

UNIVERSITÉ PARIS 7 – DIDEROT

École doctorale Savoirs scientifiques :
Épistémologie, histoire des sciences, didactique des disciplines
Laboratoire SPHERE - CNRS

THÈSE DE DOCTORAT

Discipline : Histoire et Philosophie des Mathématiques

Présentée par **FRÉDÉRIC JAËCK**

**QUELQUES ASPECTS DE L'ÉLABORATION DE
L'ANALYSE FONCTIONNELLE JUSQU'À BANACH**

RÔLE DES OPÉRATIONS DANS UNE APPROCHE GÉNÉRALE

Thèse dirigée par **KARINE CHEMLA**

Soutenue le 17 décembre 2015 à Paris

JURY :

KARINE CHEMLA (directrice)

ROSSANA TAZZIOLI (rapporteur)

JEREMY GRAY (rapporteur)

HOURYA BENIS SINACEUR

THIERRY COULHON

JEAN-JACQUES SZCZECINIARZ

UNIVERSITÉ PARIS 7 – DIDEROT

École doctorale Savoirs scientifiques :
Épistémologie, histoire des sciences, didactique des disciplines
Laboratoire SPHERE - CNRS

THÈSE DE DOCTORAT

Discipline : Histoire et Philosophie des Mathématiques

Présentée par FRÉDÉRIC JAËCK

**QUELQUES ASPECTS DE L'ÉLABORATION DE
L'ANALYSE FONCTIONNELLE JUSQU'À BANACH**

RÔLE DES OPÉRATIONS DANS UNE APPROCHE GÉNÉRALE

Thèse dirigée par KARINE CHEMLA

Soutenue le 17 décembre 2015 à Paris

JURY :

KARINE CHEMLA (directrice)

ROSSANA TAZZIOLI (rapporteur)

JEREMY GRAY (rapporteur)

HOURLA BENIS SINACEUR

THIERRY COULHON

JEAN-JACQUES SZCZECINIARZ

*Beatitudo non est virtutis praemium sed ipsa virtus
nec eadem gaudemus quia libidines coercemus sed
contra quia eadem gaudemus, ideo libidines coercere
possumus.*

— Baruch Spinoza, *Ethica*.

Remerciements

Ces années de thèse, le temps de formation et de recherche, les séminaires, les colloques ont été l'occasion de rencontres humaines très importantes.

Je remercie très vivement les membres du jury d'avoir accepté de suspendre un temps leurs propres activités de recherche pour lire ce mémoire. Si la thèse se doit d'être un travail de recherche original et non trivial, alors s'engager dans l'épaisseur d'une pensée pas tout à fait décantée et encore en mouvement demande des efforts importants.

Je remercie mes deux rapporteurs Rossana Tazzioli et Jeremy J. Gray pour avoir accepté de lire des versions en chantier et pour leur très précieux travail d'expert. Je suis aussi très reconnaissant de l'accueil bienveillant et amical qu'ils m'ont fait depuis nos premières rencontres à la faveur d'un colloque ou d'un séminaire alors que je faisais mes premiers pas en histoire des mathématiques.

Je suis très heureux de la composition de mon jury, bien sûr pour son expertise, mais aussi parce qu'il démontre la possibilité d'une grande intimité et d'échanges essentiels entre histoire, philosophie et mathématiques. Donc merci beaucoup à Hourya Benis Sinaceur, Jean-Jacques Szczeciniarz et Thierry Coulhon, j'espère que cette soutenance n'est que le début d'une longue discussion !

Enfin je remercie de tout cœur Karine Chemla dont le rôle a largement dépassé celui de directrice de thèse. Elle m'a accueilli en histoire des mathématiques d'une façon exceptionnelle et a accepté d'accompagner ma progression depuis mes premiers pas enthousiastes et peu rigoureux jusqu'à la publication de mon premier article —après mille relectures millimétriques et autant de réécritures. Tout ce qui est écrit dans cette thèse est le fruit de nos nombreuses discussions et de sa générosité intellectuelle et humaine depuis maintenant plusieurs années.

Cette thèse a aussi bénéficié de nombreuses rencontres, de la ferveur intellectuelle qui anime le laboratoire SPHERE, de l'accueil particulièrement amical que j'y ai ressenti. Les choses importantes se passent parfois dans un sourire, dans un mot, dans le foisonnement des idées, dans la possibilité de parler de mathématiques librement, de pouvoir écouter ensemble un exposé sur les mathématiques en Chine ancienne, un exposé sur les catégories ou sur Wittgenstein...

Un grand merci à Sandrine Pellé et Patricia Philippe pour leur aide administrative et organisationnelle : ce qu'elles font est essentiel.

Il y a une vie après la thèse pour dire son amour à ceux qui sont si importants et sans qui tout ceci n'aurait pas existé. Je pense à eux en écrivant ces quelques mots et nous nous retrouverons très bientôt !

Je finis ces remerciements en revenant au commencement qui fût la rencontre avec Jean-Luc Verley. Sa fille Lila avait eu la très bonne idée de réunir autour d'un repas aux douces saveurs périgourdines deux mathématiciens qui devaient bien avoir des choses à se raconter. Je me souviendrai toujours de la soirée que nous avons passée : j'ai découvert un homme passionné, érudit, très bon vivant et nous avons immédiatement sympathisé.

Je ne pourrais pas dire ici tous les thèmes que nous avons abordés, entre mathématiques, philosophie, littérature, bonne chère et autres truculences... Alors que je lui demandais par où s'engager en histoire et philosophie des mathématiques, il me cita, après un instant de réflexion car il y avait des choses importantes à ses yeux, deux noms : Hourya Benis Sinaceur et Karine Chemla.

Le lendemain matin il me téléphonait aux aurores en me donnant avec enthousiasme le numéro de Karine et en me disant de l'appeler immédiatement car elle partait en Chine !

Je lui dois cette rencontre et cette thèse.

Table des matières

	Page
Remerciements	v
Table des matières	vii
1 Introduction	11
2 Recours à l’algèbre symbolique : D. F. Gregory	19
2.1 Introduction	19
2.2 L’École Algébrique Anglaise	21
2.3 Séparation des symboles d’opération	24
2.4 L’algèbre symbolique	37
2.5 De l’algèbre symbolique à une théorie	45
Le rôle du signe +	47
Multiplicité	48
Impossible mais pas imaginaire	52
Généralité	53
2.6 Opération et sujet de l’opération	53
2.7 Conclusion	57
3 Volterra et la notion de fonction comme paradigme général	59
3.1 Introduction	59
3.2 Un calcul différentiel des substitutions	62
3.2.1 Opérations sur les substitutions	63
3.2.2 Dérivation des substitutions	65
3.2.3 Intégration des substitutions	68
3.2.4 Substitutions de plusieurs variables	70
3.2.5 Résultats phares de la théorie	72
Différentielles exactes : condition nécessaire	72
Intégration des formes différentielles totales	75
Variation de l’intégrale d’une substitution	79
Intégrales doubles et curvilignes	80
Paramètres différentiels d’ordre 2	83
3.2.6 Extension aux nombres complexes	85
3.2.7 Lien avec les équations différentielles	88
3.2.8 Conclusion	91

3.3	La « fonction de ligne » : un concept général pour l'analyse	93
3.3.1	Jellett comme source d'inspiration	95
3.3.2	Extension de la théorie de Riemann	105
3.3.3	Les « fonctions qui dépendent d'autres fonctions »	107
3.3.4	Les « fonctions de lignes »	123
3.3.5	Linéarité et simplicité	127
3.4	Conclusion	132
	Un paradigme fondé sur une notion de fonction	132
	La dimension $n \geq 2$	133
	Mécanique héréditaire	135
	Une lecture à travers le couple local / global	136
	Passage du discret au continu et du fini à l'infini	138
	Note : « variation » et « déplacement » chez Jellett.	144
4	Vers une analyse générale chez Fréchet	147
4.1	Introduction	147
4.2	Premiers contacts avec les fonctions de lignes	149
4.2.1	Hadamard sur les fonctions de lignes	149
4.2.2	Premiers travaux de Fréchet sur les fonctions de lignes	151
4.3	Extension de l'étude des fonctions	159
4.3.1	Généralisation d'un théorème de Weierstraß	159
4.3.2	La thèse de doctorat de Fréchet	161
	Ensembles abstraits, limites et opérations	162
	Les classes (L)	163
	Les opérations sur les ensembles abstraits (L)	171
	Suites d'opérations	173
	Recherche de la « classe » la mieux adaptée	177
	Les classes (V)	178
	Nouvelle définition de la continuité	180
	Application de la théorie générale	181
4.4	Les opérations linéaires	182
4.4.1	Un théorème d'Hadamard	183
4.4.2	Généralisation par extension	186
4.4.3	Modalisation du discours et éléments structurels	192
4.5	Conclusion : vers une analyse générale	198
5	Rôle de Stefan Banach	205
5.1	Introduction	205
5.2	La thèse de doctorat de Banach	208
5.2.1	Peano et les « transformations des systèmes linéaires »	209
5.2.2	Pincherle et les « opérations distributives »	214
5.2.3	Les espaces de Banach en question	223
	Les espaces vectoriels	224
	Norme, limite et complétude	227
	Sphères et topologie	231
5.2.4	Les opérations sur des ensembles abstraits	233
	Continuité	235
	Le style de Banach	236

	Théorie des opération additives	237
	Utilisation des séries	242
5.2.5	Traitement des ensembles de fonctions	243
5.2.6	Conclusion	247
5.3	Les fonctionnelles linéaires	249
5.3.1	Le principe d'extension des fonctionnelles.	250
5.3.2	Théorèmes d'extension de Hahn et de Banach	257
5.3.3	Un espace de suites numériques comme modèle	259
5.3.4	Topologie des fonctionnelles	265
6	Conclusion	273
6.1	L'analyse fonctionnelle : une approche générale	273
6.1.1	Une conception abstraite de fonction	273
	La fonction soumise aux lois de l'algèbre	273
	La fonction comme relation	277
6.1.2	Transposition de la théorie des fonctions ordinaires	278
	Dérivée et différentielle	278
	L'approche de Weierstraß comme modèle	279
6.1.3	Équations différentielles et problèmes intégraux	279
6.2	Espaces et opérations	281
6.2.1	Espaces linéaires	282
6.2.2	Éléments topologiques	283
6.3	Une histoire héréditaire de l'analyse fonctionnelle	285
	Des mathématiques et des mots	285
	L'analyse fonctionnelle	286
	Préhistoire de la théorie des opérateurs	287
	Bibliographie	289
	Références des textes de Gregory, premières éditions (Chapitre 1)	299
	Table des figures	300
A	Première lettre de Fréchet à Volterra (Automne 1903)	301
B	Lettre de Fréchet à Volterra (30/07/1913)	305

Introduction

ANALYSE FONCTIONNELLE. Dans la préface des *Leçons d'analyse fonctionnelle* de Paul Levy, publiées en 1922, Jacques Hadamard écrit :

Logiquement parlant, le calcul fonctionnel aurait dû se constituer dès la naissance même de la notion d'intégrale définie ou, plus exactement, lorsque avec cette notion fut élaborée celle même de fonction au sens de Dirichlet, la fonction étant considérée comme définie non par telle ou telle série d'opérations analytiques, mais par la connaissance de toutes ses valeurs ; l'intégrale définie fait précisément intervenir l'ensemble de ces valeurs et par là constituait un premier fait de Calcul fonctionnel.

Un peu plus loin Levy précise le processus qui a permis à ses yeux de passer des équations différentielles ou intégrales à une théorie générale :

Le calcul fonctionnel est une branche de l'analyse qui dérive du calcul différentiel et intégral de la même manière que ce calcul lui-même dérive de l'algèbre, par une généralisation progressive des opérations effectuées.

De nombreuses études d'histoire des mathématiques ont montré l'importance du calcul différentiel et intégral dans le développement de l'analyse fonctionnelle. Les problèmes y prennent la forme d'équations dont les inconnues sont des fonctions. D'autres travaux ont aussi mis en évidence les liens qu'il existe entre l'analyse fonctionnelle et les recherches en physique-mathématique où les conceptions de la physique restent prégnantes.

Afin de saisir, dans une première vision générale, le développement de l'analyse fonctionnelle on pourra consulter par exemple les deux monographies suivantes : celle de Antonie Frans Monna intitulée *Functional analysis in historical perspective*¹ publiée en 1973 et celle de Jean Dieudonné publiée en 1981 sous le titre *History of functional analysis*². Les articles « L'analyse fonctionnelle en Italie »³ de Gaetano Fichera publié en 1962, « The origins of functional analysis »⁴ de Reinhard Siegmund-Shultze publié en 2003 et « On the origins of functional analysis and the Lvov School »⁵ de Roman Duda publié en

¹[Mon73]

²[Die81]

³[Fic63]

⁴[Sie03]

⁵[Dud04]

2004 permettent d'aborder l'histoire de l'analyse fonctionnelle suivant des éclairages et des perspectives croisés. On y trouvera des éléments essentiels pour comprendre la constitution de l'analyse fonctionnelle comme nouveau champ des mathématiques ainsi qu'une bibliographie conséquente.

Cependant les citations avec lesquelles nous avons ouvert notre thèse donnent à lire une approche dans laquelle les problèmes particuliers s'inscrivent dans un processus de « généralisation progressive des opérations effectuées ». La notion de *fonction*, qui joue le rôle de variable ou d'inconnue dans les problèmes d'équations fonctionnelles, est elle-même l'objet d'un mouvement d'abstraction et de généralisation qui est au cœur de la progression générale en l'analyse fonctionnelle.

En outre la notion de fonction joue aussi parfois un rôle d'archétype et le mathématicien la convoque, sous des formes multiples, pour établir une nouvelle conception générale. Elle est donc à la fois source et objet du processus de généralisation que nous cherchons à préciser.

LES OPÉRATIONS. Il nous a donc paru nécessaire de trouver une ligne d'analyse qui ne s'appuyait pas *a priori* sur des notions —fonction, intégrale, équation fonctionnelle, etc.— qui seront l'objet de divers processus d'abstraction et de généralisation au centre de l'élaboration de l'analyse fonctionnelle.

Il nous a semblé que le mot « opération » employé par Levy pouvait être repris pour en faire un outil qui permettrait d'étudier diverses modalités de construction d'une vision générale. Le terme « opération », bien qu'il semble nous dégager d'une certaine matière trop sensible, sera pourtant parfois employé par les acteurs de notre étude et sera lui-même soumis à diverses interprétations ou processus de généralisation.

Nous le convoquons malgré ses différentes acceptions et malgré les évolutions de sens auxquelles il sera soumis car il permet de désigner un concept assez large dont les mathématiciens se saisissent et se ressaisissent à divers moments, comme s'il permettait à chaque fois de faire le pas de côté nécessaire au processus d'abstraction. De plus les mathématiciens de notre étude redéfiniront de façon explicite ce terme pour lui donner un sens particulier dans certains cas ou bien ils s'en serviront pour désigner un élément mathématique précis, ce qui en définira l'usage. C'est l'un des aspects qui distingue la notion d'*opération* des autres idées que nous avons évoquées qui ne seront pas toujours redéfinies tout en entrant dans la dynamique des approches générales que nous souhaitons étudier.

La notion d'opération permet en outre de prendre une certaine distance critique nécessaire pour étudier en détail le rôle et la nature changeante du concept de *fonction* dont nous avons souligné qu'il était central dans la conception de l'analyse fonctionnelle. Nous utiliserons donc le terme « opération » pour désigner successivement plusieurs emplois et faire apparaître des évolutions conceptuelles que nous préciserons tout au long de notre étude.

Nous pourrons ainsi par exemple parler d'opérations algébriques, d'opérations de dérivation ou d'intégration ou d'opérations linéaires, ou encore de la notion moderne d'opérat-*eur* (application linéaire sur un espace vectoriel), ces

expressions ne prenant tout leur sens que dans les textes où nous les lisons, et au terme d'une analyse précise du rôle que leur attribuent les acteurs.

Dans tous les cas une *opération* permet de mettre en relation plusieurs éléments. Elle implique donc une manière d'attribuer des rôles à ces éléments, d'une façon non uniforme en désignant des *éléments-agissants* ou des *éléments-objets* par exemple, ou encore en établissant une relation entre des éléments indifférenciés (par exemple dans une relation binaire telle que nous la concevons de nos jours pour définir un groupe). Se concentrer sur la notion d'opération permet donc en outre de s'intéresser à la nature de la relation que les acteurs installent entre les éléments, laquelle est indissociable du rôle que les mathématiciens attribuent aux éléments mis en relation.

Le terme *opération* dont nous nous saisissons désigne donc la mise en relation de certains éléments ainsi que les modalités de la relation établie. Nous ne substituons pas le mot *opération* à celui de *relation* tel qu'il est employé dans son sens le plus général en mathématiques. La *relation* mathématique permet de savoir ou de vérifier quels sont les éléments mis en relation par une méthode qui fait appel à la définition de la relation. L'*opération* comme nous la concevons désigne en un sens plus large ce qui détermine la relation, donc sa définition mais aussi comment le mathématicien la met en place et comment il l'utilise. L'*opération* indique donc en particulier la façon d'*opérer* avec la relation et sa conception repose aussi sur le pacte nécessaire établi entre le mathématicien et son lecteur —historien-mathématicien ou mathématicien.

CORPUS. Notre corpus a été sélectionné pour illustrer quatre temps essentiels dans le développement d'une vision générale en analyse fonctionnelle. Il nous a semblé que nous pouvions ainsi constituer une scène propice à étudier un certain nombre de mécanismes de sélection et d'organisation que mettent en œuvre les mathématiciens dans leurs approches générales. La notion d'*opération* nous permettra en particulier d'éclairer différents processus d'abstraction et de production d'énoncés généraux. Nous nous efforcerons tout au long de notre parcours de dégager les mécanismes dynamiques mis en œuvre par les mathématiciens qui cherchent à choisir la distance interprétative la mieux adaptée dans l'approche d'un problème. Nous analyserons les différents types de généralité que l'on peut lire dans les textes et qui sont l'œuvre de divers mouvements d'abstraction, d'organisation ou de sélection qui participent de la dynamique que nous cherchons à établir.

Nous commencerons notre étude avec certains travaux de Duncan Farquharson Gregory établis dans la première moitié du 19^e siècle dans le domaine de ce que les acteurs de son époque nommeront *l'algèbre symbolique*. Celle-ci fût essentiellement développée au Royaume-Uni au début du 20^e siècle. Il y a plusieurs raisons à ce choix qui ne constitue en aucun cas pour nous un *commencement* de l'analyse fonctionnelle. Il forme un bord, un point d'amorce de la scène que nous cherchons à mettre en place, un point indifférencié mais néanmoins unique d'une ligne que nous allons définir.

Ce premier point nous permet d'illustrer de façon précise le rôle des opérations algébriques (+, −, ×, ÷, lesquelles opèrent sur des entités plus ou moins abstraites) dans une certaine conception de l'analyse. Nous montrerons

comment une relecture abstraite de ces opérations a permis en particulier de faire évoluer le concept de fonction et le rôle qu'elles peuvent jouer dans diverses situations.

De plus Pincherle, à la toute fin du 19^e siècle, soulignera l'importance des conceptions de *l'algèbre symbolique* dans un texte qui cherche à cerner ce qu'il désigne alors par « Calcul fonctionnel » et dans lequel il tente d'établir les fondations d'un champ nouveau des mathématiques. Notre étude s'efforcera donc de déterminer les éléments qui nous semblent plus particulièrement pertinents dans une histoire de l'analyse fonctionnelle et de comprendre pourquoi Pincherle les convoque plus d'un demi-siècle plus tard dans sa propre vision.

Les textes que nous avons choisis de Gregory nous permettront d'illustrer précisément sa conception de *l'algèbre symbolique* et le rôle qu'y joue la notion de fonction. Mais ils nous permettront aussi d'étudier comment Gregory, à travers *l'algèbre symbolique*, aborde à sa façon un certain nombre de questions et développe une approche abstraite dont nous éclaireront la nature et les enjeux.

Dans un second temps nous étudierons une série de textes de Vito Volterra en mettant en évidence un certain recours à la notion de *fonction* comme fondement d'une conception abstraite et générale. Volterra a été maintes fois reconnu par les historiens des mathématiques comme l'un des pères de l'analyse fonctionnelle. Ses travaux sur les « fonctions de lignes » publiés en 1887 ont souvent été pris comme point de départ d'une conception moderne de l'analyse fonctionnelle. Ce sont en effet des travaux que l'on reconnaît bien, avec un certain recul rétrospectif, comme appartenant à l'analyse fonctionnelle et dont les outils nous semblent familiers.

Nous avons choisi de proposer ici une nouvelle approche de l'œuvre de Volterra en établissant une dynamique qui rende compte à la fois de l'appropriation de matériaux antérieurs et aussi du long travail d'abstraction, d'organisation et de généralisation que Volterra effectue avant les articles phares qui exposent les fameuses « fonctions de lignes ». En particulier nous chercherons à comprendre comment il adapte une théorie des fonctions ordinaires pour concevoir une notion de fonction généralisée. Elle lui permettra d'étudier divers problèmes où interviennent des équations fonctionnelles *a priori* très diversifiées, sous une forme unifiée. Les « fonctions de lignes » ont donc une histoire et nous montrerons qu'elles sont le fruit d'une progression vers le général que l'on peut établir à partir de la lecture des textes du mathématicien italien.

Dans une troisième partie nous montrerons comment Maurice Fréchet s'appuiera sur les travaux de Volterra sur les « fonctions de lignes » pour construire sa propre vision générale. Les travaux de Fréchet ont été très largement commentés et les études d'un point de vue historique ont en particulier mis en avant sa conception des « espaces abstraits » et le rôle important qu'y joue la notion topologique de compacité. Notre étude et la ligne de lecture que nous développons dans cette thèse permettent de revenir sur ces thèmes qui marquent notre conception de l'analyse fonctionnelle.

Nous établirons précisément deux voies ouvertes et suivies en parallèle par Fréchet dans la recherche d'un cadre général pour certains problèmes qu'il lit

en particulier chez Hadamard ou Volterra. Ses recherches l'amèneront petit à petit à dégager des notions de structure, d'espace et de topologie qui répondent chacune à une exigence précise dans sa quête d'une vision générale.

Notre étude nous amènera à reconsidérer les travaux sur les fonctionnelles et sur l'introduction d'éléments de topologie dans une organisation plus large d'une « Analyse générale », théorie dont la conception sera rattrapée par le foisonnement des travaux dans le nouveau champ de l'analyse fonctionnelle dès les premières décennies du 20^e siècle.

Enfin nous finirons notre étude avec une sélection de textes de Stefan Banach dont l'influence en analyse fonctionnelle n'est plus à démontrer tant ses idées et les problèmes qu'il a posés ont été la source d'un formidable élan qui a contribué à façonner l'analyse fonctionnelle telle que nous la connaissons de nos jours.

Nous avons choisi de nous situer en amont de son célèbre livre *Théorie des opérations linéaires*⁶ publié en français en 1932 dans lequel Banach expose, sous une forme synthétique et ré-organisée, ce qui apparaît pour la première fois comme une véritable *théorie*. Notre but essentiel n'a donc pas été d'examiner comment ce traité a été élaboré à partir des travaux antérieurs de Banach, mais il a plutôt consisté à établir la dynamique d'une approche générale à partir d'une sélection de textes antérieurs au livre sur les opérations linéaires.

Dès sa thèse de doctorat en 1920 et dans une collection de travaux qui suivirent pendant près d'une décennie, Banach s'appuiera sur une notion choisie d'« opération » pour construire une vision générale apte à concevoir d'une façon efficace de nombreux problèmes qui jusque-là recevaient des traitements différenciés.

Nous montrerons en particulier comment la notion d'espace vectoriel normé complet —les fameux *espaces de Banach*— est intimement liée à la façon dont Banach conçoit les opérations, c'est-à-dire en nos termes les *opérateurs*.

De plus il nous semble que les travaux que nous avons sélectionnés de Banach avant le livre *Théorie des opérations linéaires* permettent de voir à l'œuvre très précisément les mécanismes qui lui ont permis de forger une vision nouvelle et générale pour l'analyse fonctionnelle. Il nous a paru important d'analyser le type de généralité qu'un regard plus réflexif de Banach sur les opérations et sur les espaces (vectoriels, souvent normés, parfois complets) pouvait faire émerger dans la lecture de ses textes.

STRUCTURES. Le mot structure est omniprésent dans les publications mathématiques du 20^e siècle et il désigne un élément incontournable dans l'étude de l'élaboration de l'analyse fonctionnelle. La notion large qu'il recouvre fait apparaître des formes stables, garantes d'un fondement solide sur lequel on peut bâtir une théorie ou qui permettent d'appréhender ensemble et de façon cohérente plusieurs éléments sélectionnés. La structure permet au mathématicien de convoquer un contexte relativement clos, auto-suffisant, et constitue un paradigme qui s'est imposé pour son aptitude à dégager et à favoriser un point de vue général.

On fait référence à l'idée de structure pour désigner des assemblages conceptuels de base dans de multiples théories (par exemple la théorie

⁶[Ban32]

des groupes, la topologie, la théorie des opérateurs, etc.), mais aussi dans de nombreuses publications où le mot structure recouvre des situations extrêmement diverses. Ainsi une caractérisation de la notion de structure qui puisse rendre compte de ses aspects variés et la mise en évidence de ses modes opératoires en mathématiques ne se laisse pas saisir de façon immédiate.

La force de la structure est multiple, car elle permet au mathématicien à la fois de créer un nouveau paysage cohérent en rassemblant plusieurs éléments, et d'instaurer un certain agencement des concepts convoqués. L'organisation induite possède elle-même la double particularité de favoriser une vision plus générale en embrassant plusieurs objets à la lumière d'une nouvelle cohérence et de permettre de choisir la distance interprétative la mieux adaptée dans l'approche d'un problème.

La naissance de l'analyse fonctionnelle, comme nous l'illustrons ici, met en jeu des phénomènes très divers que nous allons analyser en particulier à travers les opérations. Nous montrerons que les structures convoquées par les mathématiciens sont souvent intimement liées à la conception des opérations, même en ce qui concerne les aspects qui ne découlent pas directement des opérations algébriques, comme ceux liés à la norme, à la distance ainsi qu'à d'autres aspects topologiques ou géométriques.

Nous étendrons le concept de structure pour rendre compte de façon précise de la dépendance entre les opérations et les espaces d'une part, et pour mieux apprécier d'autre part le rôle que les mathématiciens attribuent à certains *éléments structurels* qui jouent en deçà de la notion que l'on peut rétrospectivement concevoir comme une structure. Ces éléments nous permettront de mener une lecture qui éclaire la dynamique de diverses approches générales en analyse fonctionnelle.

MÉTHODE. Comme nous l'avons dit, notre ligne d'analyse sera construite à partir du concept d'*opération* que nous utiliserons tout au long de notre travail. Mais il nous faut ici préciser un peu notre méthode.

Nous cherchons à faire apparaître certains éléments d'une dynamique qui puisse rendre compte des approches générales dans les textes mathématiques que nous étudions. Cette dynamique est complexe et s'appréhende à différents niveaux suivant des modalités que l'on peut distinguer.

Tout d'abord la lecture d'un texte particulier permet d'analyser des éléments qui relèvent d'une dynamique interne au texte. Ceux-ci s'apprécient en suivant l'organisation matérielle et logique du discours. Il ne s'agit pas ici de dire qu'il y a une seule logique ou que les textes mathématiques sont uniformément conformes à une logique, mais d'apprécier, au plus près des énoncés et des démonstrations, comment le mathématicien organise ses idées et les enchaîne afin de nous faire lire ou *entendre* quelque chose. C'est dans cette lecture, souvent technique, que l'on peut déceler la nature du pacte qui doit s'établir entre le mathématicien qui expose ses idées et son lecteur, à qui on demande toujours une certaine acceptation.

Le type de généralité que l'on lit dans un texte mathématique dépend intimement d'une organisation *moléculaire* du texte. Cette approche nécessite de ne pas faire l'impasse sur l'étude détaillée de certaines démonstrations,

car l'énoncé d'un théorème ou même les grandes étapes d'une démonstration ne sont pas toujours suffisants pour établir ce sur quoi repose la pensée mathématique ou pour préciser comment le mathématicien fait travailler ses concepts. Il nous faudra donc à plusieurs reprises emboîter le pas de quelques illustres savants sans perdre de vue notre propre objectif d'établir des éléments d'une dynamique de la progression des idées mathématiques dans une perspective historique. Les mathématiques que nous étudions ici ne sont pas simples et nous espérons avoir suffisamment allégé la tâche du lecteur sans occulter le lieu où se joue aussi à notre sens une part de l'essentiel.

D'autre part, une approche que nous qualifierons de *systémique* permet de dévoiler, dans une vision plus large, des modes de production d'une pensée générale qui mettent en œuvre des éléments qui échappent à la lecture d'un texte suivant la procédure indiquée au-dessus.

Entrent dans cette analyse les conditions générales de production des textes mathématiques, conditions politiques ou institutionnelles par exemple, mais aussi des contextes généraux dans lesquels se développent les idées et les pratiques de groupes plus ou moins étendus de mathématiciens que l'historien identifie.

Notre approche est différente mais nous tenterons à chaque fois de replacer nos textes dans un contexte suffisamment large en nous appuyant sur des travaux qui se placent à une autre échelle qui est tout aussi essentielle pour comprendre différents types de pratiques. En particulier, à quelques exceptions près, nous ne convoquerons des textes de mathématiciens hors de notre corpus que pour souligner un contraste qui permet de mieux appréhender les choix, les relectures éventuelles, les modes de production d'une vision générale de tel ou tel auteur central de notre étude.

Enfin, une autre façon de concevoir une dynamique de progression vers une théorie générale consiste à étudier, ici à travers l'idée d'*opération*, la constitution de tout un réseau de travaux dont l'analyse fonctionnelle à chaque instant est le fruit commun. Cette approche —que l'on pourrait avec Volterra penser en termes d'*hérité* et de *fonction de ligne*— justifie le choix et l'étendue de notre corpus. Nous espérons que le lecteur y verra avec assez de consistance que les études locales au niveau d'un énoncé ou d'un texte, et les études menées suivant une ligne —un local d'un autre type— telle celle que nous l'avons choisie grâce au concept d'« opération », se font écho et tissent une scène où de multiples lignes permettent de construire une histoire dynamique des mathématiques.

Nous ouvrons notre étude de l'analyse fonctionnelle dans cette perspective.

Recours à l’algèbre symbolique en analyse : le cas de Duncan Farquharson Gregory

*It has always appeared to me that we sacrifice many of the advantages and more of the pleasures of studying any science by omitting all reference to the history of its progress.*¹

— D. F. Gregory

2.1 INTRODUCTION

Dans son « Mémoire sur le calcul fonctionnel Distributif », publié en 1897 dans les *Mathematische Annalen*, Salvatore Pincherle dresse en introduction un panorama historique de ce qui pour lui relève du « calcul fonctionnel ». Parmi les diverses sources qu’il convoque dans son exposé on trouve l’influence d’auteurs anglais du 19^e siècle qui ont développé un « calcul symbolique ou calcul des opérations »² :

La nomenclature et les méthodes [du calcul symbolique ou calcul des opérations] s’introduisent spontanément dans les études sur les opérations fonctionnelles, surtout dans celles qui ont un caractère formel ou algorithmique. Le calcul des opérations, quoique déjà ancien, est considéré avec quelque défiance par la plupart des mathématiciens du continent, tandis que de nombreux auteurs anglais s’en sont constamment occupés : le *philosophical magazine*, les *philosophical transactions* et d’autres recueils contiennent, depuis plus de quarante ans, de nombreuses notes à ce sujet. Cette défiance dépend sans doute du caractère formel que présente ce calcul, et du

¹Préface de *Examples of the processes of the differential and integral calculus*. (Trinity College, Cambridge. October, 1841) cf. [Gre41]

²[Pin97, p.329]

peu de peine qu'on s'est donné pour limiter ses résultats de façon à les rendre rigoureux ; néanmoins plusieurs de ses applications qui, interprétées convenablement, sont susceptibles de toute rigueur, offrent beaucoup d'intérêt : par exemple la déduction des intégrales d'une équation différentielle de ceux d'une autre, la décomposition en facteurs du premier membre d'une équation linéaire différentielle ou aux différences finies, à coefficients constants ou variables, etc.

Pincherle met ici en avant deux points importants : tout d'abord il estime que les travaux anglais qui traitent du calcul symbolique ou du calcul des opérations sont particulièrement adaptés pour étudier les « opérations fonctionnelles qui ont un caractère formel ou algorithmique », et d'autre part il souligne que la « rigueur » des preuves obtenues dans ces domaines repose sur une « interprétation convenable » des résultats. Ces deux observations nous semblent essentielles pour comprendre la vision que tente de construire Pincherle à la fin du 19^e siècle en dressant un panorama précis des idées qui relèvent pour lui du « calcul fonctionnel ».

Dans ce texte Pincherle fait référence à plusieurs auteurs anglais sur lesquels nous reviendrons dans la section suivante. Ils ont été identifiés par les historiens des mathématiques comme appartenant à un groupe plus vaste de mathématiciens qui ont tous œuvré au développement de l'algèbre symbolique et des domaines qui s'y rattachent. Parmi ces auteurs Pincherle mentionne Duncan Farquharson Gregory sur lequel nous allons nous appuyer pour comprendre les éléments qui ont joué un rôle décisif dans la réception de l'algèbre symbolique chez des mathématiciens qui comme Pincherle cherchent à définir le calcul fonctionnel au tournant du siècle.

