

Le Théorème de Sturm

par Sylvain BARON

Cet article est tout simplement culturel et, s'il vaut mieux en général avoir en ce domaine des ambitions un peu modernes comme nous en avons eu il y a quelques années avec un traitement APL du théorème de prolongement de Tietze-Urysohn, il est bon aussi de revenir de temps en temps sur des sujets plus anciens qui restent encore la clef de solutions numériques essentielles sur nos ordinateurs. Comme pour tous les textes culturels de ce genre, l'APL n'est pas indispensable. Il reste néanmoins un outil privilégié d'expérimentation et de compréhension incomparable.

Charles Sturm (1803-1855) est né à Genève, à l'époque chef lieu du département français du Léman (1798-1814), ville dans laquelle il fit ses études. En 1826 il détermine, avec son ami Colladon, la vitesse du son dans l'eau et obtient l'année suivante à Paris, où il réside depuis 1823, le grand prix de mathématiques pour son mémoire sur la compressibilité des liquides. En 1829 il énonce le célèbre théorème qui porte son nom, et qui fait l'objet de cet article, théorème qui donne le nombre de racines réelles d'une équation entre deux limites. A partir de 1830, il étudie avec son ami Liouville la théorie générale des oscillations et aboutit à une classe d'équations différentielles (problème de Sturm-Liouville). Dans deux grands mémoires publiés en 1836, il fait connaître des méthodes de résolution (notions de *valeurs propres* et de développement en *séries de fonctions orthogonales*) qui seront à l'origine de nombreux travaux et découvertes mathématiques. Cette même année il est élu à l'Académie des sciences et en 1840 il succède à Poisson comme professeur à la faculté des sciences et à l'École polytechnique. Ses *Cours d'analyse de l'École polytechnique* et ses *Cours de mécanique de l'École polytechnique* seront publiés après sa mort prématurée.

La résolution numérique des équations, polynômes en puissances entières de la variable x , comprend deux problèmes distincts : la *séparation* des racines et le *calcul numérique* de ces racines. On dit qu'une racine réelle est *séparée* lorsque l'on connaît deux nombres réels qui la comprennent entre eux et qui ne comprennent entre eux aucune autre racine. Le problème de la séparation des racines complexes est en un sens plus simple en raison d'un théorème de Cauchy qui donne le nombre de racines complexes enfermées dans un contour. Enfin, la résolution d'une équation qui a des racines multiples se ramène toujours à la résolution d'une ou plusieurs équations qui n'ont que des racines simples. Depuis le début de la géométrie algébrique (couramment appelée jusqu'en 1960 : géométrie analytique) inaugurée par Descartes vers 1630, on approfondissait les méthodes de séparation et de calcul numérique des racines réelles en l'absence de toute formule connue de résolution par radicaux au delà du quatrième degré. On sait puis Abel et Galois que de tels formules générales ne peuvent pas exister.

Le théorème de Sturm vient clore une longue ligne de recherches et de théorèmes, inaugurés par Descartes lui-même, pour la *séparation des racines*. Les plus célèbres sont les théorèmes de Descartes (1637), de Rolle (1690), de Budan (1811) et de Sturm (1829), sans compter les méthodes des groupements et les règles de Mac-Laurin, Newton et Lagrange pour obtenir des bornes globales entre lesquelles sont comprises toutes les racines réelles positives d'une équation ainsi que des bornes globales entre lesquelles sont comprises toutes les racines réelles négatives. Lorsque les racines sont complètement séparées, c'est ensuite un jeu d'enfant de les calculer une à une avec un ordinateur par des méthodes itératives (dichotomie, valeurs itérées - où la corde coupe l'axe - par la méthode des parties proportionnelles, valeur itérées - où la tangente coupe l'axe - par la méthode de Newton ou à l'aide de fractions continuées par la méthode de Lagrange).

De tous ces travaux que nous venons de citer, nous ne mentionnerons que le théorème de Descartes car il apporte une idée clef, et un vocabulaire associé, que tous utiliseront jusqu'à Sturm inclus. Le vocabulaire est le suivant : on dit que deux termes d'un polynôme en x à coefficients réels ordonnés suivant les puissances entières décroissantes de la variable offre une *variation* quand ils sont de signes contraires; deux termes consécutifs qui ont le même signe offrent une *permanence*.

