JEAN-LUC DORIER

Émergence du concept de rang dans l'étude des systèmes d'équations linéaires

Cahiers du séminaire d'histoire des mathématiques 2e série, tome 3 (1993), p. 159-190
<http://www.numdam.org/item?id=CSHM_1993_2_3__159_0>

EMERGENCE DU CONCEPT DE RANG DANS L'ETUDE DES SYSTEMES D'EQUATIONS LINEAIRES

Jean-Luc DORIER

1 - Introduction

Le rang est certainement l'un des concepts élémentaires les plus riches de l'algèbre linéaire en dimension finie. Invariant fondamental d'une famille de vecteurs, il permet soit d'en mesurer "la taille", en donnant le nombre minimal de ses générateurs, soit d'en évaluer le nombre de liaisons, en donnant le nombre maximal de vecteurs indépendants.

A ce deuxième aspect correspond de façon immédiate l'approche duale, qui consiste à chercher le nombre maximal de conditions indépendantes liant les vecteurs. De façon plus théorique, ce nombre est encore la dimension du sous-espace vectoriel des formes linéaires s'annulant sur la famille, c'est-à-dire du sous-espace orthogonal dans le dual. Le résultat fondamental de dualité étant : $\dim F + \dim F^* = \dim E$, où E est l'espace de référence, F l'espace engendré par la famille de vecteurs, donc dimF en est le rang, et F^* est l'orthogonal. Ainsi le rang des vecteurs et celui des formes orthogonales sont complémentaires par rapport à la dimension de l'espace de référence. La dimension de l'orthogonal, c'est aussi le nombre minimal d'équations qui permette de définir F.

On voit donc que le concept de rang revêt plusieurs aspects étroitement liés mais qui ont néanmoins une relative autonomie, qui fait que, sous certaines conditions, des aspects peuvent apparaître sans que d'autres ne soient explicités (voire explicitables, à un stade donné de l'évolution). Ainsi une analyse historique de l'émergence du concept de rang se doit-elle de prendre en compte cette multiplicité et de regarder séparément et conjointement l'évolution de chacun des aspects. Par ailleurs on ne pourra juger que le concept est arrivé à maturité que lorsque chaque aspect aura émergé et que les ponts permettant de les relier auront été dégagés.

Pour la clarté de l'exposé, j'ai retenu trois grands axes de distinction :

- L'<u>aspect direct</u> : nombre maximal de vecteurs indépendants.

- L'<u>aspect générateur</u> : nombre minimal de vecteurs qui génèrent la famille ou le même espace que la famille.

- L'<u>aspect dual</u> : nombre minimal d'équations de liaisons indépendantes.

Pour l'ensemble de ces aspects, les questions fondamentales à se poser me semblent être:

- Quand et comment chaque aspect a-t-il été pris en compte ?
- Quand et comment s'est-on posé la question de l'invariance pour chacun de ces nombres et comment l'a-t-on résolue ?
- Quand et comment a-t-on relié ces divers aspects entre eux ?

On verra quel type de réponses, ces questions vont obtenir une fois confrontées à la réalité historique. On pourra alors mesurer la distance qui sépare, du point de vue épistémologique, le processus d'émergence historique, de l'approche reconstruite dont dispose le mathématicien actuel pour accéder au concept dans son environnement "moderne".

2 - Le rang en dehors des systèmes d'équations linéaires. Positionnement de la problématique

Dans une étude plus vaste que j'ai faite, j'ai montré l'importance de l'étude des systèmes d'équations linéaires dans la genèse des concepts d'algèbre linéaire (Dorier 1990).

Pour ce qui concerne plus précisément le concept de rang, il apparaît que le cadre des systèmes d'équations linéaires est celui dans lequel la quasi totalité des idées maîtresses se sont élaborées jusqu'au début du XX^e^ siècle. En conclusion de cet article, je donnerai un aperçu rapide des dernières évolutions du concept de rang dans le cadre de la théorie axiomatique des espaces vectoriels. Pour ce qui concerne la période antérieure au XX^e^ siècle, l'autre domaine essentiel dans lequel se sont constitués les concepts d'algèbre linéaire est la géométrie. Or le rang me semble être un concept assez pauvrement représenté en géométrie, essentiellement du fait de la limitation à la dimension trois. De plus la géométrie en dimension n, où le rang a un rôle plus crucial, ne s'est vraiment développée qu'à partir de la deuxième moitié du XIX^e^ siècle et elle utilise beaucoup de résultats issus de l'étude des systèmes linéaires et de la théorie des déterminants, au point que le concept de rang apparaît comme une exportation du second cadre dans le premier.

Ce pourrait être l'objet d'un autre exposé que de voir l'évolution du concept de rang dans les autres domaines où les idées d'algèbre linéaire se sont formées. Pourtant à l'exception notoire des travaux de Grassmann (sur lesquels je reviendrai immédiatement), il me semble que l'essentiel de la genèse du concept de rang est lié à l'étude des systèmes linéaires.

L'*Ausdehnungslehre* de Grassmann représente un événement fondamental et très particulier dans l'histoire de l'algèbre linéaire (Grassmann 1844/1862). Fondamental parce que cet ouvrage fort dense contient sous une forme très spécifique de nombreux résultats d'algèbre linéaire très en avance sur leur époque. Cependant ni la version de 1844, ni celle de 1862 n'ont pu atteindre le public de l'époque. C'est pourquoi aussi géniales qu'aient pu être les idées développées par Grassmann, force est de reconnaître que ses travaux représentent un événement quasi isolé dans la genèse des concepts d'algèbre linéaire. En fait ils ont bien eu quelque influence à retardement; par exemple c'est à la suite de son analyse de l'*Ausdehnungslehre* que

Peano donne la première définition axiomatique des espaces vectoriels (encore que cette approche, elle-même, n'a pas eu beaucoup d'écho). Mais ces effets sont très limités ou tellement tardifs qu'on n'a pu que constater, a posteriori, que Grassmann avait déjà utilisé des concepts redécouverts plus tard, de façon indépendante.

On pourrait néanmoins consacrer une recherche entière au concept de rang chez Grassmann, mais celle-ci serait tout à fait disjointe de l'étude que je me propose de faire ici et vice versa. Toutefois il me semble utile ici de dire en quelques mots l'essentiel des idées de Grassmann sur le rang. Le propos de Grassmann est très général, celui-ci introduit beaucoup de vocabulaire et de notations, qu'il accompagne de nombreuses justifications et considérations paramathématiques. Il semble que ce soit une des raisons essentielles qui ont rendu son œuvre peu accessible. De plus je crois aussi que certaines notions élémentaires sont abordées sous un angle maladroit qui complique certains points ultérieurs du propos. Par exemple la notion de multiplication par un scalaire est totalement absente de la première partie sur les grandeurs d'extension. A la place on trouve les notions de *changement* et de *manière de changement*[1] qui représentent en quelque sorte des vecteurs sans longueur fixe. De même il n'y a pas toujours une distinction très nette - peut-être faute de vocabulaire adéquat - entre ce que l'on appelle maintenant les notions affines et vectorielles. Si on ajoute à ces quelques "maladresses" le fait que, par ailleurs, le propos se veut très général et englobe un champ très large de notions et de concepts, on comprend la difficulté que peut représenter un tel ouvrage, même pour un lecteur actuel et a fortiori pour un lecteur contemporain de Grassmann.

Si on fait donc abstraction des spécificités liées aux prémisses et au contexte de la théorie de Grassmann, on trouve dès les premières pages (§16 de la version de 1844), une définition de la dépendance linéaire, qui en langage moderne serait : " un vecteur dépend d'autres s'il peut s'écrire comme combinaison linéaire de ceux-ci." Et dès le § suivant Grassmann démontre que : *Chaque segment d'un système de m-ième échelon peut être représenté comme la somme de m segments qui appartiennent à m manières de changement indépendantes données du système et ce toujours d'une façon.*

On retrouve dans la deuxième partie, le même type de résultat à propos, non plus des grandeurs d'extension comme ci-dessus, mais des grandeurs élémentaires (§107). En fait ces deux résultats se traduisent en langage moderne par le théorème sur l'invariance du nombre d'éléments d'une base, à ceci près que Grassmann ne montre pas qu'il est impossible d'obtenir un système générateur de moins de m éléments. Néanmoins sa démonstration repose sur le principe fondamental de substitution graduelle des éléments d'une base par m vecteurs

---

[1] Je reprends ici les termes de la traduction de D. Flament de la version de 1844 (à paraître).

indépendants[2]. Or avec ce principe il est facile de voir qu'un système générateur contient au moins m éléments, on peut donc penser que cette remarque n'a pas échappée à Grassmann. C'est donc le principe même du rang qui est posé ici de façon presque immédiate et avec beaucoup de clairvoyance. De fait on peut dire que le concept de rang sous ses aspects direct et générateur est clairement défini par Grassmann et de façon tout à fait semblable à ce qui se fait aujourd'hui (aux différences de contexte près). Par ailleurs l'aspect dual est abordé dans la suite, bien que sous une forme assez éloignée du concept actuel.

La suite de cet exposé montrera que la position de Grassmann est tout à fait originale pour son époque et que dans le cadre des systèmes d'équations linéaires, le concept de rang prendra un tout autre aspect, son évolution étant beaucoup plus lente. En fait, la différence majeure vient de la prédominance de l'aspect dual, favorisée par l'approche à l'aide d'équations linéaires.

3 - Le paradoxe de Cramer : première prise de conscience de la dépendance linéaire, intuition de l'aspect dual du rang.

En 1750 un paradoxe dans le domaine des courbes algébriques, va amener Euler à expliciter la notion de dépendance d'équations linéaires numériques et à entrevoir l'aspect dual du rang.

a - Le paradoxe

Au XVIII^e siècle, l'étude des courbes algébriques fait de grands progrès, ainsi les deux résultats énoncés ci-dessous (en langage moderne), sont communément reconnus comme vrais, à défaut d'être entièrement démontrés :

1) Deux courbes algébriques planes d'ordres respectifs m et n se coupent en au plus mn points. Le nombre de points d'intersection est en général plus petit, quelques points s'éloignant à l'infini, ou devenant imaginaires, mais on sait, qu'en certains cas, il vaut exactement mn.

2) Il faut et il suffit de n(n+3)/2 points pour déterminer entièrement une courbe algébrique d'ordre n.

Avec la correspondance que permet la géométrie analytique entre courbes algébriques planes d'ordre n et polynômes de degré n en deux variables réelles, les deux problèmes se ramènent aisément à des questions de résolution de systèmes d'équations linéaires.

En utilisant les techniques d'élimination alors connues, MacLaurin est l'un des premier en 1720 à conjecturer le premier résultat, sans en trouver de démonstration. Euler se penche également sur la question, il entrevoit la nécessité d'introduire des points imaginaires et des points à l'infini, mais ne peut dépasser l'explicitation de cas particuliers. Bezout, qui a donné son nom au théorème, démontre le résultat d'une façon non entièrement satisfaisante du point de

[2] Ce principe, qu'il est certainement le premier à avoir énoncé, est encore utilisé aujourd'hui dans la théorie des espaces vectoriels. Cependant les mathématiciens tels que Steinitz, qui applique ce principe vers 1910 dans le cadre des extensions de corps, ne font à ce sujet aucune allusion aux travaux de Grassmann.

vue de la rigueur et seulement dans le cas où les deux courbes n'ont pas de directions asymptotiques communes. Il faudra en fait attendre les travaux de Plücker pour avoir un énoncé général et une démonstration complète de ce résultat vers 1830.

Cependant ce résultat est, dès le milieu du XVIII^e siècle, admis bien qu'indémontré. Citons sur ce point Euler (Euler 1750, p.37) : " (...) *Cette proposition, bien qu'on en trouve nulle part, à ce que je sache, une démonstration assez rigoureuse, est pourtant très certaine, comme je le montrerai ci-après ; et les conséquences, qui en ont été tirées, sont si claires qu'il ne reste pas le moindre doute de ce côté-là.*".

Cette attitude qui peut paraître choquante de nos jours était assez fréquente, encore jusqu'à une époque récente. En fait la validité des conséquences que l'on peut tirer d'une proposition suffit à en assurer le bien-fondé.