Gregory nous semble être un auteur de première importance dans cette étude et nous allons montrer que son œuvre contient de façon centrale des travaux directement liés aux deux thèmes que nous avons soulignés dans le texte de Pincherle.

Premièrement Gregory a joué un rôle principal dans une mutation de la nature même des opérations, celles-ci passant dans les travaux anglais de son époque d'une utilisation propre à l'arithmétique à une utilisation plus générale et abstraite. En particulier Gregory réinterprète des situations issues de l'analyse, provenant essentiellement des équations différentielles ou intégrales, sous forme d'opérations symboliques. Il assujettit le traitement de ces opérations à des lois indépendantes de la nature des symboles tels qu'ils sont définis dans les équations de l'analyse, là où les symboles faisaient initialement sens. Dans sa démarche Gregory pratique une forme d'abstraction et procède d'une manière formelle, deux traits qui ont joué un rôle indéniable dans la réception de ses travaux chez Pincherle.

D'autre part, et ceci rejoint le deuxième point mis en avant par Pincherle, de nombreux travaux de Gregory font une place centrale à la problématique de l'interprétation et à la justification de la rigueur des preuves issues du travail symbolique.

Dans cette partie nous rappellerons donc les principales conceptions des opérations en algèbre symbolique dans la première moitié du 19^e siècle à travers l'étude de textes de Gregory.

Cette étude sera aussi l'occasion pour nous de montrer comment Gregory s'approprie la notion d'opération et elle nous permettra d'éclairer les enjeux qui découlent pour lui du recours à l'algèbre symbolique.

Nous pourrons alors plus tard contraster le rôle que joue l'algèbre symbolique dans les travaux de Gregory et plus largement chez les mathématiciens de l'*École Algébrique Anglaise* avec le rôle que jouent diverses formes d'opérations dans la conception de l'analyse fonctionnelle jusqu'au début du 20^e siècle.

2.2 L'ÉCOLE ALGÈBRIQUE ANGLAISE

Dans le texte de Pincherle que nous avons cité pour ouvrir notre étude, celui-ci fait référence à plusieurs mathématiciens de langue anglaise qui ont été identifiés par les historiens des mathématiques sous le nom d'*École Algébrique Anglaise*³. Une analyse conséquente des enjeux mathématiques, philosophiques et sociétaux liés aux activités de ce groupe et des mathématiciens moins connus qui y sont connectés a été développée dans de nombreux travaux d'histoire des mathématiques : on peut en particulier consulter l'étude d'Elaine Koppelman « The Calculus of Operations and the Rise of Abstract Algebra »⁴, la thèse de Christiane Ruffieux « La Naissance du Concept de Structure Algébrique en Grande-Bretagne dans la première moitié du 19^e Siècle. Influence des Philosophes de l'École Écossaise du Sens Commun. »⁵, l'article de Karen Hunger Parshall, « Victorian algebra: the freedom to create new mathematical entities »⁶, et les articles de Marie-José Durand-Richard⁷ dont on trouve une sélection dans notre bibliographie.

Les études disponibles à ce jour sur l'*École Algébrique Anglaise* ont analysé en amont le travail de relecture de travaux en arithmétique et en algèbre, ainsi que le rôle joué par des problèmes que nous classerions désormais en analyse : équations différentielles et intégrales ou calcul des différences finies par exemple. En aval la grande majorité des études montrent l'utilisation de certains de ces travaux dans le développement de la logique, de la théorie des groupes et plus largement dans l'émergence de l'idée de *structure* en algèbre moderne. En particulier, le travail de Koppelman, qui a largement contribué à faire reconnaître l'importance de l'œuvre de Gregory dans les mathématiques du 19^e siècle, propose un panorama étendu qui montre le rôle important de plusieurs textes que nous étudions à nouveau en détail ici. Enfin le livre de Jeremy J. Gray et Karen Hunger Parshall intitulé « Episodes in the History of Modern Algebra (1800-1950) »⁸ propose une vision synthétique de l'histoire de

³On fait référence ici au groupe de mathématiciens identifiés par Nový [Nov68] qui inclut Charles Babbage (1791-1871), Georges Peacock (1791-1858), Augustus de Morgan (1806-1871), Duncan Farquharson Gregory (1813-1844), Georges Boole (1815-1844), William Rowan Hamilton (1805-1865), Arthur Cayley (1824-1895) et James Joseph Sylvester (1814-1897). On peut y rattacher d'autres auteurs moins connus dont on trouve une liste assez étendue dans l'article de Marie-José Durand-Richard [Dur97].

⁴[Kop71]

⁵[Ruf05]

⁶[Par11]

⁷[Dur90; Dur96; Dur97; Dur07; Dur11]

⁸[GP11]

l'« algèbre moderne » qui resitue en particulier le travail de Gregory dans une évolution de l'algèbre entre le début du 19^e siècle et la première moitié du 20^e siècle. Le second chapitre « 'Very Full of Symbols': Duncan F. Gregory, the Calculus of Operations, and the *Cambridge Mathematical Journal* »⁹ propose une étude de Sloan Despeaux qui précise le rôle de Gregory et du « calcul des opérations » dans le renouveau des mathématiques anglaises de la première moitié du 19^e siècle.

En ce qui concerne le développement de l'analyse, quelques études ont mis en avant le rôle qu'a joué Gregory en s'appuyant essentiellement sur ses deux textes les plus connus : le premier, « On the Real Nature of Symbolical Algebra »^{10,11}, publié dans les *Transactions of the Royal Society of Edinburgh*, et le second qui constitue le livre *Examples of the Processes of the Differential and Integral Calculus*¹² publié en 1841. L'article de Patricia R. Allaire et Robert E. Bradley intitulé « Symbolical Algebra as a Foundation for Calculus: D. F. Gregory's Contribution »¹³ examine l'influence des travaux de Gregory dans l'histoire de l'analyse (*calculus*) en faisant essentiellement référence à ces deux textes¹⁴.

Or il nous a semblé que d'autres textes, essentiellement publiés dans la jeune revue *Cambridge Mathematical Journal*¹⁵, permettaient de suivre plus précisément l'évolution de la pensée du mathématicien et les nécessaires tentatives de justification d'une théorie non encore totalement acceptée à son époque. Gregory y aborde à travers de nombreux articles des thèmes qui font pour nous directement écho aux idées mises en avant par Pincherle dans son introduction de 1897 et qui permettent de préciser le rôle joué par le mathématicien écossais dans le développement de l'analyse moderne et de l'analyse fonctionnelle.

Enfin, notre étude s'inscrit dans un projet de recherche plus large qui cherche à éclairer la progression et l'organisation des idées propres à la conception de l'analyse fonctionnelle telle qu'elle sera établie au début du 20^e siècle par le travail de nombreux auteurs. Les travaux qui ont permis de forger petit à petit l'analyse fonctionnelle montrent une grande diversité dans le choix du formalisme le mieux adapté pour concevoir et manipuler les éléments convoqués par les mathématiciens. C'est là un aspect majeur des textes que nous étudions, un changement de formalisme révélant toujours un repensé et un nouveau mouvement d'abstractions.

Gregory pour sa part a recours à un formalisme unique pour traiter diverses situations issues de l'arithmétique, de la géométrie ou de l'analyse —études des

⁹[Des11]

¹⁰[Gre38b, MW : p.1–13]

¹¹NDA. Nous indiquons la pagination relative aux œuvres mathématiques complètes (MW) dans l'édition de 1865. (cf. *Mathematical Writings* [Gre65])

¹²[Gre41]

¹³[AB02]

¹⁴D'autres travaux d'une visée plus large mentionnent les travaux de Gregory et leurs liens avec l'analyse. On peut lire par exemple la thèse de Maria Panteki [Pan91] et l'article de Ivor Grattan-Guinness [Gra10]. Notre démarche s'inscrit en complément de toutes ces études mais avec une focalisation différente, centrée sur la démarche de Gregory et l'analyse des diverses formes de généralité que son œuvre met en place.

¹⁵Le *Cambridge Mathematical Journal* a été fondé en 1837 et Gregory en fut l'un de ses premiers éditeurs. Voir [Cri04] pour comprendre le contexte historique et le rôle important qu'a pu jouer cette revue.

fonctions et de leurs propriétés ou problèmes que nous classerions désormais en analyse fonctionnelle. L'idée qui sous-tend cette vision générale unique est l'interprétation des diverses situations en termes d'*opérations*, représentées par des *symboles* indépendants des champs particuliers que Gregory envisage.

Les textes de Gregory permettent de voir comment il se saisit de la conception initiale d'opération en arithmétique pour la faire évoluer en un outil très général et abstrait. En particulier Gregory interprète en termes d'« opérations » (au sens où il les conçoit) ce que nous appellerions de nos jours des opérateurs, éléments essentiels de l'analyse fonctionnelle du début du 20^e siècle et fer de lance de l'œuvre de Banach.

Ces travaux de Gregory ont une place particulière dans la littérature mathématique de l'époque. Ils se situent chronologiquement entre des ouvrages phares qui assurent la transmission d'un certain type de savoir au sein des universités anglophones et en particulier à Cambridge : l'ouvrage de Peacock *A Collection of Examples of the Applications of the Differential and Integral Calculus*¹⁶ publié en 1820, suivi de celui de Gregory lui-même *Examples of the Processes of the Differential and Integral Calculus*¹⁷ de 1841. Ce dernier livre constitue officiellement, comme l'énonce Gregory en introduction, une suite à l'ouvrage de Peacock destinée aux étudiants de l'université, et où seraient ajoutées quelques démonstrations jugées importantes. En fait, le livre est publié vers la fin de la période où sont écrits les textes que nous allons étudier sur le calcul symbolique. Pourtant, Gregory n'écrira dans son manuel qu'un seul chapitre relatif à l'utilisation des outils généraux qu'il développe dans ses recherches. Il s'agit du chapitre XV du livre, intitulé « General Theorems in the Differential Calculus », où il expose son approche symbolique sous forme de théorèmes qui ressemblent plus à des réécritures de formules classiques qu'à de véritables théorèmes obtenus par l'algèbre symbolique. En particulier on n'y trouve pas trace des démarches conceptuelles qu'il développe tout au long de ses publications de recherche. Il est frappant de constater que le livre de Gregory n'utilise le formalisme général de l'algèbre symbolique que pendant quelques lignes pour rappeler les lois de commutativité, distributivité et des puissances. La suite de l'exposé est dédiée au calcul sur les différentielles et reste liée aux techniques de séparation des symboles d'opération sous une forme déjà bien connues à son époque. En particulier tout le travail qui justifie pour Gregory l'existence de preuves dans le cadre de l'algèbre symbolique est absent du livre qui suit une voie proche de l'exposé initial de Peacock.

D'autre part, rappelons pour mieux les situer que les articles de recherche de Gregory ont été publiés entre les livres que nous avons mentionnés juste au-dessus, et les ouvrages restés célèbres de George Boole *The Mathematical Analysis of Logic, Being an Essay Towards a Calculus of Deductive Reasoning*¹⁸ et *An Investigation of the Laws of Thought on which are Founded the Mathematical Theories of Logic and Probabilities*¹⁹. On ne peut pas finir de décrire ce positionnement des textes de Gregory dans la littérature mathématique de son époque sans

¹⁶[Pea20]

¹⁷[Gre41]

¹⁸[Boo47]

¹⁹[Boo54]

mentionner le livre de Peacock intitulé *Algebra*²⁰ publié en deux volumes en 1842 et 1845. Peacock tente d'y exposer une vision générale de l'algèbre : d'abord l'algèbre arithmétique (*arithmetical algebra*) dans le premier volume et l'algèbre symbolique (*symbolical algebra*) dans le second. Ce dernier volume en particulier, qui a été écrit après les travaux de Gregory, ne propose aucune discussion sur la légitimité des preuves : les théorèmes d'algèbre symbolique y sont énoncés et établis à travers un principe admis de « substitution de symboles généraux à la place des nombres »²¹.

Il s'agit donc pour nous ici de relire le travail de Gregory en ramenant au centre un certain nombre de textes périphériques aux deux textes phares le plus souvent étudiés et de mettre en évidence leur importance dans la pensée développée par le mathématicien.

2.3 SÉPARATION DES SYMBOLES D'OPÉRATION

C'est dans les *Transactions of the Royal Society of Edinburgh* en 1838 que Gregory publie son article au titre percutant : « On the Real Nature of Symbolical Algebra ». Le but est annoncé dès les premières lignes²² :

The following attempt to investigate the real nature of Symbolical Algebra, as distinguished from the various branches of analysis which come under its dominion, took its rise from certain general considerations, to which I was led in following out the principle of the separation of the symbols of operation from those of quantity.

Dans les textes de Gregory, l'expression « algèbre symbolique » exprime une vision générale qui tente d'appréhender de façon unifiée divers domaines qui seront tour à tour source d'inspiration et lieu de réinterprétation dans une nouvelle organisation que nous allons étudier. On retrouve parmi les champs envisagés l'« algèbre arithmétique », sorte de modèle initial pour établir l'algèbre symbolique, qui dicte les règles de calculs avec les nombres. Mais aussi comme le souligne Gregory dans la citation au-dessus, des domaines usuellement classés à son époque dans diverses « branches de l'analyse ». Il s'agit essentiellement ici du calcul différentiel et intégral ou du calcul des différences finies qui sont alors au centre des travaux de bon nombre de mathématiciens du continent ou de Grande-Bretagne. Dans ces domaines, Gregory est attentif à une pratique à laquelle il fait référence sous le nom de « séparation des symboles », dont il attribue d'abord la paternité à Barnabé Brisson, puis à John Herschel quelques années plus tard en citant un article publié par ce dernier dans les *Philosophical Transactions* de 1816. Koppelman²³ resitue cette technique dans un contexte historique large et mentionne en particulier les travaux de Louis François Arbogast qui, dans son livre *Du Calcul des Dérivations*²⁴ publié en 1800, avait introduit une « méthode de séparation des échelles d'opérations ». Koppelman montre que cette pratique fût jugée

²⁰[Pea42 ; Pea45]

²¹[Pea45, p.1]

²²[Gre38b, MW : p.1]

²³[Kop71, p.161 et sq.]

²⁴[Arb00]

comme novatrice et qu'elle inspira plusieurs mathématiciens qui la reprirent à leur compte et la développèrent en l'appliquant essentiellement aux équations différentielles.

Avant d'étudier la forme que prend l'algèbre symbolique telle que Gregory l'expose dans son article à la *Royal Society of Edinburgh* en 1838, nous analysons dans cette partie un certain nombre de textes préliminaires dans lesquels il forge sa propre vision et en commente divers aspects.

Lorsque, dès 1837, Gregory publie pour la première fois dans le *Cambridge Mathematical Journal* il intitule son article « On the Solution of Linear Differential Equations with Constant Coefficients » et il y présente la « méthode » et le « principe de séparation des signes d'opération » de la façon suivante²⁵ :

If we take the general linear equation with constant coefficients

$$\frac{d^n y}{dx^n} + A \frac{d^{n-1} y}{dx^{n-1}} + B \frac{d^{n-2} y}{dx^{n-2}} + \dots + R \frac{dy}{dx} + Sy = X,$$

when X is any function of x , and separate the signs of operation from those of quantity, it becomes

$$\left(\frac{d^n}{dx^n} + A \frac{d^{n-1}}{dx^{n-1}} + B \frac{d^{n-2}}{dx^{n-2}} + \dots + R \frac{d}{dx} + S \right) y = X.$$

L'étude de Koppelman montre que le traitement de l'équation différentielle est effectué ici par Gregory d'une « façon très semblable à celle de ses prédécesseurs »²⁶. On peut poursuivre l'analyse et préciser avec le texte de Gregory comment cette « séparation des signes d'opération » va donner lieu à une interprétation abstraite et à l'introduction d'une notation qui reflète cette abstraction²⁷ :

The quantity within the brackets, involving only constants and the signs of operation, may be considered as one operation performed on y , and it may be represented by

$$f \left(\frac{d}{dx} \right) y = X.$$

En observant les deux citations côte-à-côte, on peut penser que Gregory utilise le symbole f par analogie avec l'écriture des fonctions polynômiales, à des fins pratiques et uniquement pour produire une écriture plus condensée. Si l'analogie entre la somme des dérivées d'ordres multiples et l'écriture polynômiale semble bien expliquer la notation choisie par Gregory (une fonction f dont la variable serait désormais $\frac{d}{dx}$), c'est dans les utilisations de cette écriture que l'on saisit ce qu'il conçoit réellement.

Dans un premier temps, Gregory distingue dans l'expression telle qu'il l'écrit dans la dernière citation trois éléments : $f \left(\frac{d}{dx} \right)$, y et X . Il interprète alors l'égalité

²⁵[Gre37, MW : p.15]

²⁶[Kop71, p.189-190]

²⁷[Gre37, MW : p.15]

comme une simple équation mettant en jeu ces trois termes. Il s'agit là d'un processus d'abstraction dans lequel des éléments ont été regroupés, organisés et où Gregory a attribué à chacune des nouvelles entités un rôle précis : une opération unique $f\left(\frac{d}{dx}\right)$ agissant sur y , la fonction y jouant à la fois le rôle d'inconnue et d'objet de l'opération, alors que X est une fonction donnée de la variable x . Cette vision permet à Gregory de proposer une façon de résoudre l'équation lorsque l'action de l'opération $f\left(\frac{d}{dx}\right)$ peut être inversée²⁸ :

If we represent the inverse operation by the usual symbol of $\left\{f\left(\frac{d}{dx}\right)\right\}^{-1}$, and perform the operation on both sides, we get

$$\left\{f\left(\frac{d}{dx}\right)\right\}^{-1} \cdot f\left(\frac{d}{dx}\right) y = \left\{f\left(\frac{d}{dx}\right)\right\}^{-1} X,$$

or,

$$y = \left\{f\left(\frac{d}{dx}\right)\right\}^{-1} X.$$

Gregory remarque que l'« opération inverse » n'est pas souvent facile à trouver : il propose donc d'illustrer son approche et d'en préciser l'étendue en considérant d'abord une équation du premier ordre à coefficients constants qu'il reformule en « séparant les signes d'opération »²⁹ :

Let us take the equation

$$y + \frac{dy}{dx} = ax^n,$$

or,

$$\left(1 + \frac{d}{dx}\right) y = ax^n.$$

Now the inverse operation of $\left(1 + \frac{d}{dx}\right)$ is $\left(1 + \frac{d}{dx}\right)^{-1}$. Therefore

$$y = \left(1 + \frac{d}{dx}\right)^{-1} ax^n.$$

Cette première analyse est similaire au cas général exposé plus haut. On remarque au passage l'utilisation du symbole 1 qui laisse entrevoir comment est conçue ici la notion de symbole d'opération et le lien non encore explicité qu'entretiennent ces symboles avec ceux de l'arithmétique.

Mais Gregory va aller plus loin et poursuivre la résolution de l'équation en deux temps qui illustrent les choix qu'il opère.

Tout d'abord, par analogie avec un principe classique en intégration, on peut considérer que les solutions sont définies à une constante près³⁰ :

²⁸[Gre37, MW : p.15]

²⁹[Gre37, MW : p.15]

³⁰[Gre37, MW : p.16]

As in integration there must be added an arbitrary constant which vanishes by differentiation, so here we must add a function which will vanish when the operation $(1 + \frac{d}{dx})$ is performed on it. This complementary function may be found from that condition.

Mais cette *analogie* n'est pas suffisante pour Gregory qui désire ne faire intervenir jusqu'au bout que les « symboles d'opération » : il tente de rejeter tout recours à des résultats qui n'auraient pas été interprétés en ces termes. Ceci va l'amener à prolonger son analyse dans une nouvelle voie.

Puisque la « fonction complémentaire » issue de l'intégration disparaît lorsqu'on applique $(1 + \frac{d}{dx})$, Gregory peut donner y de façon explicite, c'est-à-dire en l'exprimant seulement en termes de « symboles d'opération », sans recours à une fonction étrangère à l'équation³¹ :

We may put the value of y under the form

$$y = \left(1 + \frac{d}{dx}\right)^{-1} ax^n + \left(1 + \frac{d}{dx}\right)^{-1} 0.$$

Nous retrouverons par la suite le procédé qui consiste à faire apparaître des éléments invisibles dans les calculs afin d'en donner une interprétation nouvelle en termes d'opérations uniquement. Ici, la solution y dépend donc exclusivement du calcul de l'« opération inverse » $\left(1 + \frac{d}{dx}\right)^{-1}$ sur ax^n et sur 0.

Il reste donc à Gregory à utiliser $\left(1 + \frac{d}{dx}\right)^{-1} \cdot 0$ dans son calcul³² :

Now if we treat the symbols of operations as if they were symbols of quantity, we have

$$\left(1 + \frac{d}{dx}\right)^{-1} 0 = \frac{d^{-1}}{dx^{-1}} \left(1 + \frac{d^{-1}}{dx^{-1}}\right)^{-1} 0.$$

But $\frac{d^{-1}}{dx^{-1}}$ is the same as $\int dx$. Hence

$$\left(1 + \frac{d}{dx}\right)^{-1} 0 = \left(1 + \frac{d^{-1}}{dx^{-1}}\right)^{-1} C$$

(C being the arbitrary constant arising from the integration)

$$= \left(1 - \frac{d^{-1}}{dx^{-1}} + \frac{d^{-2}}{dx^{-2}} + \dots\right) C;$$

or, performing the operation indicated,

$$= C \left(1 - x + \frac{x^2}{1.2} - \frac{x^3}{1.2.3} + \dots\right) = C e^{-x}.$$

³¹[Gre37, MW : p.16]

³²[Gre37, MW : p.16]

Hence

$$y = \left(1 + \frac{d}{dx}\right)^{-1} ax^n + C\epsilon^{-x}.$$

Now expanding the first term

$$y = \left(1 - \frac{d}{dx} + \frac{d^2}{dx^2} - \frac{d^3}{dx^3} + \dots\right) ax^n + C\epsilon^{-x}.$$

Therefore

$$y = a\{x^n - nx^{n-1} + n(n-1)x^{n-2} - \dots\} + C\epsilon^{-x}.$$

L'analyse de ces lignes est essentielle pour comprendre la démarche de Gregory et apprécier le type de transfert d'idées et de techniques qu'il met en œuvre. Le procédé repose dans cette phase sur la possibilité de *retranscription* d'un résultat obtenu dans un autre domaine (ici l'intégration) à l'aide des symboles d'opération. Cette étape a nécessité de passer d'une conception associée à la formule « en intégration on doit ajouter une constante arbitraire qui disparaît par dérivation », où l'on insiste sur l'« arbitraire » de la constante et donc sur une certaine *multiplicité*³³, à une formulation qui n'utilise qu'une seule opération agissant sur 0 et où apparemment la multiplicité disparaît. En fait on a maintenant unicité d'écriture, ce qui ne signifie pas unicité de solution. En effet à l'époque de Gregory une fonction ou une opération peut donner plusieurs valeurs quand elle est appliquée à un élément unique. On retrouvera de nombreux cas qui illustrent ces idées plus loin dans notre étude et nous précisons ce que nous désignons par *multiplicité* dans plusieurs situations.

La suite du raisonnement exposé dans le passage que nous analysons s'articule à ce qui précède et montre une nouvelle étape dans le travail symbolique. En effet, ce sont maintenant des procédures de calcul de l'algèbre arithmétique³⁴ et des développements en séries que Gregory transpose pour poursuivre son calcul. Si l'on rétablit en nos termes les manipulations qu'il

effectue, on voit que $\left(1 + \frac{d}{dx}\right)^{-1} \cdot 0 = \frac{d^{-1}}{dx^{-1}} \left(1 + \frac{d^{-1}}{dx^{-1}}\right)^{-1} \cdot 0$ s'appuie sur l'égalité

$(1 + X)^{-1} = X^{-1} (1 + X^{-1})^{-1}$ et que $\left(1 + \frac{d}{dx}\right)^{-1}$ qu'il réécrit $\left(1 - \frac{d^{-1}}{dx^{-1}} + \frac{d^{-2}}{dx^{-2}} + \dots\right)$

est inspiré du développement $(1 + X)^{-1} = 1 - X + \frac{X^2}{1.2} - \frac{X^3}{1.2.3} + \dots$. Il se dessine ici une différence essentielle par rapport à ce qui précède : il y a une véritable *transposition* des résultats car les éléments entrant dans l'égalité en algèbre arithmétique, dénotés par X , ne sont pas de la même nature que ceux de l'égalité transférée dans la nouvelle écriture où l'on a des symboles qui désignent des opérations du calcul différentiel.

³³Le terme *multiplicité* n'apparaît pas dans les textes que nous citons, et il s'agit d'un terme d'observateur tout au long de notre étude.

³⁴Dans notre étude nous désignons par algèbre arithmétique le travail symbolique sur les nombres comme il est défini par exemple dans le traité d'algèbre de Peacock en 1830. Pour celui-ci « symbols are assumed to be the general and unlimited representatives of every species of quantity. » [Pea30, preface p.vi]

Le passage de l'écriture de l'algèbre arithmétique avec les X à l'égalité utilisant les différentielles $\frac{d}{dx}$ illustre clairement ce que Gregory entendait par « traiter les symboles d'opération comme s'ils étaient des symboles de quantité ». D'abord à un niveau que l'on pourrait qualifier de formel, on voit ici le procédé de substitution ($\frac{d}{dx}$ à la place de X) qui est déjà largement employé par les adeptes de la méthode de séparation des symboles d'opération à l'époque de Gregory. Mais il y a plus dans cette citation : nous avons bien ici non pas seulement une reformulation, mais une véritable progression et un enchaînement de lignes qui aux yeux de Gregory constitue une preuve (*proof*³⁵). L'acceptation en tant que preuve du passage de l'égalité de l'algèbre arithmétique à l'égalité *transposée* est essentielle dans la démarche de Gregory qui insistera sur ce point dès ce premier article : « we require no farther proof for the proposition »³⁶.

Nous allons montrer dans la suite comment cette *transposition* n'est pas pour Gregory une simple façon de rendre compte d'une analogie entre divers champs dans lesquels on effectue des calculs, ni une extension immédiate du champ d'action de l'algèbre arithmétique qui serait fondée sur une sorte de permanence formelle³⁷. Nous verrons que les principes qui justifient la validité de la *transposition* sont précisément mis en avant et discutés par Gregory dans de nombreuses situations.

Notons au passage que dans ces transferts de procédures sur lesquels est fondé l'exposé de Gregory ne sont jamais évoquées ni les questions relatives à la commutabilité des symboles dans le premier cas, ni les questions de convergence que soulèvent pour nous les développements en séries dans le second cas.

Dans un second temps Gregory va utiliser ce premier exemple élémentaire pour développer une approche générale fondée sur l'itération du principe qu'il vient d'exposer. Il analyse à nouveau l'équation générale de départ $f\left(\frac{d}{dx}\right)y = X$ comme suit³⁸ :

The inverse operation of $f\left(\frac{d}{dx}\right)$ cannot easily be performed directly, but we conceive the operation $f\left(\frac{d}{dx}\right)$ to be made up by the combination of n binomial operations of the form $\left(\frac{d}{dx} - a\right)$; and, by what we have shown before, we can perform the inverse operation for each of these successively, and this will be equivalent to performing the whole inverse operation of $f\left(\frac{d}{dx}\right)$ at once. For, treating the operation $f\left(\frac{d}{dx}\right)$ exactly as if it were a function of x of the

³⁵Nous indiquons, lorsque cela nous semble important, entre parenthèses et en italiques le terme anglais utilisé par Gregory dans le texte en cours d'étude.

³⁶[Gre37, MW : p.25]

³⁷La position de Gregory n'est pas sans précédent : on peut la contraster avec la vision développée par Peacock dans son introduction au traité d'algèbre de 1830. (cf. [Pea30])

³⁸[Gre37, MW : p.19]

same form³⁹, we can resolve it into factors, so that it becomes

$$\left(\frac{d}{dx} - a_1\right) \left(\frac{d}{dx} - a_2\right) \left(\frac{d}{dx} - a_3\right) \&c. \left(\frac{d}{dx} - a_n\right),$$

where a_1, a_2, a_3 &c. are the roots of the equation

$$f(z) = 0.$$

Ainsi, dans le cas général, Gregory considère cette fois-ci l'expression symbolique $f\left(\frac{d}{dx}\right)$ « comme si elle était une fonction ». Les symboles f et $\frac{d}{dx}$ se voient attribués chacun un rôle spécifique dans la relation fonction/variable ce qui permet à Gregory de faire appel à un résultat établi dans le cadre de l'étude des fonctions (polynômiales), à savoir qu'elles se décomposent en facteurs sous la forme $f(X) = (X - a_1) \cdots (X - a_n)$. Comme il l'a fait auparavant, Gregory transpose cette décomposition en une expression où X est remplacé par $\frac{d}{dx}$. Il peut transformer l'équation $f\left(\frac{d}{dx}\right)y = X$ sous la forme $\left(\frac{d}{dx} - a_1\right) \left(\frac{d}{dx} - a_2\right) \left(\frac{d}{dx} - a_3\right) \cdots \left(\frac{d}{dx} - a_n\right)y = X$ et appliquer successivement les opérations inverses $\left(\frac{d}{dx} - a_i\right)^{-1}$ pour obtenir finalement une résolution explicite⁴⁰ :

$$y = \frac{\epsilon^{a_1 x} (\int \epsilon^{-a_1 x} X dx)}{(a_1 - a_2)(a_1 - a_3) \dots (a_1 - a_n)} + \frac{\epsilon^{a_2 x} (\int \epsilon^{-a_2 x} X dx)}{(a_2 - a_1)(a_2 - a_3) \dots (a_2 - a_n)} + \&c.$$

$$+ \frac{\epsilon^{a_n x} (\int \epsilon^{-a_n x} X dx)}{(a_n - a_1)(a_n - a_3) \dots (a_n - a_{n-1})}.$$

Dans cette première partie, la progression dépend donc de la réinterprétation et de la transposition de résultats et de procédures issus de domaines classiques comme l'arithmétique, l'intégration ou l'étude des fonctions en termes d'« opérations ». Mais Gregory va, dès cet article initial, s'interroger sur la légitimité de ce transfert de procédures dans les preuves ainsi établies. Dans cette phase, l'analyse qu'il propose dévoile déjà le fondement de l'« algèbre symbolique » telle qu'elle apparaîtra un peu plus tard dans son article de 1838 présenté à la *Royal Society of Edinburgh*. Il engage une réflexion sur l'analogie qui semble relier les calculs bien fondés dans leur champ d'origine et leur transposition au nouveau domaine⁴¹ :

Before leaving the subject, we would say a few words on the legitimacy of the processes employed in this method. In the preceding pages we have spoken of treating the symbols of operation like those of quantity, so that at first sight it would appear as if the principles on which the method is founded, were drawn only from analogy. But a little consideration will show that this is not really the case, and that the reasoning on which we proceed is perfectly strict and logical.

³⁹NDA: Il y a une erreur dans le texte de Gregory. C'est bien $f\left(\frac{d}{dx}\right)$ qu'il faut lire ici et non $\frac{d}{dx}$.

⁴⁰[Gre37, MW : p.21]

⁴¹[Gre37, MW : p.24]

Gregory propose alors de reconsidérer la dualité entre « symboles d'opération » et « symboles de quantité » qui nourrit l'analogie mais paraît entacher une vision « stricte et logique »⁴² :

We have spoken as if there were a distinction between what are usually called symbols of operation, and those which are called symbols of quantity. But we might with perfect propriety call these last also symbols of operation.

Dès 1971 Koppelman avait souligné l'insistance de Gregory à défendre qu'il n'y a pas nécessairement de distinction à faire entre les symboles d'opération (de l'arithmétique) et les symboles de quantité. Il nous faut maintenant préciser les étapes de la construction que propose Gregory et analyser les conséquences du point de vue qu'il adopte en termes de généralité. Le recours aux « opérations », en un sens étendu, lui permet de dégager une notion de symbole général recouvrant à la fois les deux types de symboles de l'arithmétique (opération et quantité). Le mot « opération » montre aussi que Gregory retient comme qualité commune à ces symboles la possibilité d'être combinés et d'opérer sur d'autres symboles d'opération. En reprenant les termes de Gregory⁴³ il s'agit donc de voir les « symboles de quantité » comme des « opérations » : x comme l'« opération notée (x) appliquée à l'unité » (*performed on*⁴⁴), ou encore $a(1)$ pour représenter l'opération (a) appliquée à 1, $a(x)$ pour « représenter la même opération appliquée à x » ou $a^n(x)$ l'« opération répétée n fois sur x ». Ainsi, dans cette vision les « symboles de quantité » deviennent « symboles d'opération » au même titre que le symbole de différentiation $\left(\frac{d}{dx}\right)$ ou le symbole de différence finie (Δ) ⁴⁵ par exemple. Gregory en déduit un principe de généralité qui permet de justifier les calculs effectués sur les symboles quels qu'ils soient⁴⁶ :

If, then we take this view of what are usually called symbols of quantity, we shall have little difficulty in seeing the correctness of the principle by which other operations, such as we represent by $\left(\frac{d}{dx}\right)$, (Δ) , &c., are treated in the same way as a , b , &c. For whatever is proved of the latter symbols, from the known laws of their combination, must be equally true of all other symbols which are subject to the same laws of combinations.