Le théorème de Descartes consiste dans la proposition suivante :

THÉORÈME - Dans une équation quelconque où le premier membre est un polynôme en puissances entières de x et le second membre est 0, le nombre des racines réelles positives ne peut pas

surpasser le nombre des variations du premier membre, et, quand il est moindre, la différence est toujours un nombre pair.

Ce théorème a une conséquence importante qui est le point de départ des recherches ultérieures. Considérons le polynôme $f(x)$ d'ordre m et les polynômes dérivés successifs $f'(x), f''(x), \dots$ etc. jusqu'au polynôme dérivé d'ordre m qui est une constante. La conséquence suivante du théorème de Descartes n'est valable que si les m racines du polynôme $f(x)$ sont réelles. Dans ce cas particulier le théorème de Descartes permet de démontrer la proposition suivante :

THÉORÈME - Soient $f(x) = 0$ une équation de degré m dont les m racines sont réelles, et $f'(x), f''(x), \dots$ etc. les m dérivées successives du polynôme $f(x)$.

Si, dans la suite des $m+1$ fonctions

$$f(x), f'(x), f''(x), \dots$$

on substitue successivement deux quantités réelles quelconques a et $b, b > a$, et si, après chaque substitution, on compte les variations de signes que présente la suite des résultats, le nombre des variations perdues en passant de $x=a$ à $x=b$ sera précisément égal au nombre des racines de l'équation $f(x)=0$ comprises entre a et b .

Lorsque dans le cas général $f(x)$ n'a pas toutes ses racines réelles, malgré les recherches de nombreux mathématiciens et malgré le beau théorème de Budan (1811) qui utilise aussi les variations de la suite $f(x), f'(x), f''(x), \dots$ on n'était pas parvenu à une solution complète de la question de la séparation des racines réelles d'une équation polynôme. L'idée de Sturm fut d'introduire d'autres fonctions que la simple suite des polynômes dérivés, fonctions qu'on appelle aujourd'hui (voir par exemple Dieudonné[1]) une suite de Sturm $V_1, V_2, \dots, V_j, \dots, V_r$ qui vérifient les conditions suivantes :

- 1°) les V_j ne s'annulent qu'un nombre fini de fois dans $[a,b]$
- 2°) V_0 et V_1 ne s'annulent pas simultanément
- 3°) V_r est une constante
- 4°) il existe des fonctions Q_1, \dots, Q_{r-1} continues dans $[a,b]$ telles que

$$\begin{aligned} V_0 &= V_1 Q_1 - V_2 \\ V_1 &= V_2 Q_2 - V_3 \\ &\dots\dots\dots \\ V_{r-2} &= V_{r-1} Q_{r-1} - V_r \end{aligned}$$

Sturm a d'abord communiqué son théorème avec une suite particulières de « fonctions de Sturm ». Il a ensuite vu tout le parti qu'on pouvait tirer des conditions essentielles de sa suite de fonctions telles qu'il les a restreintes ci-dessus, en construisant d'autres « suites de Sturm » et en étendant son théorème (1835) aux racines multiples et aux racines complexes. Un exposé complet figure dans le célèbre Cours d'algèbre supérieure de J.A. Serret (1865) (Serret[2]) repris dans la plupart des cours de mathématiques spéciales jusqu'en 1914, par exemple Longchamps[3].