La deuxième proposition est un problème de dénombrement, une équation polynomiale de degré n en deux variables comporte (n+1)(n+2)/2 coefficients, qui déterminent à un facteur multiplicatif près une seule équation, il faut donc n(n+3)/2 relations, donc points, pour déterminer une courbe algébrique plane d'ordre n.

Le paradoxe auquel Cramer a laissé son nom apparaît dès que $n \geq 3$. En effet on a alors $n^2 \geq n(n+3)/2$, si bien que deux courbes algébriques de même ordre semblent avoir en commun plus de points (n^2) qu'il n'en suffit pour déterminer entièrement une seule !

On sait bien sûr que le problème vient de ce que n^2 points ne suffisent à déterminer une seule courbe d'ordre n, que lorsqu'ils conduisent à des équations indépendantes, ce qui n'est pas le cas si on prend des points communs à deux courbes... Ce fait aujourd'hui élémentaire, ne semblait donc pas clairement analysé avant Euler.

MacLaurin est l'un des premiers à relever ce paradoxe en 1720. Cramer en parle également dans son traité de 1750, *Introduction à l'analyse des courbes algébriques* (Cramer 1750 / Muir, Vol.1, p. 11-14), celui-là même dans lequel il a introduit les déterminants. Il semble avoir soumis le problème à Euler, qui y consacra, la même année, un article intitulé "Sur une contradiction apparente dans la doctrine des lignes courbes" (*op. cit.*.).

b - Résolution du paradoxe par Euler

Dans cet article, Euler commence par énoncer les deux théorèmes cités plus haut. Il en développe ensuite les conséquences respectives, qui lui permettent de bien mettre en lumière le paradoxe, qu'il prend soin d'expliciter pour n=3, 4 et 5, avant de dire que la même contradiction apparaît pour tous les ordres plus élevés. Il réexamine alors les deux propositions et leurs conséquences et en déduit que le paralogisme ne peut découler que de la deuxième proposition, c'est-à-dire du fait qu'une courbe plane algébrique d'ordre n est définie par n(n+3)/2 points. Il explicite la faille en ces termes :

"(...) *en réfléchissant bien sur l'état de cette proposition, nous remarquerons qu'il peut y avoir des cas, où (nn+3n)/2 points donnés ne sont pas suffisants pour déterminer la courbe* *d'ordre n, qui peut être tirée par ces points, ou ce qui revient au même, que (nn+3n)/2 équations ne suffisent pas pour déterminer autant de coefficients* (...)".

A la suite de cette remarque, Euler va donc s'attacher à savoir sous quelles conditions n équations suffisent à déterminer autant de coefficients, ou encore à mettre en évidence les raisons pour lesquelles un système linéaire nxn peut ne pas admettre une solution unique. Il commence avec un exemple de système 2x2 : 3x-2y=5 et 4y=6x-10, dont il remarque qu'on ne peut pas en éliminer l'une des variables sans éliminer l'autre aussi. Il rattache ensuite ce fait à la proportionnalité des deux équations. Puis il énonce la condition générale pour qu'un système 2x2 détermine une solution unique : *que les deux équations soient différentes entre elles, ou que l'une ne soit pas déjà comprise dans l'autre.*

Cette formulation de la condition de liaison est intéressante, en effet elle amalgame en un seul terme (assez ambigu d'ailleurs) deux propriétés de natures a priori assez distinctes mais qui sont équivalentes quand on parle d'équations linéaires :

1 - l'inclusion de l'ensemble des solutions d'une équation dans l'ensemble des solutions de l'autre (exprimée ici en terme de conditions d'élimination)

2 - la proportionnalité des deux équations.

Il examine ensuite le cas de trois équations à trois inconnues, et montre qu'outre le cas précédent où une équation est comprise dans une autre, il se peut qu'une équation soit comprise dans les deux autres, et il donne un exemple, où comme précédemment il met en rapport les deux aspects de la dépendance en utilisant sans l'expliciter réellement le fait que si une équation est combinaison linéaire des deux autres, elle ne modifie pas l'ensemble des solutions du système.

Euler examine encore le cas de quatre équations, puis il énonce la règle générale suivante : "*Quand on soutient que pour déterminer n quantités inconnues, il suffit d'avoir n équations qui expriment leur rapport mutuel, il faut y ajouter cette restriction que toutes les équations soient différentes entre elles, ou qu'il n'y en ait aucune qui soit déjà renfermée dans les autres.*".

Avec cette proposition, le concept de dépendance linéaire est encore imparfaitement dégagé, puisque d'une part la formulation en est purement rhétorique donc n'a pas de caractère opératoire. En effet Euler ne dit pas vraiment ce qu'il entend par *une équation est comprise dans n autres*. Il semble bien que (en langage moderne) ceci puisse être traduit par : "l'ensemble des solutions de la première équation est inclus dans l'ensemble des solutions du système formé par les n autres".

Cette définition serait donc assez éloignée dans l'esprit de l'idée moderne de dépendance linéaire, bien qu'elle en soit l'équivalent. Néanmoins au niveau des exemples qu'Euler expose, on voit apparaître en liaison avec cette notion, l'idée qu'une équation s'écrit comme combinaison linéaire d'autres, mais aucun rapprochement systématique des deux aspects n'est explicité.

D'autre part Euler, dans sa définition finale, continue de distinguer le cas où deux équations sont semblables, cas qui n'a pourtant rien de spécifique, en dépit des apparences, puisqu'il entre entièrement dans la formulation générale qui suit. C'est néanmoins pour l'époque un progrès décisif que d'avoir dégagé, par rapport à une problématique, le concept d'équations dépendantes (même si le terme n'est pas prononcé) en montrant que les conditions qu'elles déterminent sur les inconnues sont alors moins nombreuses qu'il n'y paraît.

De fait la notion de dépendance d'équations ou de formes linéaires (comme on le disait à l'époque) est un concept qui s'affirmera, à la suite d'Euler, et sera couramment employé. Néanmoins il faudra longtemps avant que la caractérisation par les combinaisons linéaires s'impose face à l'idée d'inclusion des ensembles de solutions[3]. Or on va voir que cette spécificité sera un facteur de ralentissement dans l'avancée du concept de rang.

c - Une intuition du concept de rang

De plus dans la fin de cet article, Euler, revenant à son problème initial, examine les divers cas de dépendance selon l'ordre de la courbe et les positions relatives des points choisis.

Il commence avec le cas des coniques (n=2), donc d'un système de 5 équations homogènes en 6 inconnues. Par un choix de repère habile (impossible toutefois si les cinq points sont alignés, mais Euler ne le relève pas, *cf.* remarque à la fin), il se ramène à un système simplifié dont les trois premières équations lui permettent aisément de tirer trois inconnues en fonction des trois autres.

Il examine alors les deux équations à trois inconnues restantes dont il dit qu'elles "***détermineront la courbe cherchée, à moins qu'elles ne soient équivalentes.***" Pour traduire ce terme d'"équivalent", Euler dit que cela arrive si, quand dans chacune des équations on tire une inconnue en fonction des deux autres, on obtient la même valeur. Il raisonne donc bien ici sur la similitude des ensembles de solutions. Néanmoins ceci conduit à une condition de proportionnalité sur les coefficients des équations.

Ainsi on voit que pour pallier à l'inefficacité de son critère (comment déterminer en pratique si une équation est contenue dans d'autres), Euler résout la question de la dépendance en faisant "tomber" le nombre d'équations et de variables par une méthode de substitution, c'est-à-dire qu'il s'intéresse à l'ensemble des solutions; cependant en fin de compte il aboutit à une condition de proportionnalité. Il trouve ainsi une condition portant sur les coordonnées des points qui correspond au cas de l'équivalence des deux équations. Il montre enfin que cette condition signifie qu'au moins quatre des cinq points sont alignés et que les coniques ainsi définies sont des couples de droites (cas dégénéré), une passant par les quatre points alignés, la deuxième étant seulement astreinte à passer par le cinquième point.

---

[3] Pour distinguer ces deux aspects, j'utiliserai dans la suite de façon rapide les termes de "dépendance (au sens d'Euler)" et de "dépendance linéaire".

Une dernière remarque conclut ce paragraphe sur les coniques : " *si tous les cinq points donnés étaient situés en ligne droite, cette même droite avec toute autre satisferait à la question ; donc ce cas sera encore moins déterminé que le précédent.*" On voit ici qu'Euler dépasse le simple problème de la dépendance et aborde la question du nombre (maximal) de liaisons indépendantes, c'est-à-dire ce que j'ai appelé plus haut l'aspect dual du concept de rang.

Regardons ce problème en termes modernes. On dispose d'un système homogène 5x6 (5 équations, 6 inconnues). Son ensemble de solutions est donc un sous-espace vectoriel de $\mathbb{R}^5$, qui est au moins de dimension 1. Par ailleurs aucun des points n'étant confondus, et une équation algébrique comportant un terme constant, les équations ne sont pas toutes nulles et ne sont pas proportionnelles deux à deux, ainsi le système est au moins de rang 2, donc l'espace des solutions est au plus de dimension 3. Ainsi avec 5 points on détermine soit une conique unique (dimension 1), soit une infinité plus ou moins grande (dimension 2 et 3) de coniques. Le cas de quatre points alignés correspond à la dimension 2 et le cas de cinq points alignés à la dimension 3, ce qu'Euler repère en disant que ce cas est encore moins déterminé. Bien sûr sa justification repose sur une vision plus intuitive du problème, il n'empêche qu'il met bien en rapport le nombre de liaisons avec la taille de l'ensemble des solutions. Le problème de l'invariance est, ici, totalement absent, puisqu'Euler choisit un repère bien adapté et ne pose pas le problème de l'influence de ce choix sur la solution.

Pour les courbes d'ordre 3, Euler renonce à traiter tous les cas trop nombreux, ils se contente de montrer que si on prend plus ou moins de points alignés, on obtient une indétermination plus ou moins grande donnant des courbes dégénérées mais il exhibe aussi un cas d'indétermination donnant des courbes non dégénérées.

Il passe ensuite rapidement à l'ordre 4 : "*Quand deux courbes du quatrième ordre s'entrecoupent en 16 points, puisque 14 points, lorsqu'ils conduisent à des équations toutes différentes entre elles, sont suffisants pour déterminer une ligne de cet ordre, ces 16 points seront toujours tels que trois ou plusieurs des équations qui en résultent sont déjà comprises dans les autres. De sorte que ces 16 points ne déterminent (rien de) plus que s'il y en avait que 13 ou 12 ou encore moins, et partant pour déterminer la courbe entièrement, on pourra encore à ces 16 points , ajouter un ou deux points.*".

Dans cette conclusion de son article, Euler donne les premiers éléments constitutifs du concept de rang. Comme tout à l'heure c'est bien l'aspect dual qui domine, mais on peut déceler aussi l'aspect générateur, plus précisément dans la phrase : "*16 points ne déterminent rien de plus que s'ils y en avaient 13 ou 12 ou encore moins.*" Ceci revient à dire que dans le dual des formes linéaires, l'espace engendré par les équations du système ne dépasse pas la dimension 13. Ce point est d'ailleurs relié à l'aspect dual, puisqu'il est justifié par le fait que l'ensemble des solutions est indéterminé (ou plutôt de dimension au moins 2) et que 14 équations suffisent à déterminer une seule courbe (dimension 1). Donc l'ensemble des solutions étant plus large (*i. e* de dimension plus grande), les formes linéaires sont plus interdépendantes (*i. e* de rang plus faible). On notera qu'ici l'aspect générateur intervient au niveau du dual et non du primal, qui joue en fait le rôle de bidual.

d - Conclusion

Ce texte, paru la même année que celui de Cramer introduisant les déterminants (*op. cit.*), est donc riche du point de vue de l'algèbre linéaire. Il correspond à une vision très intuitive des rapports entre la taille de l'ensemble des solutions d'un système linéaire homogène et le degré de dépendance des équations. Néanmoins la définition de la dépendance n'est pas explicitement reliée à l'idée (moderne) de dépendance linéaire et met l'accent sur les conditions d'élimination ou les inclusions d'ensembles des solutions, bien que le concept de dépendance linéaire fonctionne de façon implicite au niveau des exemples.

