque. Gauss y indique tous les cas où le terme général tend vers zéro, tous les cas où la série converge, mais il utilise encore plusieurs dérivations terme à terme.

En 1816, Bolzano publie à Prague un gros livre (160 pages) consacré entièrement au théorème du binôme. Sa démonstration, très longue, énumère tous les cas : n entier, n rationnel, n réel. Il laisse volontairement de côté les cas où x et n sont complexes, et le cas d'un nombre négatif élevé à une puissance irrationnelle, estimant que les notions de nombre complexe et de nombre irrationnel ne sont pas encore assez bien définies et fondées. Le livre est intéressant surtout par sa préface, où Bolzano rejette l'utilisation des infiniment petits et leur préfère, comme le fera Cauchy en 1821, les suites convergeant vers zéro, et où il critique toutes les démonstrations précédentes, surtout les manipulations purement algébriques des séries infinies.

Il précise : le théorème du binôme exprime que " *la différence*

$$(1+x)^n - 1 - nx - n\frac{n-1}{2}x^2 - \ldots\ldots - n\frac{n-1}{2}\cdots\frac{n-r+1}{r}x^r$$

peut être rendue plus petite que tout nombre donné, si l'on prend dans la série un nombre de termes suffisamment grand ", et toutefois seulement si $|x| < 1$. Il donne aussi une définition de la continuité d'une fonction en un point, très proche de celle qu'Abel donnera dans son Mémoire de 1826.

Bolzano, mathématicien isolé, n'a cependant eu d'influence sur Cauchy. Ses livres ne seront connus en France que plus tard, dans la seconde moitié du XIXe siècle ; les deux hommes se sont pourtant rencontrés, mais après 1830, pendant l'exil de Cauchy à Prague.

Terme de cette suite de démonstrations, l'Analyse algébrique de Cauchy a donc marqué un progrès très important dans l'évolution de l'Analyse et de son enseignement. Les théorèmes qui y sont énoncés pour la première fois, concernant les séries et les fonctions continues, modifient complètement la présentation des notions anciennes et ouvrent la voie à de nouvelles améliorations au XIXe siècle.

L'étude que nous venons de faire de l'histoire de la série du binôme montre qu'elle n'a pas uniquement joué le rôle d'un énoncé de base pendant cette période. C'est aussi à l'occasion des multiples tentatives qui ont été faites pour la démontrer que sont apparues pour la première fois certaines notions fondamentales : la règle de Cauchy, la définition de la convergence d'une série, le critère de Cauchy, les théorèmes sur la multiplication des séries ont été déjà cités. C'est aussi le cas de la règle de d'Alembert : dans ses **Opuscules mathématiques** (1761), il prouve de façon très détaillée que la série du binôme converge pour $x \in]-1,1[$, en utilisant ce moyen. Enfin, "grâce" à l'erreur de Cauchy sur la somme d'une série de fonctions continues, signalée par Abel, la notion de convergence uniforme est en germe et sera dégagée quelques années plus tard.

Si l'Analyse algébrique ouvre une nouvelle période, l'étude précédente montre aussi que c'est le dernier livre marquant construit sur le plan des livres d'Analyse du XVIIIe siècle, où la série du binôme est placée en tête et permet d'obtenir les autres résultats. Il est d'ailleurs frappant de constater que cette série perd d'un coup toute son importance, dès qu'elle reçoit une démonstration rigoureuse ; elle est alors reléguée au rang d'un corollaire immédiat de la formule de Taylor. Peut-être est-ce d'ailleurs le seul énoncé mathématique à avoir ainsi changé de statut après qu'on l'ait démontré. L'intérêt marginal qu'on lui porte aujourd'hui ne permet plus de mesurer l'importance qui fut la sienne dans la construction de l'Analyse au XVIIIe siècle.