Dans ce premier article il va donc dégager trois « lois » qu'il énonce d'abord pour les « symboles » a et b , c'est-à-dire, dans le prolongement de ce qui précède, pour des symboles de quantité, puis immédiatement après pour des symboles « généraux d'opération »⁴⁷ :

⁴²[Gre37, MW : p.24]

⁴³cf. [Gre37, MW : p.24]

⁴⁴Nous signalerons lorsque nous avons traduit le mot « perform » de différentes façons suivant le contexte alors que Gregory n'emploie qu'un seul terme.

⁴⁵Ici Δ est le symbole de différence, défini par $\Delta f(x) = f(x + 1) - f(x)$.

⁴⁶[Gre37, MW : p.25]

⁴⁷[Gre37, MW : p.25]

The laws of the combination of the symbols $a, b, \&c.$ are, that

$$\begin{aligned} (1) \quad & a^m \cdot a^n x = a^{m+n} \cdot x, \\ (2) \quad & a\{b(x)\} = b\{a(x)\}, \\ (3) \quad & \text{and } a(x) + a(y) = a(x + y). \end{aligned}$$

And, if $f, f_1, \&c.$ be any other general symbol of operation (f and f_1 being of the same kind) subject to the same laws of combination, so that

$$\begin{aligned} (1') \quad & f^m \cdot f^n x = f^{m+n}(x), \\ (2') \quad & f\{f_1(x)\} = f_1\{f(x)\}, \\ (3') \quad & \text{and } f(x) + f(y) = f(x + y), \end{aligned}$$

Then, whatever we may have proved of $a, b, \&c.$ depending on these three laws, must necessarily be equally true of $f, f_1, \&c.$

Dans ces énoncés (dont nous avons numéroté les lignes) les notations employées par Gregory dévoilent une utilisation particulière des symboles. En effet, en arithmétique l'expression des lois qui définissent les opérations d'addition et de multiplication donnent aux opérands en jeu des rôles identiques. Ce n'est que dans un second temps que les symboles peuvent avoir des fonctions différenciées, par exemple x devenant l'inconnue d'une équation. Il en va différemment dans la vision proposée par Gregory où le parenthésage souligne dès le départ une évolution dans la distribution des rôles des différents symboles : certains d'entre eux sont interprétés comme opération agissant sur les autres (*performed on*). Dans notre étude, pour désigner les rôles non symétriques attribués aux symboles, nous parlerons de *symbole-agissant* et de *symbole-objet*. Les lois exposées dans la deuxième liste réservent par exemple la lettre f pour désigner les *symboles-agissant* et x ou y pour les *symboles-objets*.

Cette façon de traiter les symboles intervient de façon essentielle dans les lois listées ci-dessus. En particulier, l'utilisation de la lettre f , qui rappelle un emploi propre aux fonctions, est l'occasion pour nous de préciser la vision de Gregory. À son époque, il est usuel de noter les fonctions avec leur variable sous la forme $f(x)$, et dans ce cas nous pourrions qualifier x d'*argument muet* : couplé à f il ne sert qu'à désigner la fonction. Par contraste, dans le texte de Gregory l'opération est désignée par une seule lettre f, f_1, x ou y . Les rôles sont distribués par l'ordre d'écriture et le parenthésage : $f(x)$ désigne l'« opération » f agissant (*performed*) sur x . Dans les cas (2) et (2') c'est à nouveau cette même attribution des rôles qui est en jeu : dans le premier membre l'entité $f_1(x)$ est lue comme *symbole-objet* et sert d'argument à f , les rôles étant inversés dans l'autre membre. On ne peut pas parler ici de commutation des symboles f et f_1 , du moins pas en tant que règle régissant les symboles : la loi exprime qu'appliquer (*perform*) f_1 à x puis f au résultat obtenu revient à appliquer d'abord f à x puis f_1 à $f(x)$. Nous verrons une évolution de point de vue qui apparaîtra dans l'article de 1838 de la *Royal Society of Edinburgh* où, en faisant référence à Servois, Gregory pensera cette loi en termes de commutativité.

Enfin remarquons que les égalités (3) et (3') nécessitent que l'on puisse additionner x et y d'une part et les résultats des opérations sur ces éléments,

$f(x)$ et $f(y)$, d'autre part. Nous verrons que l'expression de cette dernière loi et le rôle de l'addition ou du symbole $+$, lequel n'est pas explicitement désigné ici comme un « symbole d'opération », seront au centre des travaux suivants de Gregory.

Après avoir isolé les lois générales que nous venons de citer, Gregory argumente que tout théorème établi pour les symboles de l'arithmétique en n'utilisant que ces lois peut être transféré rigoureusement aux autres opérations. Ceci est le cas pour les différentielles, leur addition, leur multiplication ou l'opération puissance qui se combinent selon des lois identiques à celles qui régissent les opérations correspondantes pour les nombres. Ceci permet par exemple de justifier le théorème du binôme pour $\frac{d}{dx}$ et $\frac{d}{dy}$ sous la forme suivante⁴⁸ :

The binomial theorem (to take a particular case) which has been proved for (a) and (b) is equally true for $\left(\frac{d}{dx}\right)$ and $\left(\frac{d}{dy}\right)$: so that we require no farther proof for the proposition, that when u is a function of two independent variables x and y ,

$$d^n(u) = \left(\frac{d}{dx}dx + \frac{d}{dy}dy\right)^n u = \frac{d^n u}{dx^n} dx^n + n \frac{d^{n-1}}{dx^{n-1}} \frac{d}{dy} u dx^{n-1} dy + \dots$$

Gregory utilisera ces idées⁴⁹ et les mettra en pratique dans ses publications suivantes dans le *Cambridge Mathematical Journal*. En particulier, il publie en 1838 (juste avant son article à la *Royal Society of Edinburgh*) un article intitulé « On the Solution of Linear Equations of Finite and Mixed Differences »⁵⁰ dans lequel il étudie les équations du type :

$$u_{x+n} + Au_{x+n-1} + \&c. + Ru_{x+1} + Su_x = X$$

où les lettres A, B, \dots, S sont des constantes, X est une fonction de la variable x et u_{x+k} représente (à travers des notations actuelles) la fonction $\phi(x) = u(x+k)$. Gregory réinterprète alors l'équation de départ en considérant u_{x+n} comme le résultat de l'itération d'une même opération appliquée à la fonction u_x et en s'appuyant sur la relation $\Delta u_x = u_{x+1} - u_x$ où Δ est le symbole de différence finie⁵¹ :

By a known relation we can express the quantities u_{x+n} , u_{x+n-1} , &c. in terms of u_x , and its successive differences, so that the equation may be transformed into one which may be put under the form

$$f(\Delta)u_x = X$$

⁴⁸[Gre37, MW : p.25]

⁴⁹Koppelman souligne que l'analogie entre les dérivations successives et le développement binomial remonte au moins aux travaux de Leibniz : « In a letter to Johann Bernoulli, dated 1695, Leibniz noted the analogy between 'numeros potestatum a binomio, et differentiarum a rectangulo' » [Kop71, p.157].

⁵⁰[Gre38c, MW : p.33]

⁵¹[Gre38c, MW : p.33]

which we might proceed to solve by means of the separation of the symbols. But it will be more convenient to proceed in the following manner. Since we have, generally,

$$u_{x+n} = (1 + \Delta)^n u_x,$$

we may consider $1 + \Delta$ as a separate symbol, subject to the same laws as d and Δ , the effect of which, when applied on u_x , is to convert u_x into u_{x+1} . If we represent, for the sake of shortness, $1 + \Delta$ by D , the equation will take the form

$$(D^n + AD^{n-1} + \&c. + RD + S)u_x = X.$$

La lecture des différents éléments de l'équation d'origine (les u_{x+n}) grâce à une opération unique permet à Gregory de séparer les symboles et de réinterpréter cette équation sous une forme similaire à celle obtenue dans son premier article sur les équations différentielles.

Dans le prolongement de ces idées, toujours en 1838, Gregory publie un troisième article, intitulé « On the Solution of Partial Differential Equations »⁵² dans lequel il développe un peu plus les techniques qu'il a mentionnées dans son premier article à propos des équations linéaires aux dérivées partielles qu'il envisage désormais à plusieurs variables. Nous ne nous attardons pas ici sur les démonstrations qu'il propose dans cet article et qui reposent essentiellement sur des procédés que nous avons analysés plus haut.

Enfin, mentionnons que dans ses travaux Gregory fait plusieurs fois référence au travail de Fourier qu'il tenait en haute estime. Le recueil des écrits mathématiques de Gregory, publié après sa mort en 1865, s'ouvre sur une biographie de Robert Leslie Ellis qui écrit⁵³ :

[Gregory] had long been an admirer of Fourier's great work on heat, and once, while turning over its pages, remarked to the writer, —“All these things seem to be a kind of mathematical paradise”.

Deux textes de Gregory font plus particulièrement référence au travail de Fourier sur la chaleur. Au-delà de l'admiration toujours marquée, les commentaires de Gregory permettent de préciser la forme que prennent petit à petit ses idées.

Dans une première publication intitulée « Notes on Fourier's Heat », Gregory résume la façon de procéder de Fourier⁵⁴ :

The method employed by Fourier to integrate the partial differential equations which occur in the Theory of Heat, is to assume some simple form of a singular solution, and afterwards to extend it so as to include all the circumstances of the problem.

La « méthode » de Fourier consiste donc à « étendre un cas particulier » où les solutions prennent une forme « simple » afin d'atteindre un résultat général. Mais Gregory est critique par rapport à cette approche : il pense qu'il vaut

⁵²[Gre38d, MW : p.62]

⁵³[Gre65, preface p.xxi]

⁵⁴[Gre38a, MW : p.57]

mieux partir d'une solution générale et d'en explorer les applications aux cas particuliers⁵⁵ :

But however much we may admire the means by which Fourier has overcome the difficulties of the problems he had to deal with, yet it seems more agreeable to the usual style of mathematical investigation to deduce a result by limiting the general solution by means of the conditions of the problem, than by extending a particular case.

Nous verrons dans la suite comment pour Gregory le recours à l'algèbre symbolique permet de proposer une vision de nombreux problèmes mathématiques fondée sur la primauté d'une théorie générale.

Dans un second article que nous avons déjà mentionné plus haut, « On the Solution of Partial Differential Equations », Gregory prolonge son commentaire de la méthode de Fourier⁵⁶ :

Fourier, in his *Traité de la Chaleur*, published in 1822, has shewn that the series which are obtained in the solution of several partial differential equations may be conveniently expressed by the separation of the symbols of operation from those of quantity. But though he has used this method very frequently, yet he appears to have had some unwillingness to give himself up to it entirely as a guide in his investigations, as if he were not familiar with the principles on which it is founded. His idea apparently was, that the expression which he obtained as the solutions might be conveniently expressed by separating the symbols of operation, and not that the symbolical expressions are the proper solutions of the equations, and the series merely the expansion of them.

Pour comprendre la critique de Gregory il est nécessaire de relire le traité de Fourier publié en 1822. Il faut en effet attendre la page 517 et l'article numéro 401 (sur 433 au total) pour voir apparaître une notation qui utilise la « séparation du symbole d'opération »⁵⁷ :

La série T de l'art. 399,

$$\left[v = c + tc'' + \frac{t^2}{2}c^{iv} + \frac{t^4}{2 \cdot 3 \cdot 4}c^{vi} + \frac{t^6}{2 \cdot 3 \cdot 4 \cdot 5 \cdot 6}c^{viii} + \text{etc.} \quad (T) \right]$$

qui dérive de l'équation

$$\frac{dv}{dt} = \frac{d^2v}{dx^2} \quad (a)$$

peut être mise sous cette forme $v = e^{-tD^2} \cdot \varphi x$. On développera l'exponentielle selon les puissances de D , et l'on écrira $\frac{d^i}{dx^i}$ au lieu de D^i , en considérant i comme indice de différentiation. On aura ainsi

$$v = \varphi x + t \frac{d^2}{dx^2} \varphi x + \frac{t^2}{2} \frac{d^4}{dx^4} \varphi x + \frac{t^3}{2 \cdot 3} \frac{d^6}{dx^6} \varphi x + \text{etc.}$$

⁵⁵[Gre38a, MW : p.57]

⁵⁶[Gre38d, MW : p.62]

⁵⁷[Fou22, p.517]

Suivant la même notation, la première partie de la série (X) (art. 399)

$$\left[v = a + \frac{x^2}{2}a_i + \frac{x^4}{2 \cdot 3 \cdot 4}a_{ii} + \frac{x^6}{2 \cdot 3 \cdot 4 \cdot 5 \cdot 6}a_{iii} + etc. + \right. \\ \left. xb + \frac{x^3}{2 \cdot 3}b_i + \frac{x^5}{2 \cdot 3 \cdot 4 \cdot 5}b_{ii} + \frac{x^7}{2 \cdot 3 \cdot 4 \cdot 5 \cdot 6 \cdot 7}b_{iii} + etc. \right] \quad (X)$$

qui ne contient que des puissances paires de x , sera exprimée sous cette forme : $\cos(x\sqrt{-D})\varphi t$. On développera selon les puissances de x , et l'on écrira $\frac{d^i}{dt^i}$ au lieu de D^i , en considérant i comme indice de différentiation. La seconde partie de la série (X) se déduit de la première, en intégrant par rapport à x , et changeant la fonction φt en une autre fonction arbitraire ψt . On a donc

$$v = \cos.(x\sqrt{-D}) + W,$$

et

$$W = \psi t \cdot \int_0^x dx \cos.(x\sqrt{-D}).$$

Fourier utilise donc le formalisme issu de la séparation des symboles d'opération comme moyen de représenter les solutions des équations différentielles. Mais il va aussi un peu plus loin qu'on ne le lit dans la critique de Gregory en montrant que l'on peut utiliser ces notations dans des calculs de différentiation et qu'ils permettent de « vérifier » les résultats obtenus⁵⁸ :

Ces notations abrégées et connues dérivent des analogies qui subsistent entre les intégrales et les puissances. Quant à l'usage que nous en faisons ici, il a pour objet d'exprimer les séries, et de les vérifier sans aucun développement. Il suffit de différentier sous les signes que cette notation emploie. Par exemple, de l'équation $v = e^{tD^2}\varphi x$, on déduit, en différentiant par rapport à t seulement,

$$\frac{dv}{dt} = D^2 \cdot e^{tD^2}\varphi x = D^2v = \frac{d^2}{dx^2} \cdot v^{59} ;$$

ce qui montre immédiatement que la série satisfait à l'équation différentielle (a).

Fourier commentera un peu plus loin que pour les équations « très simples » (qui sont celles dont il part pour étudier la chaleur), on peut se passer des « expressions abrégées », mais qu'en général elles « suppléent à des calculs très composés. » Dans la suite du traité sur la chaleur, Fourier n'utilisera plus ce formalisme et même si Gregory admet que Fourier a ici « ouvert une voie » (*the track Fourier opened*⁶⁰), le recours à la séparation des symboles d'opération n'est pas utilisé pour formuler le problème de départ ni même, comme le remarque Gregory, comme « guide dans ses investigations ».

⁵⁸[Fou22, p.518]

⁵⁹Il y a une coquille dans l'édition que nous citons du texte de Fourier. Il faut bien lire $D^2v = \frac{d^2}{dx^2}v$ et non comme il est écrit $D^2v = \frac{d^2}{dx^2}\varphi$.

⁶⁰[Gre38d, MW : p.62]

Dans une démarche qui contraste avec ce qu'il lit chez Fourier par exemple, Gregory cherche à montrer et à justifier que l'on peut utiliser le calcul symbolique pour reformuler des problèmes et en déduire des solutions validées par une preuve : dans ses publications il montre que l'on peut rester complètement du côté symbolique en réinterprétant convenablement les problèmes ainsi que certains résultats validés dans leur propre domaine. Dans cette première approche on a obtenu une preuve (*proof*) dès lors que les symboles employés satisfont aux lois sous lesquelles les théorèmes ont été établis dans le cadre général de l'algèbre symbolique. Ce sont ces idées que Gregory reprendra et fera évoluer dans l'article publié en 1838 dans les *Transactions of the Royal Society of Edinburgh* que nous allons maintenant étudier à la lumière des éléments que nous venons de préciser.

2.4 L'ALGÈBRE SYMBOLIQUE

L'article « On the Real Nature of Symbolical Algebra »⁶¹ de 1838 est l'une des rares publications de Gregory hors de la revue qu'il a cofondée. Gregory y expose plus précisément les principes généraux qui sous-tendaient ses premiers travaux et comment il conçoit l'utilisation des « symboles d'opération »⁶² :

The light, then, in which I would consider symbolical algebra, is, that it is the science which treats of the combination of operations defined not by their nature, that is by what they are or what they do, but by the laws of combination to which they are subject. And as many different kinds of operations may be included in a class defined in the manner I have mentioned, whatever can be proved of the class generally, is necessarily true of all the operations included under it. This, it may be remarked, does not arise from any analogy existing in the nature of the operations, which may be totally dissimilar, but merely from the fact that they are all subject to the same laws of combination.

Cet article de 1838 a été largement cité dans la littérature secondaire et nous renvoyons pour une mise en perspective plus large aux principales études que nous avons mentionnées plus haut : en particulier plusieurs passages que nous étudions sont mentionnés dans [Kop71, p.216 et sq.], [AB02, p.405 et sq.] et [Des11, p.57 et sq.]. L'optique de ces études est différente de la nôtre et nous prolongeons ici les analyses qui y sont menées en dégagant précisément ce qui relève du travail d'abstraction et engendre une vision générale dans l'œuvre de Gregory.

Pour Gregory, l'« algèbre symbolique » est donc une « science » dont les résultats découlent des « lois de combinaison » auxquelles les « opérations » sont soumises. Il définira, comme nous le verrons par la suite, cinq « classes d'opérations » organisées suivant les lois qui les gouvernent. Gregory insiste dès le départ sur ce qui légitime les résultats obtenus. En particulier, la justification qu'un même théorème est vrai pour différentes opérations dans des domaines

⁶¹[Gre38b, MW : p.1–13]

⁶²[Gre38b, MW : p.2]

variés ne découle pas d'une « analogie » qui existerait dans la « nature des opérations » ou dans leur façon d'agir, mais du fait que le théorème et sa preuve ne reposent que sur les « lois de combinaison » qui sont communes à toutes les opérations d'une classe⁶³ donnée. Il s'agit là d'un véritable *principe de généralité* que propose Gregory pour fonder une « science » rigoureuse.

De plus, Gregory rappelle que les lois envisagées ont souvent été « inspirées » (*suggested*) par celles de l'arithmétique et il précise la nature de l'abstraction à l'œuvre dans le passage de l'arithmétique à l'algèbre symbolique⁶⁴ :

The step which is taken from arithmetical to symbolical algebra is, that, leaving out of view the nature of the operations which the symbols we use represent, we suppose the existence of classes of unknown operations subject to the same laws.

Ainsi pour Gregory la nouvelle science n'est pas fondée sur une « analogie » avec les opérations algébriques de l'arithmétique qui, par une sorte de réécriture *transitive*⁶⁵, permettrait d'en définir ou d'en dévoiler le contenu. Le passage à l'algèbre symbolique nécessite une rupture (*step*) qui s'opère dans un processus d'abstraction complexe. Si la longue histoire de l'arithmétique a permis de dégager des lois qui gouvernent ses opérations sur les quantités, et de parler d'« algèbre arithmétique », la procédure d'abstraction à l'œuvre ici exprime des lois sans référence aux quantités initiales par le biais de symboles privés de leur signification originelle. L'algèbre symbolique se débarrasse donc de tout contexte dans lequel les « symboles d'opération » agiraient sur des sujets identifiés —c'est-à-dire dont la « nature » serait connue *a priori* dans tel ou tel domaine— et offre donc une vision très générale, une généralité que nous qualifierons d'*ouverte* car « on suppose l'existence de classes d'opérations inconnues qui vérifient les mêmes lois. »

Dans cette nouvelle théorie, ce qui établit la preuve se joue au niveau des symboles et son exactitude (*correctness*) est liée à la conformité des calculs aux lois qui les gouvernent⁶⁶ :

We are thus able to prove certain relations between the different classes of operations, which, when expressed between the symbols, are called algebraical theorems.

Les « théorèmes algébriques » sont écrits avec les symboles de l'algèbre symbolique et désignent des « relations » établies grâce aux seules lois qui gouvernent les « classes d'opérations ». Les lois exposées par Gregory font parfois intervenir plusieurs classes et définissent leurs possibilités de combinaisons : il dira alors que les classes sont « connectées » (*connected*) par les lois énoncées (nous verrons des exemples par la suite).

⁶³Le mot « classe » (*class*) est utilisé par Gregory pour indiquer qu'il rassemble sous une même visée un ensemble de situations ou d'éléments (ici des opérations). Le sens de ce mot sera précisé au fur et à mesure de notre commentaire.

⁶⁴[Gre38b, MW : p.2]

⁶⁵Foucault dit que « l'écriture est transitive [lorsqu'elle] est destinée à désigner, montrer, manifester hors d'elle-même quelque chose qui, sans elle, serait resté sinon caché, du moins invisible. » [Fou11]

⁶⁶[Gre38b, MW : p.2]

Enfin, Gregory explique comment l'algèbre symbolique, ou cette nouvelle « théorie », fait œuvre dans les diverses « sciences » où elle est appliquée⁶⁷ :

[I]f we can show that any operations in any science are subject to the same laws of combination as these classes, the theorems are true of these as included in the general case: Provided always, that the resulting combinations are all possible in the particular operation under consideration.

Si au départ les opérations et les lois de l'algèbre symbolique ont été suggérées (*suggested*) par des situations particulières, cette filiation est rompue pour donner naissance à une science où les théorèmes énoncent des « cas généraux » (*general case*). L'algèbre symbolique est organisée par Gregory suivant une liste de lois qui régissent les possibilités de combinaison des opérations au sein d'une même « classe » ou qui définissent les interactions possibles entre les opérations de plusieurs « classes ».

De façon contrastée avec l'emploi du mot « classe », Gregory distingue des « familles » d'opérations : il s'agit de collections d'opérations agissant dans un champ particulier qui sont soumises aux mêmes lois que celles d'une « classe » d'opérations définie en algèbre symbolique. Si parfois une « famille » désigne une collection d'opérations ayant suggéré la « classe » correspondante en algèbre symbolique (des opérations de l'arithmétique par exemple), ce n'est pas toujours le cas. Gregory illustre chaque « classe » à l'aide de plusieurs « familles » distinctes dans divers champs (en arithmétique, calcul différentiel ou géométrie).

Enfin, chaque théorème établi de façon générale en algèbre symbolique ne donne pas systématiquement un théorème dans chaque « famille » : un théorème ne sera « vrai » en ce qui concerne les « opérations » identifiées dans tel ou tel champ que si toutes les expressions obtenues par combinaison des symboles sont « possibles » dans le domaine où les symboles sont définis, où leur « nature »⁶⁸ est connue. Cette dernière nécessité que les « combinaisons soient possibles », sur laquelle Gregory reviendra dans ses travaux suivants, est ici centrale car elle lui permet de distinguer diverses « familles d'opérations » en termes de généralité⁶⁹ :

It may very well, and does actually happen, that, though each of two operations in a certain branch of science may be possible, the complex operation resulting from their combination is not equally possible. In such a case, the result is inapplicable to that branch of science. Hence we find that one family of a class of operations may have a more general application than another family of the same class.

⁶⁷[Gre38b, MW : p.2]

⁶⁸Gregory emploie le mot « nature » (*nature of the operations*) de multiples façons. Ici comme souvent, il s'agit de considérer les opérations dans un champ bien défini —arithmétique, géométrie, etc.— où les symboles désignent des opérations qui ont un sens propre dans le champ considéré. L'expression « nature d'un symbole » indique alors la connaissance du mode opératoire dans les champs particuliers.

⁶⁹[Gre38b, MW : p.2]

Cette possibilité ou au contraire l'impossibilité de donner une interprétation dans les domaines particuliers des opérations est fondamentale dans le travail de Gregory et nous y consacrerons la section suivante. Nous verrons en particulier comment la géométrie prend dans ce registre de l'interprétation une place remarquable et comment Gregory explore plusieurs « familles d'opérations » pour donner un sens à telle ou telle formulation symbolique.

Après ces réflexions préparatoires, la plus grande partie de l'article de 1838 est consacrée à exposer cinq types de lois⁷⁰ et à identifier des « familles » qui correspondent à ces lois dans divers domaines comme l'arithmétique, la géométrie, les équations différentielles ou la trigonométrie⁷¹ :

1. We assume, then, the existence of two classes of operations F and f , connected together by the following laws:

$$(1) \quad FF(a) = F(a) \quad (2) \quad ff(a) = F(a)$$

$$(3) \quad Ff(a) = f(a) \quad (4) \quad fF(a) = f(a).$$

D'une part cette citation permet de préciser ce que Gregory entend par « loi » dans ses travaux, et de distinguer cet emploi de celui qui est le nôtre quand on parle de *loi de composition* par exemple. La « loi » est en effet clairement dissociée de la notion d'« opération » : elle désigne pour Gregory une *règle* qui régit la combinaison des symboles d'une ou plusieurs classes d'opérations.

Gregory mentionne que les opérations de l'arithmétique qu'il appelle momentanément « addition » et « soustraction » satisfont aux lois rappelées plus haut, et il propose cependant de garder finalement les symboles $+$ et $-$ au lieu de F et f pour désigner de façon générale de telle opérations en algèbre symbolique. Nous verrons que Gregory reviendra sur ce vocabulaire et la notation adoptée ici en distinguant plusieurs emplois des signes $+$ et $-$ en arithmétique ou dans d'autres domaines des mathématiques. Si Gregory conserve par commodité ces symboles dans une première étape, il propose de renommer les opérations désignées par $+$ et $-$ dans le cadre général de l'algèbre symbolique « opérations circulaires ou reproduisantes » (*circulating or reproducing operations*).

D'autre part, Gregory prend pour exemple la géométrie pour illustrer le fait que l'utilisation d'un même symbole, ici $+$, n'implique pas d'analogie au niveau de la nature des opérations (entre l'arithmétique et la géométrie). En effet, les deux opérations de la géométrie auxquelles Gregory fait référence sont « le transfert d'un point le long d'une circonférence », correspondant au $+$, et « le transfert d'un point le long d'une demi-circonférence », correspondant au signe $-$ ⁷². Pour montrer que les résultats exprimés de façon générale et unique

⁷⁰Alaire et Bradley [AB02] citent aussi ces cinq lois pour illustrer les choix de Gregory et les contraster avec ceux antérieurs de Peacock. Nous analysons ici à nouveau ces lois pour mettre en évidence la sélection opérée par Gregory et les rôles précis qu'il attribue aux différents éléments impliqués dans l'écriture symbolique.

⁷¹[Gre38b, MW : p.3]

⁷²Nous précisons un peu plus loin schéma à l'appui comment Gregory conçoit l'utilisation des signes $+$ et $-$ en géométrie. Pour comprendre les idées mises en avant ici par Gregory, on peut penser à un point a sur un cercle, $+a$ son image par la rotation d'angle 2π sur le cercle (d'où

en algèbre symbolique ne peuvent pas provenir d'une simple analogie entre les opérations de l'arithmétique et celles de la géométrie, Gregory souligne qu'en géométrie les deux opérations ne peuvent pas être qualifiées d'opposées (*opposed*) l'une de l'autre, alors qu'elles le sont dans une certaine utilisation des signes $+$ et $-$ en arithmétique. Les opérations n'ont en commun que les lois qui permettent de les combiner, et pour Gregory c'est en ça qu'elles sont en « relation »⁷³, une relation qui ne découle pas d'une analogie.

Dans le deuxième groupe de lois proposées par Grégory les symboles ne sont pas disposés sur le même niveau, ce qui nous laisse entrevoir l'« inspiration » du calcul des puissances en arithmétique⁷⁴ :

2. Let us suppose the existence of operations subject to the following laws:

$$(1) f_m(a) \cdot f_n(a) = f_{m+n}(a) \quad (2) f_m f_n(a) = f_{mn}(a).$$

Dans un premier temps Gregory montre que les opérations de l'algèbre arithmétique symbolisées par a^n suivent ces lois quand on prend pour n un rationnel. Puis il explore l'étendue de cette conception symbolique à travers diverses situations. En particulier, pour Gregory, dans la notation a^m , le a et le m sont des symboles d'opération au sens général de l'algèbre symbolique (ils peuvent ne pas être des nombres). En guise d'illustration il rappelle un usage où les puissances sont conçues comme des opérations avec l'énoncé du théorème de Taylor sous la forme suivante⁷⁵ :

When, instead of supposing m to be a number integer or fractional, we suppose it to indicate any operation whatever, I do not know of any interpretation which can be given to the notation, excepting in the case when it indicates the operation of differentiation, represented by the symbol d . For we know by Taylor's theorem, that

$$\epsilon^{h \frac{d}{dx}} f(x) = f(x + h),$$

or,

$$a^{\frac{d}{dx}} f(x) = f(x + \log a).$$

Gregory insiste ici sur le fait que les symboles a et ϵ sont tous les deux lus comme des opérations et pas comme des quantités. Il y reviendra un peu plus loin dans son article en préférant reformuler le théorème de Taylor avec la notation \log^{-1} qui désignera de façon plus adéquate l'opération inverse de \log ⁷⁶ :

Properly speaking, however, the symbol ϵ ought not to be used, as it implies an arithmetical relation, and instead, we ought to employ the more general symbol of \log^{-1} .

la légitimité de l'égalité $a = +a$ admise pour l'instant par Gregory), et $-a$ le point diamétralement opposé à a .

⁷³[Gre38b, MW : p.4]

⁷⁴[Gre38b, MW : p.4]

⁷⁵[Gre38b, MW : p.5]

⁷⁶[Gre38b, MW : p.10]

On voit là encore le désir de marquer une séparation nette entre le niveau général de l'algèbre symbolique et celui de l'arithmétique qui est vue ici comme un champ d'application particulier. L'analyse du logarithme, via l'opération \log , amènera Gregory à proposer une quatrième collection de lois que nous évoquerons un peu plus bas.

Pour prolonger l'analyse de ces premières lois, Gregory propose d'explorer plus avant les situations géométriques et arithmétiques que nous avons vues plus haut. La question qui se pose alors est de savoir ce que signifient $(+)^m$ ou $(-)^m$ lorsque m est un rationnel. Lorsque m est un entier naturel, Gregory admet que $(+)^m = +$, $(+)^{\frac{1}{m}} = +$ ou $(-)^{2m} = +$ et $(-)^{\frac{1}{2m}} = -$ sont autant d'écritures qui peuvent faire sens en arithmétique⁷⁷. Mais cette vision trouve ses limites lorsqu'il s'agit de donner un sens à une expression de la forme $(-)^{\frac{1}{2q+1}}$. Gregory se sert de ce cas pour illustrer que certaines familles d'opérations peuvent donner lieu à des interprétations plus étendues que d'autres.

En effet en géométrie, quel que soit m rationnel, on peut concevoir que $(+)^{\frac{1}{m}}$ ou $(-)^{\frac{1}{m}}$ représentent ce que nous avons appelé des rotations, cette fois-ci d'angle $\frac{2\pi}{m}$ ou $\frac{\pi}{m}$ respectivement, et ainsi interpréter des notations qui ne faisaient pas sens en arithmétique. Cette réflexion initiée par Gregory dans son article à la *Royal Society of Edinburgh* sera reprise de façon centrale plus tard dans ses travaux et nous y reviendrons en détail.

La troisième collection de lois est donnée sous la forme suivante⁷⁸ :

3. I now proceed to a very general class of operations, subject to the following laws:

$$(1) \quad f(a) + f(b) = f(a + b)$$

$$(2) \quad f_1 f(a) = f f_1(a).$$

Gregory qualifera les symboles vérifiant ces lois respectivement de « distributifs » et de « commutatifs », termes qu'il emprunte à Servois qui a isolé ces propriétés dans sa propre approche du calcul différentiel⁷⁹. Parmi les exemples les plus importants envisagés par Gregory on retrouve ici l'opération de différentiation d et celle de différence Δ qui sont manipulées comme nous l'avons vu dans la section précédente. Il développe aussi un exemple issu de la géométrie, dont il nomme l'opération « *transference* » (une sorte de translation), qui donne lieu à une subtile représentation de la distributivité et de la commutativité. Nous laissons le lecteur consulter les textes de Gregory pour explorer ces exemples plus avant et nous illustrerons un peu plus bas notre propos par des situations géométriques caractéristiques de sa démarche.

⁷⁷Ces écritures auxquelles Gregory trouve un sens font émerger une nouvelle difficulté qu'il n'aborde pas dans cet article, mais sur laquelle il reviendra plus tard comme nous le verrons. En effet, d'après les égalités que nous venons de mentionner, $(+)^{2m} = +$ ainsi que $(-)^{2m} = +$, et donc $+$ possède donc deux racines paires : $(+)^{\frac{1}{2m}} = +$ mais aussi $(+)^{\frac{1}{2m}} = -$.