Quelques éléments posent les jalons pour le concept de rang. L'aspect dual et l'aspect générateur ne sont qu'ébauchés certes et aucune explicitation ne vient consolider les intuitions. De plus le problème de l'invariance n'est absolument pas soulevé, mais il faut dire que le contexte ne s'y prête pas. Par contre les deux aspects, dual et générateur, sont mis en rapport de façon très claire bien qu'implicite et intuitive.

Ce texte est le premier à discuter des conséquences de la dépendance (même si elle n'est pas encore linéaire) des équations d'un système tout en essayant de quantifier l'indétermination ainsi amenée. Euler n'élabore pas de grandes théories, mais explicite une démarche prospective et plutôt empirique pour gérer l'indétermination. On verra que ces idées vont être oubliées, pendant un siècle, face à la domination qu'exerceront les déterminants, offrant un outil plus puissant et plus sophistiqué, mais qui éloigne aussi peut-être d'une approche plus intuitive.

Par ailleurs il est à noter que l'aspect direct du concept de rang semble absent du travail d'Euler. Ce point qui pourrait choquer *a priori*, s'explique par la problématique posée par le paradoxe, l'accent est mis sur les équations et non sur les solutions (ce qui privilégie l'aspect dual) et sur la taille de l'ensemble des solutions (ce qui privilégie l'aspect générateur). Toutefois à l'origine il y a bien un problème de dépendance linéaire, ce qui laisse penser que l'aspect direct n'est peut-être pas totalement absent, mais beaucoup plus implicite. Une difficulté vient aussi de ce que nous reconnaissons des traces du concept de rang dans un contexte où celui-ci n'a pas du tout été défini ou même explicité, ce qui relègue l'aspect direct, plus descriptif qu'opératoire, au second plan.

Enfin on a vu qu'Euler utilise une méthode apparentée à celle dite du "pivot de Gauss" pour résoudre les systèmes et discuter de la dépendance des équations. Cette technique sera employée de façon implicite pendant encore longtemps, bien que là aussi les déterminants auront rapidement la suprématie. Si le nom de Gauss y est resté attaché, c'est qu'il l'a un peu plus systématisée que d'autres, sans toutefois dégager les propriétés de l'algorithme et surtout ses invariants qui s'expriment nécessairement en fonction du rang. Il faudra attendre semble-t-il le milieu du vingtième siècle, pour que les nouvelles possibilités de calculs par l'informatique avec l'essor des techniques numériques (en programmation linéaire et en analyse numérique essentiellement) conduisent à une étude théorique complète de l'algorithme de résolution du pivot de Gauss dont le rang est un pilier fondamental.

4 - Le rang dans la théorie des déterminants

En 1750 Cramer explicite le concept de déterminant (bien qu'il n'utilise pas encore ce terme) et énonce une règle de résolution des systèmes carrés (*op. cit.*). Cependant il ne s'intéresse implicitement qu'au cas des systèmes entièrement déterminés, c'est-à-dire dont le déterminant n'est pas nul, ignorant le cas de l'annulation, c'est-à-dire de la dépendance des solutions. A sa suite, la théorie des déterminants va connaître un essor important, cependant l'étude des cas autres que ceux dit de Cramer restera très peu développée pendant plus d'un siècle. Or dans le cas d'un système indéterminé, c'est-à-dire admettant une infinité de solutions (en fait plus d'une) le rang permet de mesurer le degré de l'indétermination, c'est-à-dire la "taille" de l'ensemble des solutions, ou le nombre de relations de dépendance entre les équations.

Il est frappant de voir que, pendant presque un siècle, à la suite de Cramer, les mathématiciens ne semblent pas avoir vraiment prêté intérêt au cas de l'annulation du déterminant d'un système carré. Assez rapidement après 1750, on a montré quelques résultats sur l'indétermination. Par exemple, le fait qu'un système homogène de n équations en n+1 inconnues admette une solution non nulle a été mis en évidence dès les premiers pas de la théorie des déterminants. Ce résultat se déduit aisément en effet du calcul par développement par rapport à une ligne du déterminant, qui permet de voir que les mineurs d'ordre n, calculés sur le tableau des coefficients des équations, sont une solution particulière du système. Le cas où ces mineurs sont tous nuls, se traite alors par récurrence en enlevant une équation dépendante des autres, etc.

Ce résultat peut en fait apparaître comme l'élément principal pour montrer l'invariance du rang, mais on ne pouvait pas à l'époque en mesurer toutes les conséquences. En fait il faut attendre environ 1840 pour trouver des résultats substantiels et une étude assez systématique des cas d'indétermination.

<u>a - Sylvester : les premiers pas</u>

Un des premiers mathématiciens à s'être penché sur la question est Sylvester. Dans un article de 1840 (Sylvester 1840/ Muir, vol 1, p.227-235), il commence par dire que déterminer n quantités par n équations revient à déterminer les rapports de proportionnalité de n+1 quantités par n équations homogènes, il s'intéresse donc aux systèmes homogènes admettant des solutions non nulles. On sait aujourd'hui qu'une condition nécessaire et suffisante pour qu'un système homogène aient des solutions non nulles est que son rang soit moindre que le nombre de ses inconnues (ce qui est toujours vrai quand il y a moins d'équations que d'inconnues). On a vu plus haut, que certains résultats partiels étaient connus à l'époque.

Sylvester donne tout d'abord un résultat général dont il déduira tous les autres, sous forme d'un théorème donné sans démonstration et énoncé d'une façon assez laconique. En substance, il considère un système homogène de n équations à p inconnues et il dit que toute "équation dérivée générale première" (*general prime derivative)* peut s'obtenir en prenant comme coefficient devant x_i, le déterminant ayant pour première colonne, la colonne des coefficients de x_i dans le système et (n-1) colonnes arbitraires, mais identiques pour tous les coefficients. Le terme "équation dérivée générale première" semble désigner ici toute équation pouvant être dérivée ou déduite, c'est-à-dire en fait dépendante (au sens d'Euler), des équations du système initial.

En faisant un développement, par rapport à la première colonnne, des déterminants qui forment les coefficients de l'équation obtenue par Sylvester, il est facile de voir que celle-ci est une combinaison linéaire des équations du système initial. On peut donc penser que, bien qu'il ne donne pas de démonstration, Sylvester utilise implicitement comme Euler l'équivalence entre dépendance linéaire et inclusion des ensembles de solutions. Pourtant, ici encore ce résultat n'est pas énoncé explicitement.

Dans la suite Sylvester va appliquer ce résultat, pour examiner quelques cas particuliers. Ainsi quand $n = p$, il prend comme colonnes arbitraires, celles de la matrice A du système sauf la $i^{ème}$, il obtient ainsi les n équations dérivées : $\det(A).x_i = 0$. Ainsi il déduit qu'une condition nécessaire pour qu'un système carré homogène admette une solution non nulle est que son déterminant soit nul. Cependant Sylvester ne se pose pas la question de la réciproque qui aurait été une avancée plus importante que ce résultat qu'on peut obtenir directement par la règle de Cramer.

Il regarde ensuite le cas où $p = n-1$. Il prend alors comme colonnes supplémentaires, (n-2) colonnes du système et une arbitraire. Il obtient ainsi les (n-1) équations : $D.x_i = 0$, où D est le déterminant de A augmenté d'une colonne arbitraire. Pour que le système admette donc une solution non nulle, il faut que $D = 0$, donc que les mineurs d'ordre (n-1) issus de A soient nuls puisque la dernière colonne est arbitraire. Là encore, le problème de la réciproque n'est pas évoqué. Par contre, Sylvester prend un exemple avec $n = 3$ et montre que les trois conditions se réduisent à deux. Cayley et Sylvester lui-même généraliseront bientôt ce résultat (*cf* plus bas).

Dans ce texte, Sylvester se contente de dire qu'on peut appliquer la même technique pour tout système homogène comportant plus d'équations que d'inconnues, en introduisant autant de colonnes arbitraires que nécessaire. Il s'intéresse ensuite aux systèmes où $p = n+1$, dont on sait déjà qu'ils admettent des solutions non nulles. Sylvester montre que toutes les solutions sont telles que les composantes sont dans le même rapport que les mineurs d'ordre (n-1) du système qui leur correspondent, ce qui peut permettre de montrer que l'ensemble des solutions est de dimension un si tous les mineurs ne sont pas nuls, mais Sylvester ne va pas jusque là.

Finalement, ce texte offre peu d'avancée par rapport au concept de rang. Sylvester donne des conditions nécessaires pour qu'un système homogène ait une solution non nulle (donc soit de rang inférieur au nombre d'inconnues). Sur le plan conceptuel, rien n'a vraiment progressé, mais les techniques et les règles de calcul mises en place ici vont servir dans d'autres travaux ultérieurs plus fondamentaux, et certains outils sont mis en place, c'est pourquoi ce texte est néanmoins assez important, de façon indirecte, dans la genèse du concept de rang. De plus dans le contexte de l'époque, avoir posé ce type de problème était déjà un pas en avant non négligeable.

b- Cayley et Sylvester : nombre maximal de mineurs "non coévanescents"

Dans un article de 1843 (Cayley 1843/Muir, vol.2, p.14-17), Cayley considère un tableau de (q+1) lignes et n colonnes, $q < n$, la première ligne étant constituée des inconnues x_i, les autres des coefficients Ai,...,Ki. Il explique ensuite que l'on peut extraire C_n^{q+1} mineurs d'ordre (q+1) de ce tableau qui correspondent à des fonctions linéaires en les x_i. Il dit alors, sans démonstration, que si on annule chacune de ces fonctions, les équations ainsi obtenues se réduisent à (n-q) indépendantes, et qu'alors chacune des autres équations s'écrit comme une combinaison linéaire, dont les coefficients ne dépendent pas des x_i, de ces (n-q) équations indépendantes.

Dans cette formulation, on voit que les aspects dépendance (au sens d'Euler) et dépendance linéaire sont clairement mis en rapport, ce qui est une avancée intéressante par rapport au texte précédent de Sylvester. Par contre le résultat est partiellement faux si on ne prend pas quelques précautions. C'est-à-dire que le rang peut être inférieur à (n-q), et les équations se réduisent, dans le cas le plus général, à au plus (n-q) équations indépendantes. On peut donc dire que seul l'aspect générateur du rang est pris en compte ici, alors que l'aspect direct fait défaut. Cette faute venant d'un algébriste de l'importance de Cayley, très familier des matrices et des déterminants est étonnante, elle semble indiquer une difficulté réelle dans une bonne perception du rang, à moins que ce ne soit qu'une négligence.

En 1850, Sylvester énonce un résultat qui rétablit implicitement cette erreur et s'avère être beaucoup plus général, mais est encore donné sans démonstration. Il considère une matrice de m lignes et n colonnes, dont il dénombre les mineurs d'ordre p, ceux-ci sont au nombre de $C_m^p C_n^p$. Il dit alors que le nombre maximal de mineurs d'ordre p "non coévanescents" (*uncoevanescent*), c'est-à-dire non simultanément nuls, ne peut excéder (n-p+1)(m-p+1). La démonstration de ce résultat n'est pas très complexe, quand on sait bien faire des opérations sur les colonnes des déterminants, on peut donc supposer que Sylvester la connaissait, mais qu'il n'a pas jugé bon de la donner.

On retrouve l'énoncé exact du résultat de Cayley de 1843, quand m = p ≤ n. Cependant la formulation de Sylvester ne fait plus apparaître ce résultat comme directement lié au rang, ni même à un problème de dépendance linéaire ou pas d'équations. Toutefois l'auteur utilise ce résultat pour simplifier les conditions de consistance de certains systèmes homogènes. En effet on a vu plus haut qu'il a montré en 1840 qu'un système homogène de n équations à (n-1) inconnues admet une solution non nulle, à condition que tous les mineurs d'ordre (n-1) du système soient nuls. Son nouveau résultat lui permet de ramener cette condition à l'annulation de deux mineurs seulement, (quel que soit n). Mais il ne pose pas la question du choix de ces mineurs dont l'annulation entraîne celle des autres.

Kronecker montrera peu après qu'il est suffisant de choisir (n-p+1)(m-p+1) mineurs d'ordre p, tous issus d'un même mineur d'ordre (p-1) non nul, pour que leur annulation entraîne celle de tous les mineurs d'ordre p. Ce résultat sera utilisé dans la résolution des systèmes indéterminés, notament par Baltzer (*cf.* plus bas).