Nous allons maintenant donner l'énoncé du théorème ainsi introduit par J.A. Serret: « *Aucun n'était parvenu à une solution complète de la question que nous venons de poser [la séparation des racines réelles]. L'Algèbre offrait ainsi une lacune regrettable, mais cette lacune se trouva comblée de la manière la plus heureuse par le fameux théorème de Sturm. Ce grand géomètre communiqua à l'Académie des Sciences, en 1829, la démonstration de son théorème qui constitue l'une des plus brillantes découvertes dont se soit enrichie l'Analyse mathématique* » Le théorème de Sturm comprend les deux étapes suivantes :

1 - CREATION DE LA SUITES DE FONCTIONS DE STURM - Étant donné l'équation $V=0$ dont le premier membre est un polynôme d'un degré quelconque m de l'inconnue x et qui n'a pas de racines égales, soit V_1 la dérivée du polynôme V . Effectuons la division de V par V_1 , jusqu'à ce que nous soyons arrivé à un reste de degré inférieur au degré de V_1 , changeons les signes de tous les termes de ce reste et désignons par V_2 ce qu'il devient alors. Divisons de même V_1 par V_2 , et, après avoir changer le signes du reste nous aurons un polynôme V_3 dont le degré sera inférieur au degré de V_2 . Divisons pareillement V_2 par V_3 et continuons la même série d'opérations, comme s'il s'agissait de déterminer le plus grand diviseur commun diviseur des polynômes V et V_1 , mais en ayant soin de changer les signes de chaque reste avant de le prendre pour diviseur. L'équation proposée n'ayant pas de racines égales, nous arrivons après un certain nombre $m-1$ de divisions à un reste numérique différent de zéro, et nous représentons par V_m ce reste changé de signe. Nous obtiendrons

ainsi une suite de $m+1$ fonctions, dite « *Suite de Sturm* »
 $V, V_1, V_2, \dots, V_{m-1}, V_m$
dont les degrés, par rapport à x , formeront une suite décroissante.

Cela posé, soient a et $b > a$ deux quantités réelles quelconques données, et substituons successivement a et b à x dans les fonctions de la suite précédente.

2 - THÉORÈME DE STURM - Le nombre de racines réelles de l'équation $V = 0$, comprises entre a et b , est égal à l'excès du nombre de variations que présente la suite des signes des $m+1$ fonctions de la suite de Sturm pour $x = a$, sur le nombre des variations que présente la suite des signes des mêmes fonctions pour $x = b$.

La démonstration du théorème nécessite deux lemmes et environ trois pages, nous ne la donnerons pas bien qu'elle soit très accessible. Nous allons maintenant donner les fonctions APL qui créent la suite de Sturm et celles qui donnent le nombre de variations de la suite pour une valeur donnée de la variable.

Le théorème de Sturm nécessite comme on le voit trois sous-programmes : Deux programmes pour calculer la suite de Sturm (Dérivation de polynômes et Reste de la division euclidienne d'un polynôme par un autre) et un programme pour calculer la valeur d'un polynôme pour une valeur de la variable.

Les programmes communiqués ci-dessous sont élémentaires. Ils sont écrits en APL-PLUS III, très proche de l'APL2. En ce sens, ils s'étendent naturellement aux vecteurs de polynômes (la suite de Sturm est un vecteur de vecteurs enclos) et aux opérateurs comme le produit extérieur.

Les fonctions ci-dessous supposent que les polynômes sont représentés par leurs vecteurs de coefficients, ordonnés selon les puissances décroissantes de la variable.

```

▽ r←a reste b;c
[1] A Reste de la division euclidienne
    d'un polynôme par un autre
[2] s:r←a ⋄ r←((0=ρr)/0),r←(~^a=0)/a
[3] →((ρ,r)<ρ,b)ρ0
[4] a←(1↓((ρb)†a)-b×(a[1]÷b[1])),(ρb)†a ⋄ →s
▽

```

```

▽ r←derive p
[1] A Donne le polynôme dérivé
[2] r←0 ⋄ →(1=ρ,p)ρ0
[3] r←-1↓p×Φ^-1+1ρ,p
▽

```

```

▽ r←c pol x
[1] A Valeur d'un polynôme pour x
[2] r←+/c×(Φx★^-1+1ρc)
▽

```

Nous détaillons ci-après le programme « sturm » qui fait le travail selon les principes décrits plus haut et en en suivant chaque étape :

```

▽ r←p sturm b;w
[1] A p : le polynôme
[2] A b : les bornes d'étude