⁷⁸[Gre38b, MW : p.6]

⁷⁹Servois introduit les notions de *distributivité* et de *commutativité* dans un travail sur les principes du calcul différentiel dans lequel il définit ces notions pour des fonctions et des *fonctions différentielles*. (cf. [Ser15, p.98 et p.116])

Enfin Gregory finit son article en proposant deux dernières classes d'opérations. La quatrième classe de l'article est définie par une loi qui est suggérée par le logarithme mais qui est ici exprimée dans les termes généraux de l'algèbre symbolique⁸⁰ :

4. If we define a class of operations by the law

$$f(x) + f(y) = f(xy),$$

we see that, when x and y are numbers, the operation is identical with the arithmetical logarithm. But when x and y are any thing else, the function will have a different meaning.

La dernière collection de lois est suggérée par certaines fonctions qui apparaissent dans l'étude de secteurs elliptiques⁸¹ :

5.

$$(1) \quad aF(x+y) = F(x)f(y) + f(x)F(y),$$

$$(2) \quad af(x+y) = f(x)f(y) - cF(x)F(y).$$

En particulier, lorsque les constantes a et c sont égales à 1 on retrouve des égalités vérifiées par le sinus et le cosinus d'un angle.

À partir de ces lois Gregory donne une analyse de l'égalité de De Moivre qui nous permet d'une part de mieux appréhender la nature des raisonnements auxquels il va se livrer et d'autre part d'illustrer une problématique de l'interprétation qui nous occupera dans la section suivante⁸² :

The most important theorem proved of this class of functions is that of Demoiivre, that

$$\{\cos x + (-)^{\frac{1}{2}} \sin x\}^n = \cos nx + (-)^{\frac{1}{2}} \sin nx.$$

Ici à nouveau Gregory remarque que cette expression ne fait pas sens en « algèbre arithmétique » (*arithmetical algebra*). Celui-ci rappelle un peu plus haut dans son texte que les symboles \cos et \sin peuvent représenter en géométrie des « fonctions d'angles ou de secteurs circulaires », alors que dans d'autres situations ils servent à désigner des expressions du type « $\sin x = x - \frac{x^3}{1 \cdot 2 \cdot 3} + \frac{x^5}{1 \cdot 2 \cdot 3 \cdot 4 \cdot 5}$ &c. » ou « $\cos x = 1 - \frac{x^2}{1 \cdot 2} + \frac{x^4}{1 \cdot 2 \cdot 3 \cdot 4}$ &c. »⁸³. Dans ce dernier cas Gregory conçoit la notation \cos ou \sin comme une expression abrégée de relations compliquées entre des classes d'opérations⁸⁴. C'est donc l'expression $(-)^{\frac{1}{2}}$ qui ne peut pas être interprétée en arithmétique, c'est-à-dire comme « opération » agissant sur des nombres.

Par contraste, en géométrie la combinaison de symboles fait encore une fois sens comme l'explique Gregory. Dans son texte, si a représente un segment, $a \cos x$ et $a \sin x$ représentent des segments dont les longueurs ont été modifiées

⁸⁰[Gre38b, MW : p.11]

⁸¹[Gre38b, MW : p.11]

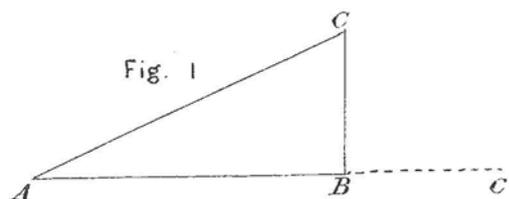
⁸²[Gre38b, MW : p.12]

⁸³[Gre38b, MW : p.12]

⁸⁴Gregory écrit : « we only know of them as abbreviated expressions for certain complicated relations between the first three classes of operations we have considered. » [Gre38b, MW : p.12]

par les facteurs $\cos x$ et $\sin x$, et l'expression composée $a\{\cos x + (-)^{\frac{1}{2}} \sin x\}$ est illustrée par une construction géométrique précise⁸⁵ :

For instance, if $AB = a \cos x$, and $BC' = a \sin x$, we must not measure it in the prolongation of AB (fig. 1), but turn it round to the position BC ; and thus, geometrically, we arrive at the point C . Also, from the relation between $\sin x$ and $\cos x$, we know that the line AC will be equal to a , and thus the expression $a\{\cos x + (-)^{\frac{1}{2}} \sin x\}$ is an operation expressing that the line whose length is a , is turned through an angle x .



Ainsi dans cette analyse, $(-)^{\frac{1}{2}}$ indique une opération qui fait tourner le segment représenté par $a \sin x$ d'un angle droit, alors que l'expression $a\{\cos x + (-)^{\frac{1}{2}} \sin x\}$ fait tourner le segment de longueur a d'un angle x . Cette façon de voir permet en outre à Gregory d'affirmer que les deux expressions $\cos \frac{2\pi}{n} + (-)^{\frac{1}{2}} \sin \frac{2\pi}{n}$ et $(+)^{\frac{1}{n}}$ « dénotent » (*indicate*) la même opération, l'une étant exprimée en coordonnées cartésiennes (*rectangular coordinates*), l'autre en coordonnées polaires.

Les lois que nous avons rappelées constituent pour Gregory les principales règles auxquelles les symboles peuvent être soumis. Il n'affirme pas qu'il s'agit là d'une liste exhaustive, et il ne propose ici que les « lois appropriées aux différentes classes d'opérations que l'on a l'habitude d'utiliser »⁸⁶. Ces « classes d'opérations » sont l'expression d'une vision abstraite élaborée à partir de la relecture de diverses situations où Gregory identifie ou fait apparaître des opérations ainsi que les lois qui les gouvernent. Nous avons mentionné que le type de généralité qui en découle est *ouvert* car, comme le remarque Gregory, on suppose qu'il peut y avoir des opérations que l'on n'a pas encore envisagées qui suivent aussi les lois de telle ou telle « classe » de l'algèbre symbolique. D'autre part, même si plusieurs « familles » particulières, issues de l'arithmétique, de l'analyse ou de la géométrie, suivent les mêmes règles générales lorsque celles-ci sont exprimées dans les termes de l'algèbre symbolique, les théorèmes obtenus peuvent ne pas être « interprétables » dans tous les champs particuliers.

Du point de vue de la généralité nous avons vu que la vision abstraite de l'algèbre symbolique faisait œuvre de façon plus ou moins étendue dans les différents domaines particuliers qu'elle recouvre. Un même théorème d'algèbre symbolique peut faire sens (*be interpreted*) dans un domaine mais pas dans un autre.

⁸⁵[Gre38b, MW : p.12]

⁸⁶[Gre38b, MW : p.3]

Enfin, parallèlement à la mise en place du cadre abstrait de l'algèbre symbolique, les travaux de Gregory mettent nettement en avant des efforts de réinterprétation d'énoncés choisis dans divers champs des mathématiques à l'aide de la nouvelle conception symbolique. Les textes que nous abordons maintenant permettent d'illustrer plus précisément ce travail de réinterprétation et les difficultés que Gregory rencontre et essaie de contourner.

2.5 DE LA CONCEPTION DE L'ALGÈBRE SYMBOLIQUE À L'ÉTABLISSEMENT D'UNE THÉORIE

La première partie de notre travail a permis de montrer comment la conception de l'algèbre symbolique élaborée par Gregory relevait d'un double travail d'abstraction et de sélection. Le mouvement d'abstraction auquel nous avons fait référence doit être précisé. L'abstraction que nous avons cherché à mettre en évidence est précisément celle qui permet à Gregory de passer des champs particuliers à une conception symbolique unique fondée sur un nombre très limité de lois (les cinq lois mises en avant que nous avons rappelées). Ce type d'abstraction a une portée générale en ce sens qu'elle cherche à embrasser les diverses opérations propres aux champs particuliers, opérations connues ou inconnues (*généralité ouverte*). Pourtant, par contraste avec cette *ouverture*, le système mis en place par Gregory est bien clos et ses travaux ne cherchent pas à étendre le nombre des « classes d'opérations ». Au prix d'efforts substantiels de réinterprétation des opérations dans les champs particuliers qu'il envisage, Gregory maintient une conception de l'algèbre symbolique fondée sur les seules cinq lois posées dans son article de 1838 « On the Real Nature of Symbolical Algebra ».

Si les articles de 1838 et en particulier celui publié à la *Royal Society of Edinburgh* établissent les énoncés des lois et les méthodes de l'algèbre symbolique, Gregory va produire plusieurs autres travaux qui sont à nos yeux incontournables pour saisir la pensée que le mathématicien développe.

En particulier Gregory publie dans le *Cambridge Mathematical Journal* en 1839 un article intitulé « On the Impossible Logarithms of Quantities »⁸⁷ qui permet d'explorer précisément sa conception de l'algèbre symbolique.

Dans ce nouveau texte Gregory essaie d'analyser une divergence de point de vue sur le problème du « logarithme des quantités négatives »⁸⁸ :

The other point which I propose to elucidate at present, and which is the chief object of this paper, is the plurality of logarithms of quantities [· · ·] This is closely connected also with the discussion concerning the logarithms of negative quantities, which attracted so much attention in the time of Euler, d'Alembert, and John Bernoulli, and the interest of which has been revived of late years by the researches of Vincent, Ohm, and Graves. Euler had apparently set the question at rest by demonstrating the existence of an infinite number of logarithms of a quantity, one only of which is possible;

⁸⁷[Gre39, MW : p.124]

⁸⁸[Gre39, MW : p.124]

and the formula he gave was that

$$\log a = L(a) + 2r\pi\sqrt{(-1)},$$

representing by $L(a)$ the arithmetical logarithm⁸⁹ of a .

Mr. Graves, by a different and very circuitous process, arrives at the result

$$\log(a) = \frac{L(a) + 2r\pi\sqrt{(-1)}}{1 + 2r'\pi\sqrt{(-1)}},$$

the logarithms being taken with respect to the base ϵ for simplicity.

The correctness of this result is doubted by Professors Peacock and De Morgan, but it is corroborated by the researches of Sir W. Hamilton and Mr. Warren, as well as of M. Ohm. It is therefore both an interesting and an important question to determine which is the correct result or at least to point out the cause of the differences between them.

Bien qu'il s'agisse ici d'un problème dans lequel le logarithme semble jouer un rôle central, c'est véritablement un travail général sur l'algèbre symbolique qui est en jeu. En particulier Gregory met en avant ce qu'il appelle des « anomalies analytiques »⁹⁰ (*analytical anomalies*) dont il attribue l'apparition à une conception ambiguë du signe $+$ en arithmétique. La conclusion de l'article, qui ne reprend pas un certain nombre d'observations fines que Gregory va développer, établit tout de même le rôle central du signe $+$ dans le problème du logarithme ou des puissances⁹¹ :

In conclusion, I will recapitulate the conclusions to which I have been led by this mode of considering the symbol $+$.

1. A simple distributive and commutative operation has only one root, but if it be compounded with $+$ it has a plurality of roots depending on the indeterminate nature of $+$.
2. If the base of a system of logarithms and the number be simple distributive and commutative operations, there is only one corresponding logarithm; but if the number of the form be $+^r y$, there is an infinite number of logarithms.
3. If the base of the system be of the form $+^r a$, we are only allowed to assign one value to r , (as otherwise we alter the system,) and then there will be no plurality of logarithms.
4. The impossible parts of the logarithms, as usually given, are the logarithm of $+$ and of $-$.

⁸⁹Lorsque a est un nombre positif $L(a)$ désigne la valeur arithmétique de x telle que $e^x = a$. Chaque nombre positif possède donc ce que Gregory nomme ici un « logarithme arithmétique ». (cf. [Pea30, Art.684, p.541])

⁹⁰[Gre39, MW : p.124]

⁹¹[Gre39, MW : p.135–136]

LE RÔLE DU SIGNE $+$. Dans un premier temps Gregory envisage le problème de la pluralité des racines d'un nombre a : alors que la puissance associe à un nombre un résultat unique, l'opération inverse semble donner de façon inexplicable plusieurs valeurs. Gregory va interpréter cette *anomalie* comme conséquence d'une mécompréhension de ce qui est réellement en jeu dans l'analyse des puissances et du signe $+$ ⁹² :

Much of the obscurity connected with the subject is due to an oversight, by which the existence of this $+$ is totally overlooked. For it is not a , but $+a$, which has a plurality of roots: and though these quantities are usually reckoned to be the same, this idea is founded on an illegitimate extension of a supposed relation in the science of numbers.

Ainsi Gregory va-t-il distinguer a et $+a$ en précisant comment ils doivent être conçus de façon non ambiguë en examinant plusieurs aspects qui permettent d'éclairer à ses yeux le problème.

Tout d'abord il va interpréter les deux éléments comme des opérations distinctes de l'arithmétique l'une étant *absolue* et l'autre *relative*⁹³ :

I say *supposed* [as above], because I hold, that even in Arithmetic a and $+a$ are different, and ought not to be confounded —the former being an absolute operation, the other always a relative one, and consequently incapable of existing by itself. But however this may be, there is no doubt that it is entirely illegitimate to suppose that in all cases a and $+a$ are the same, since generally we know not even what their meanings may be.

Cette citation est l'occasion de préciser la conception de Gregory. En effet, il ne faut pas penser qu'il cherche ici à distinguer deux types de signes $+$, l'un faisant référence à l'opération d'*affectation* alors que l'autre désignerait une opération différente. Il y a bien deux opérations, l'une notée simplement $+$ qui suit les lois de l'*affectation* que nous avons rappelées, et l'autre notée $+a$ que Gregory analyse comme une opération composée et qu'il qualifie de « relative ». Gregory utilisera parfois des notations auxiliaires pour désigner et distinguer plus facilement ces deux types d'opérations (nous verrons de nouveaux exemples par la suite.)

Poursuivant son analyse, Gregory insiste qu'en géométrie on distingue déjà a et $+a$, le premier symbole désignant un segment de droite alors que $+a$ désigne un segment orienté⁹⁴.

Une fois cette distinction entre $+a$ et a établie quant à leur nature dans plusieurs domaines, Gregory précise que les deux opérations suivent des lois différentes, ce qui, du point de vue de l'algèbre symbolique, impose de les distinguer. Gregory peut donc revenir sur l'« anomalie » des racines : a possède une unique racine en arithmétique où l'on a des quantités (positives) alors que les lois qui régissent $+$ et $-$ donnent $(+)^2 = +$ et $(-)^2 = +$ d'où il découle

⁹²[Gre39, MW : p.125]

⁹³[Gre39, MW : p.125]

⁹⁴Gregory écrit : « a represents a line considered with reference to magnitude only, $+a$ with reference both to magnitude and direction. » [Gre39, MW : p.125]

que $+$ possède deux racines. Il en est de même pour $+a$ car $(+a^{\frac{1}{2}})^2 = +a$ tout comme $(-a^{\frac{1}{2}})^2 = +a$ ⁹⁵ :

From this we see, that the reason why there is a plurality of values for the roots of a quantity, is to be found in the nature of the operation $+$; and that it is only the compound operation $+a$, which admits of this plurality, a itself having only one value for each root.

Ce n'est qu'après avoir ainsi clarifié les différents emplois du signe $+$ que Gregory aborde le problème du logarithme en explorant les différentes approches de ses prédécesseurs. Cette analyse va être pour lui l'occasion de préciser sa conception de l'algèbre symbolique et de lever les anomalies qui faisaient obstacle à l'établissement d'une science. Les problèmes qu'envisage Gregory vont lui permettre entre autres de traiter un certain nombre de situations qui relèvent pour nous de ce que nous avons déjà qualifié de *multiplicité*. L'analyse qui suit permet de préciser cet aspect essentiel du travail de Gregory et de mieux cerner la portée de la vision générale qu'il met en place.

MULTIPLICITÉ. Pour commencer Gregory relit les usages du logarithme en termes d'opérations et donne les lois qui le gouvernent⁹⁶ :

It is necessary first to lay down distinctly what is the meaning of the operation denoted by \log ; and this, according to my system, is done by defining its laws of combination. These are

$$\begin{aligned}\log x + \log y &= \log(xy), \\ \log(x^y) &= y \log x,\end{aligned}$$

where x and y are distributive and commutative operations,

$$\log a = 1,$$

which assumes the species to be that in which the base is a .

Tout va alors se jouer dans le rapport qu'entretiennent les opérations \log et $+$ et donner lieu à une analyse double dans le champ arithmétique d'une part et en algèbre symbolique d'autre part.

Gregory commence par analyser la démarche d'Euler⁹⁷ :

He substitutes for the number y the expression

$$\{\cos 2r\pi + \sqrt{(-1)} \sin 2r\pi\}y,$$

r being any integer which he considers to be equivalent to it; and then taking the logarithms with respect to ϵ , he says that

$$\log y = L(y) + \log\{\cos 2r\pi + \sqrt{(-1)} \sin 2r\pi\},$$

⁹⁵[Gre39, MW : p.127]

⁹⁶[Gre39, MW : p.129]

⁹⁷[Gre39, MW : p.130]

where $L(y)$ represents the arithmetical logarithm of y : and as

$$\cos 2r\pi + \sqrt{(-1)} \sin 2r\pi = e^{2r\pi\sqrt{(-1)}},$$

we have

$$\log y = L(y) + 2r\pi\sqrt{(-1)}.$$

Si le résultat est bien obtenu en suivant les règles de l'algèbre symbolique, Gregory souligne que certaines de ces expressions ne peuvent pas être interprétées en arithmétique. Dans la présentation de Gregory ci-dessus, chaque choix de l'entier r donne un logarithme, et seul $r = 0$ donne un logarithme « possible », c'est-à-dire un nombre (*quantity*). La partie qui apparaît lorsque r est non nul est « impossible » à interpréter en arithmétique, et l'expression résultant du calcul reste du côté de l'algèbre symbolique.

Mais ce qui fait que Gregory remet en question cette démonstration relève essentiellement de l'interprétation de l'égalité entre y et $e^{2r\pi\sqrt{(-1)}}y$ qui ne peut pas se transférer à l'algèbre symbolique sans poser problème⁹⁸ :

The correctness of this result depends essentially on the assumption that y and $\{\cos 2r\pi + \sqrt{(-1)} \sin 2r\pi\}y$ are identical: an assumption which at first it seems very natural to make, since the expression $\cos 2r\pi + \sqrt{(-1)} \sin 2r\pi$ is usually considered to be equal to unity. But if we suppose the quantities with which we are dealing to be general quantities, and not numbers merely, a numerical value of $\cos 2r\pi + \sqrt{(-1)} \sin 2r\pi$ can have no place in our investigation, and we must seek for its general algebraical meaning.

Pour contourner cet obstacle Gregory explique qu'au niveau des opérations ce sont $+$ et $e^{2\pi\sqrt{-1}}$ qui sont équivalents, car les deux suivent les mêmes lois. L'opération désignée par $e^{2\pi\sqrt{-1}}$, tout comme celle désignée par $+$, doit être traitée avec la *multiplicité* que traduit l'égalité $+ = +^r$, contrairement aux symboles tels que 1, propres à l'arithmétique, qui comportent une certaine unicité de conception liée à la notion de *quantité*.

Pour Gregory il faut donc distinguer deux situations. Au niveau des quantités (positives), l'équation exprimée avec les symboles de l'algèbre arithmétique $y = e^x$ permet de dire que y possède un logarithme arithmétique (*arithmetical logarithm*) noté $L(y) = x$. Par contraste au niveau de l'algèbre symbolique il faut distinguer y et $+y$, le premier désignant un certain type d'opération « absolue » alors que le second désigne une opération composée de $+$ et de y , et qui par conséquent possède ce que nous avons nommé une certaine *multiplicité*. Le mélange de ces deux niveaux d'interprétation est pour Gregory la cause de l'erreur qu'il décèle dans la démarche d'Euler et de plusieurs autres mathématiciens.

En réintroduisant subtilement le signe $+$ Gregory montre qu'il n'existe pas de logarithme commun aux nombres négatifs et positifs⁹⁹. Il repense ici les

⁹⁸[Gre39, MW : p.130]

⁹⁹Bien que l'on observe des choix de formalismes distincts, la discussion menée par Gregory est ici organisée de façon très proche de celle proposée par Peacock dans son ouvrage de 1830. (cf. [Pea30, p.568–572])

nombre de l'arithmétique en des termes adaptés à la généralité qu'impose l'algèbre symbolique. Dans le cas général Gregory distingue les nombres positifs, qui doivent s'écrire $+^r a$, et les nombres négatifs, qui doivent s'écrire $+^{\frac{2r+1}{2}} a$. Les premiers ont donc pour logarithme $\log a + \log(+^r) = L(a) + 2r\pi\sqrt{-1}$ alors que les autres ont pour logarithme $\log a + \log(+^{\frac{2r+1}{2}}) = L(a) + \frac{2r+1}{2}\pi\sqrt{-1}$. Dans ces expressions Gregory estime que l'on ne peut ni prendre $r = 0$ dans le premier cas, ni $r = -\frac{1}{2}$ dans le second sinon on confond encore $+a$ et a .

Cette analyse a une double conséquence : d'abord les logarithmes des nombres positifs ou négatifs ne coïncident jamais, et de plus chacune de ces expressions possède une partie « impossible » que l'on ne peut pas faire disparaître.

Gregory insiste sur le fait qu'il s'agit d'une différence majeure avec les visions de ses prédécesseurs : contrairement à Vincent ou Graves, il pense que le résultat symbolique avec toute sa multiplicité (que nous lisons à travers l'indétermination de r) est le théorème qu'il faut conserver et qu'il n'a rien d'« imaginaire »¹⁰⁰ :

It might, perhaps, have weakened [M. Vincent's] belief in the correctness of the result, if he had come to the conclusion, as he ought to have done, that the same logarithm correspond to positive, negative, and impossible quantities. These last he seems quite to have overlooked, which may have arisen from his having adopted, with many other mathematicians, the name of *imaginary* quantities.

Dans une dernière partie de son article, Gregory revient sur une démarche proposée par Peacock sur le logarithme des quantités négatives. Ce passage permet de préciser les idées de Gregory et de les contraster avec la vision de son prédécesseur¹⁰¹ :

It is somewhat remarkable, that Mr. Peacock has been led into the same error as M. Vincent and Mr. Graves, respecting the coincidence in some cases of the logarithms of positive and negative quantities. As the cause of his error has reference to the remark which I have made, and is not very easy to be detected, I shall point it out more particularly.

He considers $-a^m$ to be equivalent to $-1(+a)^m$, which gives

$$\begin{aligned} -(a)^m &= \log(-1) + \log(+a)^m \\ &= (2r + 2mr' + 1)\pi\sqrt{-1} + \log a. \end{aligned}$$

He then supposes $m = \frac{p}{2n}$ where p is prime to n , $r' = -n$ and $r = \frac{p-1}{2}$; and as these values make the multiplier of $\pi\sqrt{-1}$ vanish, he concludes that the logarithm of $-(a)^m$ coincides with that of a^m , since it becomes $m \log a$.

¹⁰⁰[Gre39, MW : p.133]

¹⁰¹[Gre39, MW : p.134]

L'approche de Peacock met elle aussi en œuvre une démarche symbolique abstraite dont le formalisme est souvent proche de celui choisi par Gregory. Les différences se situent au niveau de l'utilisation des notations, et de ce qu'elles désignent pour chacun d'eux¹⁰². Les deux mathématiciens utilisent à leur manière des symboles particuliers pour indiquer une certaine *multiplicité* qui n'apparaît pas au niveau des quantités telles qu'elles sont traitées en arithmétique.

Pour Peacock cette *multiplicité* s'écrit à travers des égalités du type¹⁰³ $1 = e^{2r\pi\sqrt{-1}}$, $-1 = e^{(2r+1)\pi\sqrt{-1}}$, $\sqrt{-1} = e^{\frac{2r+1}{2}\pi\sqrt{-1}}$ ou encore $+a = e^{2r'\pi\sqrt{-1}}a = a$. La vision de Peacock repose principalement sur l'utilisation des racines de l'unité comme élément d'*affectation* des quantités et sur leur cyclicité qui engendre des écritures multiples. Les égalités font intervenir ce qu'il appelle des éléments « imaginaires » qui permettent de mener à bien certains calculs tels qu'on les voit dans la citation précédente. Un des aspects que critique Gregory provient de la différence de nature de ce qui est désigné par le membre de droite ou de gauche des égalités, l'un appartenant à l'algèbre symbolique alors que l'autre appartient parfois à l'arithmétique et demande un effort constant de réinterprétation pour le concevoir comme symbole d'opération.

De son côté Gregory préfère éviter tout recours aux symboles désignant des quantités et insiste sur son idée d'employer le symbole $+$ et ses puissances comme seule marque d'*affectation* positive ou négative des quantités. Désormais lorsque Gregory écrit $+ = +^r$ les deux membres de l'égalité sont de la même nature, des *opérations*, et toutes les manipulations s'effectuent selon les lois de l'algèbre symbolique.

Ainsi les choix de Peacock ou de Gregory de faire porter la marque de la *multiplicité* par les symboles 1 ou $+$ respectivement, qui pourraient passer pour anodins rétrospectivement, révèlent une différence de conception de l'algèbre symbolique et de son lien avec l'arithmétique. Il y a bien une pratique commune qui consiste en une manipulation des symboles d'opération suivant des lois que l'un et l'autre reconnaissent comme valides. Mais pour Peacock un certain nombre d'écritures restent dans le domaine de l'*imaginaire* alors que Gregory rejette cette idée. On voit bien tout au long de son traité de 1830 que Peacock cherche à tout prix à éviter certaines situations en sélectionnant des cas qui lui permettent d'interpréter les résultats obtenus en termes de nombres ou de quantités. Ceci l'amène à écrire une erreur que décèle Gregory dans la citation au-dessus où Peacock oublie tout d'un coup que lorsque l'on écrit $1 = e^{2\pi r}$ on a introduit, en passant du membre de gauche à celui de droite, une certaine *multiplicité*. Dès lors chaque ligne de calcul doit rester du côté de l'algèbre symbolique et $-a^{\frac{1}{2}}$ ne fait plus univoquement référence à un seul nombre, mais à deux, l'un positif, l'autre négatif.

Gregory de son côté marque, par son choix de symboles, une nette séparation entre l'arithmétique et l'algèbre symbolique en prenant soin de ne pas écrire sans précaution d'égalité dont les membres appartiennent aux deux champs

¹⁰²Une analyse comparée des conceptions de Peacock et Gregory dépasserait largement le cadre de notre étude. Nous dégageons quelques éléments qui éclairent l'approche de Gregory par contraste avec les choix opérés par Peacock.

¹⁰³cf. [Pea30, p.568 §716] par exemple.

qu'il dissocie. Par conséquent le symbole 1 ne peut pas être conservé en algèbre symbolique car son utilisation en arithmétique lui confère un « sens limité » (*a symbol limited in its signification*). On voit donc ici que c'est un principe de généralité qui guide l'abstraction chez Gregory et qui empêche d'utiliser certains symboles en algèbre symbolique, car pour être général, un symbole ne doit pas être lié à un sens donné dans une discipline particulière.

IMPOSSIBLE MAIS PAS IMAGINAIRE. Cet article sur les logarithmes illustre bien la double difficulté que nous souhaitons mettre en valeur et permet pour nous de faire écho à l'emploi du mot « impossible » dans le titre choisi par Gregory. Pour lui, c'est au niveau symbolique que doit s'établir la « preuve » (*proof*) et il rejette le qualificatif d'*imaginaire* employé par Peacock et d'autres pour mettre en avant la possibilité ou l'impossibilité¹⁰⁴ de donner une interprétation de telle ou telle expression dans un champ particulier¹⁰⁵ :

I adhere to the name *impossible* instead of *imaginary*, because the latter involves an idea which I conceive to be very deleterious in analysis. We may be unable to perform an operation though it be by no means an *imaginary* one; and indeed all that we can say of those quantities which have this name affixed to them is, that they are *uninterpretable in arithmetic*. For this reason, if I were permitted to propose a change, I should prefer to call these quantities "operations uninterpretable in arithmetic", as this involves no theory of their nature, but only expresses what is a fact.

Pourtant les résultats obtenus, qui ne sont en rien « imaginaires » pour Gregory, peuvent parfois ne pas être interprétables en arithmétique, première impossibilité qu'il commente souvent dans ses articles sur l'algèbre symbolique et que nous avons illustrée. Son emploi du mot « impossible » est alors essentiellement associé à l'apparition, lors de l'approche symbolique, des termes comportant $\sqrt{-1}$. Lorsque ces termes ne disparaissent pas en fin de calcul¹⁰⁶ ou dans des cas particuliers, il est alors « impossible » pour Gregory de donner une interprétation au niveau de l'arithmétique de la formule symbolique. La deuxième difficulté provient de l'interprétation même des situations de l'arithmétique en termes d'opérations. Comme nous l'avons vu en suivant l'analyse de Gregory, certains symboles qui disparaissent au cours des manipulations et des calculs en arithmétique doivent être réintroduits pour se prêter à une interprétation en termes d'opérations, comme c'est le cas pour le signe + (d'affectation) qui est au centre de l'analyse proposée par Gregory. Des égalités, et plus généralement les formulations de l'arithmétique,

¹⁰⁴Peacock emploie aussi les termes « impossible » et « imaginaire » en parlant d'« opération impossible » et de « résultat imaginaire ». (cf. [Pea30, p.354 §430] par exemple.)

¹⁰⁵[Gre39, MW : p.133]

¹⁰⁶Le traitement général des symboles d'affectation dépasse le cadre de cette thèse. Notons cependant que Gregory n'emploie le terme d'affectation que pour les symboles + et -. De son côté, et en particulier dans son traité de 1830, Peacock introduit le symbole *i* pour désigner une certaine affectation que + et - ne peuvent dénoter. Il détermine ensuite que *i* doit avoir la propriété $i^2 = -1$ et propose d'utiliser $\sqrt{-1}$ comme signe d'affectation au même titre que + et -. (cf. [Pea30, p.351-352])

ne peuvent pas être directement retranscrites en algèbre symbolique sans un travail minutieux de réinterprétation et de *transposition*. On voit là une situation bien différente du cas de l'intégration que nous avons mentionné précédemment : un résultat établi en intégration (apparition d'une constante) pouvait être simplement *retranscrit* dans le formalisme de l'algèbre symbolique. Dans les exemples issus de l'arithmétique nous avons vu que Gregory mettait en avant une méthode d'interprétation particulière qui justifiait le passage à l'algèbre symbolique. Cette démarche qu'il applique dans de nombreuses situations différencie fondamentalement sa pratique de celle de la *séparation des symboles d'opération* à l'œuvre antérieurement.

GÉNÉRALITÉ. Pour finir cette section, précisons la nature de la relation entre l'algèbre symbolique et l'arithmétique. Pour l'instant, nous avons vu que Gregory considérait l'algèbre symbolique comme plus générale que l'arithmétique car elle permettait de manipuler des opérations abstraites (les symboles y sont dénués de sens) et de donner des résultats interprétables dans divers champs, en arithmétique mais aussi en géométrie par exemple. Pourtant, dans cette hiérarchie du général, l'arithmétique ne s'est pas fondue dans le formalisme de l'algèbre symbolique. Le champ arithmétique reste un lieu de l'interprétation toujours présent, un lieu où les symboles reprennent leur sens sans disparaître dans le domaine plus général de l'algèbre symbolique. Dans cette organisation, l'arithmétique ou les divers champs particuliers ne sont pas constitués comme une collection d'énoncés de l'algèbre symbolique, et ne forment pas un sous-domaine qui serait défini par l'utilisation exclusive d'une collection restreinte de symboles ou de lois.

Gregory construit tout au long de son travail de recherche une véritable méthode qui permet un va-et-vient entre l'algèbre symbolique et les divers champs d'interprétation. Son approche ne se limite pas à préciser les lois permettant de manipuler les symboles et nous avons établi que l'algèbre symbolique était générale suivant deux axes. D'une part elle permet de relire sans ambiguïté les diverses situations des champs qu'elle embrasse —en premier lieu l'arithmétique. Et d'autre part elle assure la possibilité d'exprimer ses résultats dans le vocabulaire et le formalisme des champs initiaux car les lois sur lesquelles elle s'appuie sont valables dans toutes les familles d'opérations qui suivent les mêmes lois.

2.6 OPÉRATION ET SUJET DE L'OPÉRATION

Nous avons vu dans la section précédente comment Gregory distinguait plusieurs emplois du symbole $+$ en arithmétique. En introduisant une notation préfixée (terme d'observateur) et en désignant par $A_x(a)$ l'addition de x au nombre a , ou plus précisément l'opération *relative* qui revient à ajouter x , Gregory fait apparaître les lois qui régissent cette opération de façon particulière¹⁰⁷ :

¹⁰⁷[Gre42, MW : p.240]

The distributive law, which is met with so frequently in algebraical operations, and which is usually written

$$f(a + x) = f(a) + f(x)$$

is in this notation expressed by the equation

$$f[A_x(a)] = A_{f(x)}[f(a)].$$

In like manner the index law, which is usually written

$$f^m f^n(a) = f^{m+n}(a),$$

becomes in this notation

$$f^m f^n(a) = f^{A_n(m)}(a).$$

Ici les lois sont exprimées pour les deux symboles f et A_m qui n'ont pas de sens *a priori* en arithmétique. De plus, l'ambiguïté soulignée en arithmétique, où le signe $+$ servait à indiquer deux types d'opérations, disparaît : les symboles sont « algébriquement définis » (*defined algebraically*) par l'énoncé des lois qui les gouvernent en algèbre symbolique. Pour Gregory A_m représente l'opération d'addition du nombre m telle qu'elle est utilisée en arithmétique car les mêmes lois gouvernent les opérations aussi bien en algèbre symbolique qu'en arithmétique. De son côté le symbole f utilisé pour exprimer ces lois reste purement indéterminé, ce qui illustre l'idée de Gregory selon laquelle « il y a une algèbre symbolique générale distincte de l'algèbre arithmétique »¹⁰⁸ et donne aux énoncés une certaine généralité.