On voit encore une fois que ces résultats ne sont pas fondamentaux dans la genèse du concept de rang. Ils apportent une amélioration technique non négligeable, mais ne sont pas très producteurs directement sur le plan conceptuel. On notera toutefois, surtout à travers les travaux de Cayley que les deux aspects de la dépendance (linéaire et ensembles de solutions) sont clairement mis en rapport.

c - Baltzer[4], Trudi[5] : les premiers résultats conséquents

Dans une première édition, datant de 1857, de sa *Théorie et applications des déterminants*, traduite en 1861 par J. Hoüel (Baltzer 1857), Baltzer montre comment dans un système carré de déterminant nul, si le numérateur correspondant à une des inconnues dans la règle de Cramer est nul, ceux de toutes les autres inconnues le sont aussi. Sa démonstration repose sur le fait que, si le déterminant général du système est nul, les n mineurs d'ordre (n-1) construits en supprimant une même colonne sont globalement proportionnels à ceux correspondants, obtenus en supprimant une autre colonne. Ce résultat se démontre en utilisant des arguments identiques à ceux utilisés par Sylvester pour montrer que les solutions d'un système homogène de (n-1) équations en n inconnues sont dans les mêmes proportions que les mineurs d'ordre (n-1) du système. Le développement par une colonne du déterminant permet alors de voir que si un

[4] Baltzer, Heinrich Richard (1818-1887) a d'abord été professeur au gymnasium de Dresde puis à l'université de Leipzig. Il a travaillé entre autres à améliorer la formule sur la courbure de Gauss et à montrer l'intérêt des travaux de Bolyai et Lobachevski sur les géométries non euclidiennes. Il a également codirigé des éditions des œuvres complètes de Jacobi et Möbius. Il est intéressant de noter pour l'anecdote que , dans une lettre à Möbius, Baltzer avoue son incapacité, à la suite de la lecture de l'*Ausdehnungslehre*, à entrer dans la problématique de Grassmann.

[5] Trudi, Nicolo (1811-1884). En 1851 il devient professeur à l'université de Naples et travaille essentiellement dans le domaine du calcul différentiel et intégral. En 1862, il publie, à Naples, un ouvrage de 268 pages, intitulé *Théoria de Determinanti, e loro applicazioni*, destiné à des lecteurs débutants, et dans lequel il fait preuve d'un souci de rigueur et d'exhaustivité remarquable pour l'époque.

numérateur s'annule, tous s'annulent puisqu'ils entretiennent les mêmes rapports de proportionnalité que les mineurs d'ordre (n-1).

En 1862, Trudi (Muir, vol.3, p.84-85)) démontre de la même façon le même résultat et le complète en disant qu'alors une des équations peut se déduire des autres. Il est intéressant de noter qu'il obtient ce dernier résultat, en prenant une combinaison linéaire judicieuse (il faut prendre les bons mineurs comme coefficients) des (n-1) dernières équations et en montrant qu'on obtient ainsi la première équation à condition qu'un mineur, au moins, obtenu en supprimant la première ligne et une colonne quelconque ne soit pas nul[6]. Dans cette partie de l'article, on voit donc que les deux aspects de la dépendance des équations sont vraiment posés a priori comme semblables.

Par ailleurs, dans le même traité, Trudi démontre également que si le déterminant construit sur les coefficients de (n+1) équations en n inconnues (seconds membres compris) est nul, alors une des équations se déduit des autres. Pour ce faire, il considère $x = (x_1, x_2, \ldots, x_n)$ une solution des n premières équations, puis il remplace la dernière colonne du déterminant construit sur les coefficients des équations par $col(n+1)-[x_1 col(1)+\ldots+x_n col(n)]$. Cette opération ne change pas la valeur du déterminant, il obtient donc : $dét(a_{ij}).[u_{n+1}-(a_{n+1,1}x_1+\ldots+a_{n+1,n}x_n)] = 0$, ce qui montre que x vérifie alors la dernière équation si $dét(a_{ij})\neq 0$ (le cas où tous les mineurs d'ordre n seraient nuls n'est pas envisagé). Ce qui lui permet de conclure que la dernière équation est comprise dans les n premières. Dans cette démonstration, il est clair que la notion de dépendance des équations repose sur le fait qu'une équation dérive d'autres si elle admet comme solution, au moins toute solution commune aux autres. Ce n'est pas l'idée de dépendance linéaire qui est utilisée.

Ainsi dans deux démonstrations consécutives, Trudi utilise les deux aspects de la dépendance d'équations : la dépendance linéaire dans le premier cas et l'inclusion des ensembles de solutions dans le deuxième. Implicitement ces deux aspects semblent donc être considérés comme équivalents. On retrouve donc la même attitude qu'Euler, mais ici la dépendance linéaire est plus présente et apparaît explicitement comme outil de démonstration sur un exemple général.

Si on regarde ces deux derniers résultats à la lumière des concepts d'algèbre linéaire moderne, chacun d'eux revient à chercher une condition de dépendance de n (respectivement n+1) vecteurs dans un espace de dimension n+1. Or à l'époque, la géométrie en dimension n en était à ses premiers pas et les problèmes de dépendance linéaire dans $\mathbb{R}^n$ peu étudiés. De plus identifier une équation linéaire en n inconnues, comme le (n+1)-uplet de ses coefficients (y compris le second membre) n'est pas une attitude évidente et personne n'avait encore posé le

---

[6] Le cas où cette dernière condition ne serait pas remplie n'est pas examiné, mais il serait facile de voir qu'alors une autre équation que la première serait obtenue comme combinaison linéaire des autres.

problème en ces termes. D'ailleurs la notion de dépendance (au sens d'Euler) représente bien un obstacle à cette idée dans la mesure où si la dépendance linéaire peut s'exprimer de la même façon sur $\mathbb{R}^n$ que sur des équations, l'autre aspect de la dépendance n'a pas de correspondant dans $\mathbb{R}^n$.

Ainsi s'ils montrent des avancées, ces travaux mettent aussi bien en évidence les difficultés qui restent encore à franchir. En particulier le fait de ne jamais traiter les cas "gênants" où un déterminant serait nul montre une certaine régression par rapport à ce qu'avait fait Euler plus d'un siècle plus tôt. Regardons en effet la dernière démonstration de Trudi. Le cas où $\det(a_{ij})=0$, correspond au cas où les n premières équations sont dépendantes, alors la dernière n'est pas forcément déductible des autres mais l'une des n premières l'est des autres. Si on compare avec le texte d'Euler, on voit que celui-ci avait plus de souci de ces cas particuliers et que sa perception plus intuitive des problèmes de dépendance linéaire lui permettait de mieux envisager les différents cas de figure. On peut dire, dans ce sens, que la théorie des déterminants, avec ses outils perfectionnés mais techniques éloigne d'une réalité plus intuitive et a pu, à un certain stade de la genèse, être un obstacle épistémologique à une bonne perception du concept de rang.

Dans une deuxième version datant de 1864 de son ouvrage (Muir, vol.3, p.85-86), Baltzer fait preuve d'une bien meilleure connaissance des systèmes indéterminés. Il donne en effet une règle pratique de résolution qui permet d'en bien comprendre la nature. Son théorème principal est le suivant :

"*Dans un système de n équations en n inconnues, si un mineur d'ordre k ne s'annule pas et si tous les mineurs d'ordre (k+1) construit sur celui-ci s'annulent, alors le système est consistant et (n-k)-uple indéterminé.*"

Ce résultat introduit une nouveauté de taille, car jusqu'alors les annulations de mineurs étaient exclusivement reliées à des questions de dépendance des équations, alors qu'ici intervient la taille de l'ensemble des solutions. On va voir que ce premier pas est fondamental dans la construction des différents aspects du concept de rang et dans l'élaboration de leurs différentes connexions.

La démonstration est assez simple, et depuis devenue classique. Elle consiste à résoudre par la règle de Cramer le système carré construit sur le mineur d'ordre k non nul, les inconnues non principales étant passées en second membre. On obtient ainsi k inconnues en fonctions des (n-k) autres, puis grâce à l'annulation des mineurs d'ordre (k+1), en substituant les valeurs des inconnues dans les (n-k) équations restantes, on voit aisément que ces équations sont encore vérifiées. On obtient ainsi toutes les solutions du système en fonction de (n-k) variables arbitraires (ce qui justifie et définit la dénomination "(n-k)-uple indéterminé") .

On peut compléter en disant que le résultat démontré par Kronecker, à la suite des travaux de Cayley et Sylvester (*cf.* b), permet de dire que tous les mineurs d'ordre (k+1) sont nuls et par suite que tous les mineurs d'ordres supérieurs à k s'annulent. Ainsi k apparaît comme l'ordre maximal auquel on peut trouver un mineur non nul. C'est également le rang du système. Baltzer ne parle bien sûr pas du rang, mais il n'insiste pas non plus sur l'aspect de maximalité, on peut douter cependant que ce fait lui ait échappé.

Dans ce résultat, on voit émerger les aspects dual et générateur du concept de rang. Cependant l'aspect dual n'est pas vraiment complet, parce que l'aspect direct est absent. En effet dans l'aspect dual il y a la prise en compte du nombre minimal d'équations, qui correspond à l'aspect direct sur l'espace dual, non explicité ici. L'aspect générateur n'est également que partiellement dégagé car si Baltzer exhibe ce qu'on appellerait aujourd'hui une base canonique de solutions, ce qui donne le degré de l'indétermination, rien ne prouve qu'on ne pourrait pas ramener celle-ci à un degré moindre, en exhibant une "base" contenant moins d'éléments. En fait c'est tout le problème de l'invariance qui n'est pas encore envisagé. Le nombre k est bien un invariant du système - ceci est évident grâce à Kronecker (rappelons quand même que Baltzer n'y fait pas allusion)-, mais le degré d'indétermination est encore un concept mal défini donc la question de l'invariance ne peut être posée, faute d'une vision assez large du problème. De plus une interrogation importante reste de savoir si tous les systèmes ayant le même ensemble de solutions ont même "ordre maximal de mineur non nul", c'est-à-dire si le rang est aussi un invariant de l'ensemble des solutions lui-même.

Ainsi on peut dire qu'à ce stade :

- Le rang d'un système carré est devenu un objet explicitable et dont la définition est opératoire grâce aux déterminants.

- C'est un invariant du système.

- Il est relié par l'aspect dual à l'ordre d'indétermination du système, qui correspond à l'aspect générateur du rang.

- Mais dans cette correspondance le problème de l'invariance n'est pas posé, pas plus que celui de l'invariance au niveau de l'ensemble des solutions indépendamment du système.

Notons également que Baltzer montre, dans ce même ouvrage, comment trouver les solutions d'un système indéterminé en fonction de celles du système homogène associé. Par contre il ne se pose pas le problème de la réciproque. Ceci permet toutefois de limiter la recherche de l'ordre d'indétermination aux seuls systèmes homogènes, c'est-à-dire, en langage moderne, de ramener la dimension de la variété affine à celle du sous espace vectoriel associé.

d- Dodgson et Rouché : généralisation de Baltzer

A la suite de Baltzer, plusieurs mathématiciens vont donner des règles de compatibilité et des résultats permettant d'évaluer l'ordre d'indétermination d'un système. Parmi eux, j'en distinguerai deux, pour la précision et la clarté des résultats qu'ils ont obtenus.

Dodgson, en 1867, est l'un des premiers à énoncer plusieurs résultats qui rendent compte de tous les cas de figure en fonction des dimensions d'un système d'équations linéaires

(Dodgson 1867/Muir, vol.3, p.86-90). Il lui manque malheureusement une vision d'ensemble qui permettrait de synthétiser ses résultats.

Rouché sera l'un des premiers à atteindre ce but de synthétisation en 1880 (Rouché 1880/ Muir, vol.3, p.91-92). Il s'intéresse à un système quelconque de n équations en m inconnues, dont les premiers membres ne sont pas tous nuls. Il introduit le terme de *déterminant principal*, pour désigner un des mineurs de plus grand ordre (que l'on appellera p) non nul. Les *inconnues principales* sont alors celles qui correspondent au déterminant principal, les autres étant appelées les *inconnues secondaires*.