```

```

[3]
[4]  A Le cas du degré inférieur ou égal à 1
[5]  r←-1+ρp ◊ →((ρ,p)<2)ρ0 ◊ b←2ρb
[6]
[7]  A Le polynôme et sa dérivée
[8]  w←(c p),cderive p
[9]
[10] A Construction de la suite de Sturm
[11] a:w←w,c-(((ρw)-1)▷w) reste (ρw)▷w
[12] →(1<ρ,(ρw)▷w)ρa
[13]
[14] A Les signes des valeurs des polynômes
[15] A pour les bornes ±
[16] r←xw ◊.pol b
[17]
[18] A Les différences des sommes des "variations"
[19] r←|-/+/(-1 0↓r)≠(1 0↓r)

```

▽

Afin de pouvoir faire des tests aisément il est bon d'avoir quelques outils de plus comme le produit, la somme et le quotient de deux polynômes. La fonction « coeff » offre une solution alternative à « prod » pour calculer les coefficients d'un polynôme dont on connaît les racines réelles. La fonction « coeff » met à profit les fonctions symétriques des racines; on la trouve notamment commentée dans Iverson[4].

```

▽ r←a prod b
[1]  A Produit des polynômes a et b
[2]  r←+/(1-↓ρa)φa◊.×b,0×1↓a
▽

▽ c←coeff r;t
[1]  A c : coefficients du polynôme
[2]  A dont les racines sont r
[3]  t←((ρr)ρ2)↑(↓2★ρr)-1
[4]  c←φ((0,↓ρr)◊.=+~t)+.×(-r)×.★t
▽

▽ r←a quotient b
[1]  A Quotient euclidien de deux polynômes
[2]  r←d←10
[3]  s:r←r,d ◊ →((ρ,a)<ρ,b)ρf
[4]  a←(1↓((ρb)↑a)-b×(d←a[1]÷b[1])),(ρb)↓a ◊ →s
[5]  f:r←r,(0=ρr)/0
▽

▽ r←a plus b
[1]  A Somme de deux polynômes
[2]  r←-(ρ,a)↑ρ,b
[3]  r←(r↑a)+r↑b
▽

```

Entre les fonctions "quotient", "prod", "plus" et "reste", on a bien sûr l'identité suivante :

$$p((p \text{ quotient } q) \text{ prod } q) \text{ plus } (p \text{ reste } q)$$

Calculons les coefficients p d'un polynôme dont les racines sont -0.5 2 3 3.1 et 3.2. Pour cela

utilisons la fonction « coeff » :

```
␣←p←coeff -0.5 2 3 3.1 3.2  
1 -10.8 41.77 -63.69 15.82 29.76
```

Il est ensuite aisé de mettre en oeuvre le théorème de Sturm.

```
p sturm -10 10  
5  
p sturm 3.05 3.15  
1  
p sturm 2.1 4  
3  
p sturm 1.9 4  
4
```

Nous laissons aux lecteurs désireux de voir par eux-mêmes toutes les implications de ce théorème, le soin de saisir sur leur ordinateur ces différentes fonctions et d'expérimenter ans les cas suivants :

- Utiliser la fonction « prod » pour créer des polynômes qui ont des racines complexes ou des racines multiples. Tester « sturm ».

- voir le comportement du programme quand une borne est une racine; voir de quelle « coté » tombe le compte de la racine.

- Écrire un programme, utilisant « sturm » comme sous-programme, qui par encadrement décroissant isole ET calcule les racines. Ici les encls permettent de faire des mesures massives sur des couples de bornes.

Par cette suggestion d'essais multiples, nous rejoignons notre introduction qui nous rappelait le rôle privilégié d'APL pour l'expérimentation et la compréhension de très nombreux problèmes.

Références

- [1] Dieudonné J., *Calcul infinitésimal*, Hermann, 1968
- [2] Serret J.A., *Cours d'Algèbre supérieur*, Gauthier-Villars, Tome 1, 6ème édition, 1910
- [3] Lonchamps, M.G. de, *Cours de Mathématiques Spéciales*, Tome 1, Algèbre, 1883
- [4] Iverson, K., *Notation as a tool of thought*, A source book in APL, APL Press, Palo Alto, 1981