Enfin, les deux symboles f et A_m n'ont pas exactement le même statut dans ces lignes. En arithmétique dans un calcul du type $a + x$, où $+$ indique l'addition, a et x représentent des nombres et ont des rôles similaires. Dès qu'on analyse la même expression en y voyant l'opération qui consiste à additionner a , les deux symboles a et x n'ont pas la même fonction. Ici c'est dans l'expression des lois de l'algèbre symbolique que les symboles f et A_m se voient attribué des rôles distincts, une pratique courante dans l'œuvre de Gregory et qui mérite d'être analysée plus avant.

On retrouve une différenciation des rôles joués par des symboles dans une réinterprétation de certaines *équations fonctionnelles* que l'on peut lire dans un article intitulé « On the Solution of Certain Functional Equations ». Gregory y est sensible à une approche proposée par Leslie Ellis dans un article qu'il a publié dans le *Cambridge Mathematical Journal* (vol. III p.92)¹⁰⁹ :

[Ellis's] idea is, that these classes of operations are distinguished by the *order*¹¹⁰ in which the operations are performed, so that, whereas in our ordinary equations the known operations are performed on those which are unknown, in functional equations the converse is the case, the unknown operations being performed on those which are known.

¹⁰⁸[Gre42, MW : p.236]

¹⁰⁹[Gre43, MW : p.247]

¹¹⁰En italique dans le texte de Gregory.

Ainsi, dans ces « équations fonctionnelles », interprétées en termes d'opérations, l'ordre d'écriture qui indique l'ordre dans lequel les opérations sont effectuées ne désigne pas l'élément à droite, *objet* de l'opération, comme inconnue. Poursuivant dans cette voie, Gregory distingue deux « natures de difficultés » rencontrées (*difficulty wholly different in kind*¹¹¹) et propose de prendre comme exemple des équations fonctionnelles classiques.

Pour illustrer le premier type de difficulté, Gregory rappelle l'exemple suivant : lorsque l'on écrit $\frac{dy}{dx} - ay = 0$, l'inconnue se trouve à droite : c'est la fonction y . Grâce à la distributivité telle que Gregory la conçoit et en reprenant une notation classique à l'époque pour la fonction inconnue ($\phi(x)$, avec la variable attachée à l'écriture de la fonction), on peut comme Gregory réécrire l'équation $\left(\frac{d}{dx} - a\right)\phi(x) = 0$ et une inversion de l'opération entre parenthèses donne une façon d'écrire la solution :

$$\phi(x) = \left(\frac{d}{dx} - a\right)^{-1} 0 = Ce^{ax}.$$

Pour faire apparaître le second type de difficulté, Gregory contraste son premier exemple avec l'analyse d'une équation du type $\phi(x+1) - a\phi(x) = 0$ où l'inconnue est ϕ : celle-ci ne permet pas d'après lui d'utiliser la distributivité sans nuire à la généralité de l'énoncé et des résultats obtenus. C'est dans les techniques de calcul des différences que Gregory va alors trouver les idées nécessaires pour analyser ce type d'équations fonctionnelles¹¹² :

The equation

$$u_{x+1} - au_x = 0$$

might be written

$$(1) \quad \phi(x+1) - a\phi(x) = 0,$$

in which the form of ϕ is to be determined.

Here the known operations are the subjects of the unknown, and we cannot directly disengage them; but we are enabled to do so by transforming the equation into one in which the unknown operation ϕ is the subject. For, assuming the operation D to be such that

$$D\phi(x) = \phi(x+1),$$

we are able to investigate the laws of combination of this new symbol and its various properties, so as to make it a known operation. The equation then becomes

$$D\phi(x) - a\phi(x) = 0.$$

Now we can shew that D is a distributive symbol with respect to its subject, and that it is commutative with respect to a ; we may therefore write the equation in the form

$$(2) \quad (D - a)\phi(x) = 0,$$

$$\phi(x) = (D - a)^{-1}0.$$

¹¹¹Le mot « kind » est en italiques dans le texte de Gregory.

¹¹²[Gre43, MW : p.249]

Au-delà de l'astucieuse transformation rendue possible grâce à l'analyse en termes d'opérations, il est intéressant d'observer précisément la nature des éléments manipulés par Gregory et de suivre les rôles affectés à chacun d'entre eux.

Gregory n'analyse pas systématiquement le rapport entre une fonction et sa variable du point de vue de l'algèbre symbolique dans ses articles et il faut donc lire ces lignes avec prudence. Il serait en effet hasardeux de penser dans (1) au-dessus que x ou $x + 1$ sont des symboles d'opération, et ceci malgré les exemples que nous avons rappelés plus haut sur l'utilisation du signe $+$ pour dénoter diverses opérations, ou la capacité de Gregory à analyser ces écritures en termes d'opérations dans certaines situations. Ici $\phi(x)$ désigne une fonction de la variable x (à nouveau la notation usuelle à l'époque de Gregory où la variable est accolée à la lettre qui sert à nommer la fonction) et on peut penser, en restant fidèle aux conceptions de l'époque, à $\phi(x + 1)$ comme une fonction dont la variable est $x + 1$.

Le commentaire donné par Gregory est essentiel pour comprendre la conception qui est en jeu dans l'égalité (2). En effet celui-ci précise que « l'opération inconnue ϕ est le sujet », formulation dans laquelle toute trace de x a disparu et qui contraste avec l'écriture des équations. Il faut donc distinguer deux niveaux d'analyse pour comprendre ces équations.

Au niveau symbolique x disparaît, donc avec lui tout référence au lien qu'entretiennent ϕ et x , qu'il soit lien entre la fonction et sa variable ou lien entre *symbole-agissant* et *symbole-objet*. On voit donc apparaître une façon de traiter les fonctions de manière abstraite, purement symbolique et dénuée du sens usuellement donné au mot « fonction » à son époque. Ici la fonction ϕ est conçue comme une opération non pas dans sa relation avec x mais dans son interaction via certaines lois de l'algèbre symbolique avec d'autres opérations. Dans cette analyse ϕ se voit attribuer le double rôle d'une part d'objet de l'opération $D - a$ et d'autre part d'inconnue de l'équation. Du point de vue calculatoire, et toujours au niveau symbolique, ces deux opérations sont soumises aux lois de distributivité qui assurent que $D\phi - a\phi = (D - a)\phi$.

Au niveau du champ particulier des équations fonctionnelles (l'écriture de la fonction avec sa variable accolée $\phi(x)$ témoigne du retour au champ fonctionnel) l'équation garde la forme imposée par l'analyse dans les termes de l'algèbre symbolique. Si les divers éléments retrouvent le sens qui leur est habituellement attribué dans une équation fonctionnelle, la lecture de l'équation est maintenant radicalement différente si on la compare avec la ligne (1). Dans la conception de la fonction notée $\phi(x)$ la nature de la liaison entre la fonction et sa variable reste quelconque et indéterminée. Par contre c'est désormais l'opération $D - a$ dont la nature importe ici : en effet, pour résoudre le problème Gregory précise qu'il faut savoir inverser cette opération, et la solution effective $\phi(x)$ dépend essentiellement de l'identification et du calcul de $(D - a)^{-1}$ en tant qu'opération agissant sur les fonctions.

2.7 CONCLUSION

Les dernières idées que nous venons d'évoquer offrent une place particulière à certaines opérations qui agissent sur des fonctions et elles font directement écho pour nous à la vision de l'analyse fonctionnelle que tente de développer Pincherle en 1897 dans son *Mémoire sur le calcul fonctionnel distributif*. Dans l'introduction il y insiste sur le point de vue qu'il adopte et qu'il rattache directement aux travaux des Algébristes Anglais¹¹³ :

Dans le présent Mémoire c'est la nature particulière des opérations qu'on a à exécuter sur les fonctions variables qui a ici le plus d'importance. [...] Nous remarquons sans peine, dès le début, que les opérations fonctionnelles nous offrent une très grande variété : pour ne citer que les plus communes, nous trouvons parmi elles, en outre de toutes les opérations arithmétiques, celle de dérivation, celle de différentiation finie, de substitution, etc., et celles qu'on obtient par la réitération, la composition et l'inversion de celles que nous venons d'énumérer.

Il y a différentes lectures possibles à partir de l'algèbre symbolique telle que nous venons de l'évoquer à travers l'approche de Gregory.

Pour Pincherle¹¹⁴ en 1897, l'algèbre symbolique permet de concevoir une « opération fonctionnelle » à partir d'opérations « simples » (en particulier la dérivation) que l'on compose suivant certaines lois. La « nature particulière des opérations » est ici liée aux types de compositions qui entrent en jeu dans leur conception ainsi qu'à leurs propriétés *algébriques*. Pincherle met en avant, dès les premières pages de son texte, la propriété de distributivité qui sera centrale dans son approche et qui répond pour lui à un critère de simplicité.

Mais il nous semble aussi que les efforts des Algébristes Anglais ne relèvent pas simplement de la capacité à concevoir les symboles d'opération suivant les lois qui les gouvernent, c'est-à-dire comme des éléments algébriques.

La nature indifférenciée des symboles que Gregory maintient au premier plan permet de faire évoluer les rôles que l'on attribue aux différents éléments d'une expression algébrique abstraite. Nous voyons là une idée essentielle qui permettra d'éclairer les efforts multiples au tournant du siècle pour définir le champ nouveau de l'analyse fonctionnelle dans lequel des opérations, qui prendront finalement la forme moderne d'opérateurs, viendront jouer un rôle central.

Nous verrons dans la suite comment une conception symbolique et algébrique a permis à la fois une approche calculatoire mais aussi diverses façons abstraites de concevoir les opérations fonctionnelles, en réinterprétant les rôles que peuvent jouer des éléments qui *a priori* semblaient se présenter sous une

¹¹³[Pin97, p.326]

¹¹⁴Les remarques que nous faisons ici ne résument pas l'œuvre de Pincherle, ni la teneur du texte de 1897 qui nécessiterait une analyse beaucoup plus étendue. Comme dans toute notre thèse, les extraits choisis illustrent la façon dont un mathématicien met en œuvre certains concepts dans un cadre donné. Il y a bien sûr des évolutions, qui ne ressortiront pas toujours lorsque nous nous appuyerons sur un texte convoqué pour mettre en lumière tel ou tel aspect spécifique dans l'élaboration de l'analyse fonctionnelle ou pour contraster localement deux façons de procéder.

forme unique et bien définie : une fonction et sa variable, l'intégrale ou la dérivée d'une fonction, etc.

Volterra et la notion de fonction comme paradigme général pour l'analyse fonctionnelle

Pour écrire un traité (comme pour professer un cours) ne suffit pas la connaissance scientifique du sujet, mais il faut aussi un certain sens artistique. Les livres qui ont traversé les siècles sont des œuvres d'art autant que des œuvres de sciences. EUCLIDE qui a systématisé et unifié la géométrie créée par ses devanciers, était un savant, mais il possédait le même goût artistique qui a rendu célèbres HOMÈRE et PHIDIAS.¹

— V. Volterra

3.1 INTRODUCTION

Volterra a trouvé une méthode qu'il a appliquée avec succès tout au long de sa carrière et qu'il désignera souvent comme la méthode *du passage du fini à l'infini* ou *du discret au continu*.

Probablement ces deux figures, choisies pour exprimer une certaine approche générale, ont-elles été perçues parfois comme caricaturales ou par trop éthérées, et de ce fait les concepts développés par Volterra, bien qu'ingénieux, ne figurent pas en bonne place dans la liste des idées qui ont définitivement marqué l'analyse fonctionnelle. On est tenté, au vu du formalisme que nous

¹Ces quelques mots ont été prononcés en guise de conclusion par Volterra au Congrès international des mathématiciens tenu à Strasbourg en 1920. Il y donnait alors une conférence intitulée « Sur l'enseignement de la physique mathématique et de quelques points d'analyse ». (cf. [Vol60, p.385])

avons adopté de nos jours, de penser que les conceptions de Fréchet ou de Banach par exemple furent plus efficaces et de meilleurs choix.

De fait, bien que Volterra ait laissé son nom à des problèmes que l'on étudie encore aujourd'hui —essentiellement des équations intégrales—, le formalisme qui fût le sien n'a pas été beaucoup employé en dehors de ses propres publications.

Pourtant on ne peut pas apprécier le rôle de Volterra dans la conception moderne de l'analyse fonctionnelle et l'importance de ses idées en arguant seulement que d'autres conceptions ont été reprises *a posteriori* pour développer l'analyse fonctionnelle telle que nous la connaissons.

De nombreux travaux ont d'ailleurs démontré la fertilité et le rôle des idées de Volterra dans de nombreux registres des mathématiques.

On peut tout d'abord consulter l'ouvrage d'Angelo Guerraggio et Giovanni Paoloni intitulé *Vito Volterra*² qui dresse un panorama très large et très documenté de la vie et de l'œuvre de Volterra. Dans ce livre on trouve en outre les références aux études principales qui permettent de resituer et d'apprécier l'ampleur du travail du mathématicien italien qui est l'auteur d'environ trois cents publications.

De nombreuses études analysent certains aspects des travaux de Volterra suivant un angle spécifique : ils apportent un éclairage important sur divers aspects de son œuvre et permettent de comprendre les problèmes qu'il étudie en les connectant à différentes trames historiques.

L'article de Monna « Volterra et les fonctions de lignes : un centenaire »³ permet de comprendre certains aspects liés à l'idée du *passage du fini à l'infini* et au rôle des fonctions, thèmes que notre travail cherche à aborder selon un angle nouveau.

L'article de Jeremy J. Gray « Mathematics and natural science in the nineteenth century: the classical approaches of Poincaré, Volterra and Hadamard »⁴ permet de comprendre le lien profond qui lie les mathématiques avec la physique. L'analyse de Gray étudie précisément le type de questions dont se saisissent les trois savants Poincaré, Volterra Hadamard à la frontière mouvante entre mathématiques et physique. Nous analyserons à notre tour quelques conceptions dont le développement s'appuie sur une approche propre à la physique mathématique.

Le livre *Hidden harmony—geometric fantasies*⁵ de Umberto Bottazzini et Jeremy J. Gray est essentiel pour comprendre les enjeux liés aux fonctions complexes qui seront au cœur de beaucoup de travaux de Volterra. On y trouve entre autres les liens qu'entretiennent ces fonctions avec les équations différentielles et intégrales, l'approche de Riemann ainsi que celle de Weierstraß qui sont importantes pour comprendre les questions que nous aborderons. La section 7.7.4, intitulée « Functional-Analytic Attempts », établit un lien précis entre les travaux de Volterra que nous allons étudier plus bas et le contexte général développé dans le livre.

²[GP13]

³[Mon82]

⁴[Gra01]

⁵[BG13]

On pourra aussi se référer à l'article de Monna « La théorie des fonctions, réflexions historiques »⁶ pour comprendre diverses notions de fonctions et les contraster avec celles convoquées par Volterra dans notre étude.

En ce qui concerne les équations intégrales en liaison avec celles qu'étudie Volterra et le rôle des mathématiques italiennes vers le début du 20^e siècle on pourra consulter l'article de Rossana Tazzioli « The role of differential parameters in Beltrami's work »⁷ et les travaux de Tazzioli en collaboration avec Tom Archibald « The reception of Fredholm's results on integral equations: preliminary report »⁸ ainsi que « Integral equations between theory and practice: the cases of Italy and France to 1920 »⁹.

L'étude de Jeremy J. D. Gray intitulée « The shaping of the Riesz representation theorem: a chapter in the history of analysis »¹⁰ étudie des thèmes connectés à plusieurs occasions à notre étude et comporte un chapitre dédié aux travaux de Volterra.

Enfin l'article « Mathematicians at war »¹¹ de Laurent Mazliak et Tazzioli permet d'appréhender dans un contexte plus large les conditions de production des mathématiques à l'époque de Volterra et dans un contexte tourmenté.

Nous ne citons là que quelques références qui permettent à nouveau d'appréhender le problème de l'histoire de l'analyse fonctionnelle sous des angles multiples tout à fait incontournables. On trouvera de plus dans ces ouvrages de nombreux compléments bibliographiques.

Du point de vue historique l'idée du *passage du fini à l'infini* permet de mettre au jour un certain type de généralité dans les publications de Volterra. Il nous semble que cette idée permet de relire un certain nombre de travaux en faisant apparaître des mécanismes de sélection, d'abstraction et d'organisation qui sont le support d'une progression vers une vision générale.

Notre étude vise à montrer que le passage *du fini à l'infini* opère principalement chez Volterra à travers une notion de *fonction* qu'il façonne et étend petit à petit. Nous analyserons en particulier comment Volterra cherche à étudier certaines *opérations* qui deviendront des *fonctionnelles* en transposant en plusieurs étapes le modèle de l'analyse des fonctions ordinaires. Nous montrerons que la vision qu'il développe à cette fin s'accompagne d'éléments abstraits divers —que nous qualifierions aujourd'hui de topologiques— dont nous étudierons la portée et le rôle dans la constitution d'une vision générale.

Nous contrasterons ces idées par la suite avec celles choisies par d'autres mathématiciens comme Fréchet ou Banach. En particulier nous montrerons l'importance des travaux de Volterra dans la progression qui permettra à Fréchet de construire sa propre vision et de progresser vers ce qu'il nommera « analyse générale ».

⁶[Mon78]

⁷[Taz97]

⁸[AT05]

⁹[AT14]

¹⁰[Gra84]

¹¹[MT09]

3.2 UN CALCUL DIFFÉRENTIEL DES SUBSTITUTIONS

En 1887 et 1899 Volterra publie à la Société Italienne des Sciences un mémoire en deux parties intitulé « Sui fondamenti della teoria delle equazioni differenziali lineari »¹². Volterra y expose ce qu'il nomme un « calcul infinitésimal » des substitutions qui doit servir de « fondement » à la théorie des équations différentielles linéaires. Plus précisément, et en des termes d'observateur, Volterra développe dans ce texte un véritable calcul différentiel et intégral pour les substitutions, lesquelles sont envisagées sous une forme que nous qualifierions aujourd'hui de matricielle. Volterra introduit dans la première partie de son exposé les substitutions (*sostituzione*) sous la forme de tableaux de la forme $\begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}$ où les coefficients $\alpha, \beta, \gamma, \delta$ seront des fonctions d'une ou plusieurs variables réelles plus ou moins régulières tout au long de la progression de l'exposé.

Volterra ne met pas en avant dans ce texte le lien entre le calcul des substitutions et les équations différentielles linéaires. Il y reviendra dans une publication séparée intitulée « Sulle equazioni differenziale lineari »¹³ publiée cette même année 1887 aux *Rendiconti della Reale Accademia dei Lincei*. L'organisation du discours de son mémoire, qui met en avant l'idée de substitution, accentue le statut particulier de ce texte et souligne le rôle de « fondement » que Volterra entend lui faire jouer.

Dès le départ Volterra indique que tout ce qu'il expose pour des substitutions d'ordre 2 (des matrices 2×2) reste valable pour des substitutions d'ordre n quelconque (fini). Enfin la seconde partie du mémoire sera publiée en 1899 et Volterra y envisagera des substitutions dont les coefficients sont des fonctions de la variable complexe.

Le but de cette section est d'illustrer chez Volterra une première façon d'adapter le « calcul infinitésimal » des fonctions ordinaires¹⁴ aux substitutions. Nous cherchons à mettre en évidence les éléments sur lesquels Volterra s'appuie pour transposer un certain modèle qui a fait ses preuves dans l'étude des fonctions ordinaires. Nous verrons que les éléments sectionnés par le mathématicien italien sont suffisamment adaptables pour lui permettre de développer un calcul infinitésimal des substitutions et qu'ils ont à ses yeux un caractère général qui permet la transposition. En particulier nous monterons le rôle particulier réservé par Volterra à une certaine notion d'*opération* qui apparaît comme un pivot essentiel de la vision qu'il développe dans ce texte.

Enfin nous examinerons le statut de ce texte, que Volterra désigne comme « fondement » de la théorie des équations différentielles linéaires, à la lumière des résultats qu'il expose ici mais aussi en confrontant cette idée de « fondement » avec le type de généralité que nous pouvons y lire.

¹²[Vol54, p.209]

¹³[Vol54, p.291]

¹⁴Nous utiliserons l'expression « fonction ordinaire » pour désigner des fonctions de la variable réelle ou complexe à valeurs dans \mathbb{R} ou \mathbb{C} par contraste avec d'autres utilisations et adaptations du mot « fonction » dans les textes de Volterra.

3.2.1 OPÉRATIONS SUR LES SUBSTITUTIONS

Dans l'introduction de son traité Volterra indique les notations qu'il adopte et met en avant une approche algébrique ayant un caractère abstrait. Il définit en effet une seule opération, la multiplication de deux substitutions, non pas sous sa forme algorithmique qui permet d'effectuer le calcul à partir des coefficients, mais en dégagant avant tout quelques propriétés de l'opération. En particulier Volterra insiste sur la non-commutativité de la multiplication¹⁵ :

Se sopra una variabile qualunque u si eseguiscono successivamente due sostituzioni lineari i cui coefficienti sono rispettivamente $\begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}$, $\begin{pmatrix} \alpha_1 & \beta_1 \\ \gamma_1 & \delta_1 \end{pmatrix}$, ciò equivale ad eseguire una sola sostituzione $\begin{pmatrix} \alpha_2 & \beta_2 \\ \gamma_2 & \delta_2 \end{pmatrix}$ che si usa chiamare il prodotto delle due. Si scrive

$$\begin{pmatrix} \alpha_2 & \beta_2 \\ \gamma_2 & \delta_2 \end{pmatrix} = \begin{pmatrix} \alpha_1 & \beta_1 \\ \gamma_1 & \delta_1 \end{pmatrix} \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}$$

È noto che, se

$$\begin{pmatrix} \alpha_3 & \beta_3 \\ \gamma_3 & \delta_3 \end{pmatrix} = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \begin{pmatrix} \alpha_1 & \beta_1 \\ \gamma_1 & \delta_1 \end{pmatrix}$$

si avrà in generale

$$\begin{pmatrix} \alpha_2 & \beta_2 \\ \gamma_2 & \delta_2 \end{pmatrix} \neq \begin{pmatrix} \alpha_3 & \beta_3 \\ \gamma_3 & \delta_3 \end{pmatrix}.$$

Les substitutions sont supposées de déterminant unitaire tout au long de l'exposé. Dans son introduction Volterra choisit donner un rôle particulier à la multiplication qu'il qualifie d'« opération fondamentale ». C'est à partir d'elle que l'on peut reconstruire diverses opérations « ordinaires »¹⁶ :

Dalla operazione fondamentale considerata del prodotto di due sostituzioni possono farsi dipendere come casi particolari la somma, la moltiplicazione, la sottrazione e la divisione ordinarie.

Infatti, come è noto, avremo

$$\begin{aligned} \begin{pmatrix} 1 & 0 \\ \alpha & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ \beta & 1 \end{pmatrix} &= \begin{pmatrix} 1 & 0 \\ \alpha + \beta & 1 \end{pmatrix} \\ \begin{pmatrix} \alpha & 0 \\ 0 & \frac{1}{\alpha} \end{pmatrix} \begin{pmatrix} \beta & 0 \\ 0 & \frac{1}{\beta} \end{pmatrix} &= \begin{pmatrix} \alpha\beta & 0 \\ 0 & \frac{1}{\alpha\beta} \end{pmatrix} \\ \begin{pmatrix} 1 & 0 \\ \alpha & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ \beta & 1 \end{pmatrix}^{-1} &= \begin{pmatrix} 1 & 0 \\ \alpha - \beta & 1 \end{pmatrix} \\ \begin{pmatrix} \alpha & 0 \\ 0 & \frac{1}{\alpha} \end{pmatrix} \begin{pmatrix} \alpha & 0 \\ 0 & \frac{1}{\beta} \end{pmatrix}^{-1} &= \begin{pmatrix} \frac{\alpha}{\beta} & 0 \\ 0 & \frac{\alpha}{\beta} \end{pmatrix}. \end{aligned}$$

¹⁵[Vol54, p.209]

¹⁶[Vol54, p.210]

Volterra choisit donc d'isoler tout d'abord la multiplication des substitutions qu'il présente comme élément constitutif d'une vision générale. Elle permet en effet de repenser en termes de cas particuliers et via des produits de substitutions les opérations « ordinaires », c'est-à-dire les opérations algébriques (addition, soustraction, multiplication et division) effectuées sur les éléments α ou β dont nous verrons par la suite qu'ils seront des fonctions.

Ce n'est que dans un second temps et dans une section séparée, qu'il nomme « préliminaire », que Volterra donne les procédés algorithmiques pour le calcul du produit de deux substitutions ou de l'inverse à partir des coefficients (identiques au calcul matriciel que nous connaissons).

Volterra rajoute à cette liste rappelant les méthodes de calcul, une propriété qui permettra plus tard de parler de *limite* pour les substitutions. Il commence par définir une notion qui évalue la *taille* (terme d'observateur) d'une substitution à partir de ses coefficients¹⁷ :

Se in una sostituzione qualunque

$$S = \begin{pmatrix} 1+a & b \\ c & 1+d \end{pmatrix}$$

a, b, c, d , sono minori di m , la sostituzione si dirà *minore* di m .

Il introduit ensuite une notion qui donne la possibilité de mesurer l'*écart* (terme d'observateur) entre deux substitutions¹⁸ :

Se due sostituzioni $\begin{pmatrix} a & b \\ c & d \end{pmatrix}, \begin{pmatrix} a_1 & b_1 \\ c_1 & d_1 \end{pmatrix}$ sono tali che $a - a_1, b - b_1, c - c_1, d - d_1$, sono minori di m si dirà che le due sostituzioni *differiscono meno di m* .

Enfin, en s'appuyant sur ces deux notions, Volterra montre comment on peut contrôler la *taille* d'un produit (ou d'un quotient) de deux substitutions lorsque l'on connaît l'*écart* entre les termes du produit et la *taille* de l'un des termes¹⁹ :

Se la sostituzione S (det. Δ) minore di m differisce dalla sostituzione S' per meno di ϵ , avremo che $S^{-1}S'$ e $S'S^{-1}$ saranno minori di $\frac{\epsilon(1+2m)}{\Delta}$ e reciprocamente se $S'S^{-1}$ è minore di ϵ e S' è minore di m , S e S' differiranno fra loro meno di $\epsilon(1+2m)$.

Volterra développe quelques autres propriétés liées à ces deux éléments de mesure, mais la citation ci-dessus permet de comprendre ce qui est en jeu et qui va se révéler crucial pour établir un calcul différentiel pour les substitutions. Les deux notions que nous avons appelées *taille* et *écart*, définies à partir des coefficients, désignent donc désormais des propriétés des substitutions. Ce passage à un vocabulaire qui s'émancipe de la référence explicite aux coefficients est l'expression d'un travail d'abstraction qui permet finalement de produire un énoncé où les coefficients ont complètement disparu de la scène comme dans le théorème précédent.

Le choix de Volterra de commencer un exposé qu'il conçoit comme « fondement » d'une théorie des équations différentielles linéaires en convoquant

¹⁷[Vol54, p.216]

¹⁸[Vol54, p.216]

¹⁹[Vol54, p.216]

et en organisant un certain nombre d'idées préliminaires qui correspondent à la mise en place de ce que nous nommerons des *éléments structurels*²⁰ est à nos yeux essentiel. Nous ne parlerons pas ici de *structure* car il manque un regard réflexif sur un agencement nouveau que Volterra aurait pu identifier — nous pourrions de nos jours penser en termes d' *espace de substitutions muni d'une topologie*²¹ par exemple. De même, si Volterra organise son texte autour d'une opération unique, la multiplication, celle-ci n'est pas traitée d'un point de vue qui ferait ressortir une structure algébrique, seules quelques propriétés étant sélectionnées. Cependant il reste remarquable que l'on trouve ici chez Volterra dès ce texte de 1886-87²² la mise en avant, sous une forme qui crée un point de vue particulier et une vision synthétique, à la fois d'une opération et d'une façon d'estimer l'écart entre les substitutions. Les deux éléments sont de plus clairement pensés dans leur rapport mutuel. Ces divers aspects feront étonnamment écho à des idées que nous retrouverons bien plus tard chez Fréchet ou chez Banach.

Enfin Volterra finit cette partie préliminaire avec un paragraphe sur la réduction des substitutions à une forme normale grâce à la notion de diviseurs élémentaires et en s'appuyant sur un théorème de Weierstraß sur les formes quadratiques. Nous n'entrons pas ici dans les détails pour nous concentrer sur le calcul différentiel et intégral des substitutions qui est au cœur de la progression que nous cherchons à éclairer dans le travail de Volterra.

3.2.2 DÉRIVATION DES SUBSTITUTIONS

On considère ici des substitutions de déterminants unitaires dont les coefficients sont des fonctions d'une seule variable réelle et à valeurs réelles. Volterra définit d'abord les substitutions continues : ce sont celles dont les coefficients sont tous continus.

La dérivée²³ s'obtient alors en un point x_0 avec un accroissement Δx . Pour une substitution donnée²⁴ $T = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$ Volterra introduit deux substitutions construites de façon symétrique : un rapport différentiel à gauche

$$S_{x_0+\Delta x, x_0} = \begin{pmatrix} A & B \\ C & D \end{pmatrix}_{x_0+\Delta x} \begin{pmatrix} A & B \\ C & D \end{pmatrix}_{x_0}^{-1}$$

²⁰Nous employons l'expression *élément structurel* pour indiquer la présence d'agencements, que l'on repère dans l'organisation du discours ou dans la composition des concepts, sur lesquels les mathématiciens s'appuient pour convoquer un paysage cohérent et proposer une vision générale.

²¹Il n'y a pas suffisamment d'éléments dans le texte de Volterra pour préciser ici ce que nous entendons par *espace* ou *topologie*. On peut penser à un groupe topologique par exemple.

²²Volterra précisera par la suite que ce premier texte publié en 1887 était prêt dès 1886.

²³Il s'agit d'une dérivée logarithmique (termes d'observateurs) en fait, avec deux possibilités à chaque fois, du fait de la non-commutativité de la multiplication des substitutions, qui donnent une dérivée à droite et une à gauche.

²⁴Volterra n'utilise pas en général une seule lettre pour désigner les substitutions qui sont manipulées sous leur forme de tableau carré. Nous précisons dans le cours de notre étude quand il utilisera un type de notation plus synthétique. Pour éclairer le texte nous désignons ici par T une substitution quelconque et par I la substitution $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$.

et son pendant à droite

$$S'_{x_0, x_0+\Delta x} = \begin{pmatrix} A & B \\ C & D \end{pmatrix}_{x_0}^{-1} \begin{pmatrix} A & B \\ C & D \end{pmatrix}_{x_0+\Delta x}.$$

En utilisant les règles de calcul énoncées au début du traité on peut déterminer explicitement les coefficients de ces substitutions²⁵. En suivant Volterra on a $S_{x_0+\Delta x, x_0} = I + \begin{pmatrix} \lambda & \mu \\ \nu & \rho \end{pmatrix}$ où λ, μ, ν, ρ tendent vers 0 lorsque Δx tend vers 0. Volterra dira que $S_{x_0+\Delta x, x_0}$ « tend vers l'identité » sans préciser comment il conçoit ici cette convergence. Nous pouvons repenser la convergence en termes d'écart (terme d'observateur) tel que Volterra l'a introduit précédemment : dire que les coefficients λ, μ, ν, ρ sont petits signifie bien conformément aux idées de la partie préliminaire que l'écart entre $I + \begin{pmatrix} \lambda & \mu \\ \nu & \rho \end{pmatrix}$ et I est petit.

Volterra considère alors la substitution dont les coefficients sont les taux d'accroissement :

$$\begin{pmatrix} \frac{\lambda}{\Delta x} & \frac{\mu}{\Delta x} \\ \frac{\nu}{\Delta x} & \frac{\rho}{\Delta x} \end{pmatrix}.$$

Un premier théorème énonce que cette dernière substitution tend vers une substitution limite notée $\begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}$ si et seulement si les coefficients de T sont tous dérivables en x_0 . Dans la suite Volterra travaillera sous cette hypothèse : les coefficients de T seront continus et dérivables.

L'approche que nous venons de résumer permet à Volterra de définir les éléments dont il a besoin pour son calcul différentiel. Il introduit les substitutions produits $S_{x_0+dx, x_0} = \begin{pmatrix} A+dA & B+dB \\ C+dC & D+DD \end{pmatrix} \begin{pmatrix} A & B \\ C & D \end{pmatrix}^{-1}$ et $S_{x_0, x_0+dx} = \begin{pmatrix} A & B \\ C & D \end{pmatrix}^{-1} \begin{pmatrix} A+dA & B+dB \\ C+dC & D+DD \end{pmatrix}$, nommées respectivement « différentielle à gauche » et « différentielle à droite » de T au point x_0 . Volterra les notera dans le reste de son exposé $d \begin{pmatrix} A & B \\ C & D \end{pmatrix}$ et $\begin{pmatrix} A & B \\ C & D \end{pmatrix} d$. Nous désignerons de notre côté ces différentielles par dT et Td dans la suite²⁶.