Il appelle *déterminant caractéristique* tout mineur d'ordre (p+1), obtenu en rajoutant une colonne de seconds membres et une ligne formée des coefficients d'une équation non principale (ou des zéros si p = n). La condition nécessaire et suffisante de compatibilité est alors que tous les déterminants caractéristiques soient nuls. Alors l'ordre d'indétermination est (m-n), ce qui correspond au fait que l'on calcule les valeurs des inconnues principales en fonction des inconnues secondaires qui restent arbitraires (comme l'avait fait Baltzer pour les systèmes carrés).

Cette règle apparaît ainsi comme une généralisation du résultat de Baltzer pour des systèmes rectangulaires. Elle permet donc de définir le rang d'un système quelconque. Mais les mêmes remarques que celles faites sur Baltzer s'appliquent encore dans le cas de Rouché.

e - Smith : un nouveau point de vue

Avec Smith, l'aspect direct du concept de rang au niveau de l'ensemble des solutions d'un système sera mieux pris en compte, et dans ce cadre la question de l'invariance va être partiellement posée et résolue. Dans un article de 1861, Smith considère un système homogène de n équations en n+m inconnues, de matrice A, dont les mineurs d'ordre n ne sont pas tous nuls (Smith 1861/Muir, vol.3, p.83-84). Il traduit immédiatement ce fait, en disant que le système est formé d'équations indépendantes, et que son ordre d'indétermination (*index of indeterminates*) ou encore *l'excès du nombre d'indéterminées sur le nombre d'équations réellement indépendantes*, vaut m.

On peut dire qu'ici Smith donne une définition descriptive du rang vu sous son aspect dual. On pourrait aujourd'hui le traduire en disant en langage moderne que la dimension du sous espace vectoriel des solutions dans $\mathbb{R}^n$ (ce que Smith appelle l'ordre d'indétermination) est égale à n-p, où p est le nombre maximal d'équations indépendantes (*i. e* le rang du système, vu sous son aspect direct). La prise en compte de l'invariance est contournée par l'expression "*nombre d'équations réellement indépendantes*", qui évite d'avoir à montrer que le nombre maximal d'équations indépendantes engendrées par celles du système est toujours identique. Mais par rapport aux travaux que nous venons d'examiner, il est important de noter que l'ordre d'indétermination est directement mis en relation avec le degré de dépendance des équations, ce qui est un fait nouveau. Rien ne permet de savoir à quel type de dépendance (linéaire ou au sens d'Euler) Smith se réfère, mais tout laisse à penser que celui-ci conçoit implicitement les deux aspects comme équivalents.

Il se donne ensuite une famille de r solutions. Il commence par dire qu'il est évident que si r > m, ces solutions forment une matrice X dont tout les déterminants d'ordre r sont nuls, donc que c'est une famille de solutions indépendantes. Il ne donne pas de démonstration de ce résultat. On peut se poser la question de savoir quel est la définition de l'indépendance des solutions pour Smith. En fait la suite du texte laisse à penser qu'il définit l'indépendance des solutions aussi bien que celle des équations par la non annulation d'aucun des mineurs d'ordre maximal, ce qui explique d'ailleurs ici l'absence de démonstration. Or par ailleurs quand il parle de solutions dépendantes, il le traduit immédiatement par le fait que certaines sont combinaisons linéaires d'autres. Ainsi de façon implicite, la dépendance des n-uplets apparaît sous l'aspect combinaison linéaire et annulation d'un mineur de plus grand ordre.

Pour mettre en rapport ces deux aspects, une démonstration consisterait à utiliser d'abord le fait que chaque solution du système est telle que n composantes s'expriment en fonction des m autres. Ceci est très simple à voir en regardant le système comme un système carré, puis en faisant passer m inconnues dans le second membre. L'hypothèse que tous les déterminants d'ordre n de A ne sont pas nuls permet ainsi d'appliquer la règle de Cramer, qui donnent ainsi n inconnues en fonction des m autres. Le résultat repose ensuite sur le fait que si toutes les colonnes d'un déterminant dépendent de moins de colonnes que l'ordre de la matrice alors le déterminant est nul. Le premier résultat se trouvent dans plusieurs textes légèrement postérieurs à celui de Smith, comme celui de Baltzer cité plus haut (je n'en ai personnellement pas trouver de trace avant), le second était connu quasiment depuis le début des déterminants.

On voit donc que la démonstration de ce résultat repose sur des choses tout à fait accessibles à l'époque. Savoir si Smith était en mesure de faire la démonstration, et si oui, s'il a été le premier, n'est pas, me semble-t-il, le point fondamental. Ce qui semble certain c'est que c'est bien vers 1860, que ce genre d'idée a pris forme dans le cadre de la théorie des déterminants, et que Smith est l'un des premiers à y faire référence de façon explicite. De plus ce texte montre une assez grande flexibilité dans la perception du concept de dépendance linéaire tant sur les équations que sur les solutions du systèmes, même si certaines précisions manquent encore, en particulier sur le lien entre l'annulation ou non de certains déterminants et la possiblité d'écrire certaines équations ou solutions comme combinaisons linéaires d'autres.

Par ailleurs, on voit qu'ici c'est l'aspect direct du rang qui est en jeu, toujours en liaison avec l'aspect dual toutefois. Mais il est important de noter que c'est un des premiers textes dans lesquels l'aspect direct est si nettement mis en évidence. Dans ce contexte le point le plus important me semble être le problème de l'invariance qui est très nettement mieux pris en charge que dans tout ce qui s'était fait jusqu'alors. Il est en effet clairement dit (si ce n'est démontré) que le nombre maximal de solutions indépendantes du système ne peut dépasser m.

De plus Smith montre à la fin de son article qu'il existe des familles de solutions indépendantes possédant m éléments, qu'il appelle les familles fondamentales de solutions (*fondamental set of solutions*). Pour ce faire il procède par récurrence sur l'ordre d'indétermination du système. Il commence par donner ce qui va être le résultat fondamental lui permettant de passer d'un pas au suivant. Il considère un système homogène de r équations indépendantes en (m+r) inconnues, dont il se donne une solution non nulle (la question de son existence n'est pas posée) puis il montre que si ce (m+r)-uplet n'est pas solution d'une quelconque (r+1)-ième équation, alors les (r+1) équations ainsi obtenues sont indépendantes. La démonstration élémentaire de ce résultat repose sur le développement par rapport à une ligne d'un déterminant. Ce résultat est donc une mise en rapport explicite entre la définition de l'indépendance choisie par Smith (non annulation de tous les mineurs de plus grand ordre) et celle reposant sur l'inclusion des ensembles de solutions.

On voit qu'ainsi Smith adopte une définition conjointe de l'indépendance pour les n-uplets et pour les équations linéaires homogènes qu'il met en rapport avec l'indépendance linéaire pour les n-uplets et l'indépendance au sens d'Euler pour les équations. En utilisant ce résultat fondamental, il lui suffit de prendre une solution non triviale du système de départ, et une équation dont elle n'est pas racine. Le système ainsi complété est donc encore indépendant, et son ordre d'indétermination est tombé à m-1. Grâce à l'hypothèse de récurrence il peut donc en trouver une famille fondamentale de solutions. L'application de la même règle que précédemment permet alors de montrer que cette famille augmentée de la solution de départ est encore formée de solutions indépendantes, qui constituent donc une famille fondamentale du premier système.

Dans toute cette démonstration, notons que Smith utilise la caractérisation de l'indépendance par non annulation de tous les mineurs de plus grand ordre. Pour conclure il ne reste plus qu'à montrer qu'un système d'ordre d'indétermination 1 admet une solution non triviale, ce qui, on l'a vu plus haut, était connu depuis longtemps. On peut remarquer que ce dernier résultat pourrait également permettre de justifier l'existence d'une solution non triviale dans tout système homogène redondant, ce que Smith a utilisé deux fois (sans le souligner) dans sa démonstration, mais cela ne lui semble pas devoir être justifié.

Ainsi Smith montre qu'il existe toujours dans un système homogène d'ordre d'indétermination m, une famille fondamentale de solutions possédant m éléments et que toute famille de solutions possédant plus de m éléments est liée. C'est une façon de caractériser le rang d'une famille ou la dimension du sous-espace des solutions. Ainsi dans le cadre d'un système ayant n équations indépendantes en n+m inconnues, Smith met bien en évidence l'aspect direct du rang sur l'ensemble des solutions. L'aspect dual est implicitement présent puisque m apparaît d'abord comme l'excédent d'inconnues sur le nombre d'équations indépendantes avant d'être le nombre maximal de solutions indépendantes. La question de l'invariance est très bien résolue et indépendamment du système d'équations.

f - Bilan

A ce stade, regardons les points qui restent encore à éclaircir, par rapport à la notion de rang.

Le rang n'est pas encore un concept explicité, mais il apparaît implicitement l'ordre maximal auquel on peut trouver un mineur non nul. Les connaissances de l'époque permettent donc de dire que c'est un invariant.

Cette façon de définir implicitement le rang permet de faire fonctionner une définition de l'indépendance linéaire commune aux n-uplets et aux équations, ce qui est une condition *sine qua non* pour relier tous les aspects du rang. Cela permet également de prendre en compte partiellement l'aspect générateur au niveau des solutions, mais il reste à voir que l'ordre de l'indétermination est invariant, ce qui appelle une définition du rang intrinsèque à l'ensemble des solutions.

Grâce à Smith, ce problème est partiellement résolu, quand le sous ensemble de $\mathbb{R}^{n+m}$ peut apparaître comme ensemble des solutions d'un système de n équations en (n+m) inconnues dont un mineur d'ordre n n'est pas nul. Ce qui revient à montrer que tout système de rang k est équivalent à un système de k équations idépendantes. Avec les connaissances de l'époque ceci est tout à fait réalisable.

D'un point de vue "technique", il ne manque donc rien aux mathématiciens, vers 1870, pour définir et relier tous les aspects du rang. Cependant la nécessité d'une explicitation ne se fait pas sentir immédiatement. Les questions telles que je les ai formulées plus haut ne peuvent apparaître aussi clairement qu'*a posteriori*. Pourtant certains points sont arrivés à une maturité plus grande que d'autres.

Ainsi les points faibles, me semble-t-il, restent :

- L'explicitation des conséquences de l'annulation de certains mineurs sur le nombre maximal d'équations indépendantes. Ceci permettra de compléter l'aspect dual et d'enrichir l'aspect direct.

- Le pont entre l'aspect direct et générateur, qui permet de déterminer le rang, soit en diminuant les générateurs soit en augmentant les familles libres. En particulier l'ordre d'indétermination doit pouvoir être caractérisé autant en terme de degré de dépendance qu'en terme de nombre de générateurs, ce qui n'est pas clair tant qu'on s'appuie sur les notions d'inconnues principales et secondaires, c'est-à-dire sur une base trop particulière.

- Le problème de l'équivalence entre différents systèmes et donc l'explicitation de l'invariance au niveau des solutions et non pas seulement du système.

En plus de ces quelques points techniques, un pas essentiel non encore franchi reste la reconnaissance explicite d'un concept identique de dépendance pour les équations et pour les n-uplets qui ne repose pas sur les déterminants.

On va voir maintenant comment le dépassementt de cette étape fondamentale va permettre la reconnaissance du concept de rang, comme un outil théorique puissant.

5 - Frobenius : explicitation du concept

Entre 1875 et 1879, le mathématicien allemand G. F. Frobenius réussit à dégager les principales caractéristiques du rang, il est aussi le premier à l'avoir défini en tant que notion, en lui donnant son nom actuel et il en a enfin présenté des applications à de nombreuses reprises.

En 1875, Frobenius publie un traité de presque cent pages sur le problème de Pfaff, dont les §3 et 4, tenant sur sept pages d'une grande densité, sont consacrées aux systèmes d'équations linéaires (Frobenius 1875, p. 255-261). Comme Smith, Frobenius commence par considérer un système de m équations homogènes et indépendantes en n inconnues (m < n) :

$$(10)\quad a_1^{(\mu)} u_1 + \ldots + a_n^{(\mu)} u_n = 0,\ (\mu=1,\ldots,\ m).$$

Il dit ensuite : " *soit* $a_1 u_1 + \ldots + a_n u_n = 0$, *une quelconque relation linéaire entre elles* [les inconnues u_i], *on peut alors la calculer à partir du système* (10)."

Cette réflexion peut sembler anodine, pourtant elle nous renseigne déjà sur la flexibilité dont Frobenius fait preuve dans l'utilisation de la dépendance linéaire. Cette phrase de Frobenius nous montre d'emblée, qu'il a, lui, dépassé les difficultés énoncées plus haut. En effet ce qu'il appelle une quelconque relation linéaire entre les inconnues apparaît comme une équation dépendante de celles de (10) au sens d'Euler (*i. e.* qui admet pour solutions au moins toutes celles du système (10)) ; or il nous dit qu'on peut la *calculer [rechnen]* à partir de celles de (10), c'est-à-dire l'écrire comme combinaison linéaire de celles-ci.

En fait ces quelques pages sont émaillées de nombreuses remarques apparemment anodines comme celle-ci, mais qui en disent long sur la modernité et la richesse de ce texte. Ainsi toujours en guise de préambule, Frobenius explique que si on remplace les m équations de (10) par m autres équations indépendantes construites sur les m précédentes, les mineurs d'ordre m du nouveau système, qui ne sont pas tous nuls puisque les équations sont indépendantes (voilà au passage le troisième aspect de la dépendance relié aux deux autres), sont égaux à ceux correspondants de l'ancien système, à un facteur multiplicatif non nul près. Ce résultat a été démontré par Cayley en 1843 (*op. cit.*), le facteur est en fait le déterminant de ce qu'on appelle aujourd'hui la matrice de passage. Frobenius s'intéresse ici au problème du changement de base. On va voir que la notion de base est pour lui très claire. Ce résultat préliminaire, lui permettra plus loin de montrer que les nouvelles équations forment un système équivalent au premier.

Après ce préambule, l'auteur s'attache à caractériser l'ensemble des solutions. Ainsi, il remarque tout d'abord qu'une combinaison linéaire de deux solutions de (10) en est encore une. Cette remarque, pour évidente qu'elle soit, est fondamentale pour la structure linéaire et, sans doute, Frobenius en voit -il l'importance.

Puis il considère plusieurs solutions : $A_1^{(\chi)}, \ldots, A_n^{(\chi)}, (\chi = 1,\ldots,k)$, et il énonce que *ces solutions seront dites indépendantes si* $c_1 A_\alpha^{(1)} + \ldots + c_k A_\alpha^{(k)}$ *ne peut pas s'annuler pour* $\alpha = 1,\ldots, n$, *sans que* $c_1, \ldots c_k$, *soient tous nuls, en d'autres termes quand les k formes linéaires* $A_1^{(\chi)} u_1 + \ldots + A_n^{(\chi)} u_n$ *sont indépendantes.*

On voit ici se concrétiser un aspect encore plus fort de la remarque du début de texte. Avec cette définition, Frobenius réalise deux avancées importantes. Tout d'abord les n-uplets et les équations linéaires homogènes en n inconnues sont assimilés à une même catégorie d'objets, pour leurs propriétés linéaires. C'est le concept de vecteur et, derrière lui, celui de structure algébrique, qui apparaît ici, certes de façon modeste, mais il me semble que le fait mérite qu'on le remarque. La qualité du travail de Frobenius nous laisse à penser que ce genre de "petites" généralisations ou unifications, participe beaucoup à l'avancée des concepts. En effet on va voir que, grâce à cette mise en rapport, Frobenius va pouvoir s'appuyer sur la dualité pour faire le lien entre les différents aspects du rang, chaque information sur les équations va pouvoir se traduire par un résultat de même nature au niveau des solutions et *vice versa*. Ensuite, il donne de l'indépendance une définition opératoire qui est d'ailleurs, au formalisme près, la définition actuelle. Celle-ci lui permettra beaucoup plus de souplesse dans ses raisonnements.

Si on compare avec Smith, il me semble que le point fondamental est la distance qui est prise par rapport à l'utilisation des déterminants. En effet chez Smith ceux-ci sont omniprésents et rien n'a vraiment d'interprétation hors ce contexte. Ainsi, il définit l'indépendance des n-uplets comme des équations par la non-annulation d'un mineur de plus grande taille; mais cette définition commune, si elle a un côté opératoire, est par contre très pauvre au niveau du sens, ce qui est un grave handicap quand on veut tirer des conséquences et donc énoncer des théorèmes. La définition de Frobenius est beaucoup plus souple et de toutes façons plus riche puisqu'il la relie à celle de Smith.

Frobenius énonce ensuite un résultat important sans démonstration. Il consiste en ce que les équations (10) étant indépendantes, on peut compléter le tableau des coefficients $a_\alpha^{(\mu)}$, par (n-m) lignes de n termes $U_\alpha^{(\nu)}$, de façon à ce que le déterminant D de ce tableau carré ne soit pas nul. En utilisant le théorème principal de Smith (*cf.*§3-e), cela revient à construire une forme linéaire n'ayant pas toutes les solutions de (10) pour racines, puis à faire de même avec le système augmenté d'une équation, etc., jusqu'à avoir n équations indépendantes donc de déterminant non nul. Le problème consiste donc à construire une équation en n variables n'ayant pas toutes les solutions d'un système de (m+k) équations homogènes comme racines (k = 0,..., n-m-1). Ce résultat est en fait assez simple à montrer puisqu'un système homogène ayant moins d'équations que d'inconnues admet une solution non nulle, et qu'il est ensuite très aisé de construire une forme linéaire ne s'annulant pas en ce n-uplet non nul.

Cette démonstration semble donc être assez accessible à l'époque pour ne pas trop s'interroger sur l'absence de preuve chez Frobenius. Il n'est pas impossible toutefois que celui-ci se soit appuyé sur son intuition et ne soit pas posé le problème de la nécessité d'une preuve.

Grâce à ce résultat Frobenius va construire une base de solutions et montrer que toutes les bases ont le même nombre d'éléments. Pour commencer, il note $A_\alpha^{(\upsilon)}$, le coefficient de $U_\alpha^{(\upsilon)}$, dans le développement de D. En remplaçant ensuite une ligne des $U_\alpha^{(\upsilon)}$ par une ligne de $a_\alpha^{(\mu)}$, dans D, et obtient ainsi les m×(n-m) relations :

$$(11)\ a_1^{(\mu)}A_1^{(\upsilon)} + \ldots + a_1^{(\mu)}A_n^{(\upsilon)} = 0 \qquad (\mu = 1,\ldots, m\ ;\ \upsilon = 1,\ldots, n\text{-}m).$$

Ceci donne un moyen de calculer (n-m) solutions du système (10). Ce résultat est en fait une généralistaion du même résultat quand m = n-1, qui, on l'a dit, était connu depuis longtemps.

Frobenius va ensuite s'attacher à montrer que ces (n-m) solutions sont indépendantes. Il obtiendra ainsi un résultat analogue à celui obtenu par Smith (*op. cit.*), avec en plus un procédé de construction beaucoup plus simple. De plus on verra que Frobenius va tirer beaucoup plus de conséquences de ce premier résultat. Pour montrer l'indépendance de (n-m) solutions, il utilise un résultat assez technique, déjà connu à l'époque et qui dit que les mineurs d'ordre m construits sur le tableau des $a_\alpha^{(\mu)}$ se comportent comme les mineurs d'ordre (n-m) complémentaires dans le tableau des $A_\alpha^{(\upsilon)}$ (*i. e.* sont proportionnels deux à deux).

Ainsi puisque les équations (10) sont indépendantes tous leurs mineurs d'ordre m ne sont pas nuls, donc tous les mineurs d'ordre (n-m) des $A_\alpha^{(\upsilon)}$ ne sont pas nuls non plus, et ces (n-m) solutions sont donc indépendantes.

Le théorème au centre de la preuve de Frobenius est donc très technique, alors que celui de Smith était élémentaire; par contre le processus de construction qui en découle est très simple, alors que celui de Smith était fort laborieux.

Frobenius montre ensuite que (n-m+1) solutions de (10) sont toujours dépendantes. Smith a énoncé le même résultat en 1861 (*cf.*§3-e), mais ne l'a pas démontré. J'ai envisagé plus haut comment il aurait pu le montrer avec les connaissances de l'époque. On va voir que la démonstration de Frobenius est fondamentalement différente. Il considère (n-m+1) solutions $B_1^{(\upsilon)}, \ldots, B_n^{(\upsilon)}$ ($\upsilon = 1,\ldots,$ n-m+1). Il construit alors les n valeurs : $B_\alpha = c_1B_\alpha^{(1)} + \ldots + c_{n\text{-}m+1}B_\alpha^{(n\text{-}m+1)}$ (α=1,..., n), qui fournissent une autre solution de (10). Un système homogène de n-m équation en (n-m+1) inconnues ayant toujours une solution non nulle, on peut donc trouver $c_1,\ldots, c_{n\text{-}m+1}$, non tous nuls tels que $B_\alpha = 0$; α = m+1,..., n. Les équations de (10), correspondent alors à un système m×m, dont les inconnues sont les $B_\alpha^{(\upsilon)}$ (α=1,...m). En supposant enfin que c'est le premier mineur du système (10) qui est non nul, on trouve que ces inconnues sont forcément nulles donc que tous les B_α sont nuls, et donc les (n-m+1) solutions de départ sont dépendantes.

Cette démonstration qui repose sur des arguments fort simples est très élégante. Par rapport à celle que j'ai supposée plus haut dans le cas de Smith (*cf.*§3-e), elle a l'avantage de ne pas utiliser de "base canonique", et revêt donc un caractère plus intrinsèque, la seule hypothèse utilisée étant l'existence d'un mineur d'ordre m non nul. Bien sûr, elle se généralise immédiatement à un système quelconque ayant un mineur d'ordre n non nul, bien que Frobenius ne le dise pas (il ne s'occupera de ce cas qu'au §4). Il en déduit par contre immédiatement que toutes les solutions du système (10) s'écrivent comme combinaisons linéaires des (n-m) solutions indépendantes exhibées plus haut. Bien sûr ce résultat découle aisément du précédent, en remarquant que si on rajoute une solution quelconque aux (n-m) indépendantes, on a un système dépendant. Toutefois il faut quand même montrer (ce que Frobenius ne fait pas) que comme les (n-m) premières sont idépendantes, le coefficient de la dernière dans la condition de liaison n'est pas nulle, ce qui permet d'écrire ce dernier n-uplet comme combinaison linéaire des autres. On peut s'interroger pour savoir si Frobenius n'a vraiment pas vu la difficulté, ou, s'il l'a remarquée, pourquoi il n'a pas pris la peine de la préciser.

Néanmoins avec ce dernier résultat, on peut ditre que Frobenius a fait le pont entre l'aspect direct et l'aspect générateur du rang. Cependant, il ne dit pas qu'on ne peut pas trouver une famille génératrice de moins de (n-m) solutions, et ce point demanderait quelques éclaircissements, qui ne seront jamais fournis par Frobenius bien que, comme nous le verrons plus loin, il avait les moyens de le faire sans beaucoup de peine[7]. De même l'invariance de m n'est pas réglée puisqu'elle est mise au niveau des hypothèses en quelque sorte. Ce point sera éclairci au §4.

Enfin au delà du rang, qui n'est toujours pas une notion définie, c'est la notion de base qui est ici explicitée (la famille fondamentale de solutions de Smith), bien que Frobenius ne voit pas l'utilité de lui donner un nom.

En passant il donne ensuite un résultat sur le "changement de base". Il considère en effet (n-m) solutions indépendantes du système (10). Celles-ci s'expriment en fonction des (n-m) solutions particulières construites plus haut et les mineurs d'ordre (n-m) sont tous proportionnels, entre les deux familles de solutions, donc se comportent également comme les mineurs complémentaires d'ordre m du système (10). Ce résultat élémentaire permet de donner une propriété des "bases" de solutions et permet églement de relier deux systèmes ayant les même solutions. C'est l'analogue de ce qui a été dit au début sur le changement des équations.

---

[7] On peut remarquer qu'il y a, ici, la même faiblesse (ou négligence ?) que dans l'*Ausdehnungslehre* de Grassmann.