La substitution limite $\begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}$ est appelée « dérivée à gauche au point x_0 » et sera notée $\frac{d}{dx}T$. La dérivée à droite sera notée $T\frac{d}{dx}$. Enfin en calculant explicitement les coefficients des diverses substitutions produits, Volterra montre que $\frac{d}{dx}T = \begin{pmatrix} A' & B' \\ C' & D' \end{pmatrix} \begin{pmatrix} A & B \\ C & D \end{pmatrix}^{-1}$ où A', B', C' et D' désignent les dérivées ordinaires des fonctions. Avec nos notations on a donc $dT = I + \begin{pmatrix} dA & dB \\ dC & dD \end{pmatrix} \begin{pmatrix} A & B \\ C & D \end{pmatrix}^{-1}$.

Volterra choisit d'écrire cette relation sous la forme $d \begin{pmatrix} A & B \\ C & D \end{pmatrix} = \frac{d}{dx} \begin{pmatrix} A & B \\ C & D \end{pmatrix} dx$. On voit ici que cette égalité et le dx qui suit la dérivée doivent

²⁵Dans la suite nous montrons la démarche de Volterra pour la dérivée à gauche, les calculs étant similaires pour la dérivée à droite.

²⁶Nous préciserons quand les notations synthétiques sont celles de Volterra.

être interprétés en suivant précautionneusement les idées exposées plus haut. En particulier il ne faut pas penser pour une substitution $U = \begin{pmatrix} u_1 & u_2 \\ u_3 & u_4 \end{pmatrix}$ donnée Udx comme la multiplication terme à terme d'une substitution par un scalaire ou un infiniment petit. La notation Udx désigne dans ce texte de Volterra la substitution $\begin{pmatrix} 1 + u_1 dx & u_2 dx \\ u_3 dx & 1 + u_4 dx \end{pmatrix}$. Ici, bien que l'exposé fasse apparaître les coefficients des substitutions pour introduire différentes définitions ou certains calculs, les notations empruntées au calcul différentiel des fonctions ordinaires (d ou $\frac{d}{dx}$ par exemple) nécessitent une interprétation particulière et dénotent une certaine conception abstraite.

Enfin cette partie sur la dérivation se prolonge en une collection de résultats qui précisent les propriétés de la dérivation et qui établissent des formules de dérivation d'un produit ou d'une substitution à coefficients constants par exemple.

En particulier, Volterra désigne le théorème suivant comme « fondamental » dans la théorie des équations différentielles²⁷ :

TEOREMA II. La derivata a sinistra di una sostituzione (det. 1) non cambia se si moltiplica A SINISTRA la sostituzione per una sostituzione costante (det. 1)

Nous n'insistons pas davantage sur les théorèmes relatifs à la dérivation : ce qui précède illustre précisément la façon d'opérer de Volterra qui adapte ici une partie de la théorie du « calcul infinitésimal » pour les fonctions ordinaires au cas des substitutions. Son approche combine une vision que l'on pourrait qualifier de classique à son époque et une utilisation de motifs que l'on retrouve dans l'étude tout aussi classique des équations différentielles linéaires. D'un côté Volterra s'appuie sur une notion de fonction pour laquelle on sait définir, via un accroissement de la variable, d'abord une dérivée puis une différentielle. D'un autre côté la multiplication et la dérivée que nous avons qualifiée de *logarithmique* font fortement penser à des équations du type $y' = Ay$ pour des fonctions ordinaires. Nous reviendrons sur cet aspect en nous appuyant à nouveau sur un texte de Volterra un peu plus tard.

Enfin notons que le calcul différentiel mis en place par Volterra repose sur l'utilisation du rapport de la variation de la substitution par la variation de la variable x . Nous pouvons préciser cette approche. Suivant une analyse classique à l'époque (que l'on peut lire par exemple dans le cours de Serret de 1868²⁸), la dérivée peut se concevoir comme rapport de la différentielle d'une fonction (ordinaire) par la différentielle de la fonction $x \mapsto x$. Cette façon de voir donne une place particulière à la fonction $x \mapsto x$ et à sa différentielle notée dx qui est prise comme élément de référence. Chez Volterra on peut relire l'expression $d \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \frac{d}{dx} \begin{pmatrix} a & b \\ c & d \end{pmatrix} dx$ suivant la même analyse, en pensant dx comme différentielle d'une fonction ordinaire particulière. Bien que Volterra n'aborde pas le sujet sous cet angle, toute la théorie qu'il développe s'appuie sur

²⁷[Vol54, p.224]

²⁸cf. [Ser68, p.26]. Il s'agit de la première édition du cours de Serret professé à la Sorbonne. Il y a plusieurs rééditions successives qui exposent toujours de la même façon l'approche des différentielles.

une référence aux fonctions ordinaires et la notation dx qu'il conserve dans ce cas pour exprimer la différentielle reste en quelques sortes étrangère à la théorie des substitutions.

L'égalité du dessus, que nous notons $dT = \frac{d}{dx}Tdx$, met en jeu des éléments de nature différentes qui impliquent qu'elle doit être conçue via une interprétation particulière. D'une part elle dénote sous forme synthétique le résultat d'une démarche liée à la dérivation, mais d'autre part le choix de la notation dx renvoie à des considérations liées à la théorie des fonctions ordinaires.

Nous insistons sur ce point car plusieurs années plus tard Fréchet reviendra, à propos de certains travaux de Volterra, sur le rôle particulier donné à la fonction $x \mapsto x$, pour à son tour analyser diverses pistes de généralisation liées au choix d'une fonction ou d'un type de fonction de référence dans le calcul différentiel. Nous verrons en outre par la suite que la question qui consiste à choisir relativement à quoi on doit comparer la variation d'une grandeur dans un rapport différentiel est essentielle dans plusieurs travaux de Volterra.

3.2.3 INTÉGRATION DES SUBSTITUTIONS

Dans une deuxième section Volterra développe une notion d'intégration pour les substitutions qui va compléter l'outillage nécessaire à une théorie générale des équations différentielles linéaires. Pour obtenir une opération inverse de la dérivation il lui faut choisir une version adaptée de l'intégration. Celle-ci, conformément à l'approche riemannienne, repose sur une division de plus en plus fine de l'intervalle d'intégration et la convergence (dans le cas intégrable) de produits de substitutions.

Rappelons la présentation de Volterra²⁹ :

1. Abbiassi una sostituzione $\begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}$ funzione di un argomento reale x variabile in un intervallo $(p \cdots q)$, $p < q$. Dividiamo questo intervallo in n parti h_1, h_2, \cdots, h_n (a cominciare dal limite inferiore p dell'intervallo) e denotiamo con $\alpha_i, \beta_i, \gamma_i, \delta_i$ valori di $\alpha, \beta, \gamma, \delta$ qualunque compresi fra il limite superiore e l'inferiore di queste quantità nell'intervallo h_i .

Formiamo i prodotti delle sostituzioni

$$A_n = \prod_1^n \begin{pmatrix} 1 + \alpha_i h_i & \beta_i h_i \\ \gamma_i h_i & 1 + \delta_i h_i \end{pmatrix} = \begin{pmatrix} 1 + \alpha_n h_n & \beta_n h_n \\ \gamma_n h_n & 1 + \delta_n h_n \end{pmatrix} \cdots \begin{pmatrix} 1 + \alpha_1 h_1 & \beta_1 h_1 \\ \gamma_1 h_1 & 1 + \delta_1 h_1 \end{pmatrix}$$

$$A'_n = \prod_n^1 \begin{pmatrix} 1 + \alpha_i h_i & \beta_i h_i \\ \gamma_i h_i & 1 + \delta_i h_i \end{pmatrix} = \begin{pmatrix} 1 + \alpha_1 h_1 & \beta_1 h_1 \\ \gamma_1 h_1 & 1 + \delta_1 h_1 \end{pmatrix} \cdots \begin{pmatrix} 1 + \alpha_n h_n & \beta_n h_n \\ \gamma_n h_n & 1 + \delta_n h_n \end{pmatrix}$$

Se, facendo impiccolire indefinitamente tutti gli intervalli h_1, h_2, \cdots, h_n , A_n e A'_n tendono verso due sostituzioni limiti S e S' , queste si chiameranno rispettivamente l'integrale sinistro e l'integrale destro di $\begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}$ nell'intervallo $(p, \cdots q)$ e si adotteranno i simboli

$$\text{(Integrale sinistro)} \quad S = \int_p^q \begin{pmatrix} 1 + \alpha dx & \beta dx \\ \gamma dx & 1 + \delta dx \end{pmatrix} = \int_p^q \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} dx,$$

²⁹[Vol54, p.235]

$$\text{(Integrale destro)} \quad S' = \begin{pmatrix} 1 + \alpha dx & \beta dx \\ \gamma dx & 1 + \delta dx \end{pmatrix} \int_p^q = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} dx \int_p^q.$$

La section précédente permet de comprendre les notations employées ici. Volterra poursuit en donnant une condition nécessaire et suffisante pour que le produit converge. Il s'appuie sur la notion d'oscillation qu'il adapte aux substitutions³⁰ :

Se una sostituzione $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ è funzione di x in un intervallo $(p \cdots q)$, la massima oscillazione dei suoi elementi si chiamerà *l'oscillazione* della sostituzione nell'intervallo.

Le critère d'intégrabilité établi par Volterra pour les substitutions est alors énoncé de la façon suivante³¹ :

TEOREMA I. Se D_i denota la oscillazione della sostituzione finita

$$\begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}$$

nell'intervallo h_i e se

$$\lim \sum_1^n h_i D_i = 0 \quad \sum_1^n h_i = q - p$$

per l'impiccolire indefinito di tutte le h_i , esisteranno e saranno finiti l'integrale destro e l'integrale sinistro della sostituzione $\begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}$ nell'intervallo totale $(p \cdots q)$ ed in una porzione qualunque di questo intervallo.

Come conseguenza si deduce:

Nel caso in cui $\begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}$ sia sempre finita e continua o abbia un numero finito di punti di discontinuità essa sarà sempre integrabile tanto a destra quanto a sinistra.

Volterra établit ensuite un certain nombre de théorèmes pour le calcul des intégrales en suivant un programme adapté de l'étude de l'intégration des fonctions ordinaires dans la lignée des travaux de Riemann : relation de Chasles, continuité par rapport aux bornes d'intégration, intégration par parties, changement de variables entre autres.

En particulier Volterra avait décrit l'intégration dans son rapport avec la dérivation dès l'introduction de son mémoire : « Dimostrerò che le due operazioni della derivazione e della integrazione delle sostituzioni possono ritenersi come operazioni inverse »³². Deux théorèmes viennent préciser cet aspect. Le premier complète un théorème qui donnait une particularité des substitutions dérivées (leur trace est nulle) et en fait désormais une condition nécessaire et suffisante qui détermine une large collection de substitutions intégrables³³ :

³⁰[Vol54, p.216]

³¹[Vol54, p.242]

³²[Vol54, p.211]

³³[Vol54, p.245]

La condizione necessaria e sufficiente affinché una sostituzione continua $\begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}$ sia derivata di una sostituzione (det. 1) è che sia $\alpha + \delta = 0$.

Un grand nombre de développements ultérieurs se feront sous cette hypothèse de trace nulle.

Le second théorème explicite le lien entre dérivation et intégration³⁴ :

TEOREMA VI. - Abbiasi $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ (det. 1) e

$$\frac{d}{dx} \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix},$$

si avrà

$$\int_p^q \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} dx = \begin{pmatrix} a & b \\ c & d \end{pmatrix}_q \begin{pmatrix} a & b \\ c & d \end{pmatrix}_p^{-1}$$

supponendo $\begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}$ continua.

Enfin, après les deux premières parties qui introduisent la dérivation et l'intégration des substitutions d'une variable réelle, Volterra généralise ces notions au cas des substitutions de plusieurs variables. Le processus de généralisation par le biais de l'augmentation du nombre de variables des fonctions est central chez Volterra. Il est ici élémentaire mais il donne naissance à quelques changements d'optique qui participent de l'élaboration d'une vision synthétique et abstraite. Nous reviendrons à nouveau dans la suite sur des démarches du même type et nous les contrasterons avec le processus de passage à une *infinité* de variables.

3.2.4 SUBSTITUTIONS DE PLUSIEURS VARIABLES

L'étude des substitutions de plusieurs variables est développée en deux parties, l'une se concentrant sur les idées relatives à la dérivation, l'autre étant dédiée aux résultats d'intégration. Nous résumons rapidement les principales notions mises en place et les résultats importants obtenus. Nous utilisons ici encore des notations simplifiées en dehors des citations du texte de Volterra.

Tout d'abord Volterra définit les dérivées partielles et la différentielle totale d'une substitution. Soit $T = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$ une substitution dont les coefficients sont des fonctions de n variables possédant toutes des dérivées partielles par rapport à chacune des variables. Alors, en suivant les idées de Volterra, on peut écrire le produit $T_{x_1+dx_1, x_2+dx_2, \dots, x_n+dx_n} T_{x_1, x_2, \dots, x_n}^{-1}$ en laissant tomber les termes d'ordres supérieurs à dx_1, \dots, dx_n sous l'une des trois formes suivantes qui définissent la « différentielle totale à gauche » notée dT ³⁵ :

$$(I) \quad \begin{pmatrix} 1 + DdA - CdB & AdB - BdA \\ DdC - CdD & 1 + AdD - BdC \end{pmatrix}$$

³⁴[Vol54, p.245]

³⁵[Vol54, p.253]

$$(II) = \left\{ \begin{array}{cc} 1 + \sum_1^n \left(D \frac{dA}{dx_i} - C \frac{dB}{dx_i} \right) dx_i & \sum_1^n \left(A \frac{dB}{dx_i} - B \frac{dA}{dx_i} \right) dx_i \\ \sum_1^n \left(D \frac{dC}{dx_i} - C \frac{dD}{dx_i} \right) dx_i & 1 + \sum_1^n \left(A \frac{dD}{dx_i} - B \frac{dC}{dx_i} \right) dx_i \end{array} \right\}$$

$$(III) = \prod_1^n \left\{ \begin{array}{cc} 1 + \left(D \frac{dA}{dx_i} - C \frac{dB}{dx_i} \right) dx_i & \left(A \frac{dB}{dx_i} - B \frac{dA}{dx_i} \right) dx_i \\ \left(D \frac{dC}{dx_i} - C \frac{dD}{dx_i} \right) dx_i & \left(A \frac{dD}{dx_i} - B \frac{dC}{dx_i} \right) dx_i \end{array} \right\}.$$

Les substitutions qui composent le produit sont les « dérivées partielles » de T :

$$\frac{d}{dx_i} T = \left\{ \begin{array}{cc} \left(D \frac{dA}{dx_i} - C \frac{dB}{dx_i} \right) dx_i & \left(A \frac{dB}{dx_i} - B \frac{dA}{dx_i} \right) dx_i \\ \left(D \frac{dC}{dx_i} - C \frac{dD}{dx_i} \right) dx_i & \left(A \frac{dD}{dx_i} - B \frac{dC}{dx_i} \right) dx_i \end{array} \right\}$$

$$= \begin{pmatrix} \frac{dA}{dx_i} & \frac{dB}{dx_i} \\ \frac{dC}{dx_i} & \frac{dD}{dx_i} \end{pmatrix} \begin{pmatrix} A & B \\ C & D \end{pmatrix}^{-1}.$$

Avec le même type de notation qui indiquait le lien entre différentielle et dérivée, Volterra nomme la substitution

$$d_{x_i} \left\{ \begin{array}{cc} 1 + \left(D \frac{dA}{dx_i} - C \frac{dB}{dx_i} \right) dx_i & \left(A \frac{dB}{dx_i} - B \frac{dA}{dx_i} \right) dx_i \\ \left(D \frac{dC}{dx_i} - C \frac{dD}{dx_i} \right) dx_i & 1 + \left(A \frac{dD}{dx_i} - B \frac{dC}{dx_i} \right) dx_i \end{array} \right\}$$

« différentielle partielle à gauche ». On a donc $dx_i T = \frac{d}{dx_i} T dx_i$ et on obtient la « formule fondamentale » de Volterra que nous écrivons avec nos notations $dT = \prod_{i=1}^n \frac{d}{dx_i} T dx_i$. Les mêmes notions peuvent être définies à droite, et Volterra les dénote encore ici avec un symbole de différenciation apposé à droite de la substitution.

Volterra prolonge alors ces idées en définissant les différentielles d'ordres supérieurs. La différentielle à gauche d'ordre 2 s'obtient en prenant la différentielle à gauche de la différentielle à gauche et en laissant tomber les termes d'ordre supérieur à 2. Il montre que l'on obtient en général $d^2 T = \begin{pmatrix} 1+u_1 & u_2 \\ u_3 & 1+u_4 \end{pmatrix}$ où u_1 , u_2 , u_3 et u_4 sont obtenus à partir des coefficients de la substitution initiale T et de leurs différentielles d'ordre 2. Plus précisément on peut calculer :

$$d^2 T = \left\{ \begin{array}{cc} 1 + \frac{1}{2} (Bd^2C - Ad^2D - Cd^2B + Dd^2A) & (Ad^2B - Bd^2A) \\ (Dd^2C - Cd^2D) & 1 + \frac{1}{2} (Cd^2B - Dd^2A - Bd^2C + Ad^2D) \end{array} \right\}.$$

On obtient la notion de dérivée partielle à gauche d'ordre 2, notée $\frac{d^2}{d(x_r x_s)} T = \frac{d^2}{d(x_s x_r)} T$ en divisant les coefficients de la substitution $\begin{pmatrix} u_1 & u_2 \\ u_3 & u_4 \end{pmatrix}$ par $dx_r dx_s$. Volterra

utilise alors des notations similaires à celles employées pour les différentielles d'ordre 1 et il écrit $\begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} dx_i^2$ pour $\begin{pmatrix} 1 + \alpha dx_i^2 & \beta dx_i^2 \\ \gamma dx_i^2 & 1 + \delta dx_i^2 \end{pmatrix}$ et $2 \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} dx_r dx_s$ pour $\begin{pmatrix} 1 + 2\alpha dx_r dx_s & 2\beta dx_r dx_s \\ 2\gamma dx_r dx_s & 1 + 2\delta dx_r dx_s \end{pmatrix}$. Avec ces notations on obtient l'« expression générale de la différentielle seconde » :

$$d^2T = \prod_{i=1}^n \frac{d^2}{dx_i^2} T dx_i^2 \cdot \prod_{r,s} 2 \frac{d^2}{d(x_r x_s)} T dx_r dx_s.$$

On peut alors aussi écrire $\frac{d^2}{dx_i^2} T dx_i^2 = d_{x_i}^2 T$.

Bien que Volterra adapte ici des idées classiques pour les fonctions à valeurs réelles, il apparaît des différences sensibles qu'il met en évidence³⁶ :

[...] avremo in generale che

$$\frac{d}{dx_s} \frac{d}{dx_i} \begin{pmatrix} A & B \\ C & D \end{pmatrix} \neq \frac{d}{dx_i} \frac{d}{dx_s} \begin{pmatrix} A & B \\ C & D \end{pmatrix} \neq \frac{d^2}{d(x_i x_s)} \begin{pmatrix} A & B \\ C & D \end{pmatrix} = \frac{d^2}{d(x_s x_i)} \begin{pmatrix} A & B \\ C & D \end{pmatrix}$$

Nel nostro caso quindi non vale la legge di permutabilità nella derivazione, né la legge di formazione del differenziale secondo analoga a quella per le funzioni.

Après avoir mis en place ces différents outils de dérivation et d'intégration Volterra dédie le reste de son exposé, jusqu'à la fin de cette première partie, à l'établissement d'une collection de théorèmes qui correspondent aux résultats classiques du calcul différentiel pour les fonctions à valeurs réelles. Nous dégagons les résultats principaux dans la section suivante car ils permettront dans la suite de bien cerner le type d'objectif que cherchera à atteindre Volterra dans des travaux où il tente à nouveau à dégager une vision générale pour l'analyse fonctionnelle.

3.2.5 RÉSULTATS PHARES DE LA THÉORIE

Le résumé de ces résultats permet aussi de comprendre la teneur et l'étendue de ce que Volterra conçoit comme « fondement de la théorie des équations différentielles linéaires ».

DIFFÉRENTIELLES EXACTES : CONDITION NÉCESSAIRE. Le premier théorème de cette liste est important car il fait écho à un résultat central de la théorie des fonctions ordinaires mais aussi parce qu'il est démontré de deux façons différentes par Volterra, deux manières qui manifestent à nos yeux une évolution vers un point de vue général.

Commençons par rappeler le problème posé³⁷ :

Dato un prodotto di sostituzioni infinitesime della forma

$$(6) \quad \prod_1^n \begin{pmatrix} 1 + \alpha_i dx_i & \beta_i dx_i \\ \gamma_i dx_i & 1 + \delta_i dx_i \end{pmatrix} = \prod_1^n \begin{pmatrix} \alpha_i & \beta_i \\ \gamma_i & \delta_i \end{pmatrix} dx_i$$

³⁶[Vol54, p.257]

³⁷[Vol54, p.258]

determinare a quali condizioni esse debbono soddisfare, affinché il prodotto rappresenti il differenziale esatto totale sinistro di una sostituzione.

In altri termini: Data una espressione differenziale, trovare la condizione affinché essa sia un differenziale esatto totale sinistro.

Il s'agit donc d'établir une condition nécessaire pour qu'un produit de différentielles soit la différentielle totale d'une substitution.

Nous suivons tout d'abord la démonstration qui figure dans le corps du texte de Volterra et qui s'appuie sur le calcul explicite de certains coefficients de substitutions.

On cherche une substitution T qui vérifie $dT = \prod_{i=1}^n \begin{pmatrix} \alpha_i & \beta_i \\ \gamma_i & \delta_i \end{pmatrix} dx_i$. En écrivant $\frac{d^2}{d(x_i x_s)} T = \frac{d^2}{d(x_s x_i)} T$ on obtient par différence terme à terme des coefficients une substitution nulle. Plus précisément Volterra introduit ici un *premier paramètre différentiel* et une notation valable pour toute paire de substitutions :

$$\Delta' \left\{ \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}, \begin{pmatrix} \alpha_1 & \beta_1 \\ \gamma_1 & \delta_1 \end{pmatrix} \right\}_{x,y} = \left\{ \begin{array}{l} \frac{d\alpha_1}{dx} - \frac{d\alpha}{dy} + (\beta_1\gamma - \beta\gamma_1) \quad \frac{d\beta_1}{dx} - \frac{d\beta}{dy} + 2(\alpha_1\beta - \alpha\beta_1) \\ \frac{d\gamma_1}{dx} - \frac{d\gamma}{dy} + 2(\delta_1\gamma - \delta\gamma_1) \quad \frac{d\delta_1}{dx} - \frac{d\delta}{dy} + (\beta_1\gamma - \beta\gamma_1) \end{array} \right\}.$$

On a donc avec les notations précisées au-dessus :

$$\Delta' \left\{ \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}, \begin{pmatrix} \alpha_1 & \beta_1 \\ \gamma_1 & \delta_1 \end{pmatrix} \right\}_{x,y} = \frac{d^2}{d(x_i x_s)} T - \frac{d^2}{d(x_s x_i)} T.$$

Une version similaire notée $\Delta'' \left\{ \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}, \begin{pmatrix} \alpha_1 & \beta_1 \\ \gamma_1 & \delta_1 \end{pmatrix} \right\}_{x,y}$ s'obtient en effectuant les mêmes calculs pour les dérivées et différentielles à droite³⁸.

Grâce à ces notations Volterra obtient un critère d'intégrabilité des différentielles dans le cas de deux substitutions³⁹ :

La condizione affinché

$$\begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} dx \begin{pmatrix} \alpha_1 & \beta_1 \\ \gamma_1 & \delta_1 \end{pmatrix} dy$$

sia un differenziale *sinistro* verrà espressa da

$$\Delta' \left\{ \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}, \begin{pmatrix} \alpha_1 & \beta_1 \\ \gamma_1 & \delta_1 \end{pmatrix} \right\}_{x,y} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix};$$

mentre la condizione affinché sia un differenziale destro sarà data da

$$\Delta'' \left\{ \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}, \begin{pmatrix} \alpha_1 & \beta_1 \\ \gamma_1 & \delta_1 \end{pmatrix} \right\}_{x,y} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

³⁸On peut y penser, bien que Volterra ne l'indique ainsi pas dans son texte, sous la forme suivante : $\Delta'' \{U, V\}_{x,y} = \Delta' \{V, U\}_{y,x}$ où les coefficients de U et V sont des fonctions de x et y .

³⁹[Vol54, p.260]

Volterra propose alors une deuxième approche du même problème qui apparaît sous forme de note à la fin de ce premier mémoire. Nous examinons maintenant la nouvelle preuve car elle contraste avec celle que nous venons de rappeler et illustre une évolution dans le travail de Volterra sur laquelle nous reviendrons plus avant dans la suite.

Les notations de Volterra sont ici beaucoup plus synthétiques. Il considère un produit de « substitutions infinitésimales » qu'il note $\prod_1^n X_i dx_i$. Puis il introduit pour une substitution donnée S la substitution S'_{x_r} obtenue en dérivant chaque coefficient par rapport à l'une des variables, ici x_r . Volterra développe alors sa preuve sans passer par l'écriture explicite des coefficients des substitutions en jeu⁴⁰ :

[...] dovremo avere

$$X_r = \frac{dS}{dx_r} = S'_{x_r} S^{-1}$$

onde

$$S'_{x_r} = X_r S$$

e derivando rispetto ad x_s ,

$$(S'_{x_r})'_{x_s} = (X_r)'_{x_s} S + X_r S'_{x_s} = (X_r)'_{x_s} S + X_r S'_{x_s} X_r X_s S = [(X_r)'_{x_s} + X_r X_s] S;$$

scambiando x_r con x_s , il primo membro non deve cambiare, quindi

$$[(X_r)'_{x_s} + X_r X_s] S = [(X_s)'_{x_r} + X_s X_r] S$$

vale a dire

$$(X_r)'_{x_s} - (X_s)'_{x_r} + X_r X_s - X_s X_r = \Delta'(X_s, X_r)_{x_s, x_r} = 0.$$

On voit donc apparaître ici une nouvelle façon de présenter la démonstration qui repose sur des résultats intermédiaires exprimés de façon synthétique, alors qu'ils ont été eux-mêmes démontrés en utilisant une écriture explicite avec les coefficients des substitutions. Volterra adopte pleinement cette démarche et préfère même introduire la substitution S' par une phrase en italien⁴¹ plutôt que par une écriture sous forme de tableau des coefficients dérivés. Cette façon de procéder simplifie grandement la preuve et met en avant une sorte d'évidence calculatoire qui s'appuie sur une démarche algébrique classique.

Dans la nouvelle approche les quatre opérations algébriques sont présentées sur un même plan dans des procédures propres à la résolution des équations algébriques. Dans la phase initiale de la théorie Volterra avait attribué un rôle particulier à la multiplication qui lui permettait de construire un *calcul des substitutions*. Cette approche fondée sur une opération unique a rendu possible l'adaptation des idées classiques du calcul différentiel et intégral dans l'optique

⁴⁰[Vol54, p.289]

⁴¹Volterra commence sa preuve par « Denotiamo con S'_{x_r} la sostituzione che si ottiene da S derivandone tutti gli elementi rispetto ad x_r . » [Vol54, p.289]

de résoudre des équations différentielles linéaires. Ici il en va différemment et les opérations ne servent pas à *composer* de nouveaux éléments de la théorie, elles retrouvent leurs rôles d'opérations algébriques et permettent d'effectuer des calculs sur les substitutions de façon abstraite et synthétique. Les quatre opérations sont bien elles-mêmes définies à partir des coefficients, mais ceux-ci restent totalement hors de la scène que propose Volterra dans cette reprise du théorème d'intégrabilité des différentielles qui s'appuie sur une démarche abstraite.

Un dernier aspect est induit par le choix d'une formulation synthétique ici, il s'agit du rapport à la dimension ou plus précisément à l'*ordre* des substitutions. Volterra avait indiqué que son travail, exposé dans le cadre des substitutions d'ordre 2 pour des raisons de commodité et de clarté, restait valide et s'adaptait facilement au cas des substitutions d'ordre quelconque n fini. Il présentera cette extension dans un dernier paragraphe intitulé « Generalizzazione al caso di una sostituzione qualunque »⁴² en utilisant des tableaux de taille $n \times n$ et en exposant à nouveau le travail au niveau des coefficients, pour enfin obtenir des résultats exprimés sous une forme plus synthétique et dégagée de la référence à la dimension.

La note que nous avons mentionnée au-dessus reste donc le seul passage où Volterra s'affranchit complètement du début à la fin de la nature matricielle des substitutions et donc de leur *ordre*. Seule la référence aux variables x_s rappelle qu'il y a des fonctions ordinaires sous-jacentes dans les calculs ainsi exposés de façon abstraite. Nous verrons par la suite un nouveau pas dans la démarche générale de Volterra qui repensera la dépendance globale à une ou à des fonctions ordinaires dans les problèmes d'analyse et en particulier dans l'étude des équations différentielles.

INTÉGRATION DES FORMES DIFFÉRENTIELLES TOTALES. Dans un nouveau paragraphe, Volterra établit la réciproque du théorème précédent. Il s'agit donc ici de considérer une expression de la forme $\prod_{i=1}^n X_i dx_i$ dans laquelle les X_i vérifient $\Delta'(X_i, X_s)_{(x_i, x_s)} = 0$ pour tout $1 \leq i, s \leq n$. La progression se déroule alors en plusieurs temps qui font apparaître des variations dans la façon d'écrire les démonstrations.

Tout d'abord Volterra développe une série de quatre théorèmes préliminaires dont les énoncés font partiellement apparaître l'écriture explicite (avec les coefficients) de certaines substitutions. Nous citons ici le premier théorème qui illustre un type de formulation mixte⁴³ :

TEOREMA 1. Siano X (som. 0), Y (som. 0) due sostituzioni funzioni di x e y e sia

$$S = \int X dx, \quad X_1 = \frac{dS}{dy}.$$

Si ponga

$$\begin{pmatrix} M & N \\ P & Q \end{pmatrix} = S^{-1}(Y - X_1)S.$$

⁴²[Vol54, p.280]

⁴³[Vol54, p.289]

Avremo

$$\begin{pmatrix} \frac{dM}{dx} & \frac{dN}{dx} \\ \frac{dP}{dx} & \frac{dQ}{dx} \end{pmatrix} = S^{-1}[\Delta(X, Y)_{x,y}]S.$$

En particulier Volterra n'utilise pas la notation M'_x que nous avons déjà vue et préfère la forme explicite qui donne les coefficients. C'est aussi le cas pour les deux théorèmes suivants, mais pas pour le quatrième dont l'énoncé n'utilise que des notations synthétiques.

Plus précisément la démonstration du premier théorème, ainsi que celle des trois suivants, s'appuie essentiellement sur une analyse au niveau des coefficients des substitutions, et les expressions synthétiques, plus ou moins présentes en fonction des énoncés, ne servent qu'à désigner les calculs ou les résultats obtenus dans les démonstrations en utilisant les coefficients.

Dans un second temps Volterra établit le résultat principal de cette partie dans le cas de deux variables indépendantes. Le problème est posé en début de chapitre dans un vocabulaire synthétique et appelle une résolution par *construction*⁴⁴ :

Risolviamo ora la questione: Data la espressione differenziale
 $[dS = \prod_1^n X_i dx_i]$ i cui fattori soddisfano le relazioni $[\Delta'(X_i, X_s)_{(x_i, x_s)} = 0$
 $i, s = 1, 2, \dots, n]$, costruire, se è possibile, la sostituzione integrale S .

Les théorèmes 1 à 4 que nous avons mentionnés plus haut apparaissent alors comme des résultats préliminaires techniques sur lesquels Volterra va s'appuyer pour étudier le problème qu'il a posé. Le résultat n'est pas énoncé explicitement sous forme de théorème, il apparaît naturellement comme conclusion d'une analyse, ce qui accentue à nos yeux l'importance du processus de construction de la solution.

Nous pouvons dégager ici une organisation à deux niveaux. D'une part les théorèmes effectivement désignés comme tels par Volterra sont fondés sur l'analyse des coefficients et contiennent des expressions explicites des substitutions. D'autre part un autre niveau du discours donne l'analyse du problème et le raisonnement général. Volterra s'appuie alors sur les théorèmes préliminaires (et quelques autres démontrés dans les parties précédentes) en reprenant l'aspect le plus synthétique de leur énoncé. Nous citons la démonstration qui est assez courte et qui illustre bien notre propos et le contraste que nous voulons faire apparaître avec le type de progression que nous avons étudiée jusqu'ici⁴⁵ :

Si formi

$$S_1 = \int X dx,$$

in cui nell'eseguire la integrazione deve supporre y costante. Avremo

$$\frac{dS}{dx} = \frac{dS_1}{dx} = X,$$

⁴⁴[Vol54, p.260]

⁴⁵[Vol54, p.265]

onde per un noto teorema (§1, teorema Xiii, p. 231) S e S_1 dovranno essere eguali a meno di un *fattore a destra* indipendente dalla x .

Ne segue che

$$S = S_1 T,$$

essendo T funzione della sola y . Quindi

$$S_1^{-1} S = T,$$

e derivando rispetto ad y , avremo

$$\frac{d}{dy}(S_1^{-1} S) = \frac{dT}{dy},$$

o anche, per un noto teorema (§1, teorema X, formula (14), p. 229), l'equazione precedente potrà scriversi

$$S_1^{-1} \left(\frac{dS}{dy} - \frac{dS_1}{dy} \right) S_1 = \frac{dT}{dy},$$

vale a dire

$$(2) \quad S_1^{-1} \left(Y - \frac{dS_1}{dy} \right) S_1 = \frac{dT}{dy}.$$

La démonstration se poursuit avec un changement de type de notations utilisées. Alors que Volterra utilise une écriture essentiellement synthétique avant ce point, les exceptions qui apparaissent maintenant permettent de préciser sa conception des substitutions et la nature de l'abstraction à l'œuvre :

Bisognerà dimostrare che essendo soddisfatta la (I) $[\Delta'(X, Y)_{xy} = 0]$ il primo membro della precedente equazione è indipendente da x . Perciò osserviamo che ponendo il primo membro eguale a $\begin{pmatrix} M & N \\ P & Q \end{pmatrix}$, pel teorema I di questo paragrafo, si avrà

$$\left\{ \begin{array}{cc} \frac{dM}{dx} & \frac{dN}{dx} \\ \frac{dP}{dx} & \frac{dQ}{dx} \end{array} \right\} = S_1^{-1} (\Delta'(Y, X)_{y,x}) S_1$$

onde per la (I)

$$\left\{ \begin{array}{cc} \frac{dM}{dx} & \frac{dN}{dx} \\ \frac{dP}{dx} & \frac{dQ}{dx} \end{array} \right\} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

Questo ci dimostra che il primo membro della (2) è indipendente da x . Integrando avremo

$$T = \int S_1^{-1} \left(Y - \frac{dS_1}{dy} \right) S_1 dy$$

e per conseguenza

$$T = S_1 \int \left\{ S_1^{-1} \left[Y - \left(\frac{dS_1}{dy} \right) \right] S_1 \right\} dy.$$

La première exception à l'écriture synthétique concerne la substitution dont les coefficients sont les dérivées des coefficients d'une substitution initiale T . Volterra n'utilisera en fait la notation T' dans son mémoire qu'à deux occasions : une fois dans la note insérée à la fin de son texte et à laquelle nous avons déjà fait allusion et enfin à la toute fin de son mémoire lorsqu'il qu'il aborde les substitutions d'ordre quelconque. Ici il réécrit chaque substitution de ce type en indiquant les coefficients dérivés.

La deuxième exception, qui concerne la substitution nulle (terme d'observateur), est importante dans notre étude qui cherche à préciser comment les mathématiciens que nous étudions s'appuient sur différentes conceptions d'opérations pour produire une vision générale.

Nous avons vu que Volterra faisait des calculs sur les coefficients des substitutions, et qu'il notait souvent les résultats obtenus en employant au moins partiellement des notations synthétiques. Mais lorsque un calcul donne pour chacun des coefficients 0, alors il l'indique quasi systématiquement en écrivant un tableau de zéros, comme dans l'égalité $\Delta'(X, Y)_{x,y} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$. Il faut en fait attendre une des dernières pages de son mémoire pour qu'il introduise une notation synthétique pour la substitution à coefficients tous nuls⁴⁶ :

[...] le condizioni affinché $[\prod_1^n S_r dx_r]$ sia un differenziale esatto si esprimeranno con

$$\Delta'(S_r, S_s)_{x_r, x_s} = 0 \quad (r, s = 1, 2, \dots, n),$$

ove col simbolo 0 si intende di rappresentare una sostituzione i cui elementi sono tutti nulli.

Cette notation, qui désigne ce que nous nommerions désormais la *substitution nulle*, n'est pas présentée par Volterra comme élément particulier d'une vision algébrique ou même dans son rapport aux opérations sur les substitutions. Elle sert simplement à désigner une substitution particulière, avec tous ses coefficients nuls. Une seule exception à cet emploi apparaît encore une fois dans la note ajoutée à la fin de l'ouvrage où 0 désigne toujours la *substitution nulle* qui cette fois-ci est le résultat d'un calcul mené exclusivement en utilisant des notations synthétiques et que l'on pourrait résumer en nos termes par : $A = B \implies A - B = 0$ où A et B sont des substitutions. Dans ce dernier exemple 0 ne sert donc pas seulement à désigner la *substitution nulle*, il est un élément d'un certain calcul algébrique des substitutions qui permet à Volterra de manipuler l'égalité et de faire la différence des deux membres par exemple.

Remarquons encore que Volterra reprendra ces questions dans un paragraphe ultérieur intitulé « Generalizzazione al caso di una sostituzione qualunque », où les substitutions seront de taille $n \times n$. On y observe une utilisation de plus en plus prononcée des notations synthétiques qui accompagnent le passage au cas général. Nous ne nous étendrons pas davantage sur le paragraphe qui concerne les substitutions d'ordre n pour reprendre la chronologie des résultats importants établis par Volterra. Plusieurs points que nous allons maintenant évoquer sont essentiels pour comprendre la progression

⁴⁶[Vol154, p.287]

des idées du mathématicien italien, non seulement dans ce texte mais aussi dans d'autres textes que nous étudierons et qui nous permettront de montrer l'élaboration d'un point de vue de plus en plus général et abstrait.

VARIATION DE L'INTÉGRALE D'UNE SUBSTITUTION. Le problème de la variation d'une intégrale consiste à étudier comment évolue l'intégrale d'une fonction lorsqu'on fait subir à cette dernière une petite perturbation. C'est un thème essentiel dans le développement de l'analyse fonctionnelle et il est important d'étudier les diverses façons d'aborder ce sujet dans les travaux de Volterra. Le mémoire de 1887 dédie une section à ce problème revisité grâce à la notion d'intégrale d'une substitution introduite précédemment.

Volterra définit tout d'abord la variation d'une substitution⁴⁷ :

Se agli elementi di una sostituzione $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ (det. 1) diamo delle variazioni infinitesime $\delta a, \delta b, \delta c, \delta d$, si chiamerà *variazione sinistra* della sostituzione, la sostituzione

$$\delta \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a + \delta a & b + \delta b \\ c + \delta c & d + \delta d \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1}$$

e *variazione destra*, la sostituzione

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \delta = \begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} \begin{pmatrix} a + \delta a & b + \delta b \\ c + \delta c & d + \delta d \end{pmatrix}.$$

Il nous faut ici préciser la nature des accroissements $\delta a, \delta b, \delta c$ et δd que Volterra utilise. Contrairement à ce qui est en jeu dans la dérivation par exemple, qui repose sur un accroissement de la variable d'une fonction ordinaire qui lui-même donne un accroissement de la valeur de la fonction, il s'agit ici d'un accroissement de fonction. Volterra redéfinira plusieurs fois la nature d'un tel accroissement dans des articles publiés à la même époque que nous étudierons plus loin⁴⁸, mais dans ce texte particulier sur les substitutions il ne trouve pas nécessaire de préciser le concept.

Une fois cette définition posée Volterra peut étudier ce qu'il nomme « variation d'une intégrale ». Pour cela, étant donnée l'intégrale $S = \int_p^q \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ pour laquelle $a + d = 0$, on affecte simultanément p, q , et chaque élément a, b, c, d d'une perturbation telle que $\delta(a + d) = 0$. Volterra obtient ainsi le théorème suivant⁴⁹ :

TEOREMA II.

$$\begin{aligned} \delta S &= \delta \int_p^q \begin{pmatrix} a & b \\ c & d \end{pmatrix} dx \\ &= \begin{pmatrix} 1 + a\delta q & b\delta q \\ c\delta q & 1 + d\delta q \end{pmatrix}_q S \begin{pmatrix} 1 - a\delta p & -b\delta p \\ -c\delta p & 1 - d\delta p \end{pmatrix}_p S^{-1} \cdot \left\{ T \begin{pmatrix} \delta a & \delta b \\ \delta c & \delta d \end{pmatrix} T^{-1} \right\} dx \int_p^q \end{aligned}$$

⁴⁷[Vol54, p.268]

⁴⁸Pour l'instant, au vu des textes de Volterra, on peut concevoir l'accroissement sous la forme suivante : pour chaque fonction f définie et continue sur un intervalle (p, q) et pour un $\epsilon > 0$ donné, une perturbation φ est elle-même une fonction continue dont la valeur absolue n'excède pas ϵ sur (p, q) . Dans ce cas $\delta f = \varphi = \epsilon\psi$ (avec des notations de Volterra) est donc une fonction continue que l'on peut choisir aussi petite qu'on le désire.

⁴⁹[Vol54, p.268 et 269]

essendo

$$T = \int_x^q \begin{pmatrix} a & b \\ c & d \end{pmatrix} dx.$$

La démonstration⁵⁰ permet de préciser comment Volterra utilise sa définition générale de variation d'une substitution pour étudier la variation de la substitution particulière S qui est une intégrale de substitution. Les premières lignes de la démonstration révèlent la conception de Volterra et montrent comment il interprète la notion de variation en faisant intervenir cette fois-ci tous les paramètres de l'intégrale⁵¹ :

Infatti

$$\delta S = \left\{ \int_{p+\delta p}^{q+\delta q} \begin{pmatrix} a + \delta a & b + \delta b \\ c + \delta c & d + \delta d \end{pmatrix} dx \right\} \left\{ \int_p^q \begin{pmatrix} a & b \\ c & d \end{pmatrix} dx \right\}^{-1}.$$

Volterra considère que l'expression $\int_p^q \begin{pmatrix} a & b \\ c & d \end{pmatrix} dx$ dépend non seulement de la substitution à l'intérieur de l'intégrale, et donc de chacun de ses coefficients, mais aussi de ses deux bornes. On a donc un calcul de la variation qui s'appuie à la fois sur la perturbation des différents paramètres en jeu ainsi que sur une forme spéciale adaptée aux substitutions. Avec nos notations on pourrait résumer cette approche sous la forme suivante : si $T = T(\alpha_1, \dots, \alpha_n)$ est une substitution qui dépend des paramètres $\alpha_1, \dots, \alpha_n$, alors $\delta T = T(\alpha_1 + \delta\alpha_1, \dots, \alpha_n + \delta\alpha_n) \cdot T^{-1}$. Comme on pourrait le voir dans la démonstration du théorème, Volterra utilise le symbole δ pour désigner des accroissements de natures différentes qui seront pourtant traités suivant les mêmes règles qui lui permettent de négliger les termes à partir d'un certain d'ordre.

L'étude des variations est importante dans une approche historique comme nous la menons car elle permet de distinguer les éléments auxquels le mathématicien attribue un rôle de variable et de comprendre comment il conçoit leur perturbation. Dans le travail de Volterra tel que nous l'analyserons dans la suite nous constaterons une évolution importante dans la conception de ce qui joue le rôle de la variable ou des éléments variables. Le cas que nous venons d'étudier permettra par contraste de préciser les idées de Volterra.

INTÉGRALES DOUBLES ET CURVILIGNES. Les intégrales doubles ou curvilignes mettent en jeu de nouveaux éléments géométriques qui seront à la fois importants dans le développement de la théorie pour les fonctions complexes mais aussi dans l'évolution de la conception générale de l'analyse telle que Volterra va la concevoir petit à petit. Nous employons ici le mot *géométrique* dans un sens large que les textes de Volterra vont nous permettre de préciser petit à petit.

Volterra introduit d'abord la notion d'intégrale double à parti d'un domaine (*campo*) σ du plan repéré tel que toute droite parallèle à l'axe des abscisses ne rencontre σ qu'en au plus deux points. Le « contour » (*contorno*) de σ est supposé

⁵⁰Nous ne redonnons pas ici la démonstration qui repose sur une utilisation astucieuse de la relation de Chasles pour les intégrales de substitutions.

⁵¹[Vol54, p.269]

être une ligne continue notée s . Le domaine plan σ est toujours représenté dans le premier quadrant du repère, et Volterra désigne par y_0 et y_1 « les distances minimales et maximales entre les points de s et l'axe des x . » Enfin une droite parallèle à l'axe des x coupe s en deux points M et N de même ordonnés y . Puis, étant donnée une substitution $\begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}$ dont les coefficients sont des fonctions de deux variables, Volterra considère la substitution dont les coefficients sont les intégrales des coefficients initiaux⁵² :

$$\begin{pmatrix} \alpha_1 & \beta_1 \\ \gamma_1 & \delta_1 \end{pmatrix} = \left\{ \begin{array}{cc} \int_M^N \alpha dx & \int_M^N \beta dx \\ \int_M^N \gamma dx & \int_M^N \delta dx \end{array} \right\}.$$

Enfin, en intégrant cette fois-ci au sens des substitutions et par rapport à la seconde variable y on obtient une nouvelle substitution $\begin{pmatrix} A & B \\ C & D \end{pmatrix} = \int_{y_0}^{y_1} \begin{pmatrix} \alpha_1 & \beta_1 \\ \gamma_1 & \delta_1 \end{pmatrix}$ que Volterra nomme « intégrale double »⁵³ :

Essa si chiamerà l'integrale doppio sinistro della sostituzione $\begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}$ eseguita prima rispetto ad x , poi rispetto ad y ed estesa a tutti i punti del campo σ . Si scriverà

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix} = \iint_{\sigma} \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} dx dy.$$

Volterra souligne que l'on peut procéder de même en intégrant (toujours à gauche) cette fois-ci d'abord par rapport à x puis par rapport à y lorsque les droites parallèles à l'axe des y ne rencontrent le contour s qu'au plus en deux points. On obtient alors une nouvelle intégrale double et Volterra souligne qu'en général $\iint_{\sigma} \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} dy dx \neq \iint_{\sigma} \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} dx dy$.

En ce qui concerne l'« intégrale curviligne » (*integrale curvilineo*), les choses sont présentées de façon beaucoup plus rapide par Volterra qui s'appuie sur la notion d'intégrale d'une substitution pour introduire $\int_A^B \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} ds$ pour deux points A et B de la courbe s . Bien qu'il ne précise pas sa conception on peut penser que cette intégrale est pensée grâce à une paramétrisation de la courbe s qui ramène immédiatement le concept à celui d'intégrale de substitution.

Enfin Volterra introduit une notation particulière en désignant par

$$\int_A^B \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} dx \cdot \begin{pmatrix} \alpha_1 & \beta_1 \\ \gamma_1 & \delta_1 \end{pmatrix} dy$$

l'intégrale curviligne suivante :

$$\int_A^B \left\{ \begin{array}{cc} \alpha \frac{dx}{ds} + \alpha_1 \frac{dy}{ds} & \beta \frac{dx}{ds} + \beta_1 \frac{dy}{ds} \\ \gamma \frac{dx}{ds} + \gamma_1 \frac{dy}{ds} & \delta \frac{dx}{ds} + \delta_1 \frac{dy}{ds} \end{array} \right\}.$$

⁵²[Vol54, p.270]

⁵³[Vol54, p.270]

C'est dans ce cadre que Volterra propose une version généralisée de ce que nous pourrions qualifier de *formule de Green*, qu'il nomme simplement « relation entre intégrale curviligne et intégrale double ». Le contexte géométrique est illustré par une figure⁵⁴ :

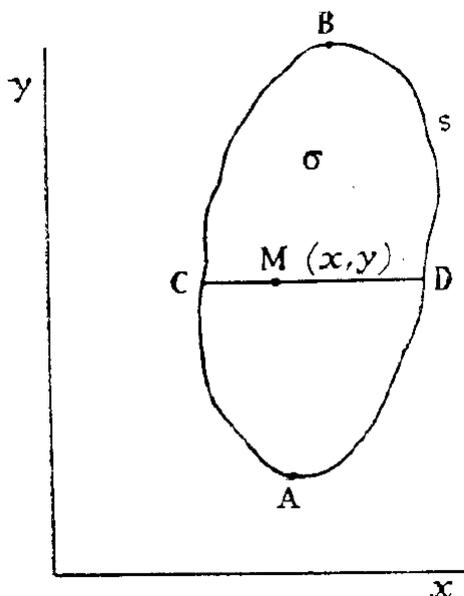


Fig. 4.

TEOREMA I. Siano X e Y (som. 0) due sostituzioni funzioni di x e di y definite nell'interno di un dato campo σ del piano xy e al contorno s (incontrato dalle parallele ad x al più in due punti).

Preso un punto qualunque M (di coordinate x, y) entro σ , si formi

$$\int_C^M X dx \cdot \int_A^C X dx \cdot Y dy = S_1$$

essendo MC parallela ad x , A il punto di s più vicino all'asse x . Avremo che

$$\int_{\sigma} (S_1^{-1} \Delta'(X, Y)_{x,y} S_1) dx dy = \int_s X dx \cdot Y dy$$

supponendo che la integrazione lungo la linea s sia eseguita partendo da A e ritornando allo stesso punto lasciando sempre a sinistra il campo σ .

La démonstration du théorème est écrite dans une notation synthétique à l'exception des deux cas que nous avons déjà soulignés : l'écriture de la substitution $\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$ et de la substitution composée des dérivées des coefficients

$$\begin{pmatrix} \frac{dM}{dx} & \frac{dM}{dx} \\ \frac{dM}{dx} & \frac{dM}{dx} \end{pmatrix}.$$

⁵⁴[Vol154, p.272]

Volterra en déduit alors un important théorème de monodromie qui indique que, sous les hypothèses adéquates de régularité et de connexité, l'intégrale curviligne d'une substitution ne dépend que des extrémités de la ligne d'intégration⁵⁵ :

TEOREMA II. Se X e Y (som. 0) sono due sostituzioni funzioni di x e y definite entro un campo semplicemente connesso Σ nel piano delle x, y ed esse e le derivate dei loro elementi sono finite e continue; e se

$$\Delta'(X, Y)_{x,y} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

avremo, qualunque sia la linea s chiusa che si prenda entro Σ ,

$$\int_s X dx \cdot Y dy = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Là encore la démonstration est donnée en désignant de façon synthétique les substitutions et une nouvelle exception à cet aspect synthétique apparaît avec la substitution $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$. Celle-ci intervient à la fois comme résultat d'un calcul et pour exprimer l'indépendance de ce calcul par rapport au choix de la ligne d'intégration. Mais Volterra ne met jamais en avant de propriété calculatoire de cette substitution et ne trouve donc pas nécessaire de la désigner de façon spéciale. Ces précisions relatives au choix des notations plus ou moins synthétiques permettent de comprendre la conception de Volterra. Qualifier dès l'abord en des termes algébriques (d'aujourd'hui ou de l'époque de Volterra) les éléments impliqués dans les calculs ne permet pas de rendre compte de la façon dont Volterra conçoit et manipule les substitutions, leurs opérations et leur multiplication qui joue un rôle particulier comme nous l'avons vu.

PARAMÈTRES DIFFÉRENTIELS D'ORDRE 2. Dans cette section Volterra étend encore sa théorie en mettant en place des outils qui serviront entre autres au développement d'une version complexe.

La notion de « paramètre différentiel », d'ordre un ou deux, a été introduite par Lamé dans l'étude mathématique de la théorie de l'élasticité. Dans [Taz97] Rossana Tazzioli précise le rôle important de ces paramètres depuis leur introduction par Lamé⁵⁶ :

Curvilinear coordinates and differential parameters were both introduced by Lamé, who also attached a geometric and physical meaning to these expressions. [...]

In order to deduce the equations of elastic equilibrium for the ether, Lamé introduced two differential invariant expressions of a function F , which he called the first and second differential parameters,

$$(1.1) \quad \Delta_1 F = \sqrt{\left(\frac{\partial F}{\partial x_1}\right)^2 + \left(\frac{\partial F}{\partial x_2}\right)^2 + \left(\frac{\partial F}{\partial x_3}\right)^2}$$

⁵⁵[Vol54, p.274]

⁵⁶[Taz97, p.26-27]

$$(1.2) \quad \Delta_2 F = \frac{\partial^2 F}{\partial x_1^2} + \frac{\partial^2 F}{\partial x_2^2} + \frac{\partial^2 F}{\partial x_3^2},$$

where x_1, x_2, x_3 are Cartesian orthogonal coordinates. He then expressed (1.1) and (1.2) in curvilinear coordinates and stressed their important role in mathematical physics.

Ces idées et les notations furent reprises par de nombreux mathématiciens qui les firent évoluer et les généralisèrent à différentes situations. En particulier l'étude de Tazzioli [Taz97] se concentre sur les travaux de Beltrami qui donne une importance notable aux paramètres différentiels dont il revisite la définition. Tazzioli montre que Beltrami généralise la notion de paramètre différentiel qu'il utilise dans diverses études de géométrie différentielle mais aussi dans des problèmes de théorie du potentiel ou de la théorie de l'élasticité⁵⁷ :

By means of differential parameters, Beltrami expressed many geometric, analytic, and mechanical quantities in order to interpret these quantities as intrinsic elements of a surface and, therefore, to point out their independence of coordinates.

Volterra, très intéressé par ces questions, cherche un moyen d'étendre la notion de paramètre différentiel avec pour objectif premier de généraliser la notion de « fonction conjuguée ». Nous verrons à d'autres moments et dans d'autres textes comment diverses notions revisitées de paramètres différentiels interviennent de façon importante dans le travail de Volterra et sa recherche d'une vision générale pour l'analyse.

Il y a plusieurs façons de généraliser une théorie, et dans le mémoire de Volterra que nous étudions, c'est sur la notion de substitution que s'appuie cette généralisation. Là encore nous avons vu qu'il y avait plusieurs voies de généralisation, l'une mettant en jeu le nombre de variables des coefficients de la substitution alors que l'autre joue sur l'ordre de la substitution, c'est-à-dire le nombre des coefficients qui la composent.

Ici Volterra considère les substitutions dont les coefficients sont des fonctions de 2 variables et étend aux substitutions d'ordre quelconque (fini) la notion de paramètre différentiel d'ordre 2.

En partant d'une substitution de déterminant unitaire $S = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$ dont les coefficients sont des fonctions de 2 variables, Volterra introduit les substitutions $X = \frac{dS}{dx}$ et $Y = \frac{dS}{dy}$ qui lui permettent de concevoir le « paramètre différentiel de second ordre à gauche » qu'il définit par l'égalité $\Delta'_2 S := \Delta'(-Y, X)_{x,y}$ ⁵⁸. Comme toujours dans cette théorie il y a un énoncé analogue en prenant les dérivées à droite.

Les coefficients de la substitution Δ'_2 sont explicités en fonction des dérivées des coefficients de la substitution S , mais les résultats de ce paragraphe sont surtout obtenus en utilisant l'égalité exprimée sous forme synthétique $\Delta'_2(X, Y) = \Delta'(-Y, X)$ qui ramène au premier paramètre différentiel et à la caractérisation des formes exactes que nous avons rappelée plus haut.

⁵⁷[Taz97, p.40]

⁵⁸Volterra rappelle que l'on a, avec les substitutions considérées, $\Delta'(X, Y) = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$.

Volterra obtient entre autres l'équivalent des fonctions conjuguées sous une forme généralisée⁵⁹ :

TEOREMA IV. Se si ha la sostituzione S (det. 1) tale che

$$\Delta'_2 S = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

esisterà sempre una sostituzione coniugata S_1 (det. 1) tale che

$$\Delta'_2 S_1 = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

e

$$\frac{dS}{dx} = \frac{dS_1}{dx} \quad , \quad \frac{dS}{dy} = -\frac{dS_1}{dy}.$$

Il s'agit d'un résultat important dans la vision de Volterra qui considère que la notion de « fonction conjuguée » est fondamentale comme point d'appui pour développer une théorie généralisée des fonctions. Nous verrons par la suite comment il proposera une nouvelle façon de penser l'extension du concept de fonction et s'appuiera à nouveau sur la notion de paramètre différentiel et de fonction conjuguée.

3.2.6 EXTENSION AUX NOMBRES COMPLEXES

Nous avons rappelé que le traité de Volterra avait été publié en deux parties. Dans la seconde partie Volterra expose sa théorie dans le cas des fonctions de la variable complexe. Ce deuxième pan du travail qui étoffe le fondement de la théorie des équations différentielles linéaires⁶⁰ n'est publié qu'en 1889 mais Volterra précise en note de bas de page que cette seconde partie était prête dès le printemps 1887. Il rajoute que la majorité des résultats avaient d'ailleurs été publiés dans une note aux *Rendiconti del Circolo Matematico di Palermo* au mois de mars 1888. Dans cette note Volterra souligne que le premier mémoire exposait en fait une « étude préliminaire à la théorie des substitutions dont les coefficients sont des fonctions de la variable complexe »⁶¹. Nous allons maintenant voir comment Volterra s'appuie sur l'étude du cas réel pour développer la théorie dans le cas de la variable complexe. Ce prolongement nous permet aussi de mieux cerner la portée générale de la version réelle de la théorie.

Pour commencer Volterra remarque que tous les résultats et les outils développés pour les substitutions de la variable réelle restent valides que l'on prenne comme coefficients des fonctions à valeurs dans \mathbb{R} ou dans \mathbb{C} .

Pour adapter la théorie aux substitutions dont les coefficients sont des fonctions de la variable complexe, Volterra considère les fonctions de la variable $z = x + iy$ comme fonctions des deux variables x et y , se ramenant ainsi à l'étude du cas de plusieurs variables réelles. En s'appuyant sur les propriétés des coefficients et sur la notion de dérivée partielle d'une substitution S Volterra

⁵⁹[Vol54, p.280]

⁶⁰[Vol56, p.383-451]

⁶¹[Vol54, p.351]

montre que les coefficients de la substitution $\frac{\partial S}{\partial x}$ coïncident avec ceux de $\frac{\partial S}{\partial y}$ divisés par le complexe i . Cette propriété sera « représentée symboliquement » par l'égalité⁶² :

$$\frac{\partial S}{\partial x} = \frac{\partial S}{i\partial y}.$$

Volterra définit alors la dérivée complexe comme la valeur commune de ces deux dérivées partielles et il note $\frac{\partial S}{dz} = \frac{\partial S}{\partial x} = \frac{\partial S}{i\partial y}$. Comme S est inversible et d'après la définition de la dérivée des substitutions de la variable réelle on peut écrire, comme le souligne Volterra, $\frac{\partial S}{\partial x} = T_1 S^{-1}$ et $\frac{\partial S}{\partial y} = T_2 S^{-1}$ où T_1 et T_2 sont les substitutions composées des coefficients dérivés de S par rapport à x et à y respectivement. L'égalité $T_1 = \frac{1}{i} T_2$ se lit aussi au niveau des coefficients et exprime que les coefficients de S (de la variable complexe) sont des « fonctions monogènes ».

En reprenant le formalisme de la première partie et en notant α_{ij} les coefficients de $\frac{\partial S}{\partial x}$ Volterra définit la différentielle à gauche⁶³ :

$$dS = \begin{pmatrix} 1 + \alpha_{11}dz & \alpha_{12}dz & \dots & \alpha_{1n}dz \\ \alpha_{21}dz & 1 + \alpha_{22}dz & \dots & \alpha_{2n}dz \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_{n1}dz & \alpha_{n2}dz & \dots & 1 + \alpha_{nn}dz \end{pmatrix} = \frac{dS}{dz} \cdot dz.$$

Puis il indique que les règles de calcul qu'il a démontrées pour la dérivation des substitutions de la variable réelle restent ici valables.

Enfin, pour introduire l'intégrale dans le cas de la variable complexe en suivant le schéma classique pour les fonctions ordinaires, il faut séparer la partie réelle et la partie imaginaire de la différentielle, ce que Volterra accomplit en s'appuyant sur un produit de substitutions. Il écrit d'abord pour une substitution T de coefficients $\alpha_{i,j}$ à variables complexes⁶⁴ :

$$\begin{aligned} Tdz &= \begin{pmatrix} 1 + \alpha_{11}dz & \alpha_{12}dz & \dots & \alpha_{1n}dz \\ \alpha_{21}dz & 1 + \alpha_{22}dz & \dots & \alpha_{2n}dz \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_{n1}dz & \alpha_{n2}dz & \dots & 1 + \alpha_{nn}dz \end{pmatrix} \\ &= \begin{pmatrix} \alpha_{11}dz & \alpha_{12}dz & \dots & \alpha_{1n}dz \\ \alpha_{21}dz & \alpha_{22}dz & \dots & \alpha_{2n}dz \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_{n1}dz & \alpha_{n2}dz & \dots & \alpha_{nn}dz \end{pmatrix} dx \cdot \begin{pmatrix} i\alpha_{11}dz & i\alpha_{12}dz & \dots & i\alpha_{1n}dz \\ i\alpha_{21}dz & i\alpha_{22}dz & \dots & i\alpha_{2n}dz \\ \vdots & \vdots & \ddots & \vdots \\ i\alpha_{n1}dz & i\alpha_{n2}dz & \dots & i\alpha_{nn}dz \end{pmatrix} dy \\ &= T_1 ds \cdot T_2 dy. \end{aligned}$$

⁶²[Vol56, p.394]

⁶³[Vol56, p.395]

⁶⁴[Vol56, p.396]

Cette décomposition reprend les notations de la partie 1 où, cette fois-ci avec des notations modernes, $T_1 dx = I + (\alpha_{i,j} dx)_{i,j}$ et l'égalité se conçoit en laissant tomber dans le produit les termes d'ordre 2 en $dx dy$.

Volterra peut alors définir l'intégrale le long d'une ligne s sous la forme $\int_s T dz := \int_s T_1 dx \cdot T_2 dy$. Les substitutions qui interviennent dans le membre de droite ont des coefficients (à valeurs dans \mathbb{C}) dont la variable est réelle et l'intégrale est donc bien définie au sens de la première partie du mémoire de Volterra.

La dérivation et l'intégration des substitutions de la variable complexe sont donc introduits grâce à des processus qui font intervenir des calculs que Volterra exprime en s'appuyant sur la première partie de son travail avec des notations souvent synthétiques, c'est-à-dire sans recours aux coefficients. Aucune des deux opérations n'est présentée directement sous la forme d'une matrice dont les coefficients seraient calculés grâce à des notions de calcul différentiel ou intégral des fonctions ordinaires. En ce sens on peut donc dire que le calcul différentiel et intégral des substitutions est présenté de façon autonome dans le mémoire de Volterra, autonomie qui entre en ligne de compte dans notre appréciation du type de généralité que Volterra met à l'œuvre.

Mais le texte de Volterra propose aussi un autre type de définition qui est fondé sur des propriétés des coefficients sans mettre en jeu d'analyse au niveau des substitutions. En effet, rapidement après avoir introduit la dérivation et l'intégration, Volterra définit une « substitution holomorphe » dans un « champ » σ (*campo*) lorsque tous les coefficients, qui sont des fonctions classiques de la variable complexe, sont holomorphes.

On retrouve ce type de définition ensuite tout au long du texte de Volterra qui définira les propriétés des substitutions en s'appuyant sur une propriété partagée par tous leurs coefficients. Ainsi une substitution est « monodrome » sur une surface de Riemann si chacun des coefficients est monodrome, elle est « régulière » ou « algébrique » si chacun des coefficients n'a d'autres singularités que ses pôles, et ainsi de suite.

Le texte exposé dans cette partie sur les substitutions de la variable complexe fait donc apparaître une pensée qui s'appuie sur une certaine hiérarchisation qui induit une vision abstraite. Dans la première partie, ce sont essentiellement des calculs effectués sur les coefficients qui ont permis d'obtenir des résultats sur les substitutions. Nous avons montré que ces résultats étaient alors souvent exprimés sous une forme qui ne faisait plus référence aux coefficients, que nous avons nommée *forme synthétique*. Le passage à la variable complexe prolonge la progression et installe un niveau d'abstraction supplémentaire. Les conditions exprimées en termes de coefficients (holomorphie etc.) garantissent juste les conditions d'applicabilité des théorèmes de la première partie, lesquels seront utilisés désormais sous leur forme synthétique pour développer les nouvelles notions et les théorèmes relatifs à la variable complexe. Certaines hypothèses s'expriment donc en termes de coefficients et définissent des classes de substitutions (mots d'observateur) pour lesquelles Volterra développe des théorèmes dont les énoncés et les démonstrations n'utilisent plus que des notations synthétiques.

Cette façon de procéder n'est cependant pas totalement uniforme et une partie de l'exposé s'appuie encore sur une analyse des coefficients comme dans la première partie pour démontrer des résultats en effectuant des calculs explicites sur les éléments de la substitution.

Volterra utilise ce type de raisonnement par exemple dans une démonstration du « théorème de Cauchy relatif aux substitutions »⁶⁵ :

TEOREMA I. Se T (som. 0) è una sostituzione olomorfa di z entro un campo semplicemente connesso σ , e s è una linea chiusa entro σ , avremo

$$\int_{\sigma} T dz = I.$$

La démonstration proposée par Volterra repose sur le fait que l'on peut écrire $T dz = T_1 dx + T_2 dy$ et sur la condition d'holomorphie des coefficients qui implique que $\Delta(T_1, T_2)_{x,y} = 0$. Volterra peut alors appliquer un théorème de la première partie sur les substitutions de deux variables réelles⁶⁶ :

Quindi applicando un teorema dell'art. 4, § 6, della prima parte, avremo che, presa entro σ una linea s chiusa qualunque, $\int_s T dz$ sarà eguale alla identità.