Ensuite Frobenius regarde des problèmes de dualité (sans utiliser ce terme). Il commence par mettre en évidence la notion de systèmes *adjoints*.. Pour ce faire il considère une famille quelconque de (n-m) solutions indépendantes du système (10), qu'il note comme précédemment : $A_1^{(\upsilon)},\ldots, A_n^{(\upsilon)}$ (υ = 1,..., n-m). Grâce à la correspondance faite entre équations linéaires et n-uplets, Frobenius considère les (n-m) équations homogènes indépendantes formées à partir de ce tableau de nombres. Alors les coefficients du système (10) apparaissent comme m [=n-(n-m)] solutions indépendantes de ce dernier système.

Il dit que les deux systèmes d'équations ainsi formés sont appelés des systèmes *adjoints [zugeordnet ou adjungirt]*. Et qu'il sont liés par les relations :

$$(11^*) \quad a_1A_1 + \ldots + a_nA_n = 0$$

C'est ici un point très important pour l'aspect dual du rang. Ce résultat donne en effet la relation qui lie la dimension d'un sous-espace vectoriel et celle de son orthogonal, et le principe de réversibilité est bien mis en évidence. Dans cette mise en parallèle, il est clair que Frobenius joue constamment avec la possibilité qu'il s'est donnée d'interpréter n valeurs commme un n-uplet ou une équation linéaire homogène (ou une forme linéaire).

De plus le fait que les relations (11*), peuvent être lues soit comme m relations indépendantes de liaisons entre les solutions de (10), soit comme (n-m) relations indépendantes de liaisons entre les équations de (10), permet de donner toute son effectivité à l'aspect dual du rang non seulemnt au niveau du primal mais aussi au niveau du dual. Frobenius fera largement usage de cette caractérisation dans la suite de ses travaux. Ce résultat est en fait un pas décisif dans la genèse du concept de rang, il permet en grande partie de combler les lacunes que j'avais mis en évidence à la fin du paragraphe précédent. De plus il permet, bien que Frobenius n'en parle pas, de montrer qu'une "famille génératrice" de solutions de (10) comporte au moins (m-n) éléments. En effet s'il existait une "famille génératrice" de moins de solutions on pourrait en extraire une "base" de k < (n-m) solutions, alors tout système adjoint serait de rang (n-k) > m, ce qui est impossible car tous les systèmes adjoints sont de même rang donc de rang m comme (10).

A la fin du §3, Frobenius met en application les notions qu'il vient d'introduire pour dégager un critère pour que toute famille de (n-m) solutions indépendantes du système (10), restent indépendantes si on les tronque à leur k premières coordonnées. Il commence par dire que ce résultat n'a d'intérêt que si k ≥ (n-m), mais n'en donne pas de raison. En fait si k < (n-m), le système des solutions tronquées n'est jamais indépendants car on a (n-m) vecteurs dans $\mathbb{R}^k$. La caractérisation par les mineurs non nuls permet bien de voir que le rang d'un système ne peut dépasser ni le nombre d'équations ni le nombre d'inconnues. Dans le cas où k ≥ (n-m), Frobenius établit que les solutions tronquées sont indépendantes si et seulement si dans le système (10) tronqué des k premières colonnes tous les mineurs d'ordre (n-k) ne sont pas nuls. Cette démonstration est intéressante en ce qu'elle utilise les résulats précédents et se construit en faisant jouer tous les aspects mis en place dans le début.

Dans le §4, Frobenius commence par élargir ce qu'il a fait au §3 à des systèmes quelconques. Ainsi, il considère un système : $a_1^{(\mu)} u_1 + \ldots + a_n^{(\mu)} u_n = 0$, $(n=1,\ldots, p)$, quelconque, c'est-à-dire qu'il supprime l'hypothèse d'indépendance des équations et la condition : $p < n$. En citant Kronecker et Baltzer, il rappelle que si un mineur d'ordre m n'est pas nul (il suppose que c'est le premier) et si tous les mineurs d'ordre (m+1) construits sur celui-ci s'annulent, alors tous les mineurs d'ordre (m+1) s'annulent et les m premières équations ont exactement les mêmes solutions que le système général. Il en déduit donc, en appliquant les résultats du §3 à ce système réduit, que le système admet (n-m) solutions indépendantes et que ce nombre est maximal. Enfin il montre que si un système en n inconnues admet (n-m) solutions indépendantes alors tous ses mineurs d'ordre m+1 sont nuls. En effet sinon, il y a m+1 équations du système qui n'ont en commun pas plus de n-m-1 solutions indépendantes et donc le système entier ne peut en avoir plus.

Faisons un bilan. Je définis, pour des raisons pratiques mais en rapport avec ce qu'a fait Frobenius, le rang d'un système d'équations linéaires homogènes en n inconnues, comme l'ordre maximal m auquel on peut trouver un mineur non nul. Cette définition, en germe dans le travail de Frobenius, est, par nature, axée sur l'aspect dual du rang. Ainsi, Frobenius a montré que (n-m) est exactement le nombre maximal de solutions indépendantes du système, ces solutions permettant de générer toutes les autres (aspect direct). Puis, grâce à la prise en compte de la dualité, il a montré également que m est le nombre maximal de relations de dépendance linéaire, indépendantes, que les solutions peuvent vérifier, ou encore que c'est le nombre maximal d'équations indépendantes ayant le même ensemble de solutions (aspect dual). Enfin, il a utilisé le principe de réversibilité, qui revient en langage moderne à dire que le bidual est le primal et que l'orthogonal de l'orthogonal est le sous espace vectoriel de départ. Il a montré, ainsi que (n-m) est le nombre maximal de relation indépendantes qui lient les équations du système. Avec ce même principe il induit implicitement que m est le nombre minimal d'équations qu'il faut pour définir cet ensemble, ou que c'est le complémentaire dans n du nombre minimal de solutions qui permettent d'engendrer toutes les autres (aspect générateur).

Cette description est un peu simpliste car, par exemple, l'aspect générateur apparaît déjà au début en liaison avec l'aspect direct, mais il n'est (implicitement) parachevé qu'en fin de course. La possibilité d'assimiler une équation à un n-uplet (et même plutôt l'inverse) a été un des éléments décisifs qui a permis à Frobenius, essentiellement en jouant sur les rapports de dualité, d'arriver à cette explicitation quasi-exhaustive des différents aspects du rang.

Dans la suite du §4, Frobenius utilise les résultats précédents pour redémontrer d'une façon plus élégante et plus rapide deux théorèmes déjà connus. Enfin, il montre qu'une matrice symétrique est toujours de rang pair et qu'une matrice antisymétrique est toujours de rang impair (en langage moderne). Dans ces trois démonstrations, on peut déjà voir la performance de l'outil, dégagé, même implicitement, au début de l'article.

C'est en 1879, que Frobenius définit explicitement le rang pour la première fois. Il s'intéresse alors aux systèmes d'équations différentielles linéaires homogènes (Frobenius 1879). Il définit celui-ci en rapport à un déterminant, ce qui pour l'époque revient à la notion de rang d'une matrice carrée. Ainsi l'assimilation forme linéaire / n-uplet est consommée, ce qui permet de ramener une famille de n-uplets ou d'équations à un tableau de nombres. Néanmoins cette définition est restrictive par rapport à la réalité puisque les tableaux rectangulaires ne sont pas envisagés. Il semble que Frobenius se soit limité au cas carré parce que c'est le seul qui lui est utile dans la suite de son article. Ses travaux antérieurs montrent bien de toutes façons qu'une telle définition s'étend sans difficulté aux matrices rectangulaires.

La définition que donne Frobenius est la suivante :*Quand dans un déterminant tout mineur d'ordre (m+1) est nul, et que les mineurs d'ordre m ne sont pas tous nuls, j'appelle m, le rang du déterminant.* (*op.cit.* p.435).

Du point de vue de la rigueur, on peut reprocher à Frobenius de ne pas envisager le cas de la matrice nulle, mais ce n'est, me semble-t-il, qu'un détail. Par ailleurs le fait de définir le rang en suppose l'invariance, or la formulation choisie laisse des doutes à un lecteur non averti. Mais, si tous les mineurs sont nuls à un certain ordre alors ils le sont aussi à tout ordre supérieur. Ce résultat élémentaire règle le problème de l'invariance, car si tous les mineurs d'ordre (m+1) sont nuls tous les mineurs d'ordre supérieur aussi, et si un mineur d'ordre m n'est pas nul, à aucun ordre inférieur tous les mineurs pourraient être nuls. L'auteur ne s'attache pas à expliciter ce petit raisonnement, d'ailleurs fort connu à l'époque.

Dans la suite de l'article, Frobenius n'utilise que des matrices symétriques et antisymétriques, dont il avait déjà montré dans l'article précédemment cité que les rangs sont respectivement pair et impair. Néanmoins si le rang apparaît comme un moyen commode de traiter certains problèmes, aucun résultat nouveau sur le rang n'est énoncé par Frobenius dans cet article. En fait l'énoncé de cette définition n'est qu'une mise au point un peu tardive d'un problème entièrement réglé en 1875. On peut dire qu'à cette date, à des détails près, l'essentiel du travail sur les propriétés élémentaires du rang et sur l'explicitation de ses différents aspects, de leur invariance et de leurs interconnexions a été accompli. On a vu que ce travail a demandé beaucoup de temps, plus d'un siècle, et qu'il s'est mené en étroite corrélation avec la théorie des déterminants. Ceux-ci ont été l'outil fondamental dans ce long processus d'émergence. La dualité a également joué un rôle fondamental, même si ce concept n'est pas encore explicite au moment où le rang émerge. En fait, une des idées maîtresses qui a permis de voir dans tous les aspects, dégagés ici et là de façon éparse, une unité, est l'assimilation n-uplet / forme linéaire qui permet de mettre sur un même plan les relations linéaires dans le primal et dans le dual.

Cette idée fondamentale n'est pas apparue avant Frobenius, or on peut voir dans les travaux de ce dernier en quoi cette idée a pu être source de découvertes et d'explicitations intéressantes.

6 - Quelques aspects du rang après Frobenius

A la suite de Frobenius, le rang d'un système linéaire (que l'on nomme aussi parfois la caractéristique d'une matrice) devient une notion très employée, dans tout ce qui touche à ce que l'on appelle aujourd'hui l'algèbre linéaire de dimension finie. De nombreux mathématiciens ont participé à la finition de l'édifice dont les fondations avaient été bien installées par Frobenius. Mais en fait le travail de ce dernier contient déjà, même si ce n'est pas toujours de façon entièrement explicite, les idées essentielles, si bien qu'on peut dire que peu de résultats nouveaux sur la nature même du rang ont vu le jour après 1880, dans le cadre de l'étude des systèmes d'équations, de $\mathbb{R}^n$, des matrices ou de la théorie des déterminants. Ainsi, dans ce paragraphe je me contenterai de citer quelques prolongements qui ont élargi certains aspects du rang en agissant à des niveaux assez fondamentaux. Mon choix est bien sûr subjectif et ne saurait être exhaustif.

a - Capelli[8] et Garbieri[9]

Dans leur *Corso di Analisi Algebrica*, publié en 1886 à Padoue, ces deux mathématiciens italiens consacrent le quatrième chapitre (90 pages) à la théorie pure des déterminants. Dans ce travail, qui s'inscrit dans la lignée d'exposés similaires de certains de leur compatriotes (Rubini en 1863 ou Janni en 1876), ils marquent toutefois une avancée substantielle. En effet, ils introduisent une nouvelle approche du rang à travers une méthode d'élimination équivalente à celle aujourd'hui connue sous le nom de "pivot de Gauss" (Muir, vol.4, p.102-103).

Ils considèrent un système linéaire homogène de m équations en n inconnues de rang k. Ils montrent alors, à l'aide de calculs assez astucieux utilisant des substitutions dont les coefficients sont des mineurs du système, que ce système est équivalent à un système de k équations, triangulaire et dont aucun terme diagonal n'est nul (*i. e.* dans la $p^{\text{ème}}$ équation, $1 \leq p \leq k$, les coefficients de x_i pour $i \leq p-1$ sont nuls et celui de x_p ne l'est pas). Un tel système triangulaire est évidemment de rang k (il a k équations et son premier mineur d'ordre k est le produit de termes non nuls), donc si réciproquement un système est équivalent à un système triangulaire de k équations dont chaque coefficient diagonal est non nul, il est de rang k. On voit ici apparaître un autre type d'invariant qui est encore le rang du système.

Ce résultat a depuis trouvé de nombreuses applications, il est en particulier au centre de tous les problèmes de résolution effective de système (en analyse numérique ou en

---

[8] Capelli Alfredo, né à Milan en 1855, enseigne d'abord à l'université de Rome de 1877 à 1881 puis devient professeur à l'université de Palerme.

[9] Garbieri Giovanni, né en 1847 à Bologne, enseigne d'abord à l'université de Padoue puis devient professeur en 1902 à l'université de Gênes.

programmation linéaire, par exemple). Notons toutefois que la démonstration des deux mathématiciens italiens n'est pas tout à fait celle que l'on rattache habituellement aujourd'hui à la méthode de Gauss, elle n'intervient pas non plus dans le contexte où cet algorithme est essentiellement utilisé de nos jours.

Dans le même texte, les auteurs montrent que si les mineurs de plus grand ordre d'un tableau rectangulaire sont tous nuls, non seulement les lignes mais aussi les colonnes du tableau sont liées par une relation de dépendance linéaire. Capelli et Garbieri donnent un moyen effectif de calculer les coefficients d'une telle relation de liaison, à l'aide de mineurs du système.De plus ils montrent que, si dans un tel tableau on supprime une ligne ou une colonne, puis qu'on recommence (en ne supprimant soit que des lignes soit que des colonnes), jusqu'à obtenir un tableau dont les mineurs d'ordre maximal ne sont pas tous nuls, le plus petit entre le nombre de lignes et le nombre de colonnes du tableau final est bien sûr égal au rang. Ceci montre qu'en fait le rang des lignes du tableau est aussi le rang des colonnes. Le premier résultat était connu, puisque les lignes sont les équations du système, le deuxième est nouveau. C'est en fait un résultat de dualité fort intéressant ; en langage moderne, on dirait qu'une matrice a même rang que sa transposée.

Enfin dans un autre article de 1888, Capelli utilise le rang pour énoncer un théorème aujourd'hui connu sous le nom de théorème de Capelli-Kronecker, ce dernier en ayant proposé une démonstration à peu près à la même époque (Muir, vol.4, p.104-105).

Ce thèorème utilise le rang pour énoncer de façon fort simple un critère de compatibilité d'un système linéaire non homogène : *Pour qu'un système d'équations linéaires non homogènes soit compatible, il est nécessaire et suffisant que les tableaux augmenté* [*i. e.* avec les seconds membres] *et non augmenté aient même rang.*

Si on lit le système linéaire sous la forme : $x_1C_1 + x_2C_2 + \ldots + x_nC_n = B$, où C_i représente le vecteur colonne des coefficients de x_i et B le vecteur colonne des seconds membres, il est clair que la condition du théorème revient à ce que le vecteur B soit dans l'espace engendré par les C_i, ce qui est bien sûr un critère de compatibilité. Ni Capelli, ni Kronecker, n'ont pu raisonner ainsi. Il faudra attendre encore plus d'un demi-siècle pour que ce type d'idées apparaissent. En fait la démonstartion de Capelli repose entièrement sur la théorie des déterminants et est, dans ce contexte, très technique.

En 1905, Frobenius redonne une démonstration du même théorème, basée sur un lemme plus général (Frobenius 1905). Cette démonstration utilise encore des résultats très techniques sur les déterminants.

b - Le rang dans la théorie axiomatique des espaces vectoriels

Les premières définitions axiomatiques des espaces vectoriels sont tout à fait contemporaines de l'émergence du rang (Peano 1888; Burali- Forti, Marcolongo 1909). Pourtant dans ce domaine, les choses avancent très peu. En fait il faut attendre plus de trente ans, pour que cette définition soit vraiment utilisée, et que du même coup les premiers résultats consistants apparaissent (*cf.* Dorier 1990).

En 1920 Banach est un des premiers à utiliser une approche axiomatique des espaces vectoriels, mais comme il s'intéresse à des questions d'analyse fonctionnelle, il traite essentiellement avec des espaces de dimension infinie et les questions de rang (fini) ne sont pas au centre de ses préoccupations (Banach 1920).

En fait l'idée de rang, sous son aspect "dimension d'un sous-espece vectoriel", hormis dans les travaux de Grassmann (dont nous avons parlé en introduction), semble bien être née de la théorie des extensions de corps. B. L. van der Waerden m'a confié dans une lettre que les travaux de Peano ou de Burali-Forti et Marcolongo n'étaient pas connus en Allemagne au début du siècle. Par contre depuis environ 1900, sous l'impulsion entre autres des travaux de Hilbert, les mathématiciens étaient familiers du concept d'extension de corps et considéraient comme évident que le degré d'une extension est un invariant. La première démonstration connue de ce résultat est due à Steinitz et date d'environ 1910[10], elle repose sur le principe de substitution un par un des éléments de la première base par les éléments de la seconde (Steinitz 1930)[11]

En 1930, van der Waerden publie la première version de sa *Moderne Algebra.* C'est un des premiers ouvrages où l'on trouve une approche du rang dans le domaine très général des espaces vectoriels et même des modules définis axiomatiquement. Les théorèmes et les démonstrations de van der Waerden sont tout à fait identiques à ce qu'on fait encore aujourd'hui. Après avoir défini la dépendance puis l'indépendance linéaires, il montre, de la même façon que Steinitz l'avait fait pour les extensions de corps, que, dans un espace vectoriel, deux familles libres équivalentes (*i.e.* chaque vecteur d'une famille s'exprime comme combinaison linéaire des vecteurs de l'autre) ont le même nombre d'éléments. Il est alors facile de définir le rang. Van der Waerden s'intéresse aussi aux systèmes d'équations linéaires sur un corps quelconque. Pour démontrer et énoncer la plupart des théorèmes connus jusque-là avec des déterminants, il utilise les résultats précédents.

7 - Conclusion

Avec cette dernière étape décisive, le concept de rang, comme la théorie des systèmes linéaires, se trouvent affranchis de l'utilisation des déterminants. Les mathématiciens réalisent alors que la dépendance linéaire exprimée sous sa forme moderne est un concept fondamental, qui se passe de l'utilisation des déterminants. Le rang et la dimension se trouvent ainsi devenir

---

[10] La date de 1910 fait référence à un article dans le *Journal für die reine und angewandte Mathematik, 137-309,* alors que la référence citée est un livre publié en 1930 et réédité en 1950, donc plus facile d'accès, et qui comporte la même démonstration.

[11] J'ai dit, plus haut, que ce même principe avait été appliqué par Grassmann dans son *Ausdehnungslehre* en 1844, mais Steinitz n'y fait pas référence. Il semblerait donc que ce principe ait été retrouvé de façon indépendante.

des concepts clés constitutifs de la théorie des espaces vectoriels. Il n'en reste pas moins que presque deux siècles se sont écoulés depuis Euler et le paradoxe de Cramer, et que, suite à ce long processus de maturation, le rang se trouve intimement mêlé à cette approche maintenant dépassée.

Rapidement après les années trente, la théorie axiomatique des espaces vectoriels va s'imposer dans de nombreux domaines, reléguant les déterminants au rang de simple application, et dont l'utilité ne sera guère reconnue que pour des calculs effectifs, avant que là aussi, l'algorithme de Gauss et l'utilisation de l'informatique ne les supplantent. Ainsi aujourd'hui, l'environnement et les outils accompagnant le concept de rang sont assez différents de ceux qui ont participé à sa genèse. De fait la nature même du rang a évolué. En particulier l'aspect dual est certainement passé au second plan par rapport aux aspects direct et générateur, alors qu'on l'a vu, ce fut longtemps l'aspect dominant et un instrument de progrès dans le processus d'émergence.

Si on examine sommairement cette question sur le plan de l'enseignement, on peut se demander dans quelle mesure on n'aurait pas intérêt à prendre mieux en compte la dimension historique. Tout ce qui touche à la dualité est en effet réputé difficile, tant semble-t-il du côté enseignant qu'étudiant, et il est vrai qu'une approche théorique complète de la dualité est un saut conceptuel susbstantiel, pas toujours facile à franchir. Mais n'y aurait-il pas une possibilité, à travers l'étude des systèmes linéaires, pour permettre aux étudiants une approche plus intuitive, bien que partielle, de ces questions de dualité ? C'est peut-être une question naïve, mais le processus historique que je viens d'examiner me semble fortement inciter à y regarder de plus près.

Bibliographie

Baltzer 1857, *Théorie et applications des déterminants,* Traduit de l'allemand par J. Hoüel, Mallet - Bachelier, imprimeur - libraire, Paris, 1861.

Banach 1920, "Sur les opérations linéaires dans les ensembles abstraits et leur application aux équations intégrales", in *Œuvres*, Editions scientifiques de Pologne, Varsovie, 1967-79, Vol.II, p.306-348.

Burali-Forti, Marcolongo 1909, *Omografie vettoriali con applicazioni alle derivate rispetto ad un punto e alla fisica matematica,* G B Pretrini di Giovani Gallazio, Turin.

Cayley 1843, "Chapters in the analytical geometry of (n) dimensions", in *Mathematical Papers*, Johnson reprint, 1963, Vol. 1, p.55-62.

Cramer 1750, *Introduction à l'analyse des courbes algébriques*, Genève.

Dodgson 1867, *An elementary treatise on determinants with their application to simultaneous linear equations and algebraical geometry*, London, 1867.

Dorier 1990, *Analyse historique de l'émergence des concepts élémentaires d'algèbre linéaire*, Cahier Didirem n°7, IREM de Paris VII.

Euler 1750, "Sur une contradiction apparente dans la doctrine des lignes courbes", in *Opera Omnia* (1911-), s. 1, t. XXVI, p.33-45.

Frobenius 1875, "Über das Pfaffsche Problem", in *Gesammelte Abhandlungen*, édité par J P Serre, Springer-Verlag, Berlin-Heidelberg-New York, 1968, vol.1, p.249-334.

Frobenius 1879, ""Über homogene totale Differentialgleichungen", in *Gesammelte Abhandlungen*, édité par J P Serre, Springer-Verlag, Berlin-Heidelberg-New York, 1968, vol.1, p.434-453.

Frobenius 1905, "Zur Theorie der linearen Gleichungen", in *Gesammelte Abhandlungen*, édité par J P Serre, Springer-Verlag, Berlin-Heidelberg-New York, 1968, vol.3, p.349-354.

Grassmann 1844/1862, "Die lineale Ausdehnungslehre", in *Mathematische und Physikalische Werke*, F. Engel ed., Leipzig 1894-1911.[12]

Muir 1906, *The theory of determinant in the historical order of development* (4 vol.), Dover, New York, 1960.

Peano 1888, *Calcolo geometrico secundo l'*Ausdehnungslehre *di H. Grassmann e preceduto dalle operazioni della logica deduttiva*, Fratelli Bocca editori, Turin, 1888.

Rouché 1880, "Note sur les équations linéaires", *Journal de l'Ecole Polytechnique*, 48e Cahier, t. XXIX, 1880.

Smith 1861, "On systems of linear indeterminates equations and congruences", in *Collected Mathematical Papers*, Chelsea, New York, 1965, t.1, p.367-409.

Steinitz 1930, Algebraische Theorie der Körper (Hasse-Baer, De Gruyter, Berlin, Leipzig, 1930), Chelsea, New York, 1950.

Sylvester 1840, "Derivation of coexistence, Part II, being the theory of simultaneous simple homogeneous equations", in *Collected Papers*, Chelsea Publishing Co, New York, 1973, vol.1, p.47-53.

Sylvester 1850, "Additions to the articles *On new class of theorems* and *On Pascal's theorem*", in *Collected Papers*, Chelsea, New York, 1973, vol.1, p.145-151.

Van der Waerden 1930, *Modern Algebra*, Trad. anglaise de la 2e édition allemande, par F. Blum, Frederick Ungar Pub. Co., New York, 1949, vol. 1, p.99-106.

(DidaTech, LSD2, BP 53X, 38 041 Grenoble Cedex)

---

[12] Une traduction en français de la version de 1844 est en préparation par Dominique Flament, je me suis appuyé sur une version provisoire de ce texte. Je tiens à remercier ici D. Flament de son aide précieuse.