La progression des idées est encore davantage mise en évidence lorsque l'on note les légères différences entre la formulation initiale du théorème dans la première partie et celle qu'utilise ici Volterra. D'une part l'égalité $\Delta(X, Y) = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$ est reformulée $\Delta(X, Y) = 0$. D'autre part la substitution qui apparaissait sous la forme $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ est désormais notée I et nommée « identité ».

Bien que la suite de l'exposé de Volterra reprenne une approche qui met à nouveau en jeu les coefficients des substitutions, un certain nombre de théorèmes obtenus ici seront aussi formulés et démontrés de façon plus synthétique dans la note que nous avons déjà citée « Sulla teoria delle equazioni differenziali lineari »⁶⁷. En particulier Volterra y donne une approche qui se sert de la « réduction des substitutions à une forme normale », qu'il a développée dans la première partie à partir d'un théorème de Weierstraß, avec laquelle il établit le théorème de Cauchy et énonce une théorie des résidus sous une forme synthétique.

Ces exemples illustrent la progression de Volterra dans l'étude des substitutions où il organise les idées en partant d'une conception où les coefficients lui permettent de définir les premières notions et les premières propriétés des substitutions, vers une formulation de plus en plus *synthétique* qui marque une approche abstraite.

3.2.7 LIEN AVEC LES ÉQUATIONS DIFFÉRENTIELLES

Le texte de Volterra publié sous la forme d'un mémoire en deux parties constitue à ses yeux un véritable « fondement de la théorie des équations

⁶⁵[Vol56, p.401]

⁶⁶[Vol56, p.401]

⁶⁷[Vol54, pp.351–355]

différentielles linéaires » comme le précise le titre. Nous avons rappelé l'optique générale adoptée par Volterra dans ce texte et nous avons étudié plus particulièrement la mise en place d'un calcul différentiel général pour les substitutions. Néanmoins il nous reste à préciser la nature exacte de ce texte dans son rôle de « fondement ».

Un des aspects importants du point de vue que développe Volterra est de considérer la dérivation et l'intégration comme deux opérations inverses l'une de l'autre. Cette idée est présente aussi bien dans le cas de la variable réelle que dans le cas de la variable complexe. Si on la pense comme centrale elle permet de concevoir l'équilibre du texte de Volterra et la cohérence des thèmes que nous avons évoqués plus haut pour fonder une théorie.

Cependant à aucun moment Volterra ne dédie de partie de son mémoire aux équations différentielles linéaires à proprement parler. Dès le premier volet de son étude Volterra dégage dans une longue introduction les idées principales qu'il va développer et précise le plan : là encore il n'y a aucun résultat relatif aux équations différentielles linéaires. Volterra dans son exposé dissocie totalement ce qu'il nomme « fondement », qui est en ses termes un « calcul infinitésimal » des substitutions, de la « théorie des équations différentielle linéaires » qui s'appuiera dessus. En fait Volterra publiera de façon séparée une note intitulée « Sulle equazioni differenziali lineari »⁶⁸, parue la même année 1887, dans laquelle il précise son point de vue. Ce texte permet de mieux comprendre la notion de « fondement » que met en avant le titre du mémoire.

Dans cette note Volterra prend comme point de départ un texte de Riemann sur les séries de Gauss publié en 1857 sous le titre « Beiträge zur Theorie der durch die Gauss'sche Reihe $F(\alpha, \beta, \gamma, x)$ darstellbaren Functionen »⁶⁹. Riemann y développe une approche qui met en avant des substitutions linéaires, qui sera reprise par de nombreux mathématiciens, et que Volterra convoque à son tour⁷⁰ :

Se si segue la teoria delle equazioni differenziali nel suo svolgersi, si può notare che già nelle Memorie di RIEMANN si manifesta il suo stretto legame colla teoria delle sostituzioni, e che i successivi lavori sullo stesso argomento hanno sempre più posto in evidenza tale relazione. Nei lavori di FUCHS, KLEIN, POINCARÉ, JORDAN e di molti altri, si ha continuamente ricorso alla teoria delle sostituzioni per trattare delle questioni sulle equazioni differenziali.

Mais Volterra voit dans l'utilisation des substitutions un outil essentiel qui dépasse la conception de ses prédécesseurs et qu'il va concevoir comme fondement d'une théorie générale. Pour lui il y a un lien intime entre la théorie des équations différentielles linéaires et celle des substitutions qui est révélé par le « calcul infinitésimal » des substitutions⁷¹ :

Si possono infatti trovare due operazioni infinitesimali sulle sostituzioni (i cui elementi si immaginano variabili) analoghe alla derivazione e alla integrazione ordinarie, le quali danno direttamente

⁶⁸[Vol54, p.291-293]

⁶⁹[Rie76, p.61-78]

⁷⁰[Vol54, p.291]

⁷¹[Vol54, p.292]

il passaggio dagli integrali fondamentali di una equazione differenziale lineare ai suoi coefficienti, e inversamente dai coefficienti agli integrali fondamentali.

Dans cette conception fondée sur l'intégration et la dérivation des substitutions vues comme deux opérations infinitésimales, Volterra insiste sur la relation qui les unit et qui constitue un outil puissant d'analyse des problèmes liés aux équations différentielles linéaires⁷² :

Le operazioni così stabilite di derivazione e di integrazione sono inverse una dell'altra, vale a dire se si integra a destra una sostituzione, e poi considerando la sostituzione integrale come funzione del limite superiore dell'intervallo di integrazione, si deriva a destra, si ritrova la sostituzione primitiva, e lo stesso vale per le integrazioni e derivazioni a sinistra.

Une fois rappelée l'existence de ces deux opérations, dont il ne donne pas les détails de construction ici, Volterra met en avant seulement un résultat qu'il qualifie de « fondamental »⁷³ :

Come teorema fondamentale si ha che la derivata a destra di una sostituzione non varia se si moltiplica a sinistra la sostituzione per una sostituzione costante, e che tutte le sostituzioni che hanno per derivata a destra una stessa sostituzione, debbono differire per sostituzioni costanti che moltiplicano a sinistra.⁷⁴

Enfin Volterra précise le lien entre un calcul différentiel et intégral des substitutions et les équations différentielles linéaires⁷⁵ :

La integrazione di una equazione differenziale lineare omogenea di un ordine qualunque può ridursi alla integrazione di una sostituzione. Così l'integrale sinistro della sostituzione

$$T = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & 1 \\ p_n & p_{n-1} & p_{n-2} & \cdots & p_2 & 0 \end{pmatrix}$$

ove p_2, p_3, \dots, p_n sono funzioni di x , è la sostituzione

$$S = \begin{pmatrix} v_1 & v_2 & v_3 & \cdots & v_{n-1} & v_n \\ v'_1 & v'_2 & v'_3 & \cdots & v'_{n-1} & v'_n \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ v_1^{(n-1)} & v_2^{(n-1)} & v_3^{(n-1)} & \cdots & v_{n-1}^{(n-1)} & v_n^{(n-1)} \end{pmatrix}$$

⁷²[Vol54, p.292]

⁷³[Vol54, p.293]

⁷⁴En nos termes, si C est une substitution constante et S est dérivable (toutes de déterminant

1) alors $\frac{d}{dx}SC = \frac{d}{dx}S$. Rappelons que l'on a aussi $\frac{d}{dx}CS = C\frac{d}{dx}SC^{-1}$. Et si T_1 et T_2 vérifient

$\frac{d}{dx}T_1 = \frac{d}{dx}T_2$ alors il existe une substitution constante C telle que $T_1 = T_2C$.

⁷⁵[Vol54, p.293]

ove v_1, v_2, \dots, v_n rappresentano un sistema di integrali fondamentali della equazione differenziale

$$y^{(n)} = p_2 y^{(n-2)} + p_3 y^{(n-3)} + \dots + p_{n-1} y' + p_n y.$$

La notion de substitution permet donc de relier via un calcul infinitésimal les coefficients de l'équation différentielle avec un système fondamental de solutions. La relation, que Volterra résume sous la forme $S = \int T dx$ et $T = \frac{dS}{dx}$, permet de concevoir qu'une théorie des équations différentielles linéaires peut être *fondée* sur un calcul différentiel des substitutions. La méthode consiste donc à essayer de généraliser ou d'adapter les résultats du calcul différentiel et intégral des fonctions au cas des substitutions. Chacun des nouveaux théorèmes obtenus pourra être interprété en termes d'équation différentielle linéaire.

3.2.8 CONCLUSION

Cette étude nous a tout d'abord permis d'analyser précisément les éléments que convoquait Volterra et de mettre en évidence comment il les organisait en un calcul différentiel et intégral pour les substitutions. Nous avons fait apparaître deux aspects importants.

D'une part Volterra choisit d'isoler une seule opération pour les substitutions, la multiplication, qu'il envisage sous un aspect général. Celle-ci est générale car elle permet de concevoir les quatre opérations usuelles algébriques sur les fonctions comme des cas particuliers de la multiplication des substitutions. Mais elle est aussi générale car elle permet de *composer* de nouvelles substitutions, et de définir des relations entre diverses substitutions. Nous avons vu que l'on ne pouvait pas en déduire de *structure algébrique* au niveau des substitutions bien que l'on ait ici des ingrédients que nous pouvons relier à des éléments ou des pratiques de l'algèbre. Un certain nombre d'aspects essentiels pour définir à proprement parler une structure algébrique sont en effet absents.

Des éléments typiques, qui sont pourtant employés, comme la substitution $\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$ ou $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, ne sont pas identifiés dans leur rôle algébrique c'est-à-dire dans le rôle particulier qu'ils jouent par rapport à la multiplication.

Il manque aussi un regard réflexif de Volterra sur lequel on pourrait s'appuyer pour penser en termes de structure dans ce texte. En particulier Volterra n'isole pas de propriétés particulières de l'ensemble des substitutions et de la multiplication qu'il sélectionne.

Mais ici, comme souvent, nous avons aussi montré que l'on était au cœur d'un processus évolutif organisé par Volterra et que les manques que nous avons soulignés juste avant tendaient à être comblés petit à petit, ce mouvement allant de pair avec l'emploi de plus en plus fréquent de notations que nous avons qualifiée de *synthétiques*. En particulier les deux substitutions que nous avons isolées vont être finalement désignées par les symboles 0 et I , et Volterra va préciser leurs rôles, non pas dans la relation à la multiplication, mais pour signifier que deux substitutions coïncident ou qu'un résultat dont la forme est une substitution est constant.

D'autre part Volterra développe un calcul qu'il nomme d'abord « infinitésimal » subtilement adapté de celui des fonctions à partir des notions de dérivée et d'intégrale. Bien qu'un enjeu important soit de penser les deux opérations comme inverse l'une de l'autre, leur développement fait apparaître des idées propres à l'une ou l'autre.

La dérivation qui est introduite en premier a nécessité de définir une façon d'évaluer ce que nous avons appelé la *taille* d'une substitution et l'*écart* entre deux substitutions, qui sont tous deux définis à partir des coefficients des substitutions. C'est sur ces notions que Volterra s'appuie pour définir la valeur limite d'une substitution qui joue le rôle de taux d'accroissement et donne la dérivée dans un premier temps, et la différentielle ensuite.

De plus la notion de taux d'accroissement fait apparaître un pseudo rapport différentiel qui met en jeu la fonction particulière $x \mapsto x$ et sa différentielle dx .

De son côté l'intégration est définie suivant une approche riemannienne adaptée à la multiplication⁷⁶. Elle est fondée sur une subdivision de plus en plus fine de l'intervalle d'intégration et par la construction d'un produit. Chaque intervalle de la subdivision donne un facteur qui est construit à partir de l'évaluation des coefficients sur cet intervalle et de la longueur de l'intervalle.

Nous avons aussi mis en évidence plusieurs aspects qui ont trait au général dans ce texte. La généralité s'y exprime de différentes façons et dans différentes directions que Volterra explore.

Une première voie de généralisation consiste à envisager des substitutions d'ordre quelconque fini n . Nous avons montré que le passage de l'étude des substitutions d'ordre 2, où l'on peut expliciter les coefficients, à l'ordre n quelconque était couplé avec l'introduction de notations *synthétiques* où les substitutions sont désignées par une lettre sans que soit fait référence aux coefficients.

Une seconde voie de généralisation est celle qui consiste à prendre des coefficients de plus en plus généraux. D'abord Volterra envisage rapidement que les coefficients puissent prendre leurs valeurs aussi bien dans \mathbb{R} que dans \mathbb{C} . Les résultats établis dans un cas restent valides dans l'autre. Puis Volterra, dans l'ordre, envisage des substitutions dont les coefficients sont des fonctions de plusieurs variables, puis complexes.

L'étude de la variable complexe, comme nous l'avons vu, se ramène principalement au cas de deux variables. Cependant encore une fois le passage aux nombres complexes comme variables s'accompagne d'une écriture beaucoup plus synthétique et petit à petit les énoncés ne font plus appel qu'à des notations où la variable est dénotée z et ne met plus en évidence telle ou telle interprétation en termes de paires de nombres réels $x + iy$.

Enfin, du point de vue de l'observateur, le texte de Volterra se présente comme une adaptation du calcul différentiel et intégral des fonctions ordinaires aux substitutions. On peut donc le concevoir comme une extension du domaine de l'étude des fonctions à de nouveaux éléments que sont les substitutions. Cette extension mise en place par Volterra, qui ne redéfinit pas la notion de fonction

⁷⁶De nombreux auteurs mentionnent que c'est Volterra qui a inventé ici ce que l'on nomme désormais l'*intégration produit*. Volterra a écrit un livre en collaboration avec Hostinsky dès 1938 sur le sujet [VH38]. On peut aussi consulter [DF11] pour apprécier l'étendue des développements de cette théorie assez peu connue.

(il n'y a pas en particulier ici de fonctions à valeurs substitutions), redéfinit le domaine d'interprétation et les frontières d'un paradigme fonctionnel fondé sur certains éléments qui ont été sélectionnés par le mathématicien : opérations algébriques, notion d'*écart*, dérivée, intégrale. Ces éléments définissent le type de généralité que Volterra installe dans son travail sur les substitutions. L'objet de la suite de notre étude est de mettre au jour une évolution de la vision de Volterra qui s'accompagne d'une évolution du type de généralité que ses textes permettent d'apprécier.

3.3 LA « FONCTION DE LIGNE » : UN CONCEPT GÉNÉRAL POUR L'ANALYSE

La notion de *fonction de ligne*⁷⁷ est une notion phare de l'œuvre de Volterra qu'il introduit en 1887 dans une série de 5 notes que nous allons étudier. On la retrouve dans la plupart de ses publications par la suite où il s'en sert pour revisiter et généraliser des conceptions mathématiques dans des champs *a priori* très divers.

En fait, bien que l'expression « fonction de ligne » devienne rapidement célèbre dans la communauté mathématique à l'époque de Volterra, nous verrons que le mathématicien italien définit d'abord une notion de *fonction qui dépend d'autres fonctions* dont la *fonction de ligne* est une version géométrique.

Avant de préciser les diverses notions que Volterra dégage dans les publications de 1887, on peut penser, le temps de cette introduction, aux *fonctions de lignes* comme des fonctions qui, à des lignes tracées sur une surface ou à des fonctions définies sur un intervalle, font correspondre un nombre réel ou complexe. Nous préciserons ces idées avec les textes de Volterra en analysant la nature de la correspondance et en étudiant comment le mathématicien pense les lignes ou les fonctions auxquelles il associe des valeurs par la *fonction de ligne*⁷⁸.

Dans des commentaires ultérieurs à ces publications initiales, Volterra dira avoir commencé à élaborer l'idée de *fonction de ligne* dès 1883⁷⁹. Il est difficile d'identifier dans les textes publiés entre 1883 et 1887 ce que nous pourrions véritablement établir comme embryon des idées qui conduisirent Volterra à la notion de *fonction de ligne*. Plusieurs études qui mentionnent les travaux de Volterra sur les *fonctions de lignes* notent qu'il travaillait depuis le début de sa carrière sur des problèmes que nous pouvons désormais rattacher à l'idée de

⁷⁷Nous choisissons ici d'écrire l'expression *fonction de ligne* avec le mot *ligne* au singulier. Ce n'est pas le choix initial de Volterra, ni le choix de beaucoup de mathématiciens qui s'approprièrent la notion de Volterra. Dans sa première utilisation Volterra écrit « fonction des lignes » (*funzione delle linee*). Ce n'est que plus tard que Volterra utilisera l'expression « fonction de ligne » (*funzione di linea*) quand il voudra désigner un concept général (cf. [Vol57, p.549]). On retrouve cette forme en français dans un texte publié en 1914 : « Nous allons passer à une nouvelle classe de problèmes qui ressortent du concept fondamental de fonction de ligne. » (cf. [Vol60, p.34])

⁷⁸Nous n'employons pas ici le terme de *fonctionnelle* qui sera proposé pour la première fois par Hadamard car il induit chez le lecteur de notre époque une certaine vision des éléments en jeu qui sont justement au cœur de notre analyse et que Volterra lui-même va définir et faire évoluer.

⁷⁹Par exemple, Volterra écrira en 1921 : « Ho immaginato da prima, nel 1883, quando ho cominciato a dare una forma concreta alle mie idee, il concetto di funzione che dipende da tutti i valori di un'altra funzione [...] » [Vol60, p.393]

fonction de ligne. Il est souvent hasardeux d'y voir plus qu'un contexte général propice à penser en termes de fonctions, de courbes souvent présentes dans les études de physique mathématique et des problèmes d'équations différentielles ou intégrales qui s'y rattachent.

Les thèmes mathématiques qui intéressent alors Volterra ont été décrits par exemple dans l'article « Lines on the horizon » de Guerraggio, Jaëck et Mazliak⁸⁰ : on y retrouve les idées de Volterra d'abord sur l'étude des fonctions, puis sur des problèmes ayant trait à la physique mathématique, aux équations aux dérivées partielles, à l'intégration ou aux variations. Volterra lui-même commentera *a posteriori* l'existence sous-jacente des *fonctions de lignes* dans de nombreuses situations de la physique et décrira ses propres travaux dans une vision unifiée où la notion de *fonction de ligne* semble devoir s'imposer naturellement. Dans leur livre intitulé *Vito Volterra*⁸¹ Guerraggio et Paolini mettent en évidence cet aspect en rappelant un commentaire de Volterra de 1921 que nous citons à notre tour⁸² :

Ero compreso dalla necessità di considerare le funzioni di linee, poichè molti fenomeni naturali conducono all'esame di quantità che dipendono da un numero infinito di variabili. Anche molti problemi di analisi portano alle stesse quantità.

Il loro concetto e la loro definizione si presentavano dunque naturalmente, ed io ho pensato che sarebbe stato utile di studiarle in modo particolare. Essi dovevano costituire una categoria di enti di cui si sarebbe potuto ottenere delle proprietà comuni e che si potevano considerare nel loro insieme.

Outre ce contexte général qui s'appuie largement sur des visions rétrospectives, on peut préciser les idées de Volterra avant 1887 et relever une remarque précise qui détermine une vision particulière dans son approche de la physique mathématique. Dans une note intitulée « Sull'equilibrio delle superficie flessibili ed inestendibili », publiée entre 1883 et 1884 en deux parties où il prolonge des travaux de Beltrami, Volterra écrit⁸³ :

È noto il vantaggio che si ha in ogni questione di Meccanica prendendo in considerazione gli elementi caratteristici dello spostamento infinitamente piccolo del sistema di cui si studia il moto o l'equilibrio, ossia gli elementi necessari e sufficienti a determinare un tale spostamento. Ho perciò cominciato da tale ricerca che rientra nella cinematica di una superficie flessibile e inestendibile.

Volterra s'aperçoit ici qu'un certain nombre de problèmes de physique mathématique peuvent être envisagés du point de vue d'une « cinématique » qui s'appuie sur une notion de « déplacement » infinitésimal. De nos jours, une approche abstraite nous permet de ramener ces idées à une vision uniforme par le truchement de fonctions définies sur et dans des espaces munis de structures adéquates.

⁸⁰ [GJM15, à paraître]

⁸¹ [GP13]

⁸² [Vol60, p.393]

⁸³ [Vol54, p.180]

Mais ces idées modernes obscurcissent un peu ce que Volterra entend précisément par « cinématique ». C'est dans une référence qu'il fait à un texte de John Hewitt Jellett que l'on trouve des éléments qui permettent de préciser la pensée de Volterra. Il s'agit pour nous d'un point crucial de l'œuvre de Volterra et un des premiers éléments tangibles d'une inspiration pour ses travaux futurs sur les *fonctions de lignes* qui intervient précisément autour de l'année 1883.

3.3.1 JELLETT COMME SOURCE D'INSPIRATION

Volterra utilise les symboles δx , δy et δz pour désigner des « déplacements » en faisant explicitement référence à un emploi que l'on trouve chez Jellett⁸⁴ :

Jellett chiamando δx , δy , δz le componenti dello spostamento infinitamente piccolo di un punto di una superficie flessibile e inestendibile $z = z(xy)$ [...]

La référence à Jellett est importante ici car elle permet de mieux cerner la notion de « déplacement » (*spostamento*) à laquelle Volterra fait référence et elle va de plus éclairer très précisément plusieurs notions clés qu'il convoquera par la suite.

Jellett a publié en 1850 un traité intitulé « An elementary treatise on the calculus of variations » qui nous permet de distinguer diverses notions d'*accroissements*.

Il organise son texte en définissant tout d'abord ce qu'il entend par « fonction »⁸⁵ :

One variable quantity is said to be a function of any number of others, when there exists between them a certain relation, rendering the value of the first dependent on the values of the others, which are termed independent variables.

Cette définition posée, Jellett précise sa conception des divers éléments mis en jeu dans le concept de fonction⁸⁶ :

From this definition it is evident that the value of a dependent variable or function depends upon two different things, viz.: 1. the values of the independent variables; 2. the form of the function, or nature of the relation by which it is connected with them. A change in either of these will change the value of the function.

La notion de fonction repose donc ici sur l'idée de dépendance qui attribue des rôles différents à des « quantités », les unes dépendant des autres (dans un certain ordre). Cette organisation permet ainsi de distinguer la notion de « variable indépendante », élément de référence dans la dépendance, et celle de « variable dépendante ». De plus Jellett introduit l'expression « forme de la fonction » (*form of the function*) pour désigner la « nature de la relation » ainsi établie. C'est cette « forme » que désigne pour Jellett le symbole φ dans l'écriture $u = \varphi(x_1, x_2, \dots)$.

⁸⁴[Vol54, p.180]

⁸⁵[Jel50, p.1]

⁸⁶[Jel50, p.1]

En s'appuyant sur la distinction constant/variable qui est commune lorsque l'on considère des quantités, Jellett peut alors proposer une distinction entre deux types de fonctions à partir de la notion de « forme »⁸⁷ :

Functions may, therefore, be divided into 1. Determinate, or those whose form is not supposed to change; 2. Indeterminate, or those whose form is variable. This division is analogous to that of ordinary quantities into constant and variable.

Il y a donc deux types d'études des fonctions que l'on peut mener, suivant que l'on fait varier la variable indépendante ou la nature de la fonction, c'est-à-dire sa « forme »⁸⁸ :

The changes in value of which determinate functions are susceptible, arising solely from a change in the value of some one or more of the independent variable, have been already fully discussed in the Differential Calculus.

But the changes in value of which indeterminate functions are susceptible, arising, as they do, from a different cause, require to be treated in a different way, the rules of the Differential Calculus not being (as will be seen hereafter) universally applicable to them. It is with such changes in value, those, namely, which arise from a change in form, that we shall be, in our present subject, principally⁸⁹ concerned, inasmuch as these changes only are peculiar to the Calculus of Variations.

Jellett prolonge alors l'idée que la fonction puisse à son tour devenir élément variable en introduisant un nouveau type de dépendance⁹⁰ :

It is evident that the form of one function may be so connected with the form or forms of one or more others, that if the form or forms of the latter be determined, that of the first is determined also.

Cette nouvelle « connexion » ou « relation » attribue alors des rôles distincts à plusieurs fonctions, la forme des unes étant dépendantes de la forme des autres. Pour les différencier Jellett utilise les expressions « fonction primitive » et « fonction dérivée » :⁹¹ :

This species of relation between functions may be denoted by giving the name of *primitive* functions to the latter, whose forms are independent, and that of *derived* function to the former, whose form depends on those of the primitives. We shall denote it analytically by the symbols $F, F', & c.$; i.e. we shall use the symbol $F \cdot \varphi$ to represent a function whose form depends on that of the function φ .

⁸⁷[Jel50, p.1]

⁸⁸[Jel50, p.1]

⁸⁹Jellett précise ici que les problèmes de Calcul des Variations amènent à considérer à la fois des variations de la variable indépendante mais aussi de la "forme" en mettant en jeu des "accroissements" (*increments*) spécifiques aux deux types d'éléments variables.

⁹⁰[Jel50, p.1]

⁹¹[Jel50, p.2]

Les expressions « fonction primitive » ou « fonction dérivée » ont été empruntées à l'étude de la dérivée et de l'intégration des fonctions ordinaires⁹² vues sous un angle particulier de dépendance entre fonctions, comme « la forme d'un coefficient différentiel d'une fonction quelconque dépend, et peut être déduit, de la forme de la fonction elle-même »⁹³.

Il y a deux niveaux d'analyse qui sont impliqués de façon intriquée dans les idées exposées par Jellett, suivant que l'on considère des « fonctions » ou des « formes ». La « connexion » des « formes » induit une « relation » entre les fonctions. La notation F que Jellett qualifie d'« analytique » désigne la « relation » entre les fonctions. La notation plus explicite $F \cdot \varphi$ permet de distinguer les rôles joués par les fonctions mises en « relation », l'une étant dite « fonction primitive » alors que l'autre est nommée « fonction dérivée ». Mais le symbole $F \cdot \varphi$ sert aussi à « représenter une fonction dont la forme dépend de celle de φ » (*to represent a function whose form depends on that of the function φ*).

Pour préciser encore les choses Jellett écrit en note de bas de page⁹⁴ :

[Let $v = F \cdot u$]. It must be carefully kept in mind that v is not a function of u , as the symbol F denotes a relation between *forms*, not between *magnitudes*.

Nous verrons par la suite comment en plusieurs points les idées de Volterra exposées en 1887 font écho pour nous à ce texte de Jellett et aux notions que nous isolons ici. En particulier nous montrerons le rapprochement que l'on peut établir entre ce texte de 1850 et l'introduction de la notion de « fonction qui dépend d'autres fonctions » par Volterra quelques décennies plus tard.

Enfin ces précisions de vocabulaire et de concepts permettent à Jellett de définir de façon abstraite, c'est-à-dire en s'appuyant sur les notions de « fonction primitive » et de « fonction dérivée » qu'il vient de définir, l'objet du Calcul des Variations⁹⁵ :

Now if the form of one or more of the primitive functions be supposed to change, it is evident that the form of the derived function will undergo a corresponding change; and if the relation between the forms of the primitive and derived functions be supposed invariable, the change in the form of the latter will not be arbitrary, but connected by a fixed relation with the change in the form of the former. To deduce this relation, or, in other words, "to investigate the change in a derived function, consequence of a change in the form of its primitive," is the object of the Calculus of Variations.

Jellett précise alors la notion d'« accroissement » (*increment*) ou de « variation ». Comme dans le calcul différentiel, il est nécessaire pour le calcul des variations de définir un « accroissement » de la variable (ici une « forme » de

⁹²Le lecteur actuel pourrait s'y perdre. Pour Jellett rappelons, pour fixer les idées, que $F \cdot u$ dénote ce que nous nommerions une fonction composée, F est la « fonction dérivée », u la « fonction primitive ».

⁹³Jellett écrit : « the form of the differential coefficient of any function depends on, and may be deduced from, the form of the function itself. » [Jel50, p.2]

⁹⁴[Jel50, p.2]

⁹⁵[Jel50, p.2]

fonction) susceptible de devenir aussi petit qu'on le désire. Dans le texte de Jellett un accroissement d'une « fonction primitive » $u = \varphi(x_1, x_2, \dots)$ est écrit $i\psi'(x_1, x_2, \dots)$ et donne une fonction u' , ce qu'il exprime sous la forme $u' - u = i\psi'(x_1, x_2, \dots)$. La seule restriction est ici que la fonction ψ ne prenne que des valeurs finies pour chaque valeur (x_1, x_2, \dots) impliquées dans le problème étudié. Le symbole i désigne une « quantité constante » qui sera prise aussi petite qu'on le désire. Donc un « accroissement » de fonction est présenté sous une forme particulière mais reste, en accord avec les termes de Jellett, « parfaitement arbitraire ».

La notion d'accroissement ainsi précisée permet de définir la variation comme le terme de premier ordre relatif à i ⁹⁶ :

Let $u = \varphi(x_1, x_2, \dots)$, an indeterminate function of any number of variables, and let $v = F \cdot u$, a function derived from u , i.e. a function whose form depends on the form of u . Let it be required to find the variation of v , i.e. the change which v undergoes in consequence of an indefinitely small change in the form of u .

Substitute, according to the method [mentioned above], for

$$\varphi(x_1, x_2, \dots), \varphi(x_1, x_2, \dots) + i\psi(x_1, x_2, \dots),$$

and let the operation denoted by F be performed on the function so changed, sufficiently far to obtain the coefficient of the first power of i . If this coefficient be denoted by w , iw will be the required variation.

Cet énoncé laisse apparaître que Jellett considère des fonctions qui sont développables en séries et la variation apparaît ici comme un terme du développement en série. Il s'agit néanmoins d'une vision générale qui transpose et adapte des notions classiques du calcul différentiel à un terrain nouveau⁹⁷ et à des accroissements qui sont eux-mêmes des fonctions.

Dans un texte publié en 1860 intitulé « Die Grundlehren der Variationsrechnung »⁹⁸, qui est essentiellement une reprise en allemand du texte de 1850, Jellett emploiera la notation $F\{\varphi(x_1, x_2, \dots) + i\psi(x_1, x_2, \dots)\}$ pour désigner le résultat de la substitution mentionnée au-dessus. Cette même notation apparaissait dans le texte de 1850 seulement dans le cas où la fonction F était distributive (nous y reviendrons un peu plus tard). Nous pourrions comparer ces notations que nous utiliserions encore de nos jours avec celles que choisira Volterra par la suite pour désigner une « fonction qui dépend d'autres fonctions ».

Finalement l'introduction de nouveaux éléments variables va amener Jellett à utiliser plusieurs notations pour distinguer les divers types d'accroissements qui interviennent. Il notera en particulier δ l'accroissement qui concerne la « forme » d'une fonction. Conformément aux habitudes du calcul différentiel il notera d l'accroissement d'une fonction lié à une variation de ses variables

⁹⁶[Jel50, p.3-4]

⁹⁷Nous ne discutons pas ici de la nouveauté des idées de Jellett à son époque. Il s'agit pour nous simplement d'éclairer par la suite les travaux de Volterra qui fait référence à Jellett et de montrer précisément les éléments qu'il a pu lire chez ce dernier. Nous nous attacherons à montrer le type d'évolution que l'on peut établir entre les deux auteurs et nous analyserons en quoi les différences attestent l'émergence d'un certain type de généralité chez Volterra.

⁹⁸[Jel60, p.3]

indépendantes et l'accroissement d'une variable indépendante x sera notée dx . Et enfin il notera D l'accroissement total d'une fonction (*total increment*) lié à la variation de tous les éléments dont dépend la fonction, « forme » et variables indépendantes.

Le texte de Jellett permet alors de préciser la façon d'employer ces divers accroissements à travers plusieurs exemples que nous résumons ici. La méthode de base consiste à additionner les accroissements obtenus suivant un principe « parfaitement analogue au principe de superposition des petits mouvements en mécanique »⁹⁹ qu'invoque Jellett.

1. Si $u = f(x_1, x_2, \dots)$ est conçue comme « fonction déterminée » (*determinate function*) des variables indépendantes x_1, x_2, \dots alors sa « forme » ne varie pas et la variation totale s'écrit de façon classique¹⁰⁰ :

$$Du = du = \frac{du}{dx_1} \cdot dx_1 + \frac{du}{dx_2} \cdot dx_2 + \&c. + \frac{du}{dx_n} \cdot dx_n.$$

2. Lorsque la fonction notée $u = \varphi(x_1, x_2, \dots)$ par Jellett est une « fonction primitive indéterminée » (*a primitive indeterminate function*), cela signifie que sa « forme » varie et on doit prendre en compte à la fois la variation des variables indépendantes x_1, x_2, \dots et la variation de la forme de u qui est du type $\delta u = i\varphi$. On obtient donc¹⁰¹ :

$$Du = \delta u + du$$

ou encore :

$$Du = \frac{du}{dx_1} \cdot dx_1 + \frac{du}{dx_2} \cdot dx_2 + \&c. + \frac{du}{dx_n} \cdot dx_n + i\psi(x_1, x_2, \dots).$$

Ici $i\psi$ représente la « variation » de u qui n'est autre que l'accroissement de φ .

3. Si on considère une fonction $V = f(x_1, x_2, \dots, u_1, u_2, \dots)$ où x_1, x_2, \dots sont des variables indépendantes, et u_1, u_2, \dots des fonctions indéterminées de ces mêmes variables alors on a affaire à une situation où les diverses notions sont imbriquées. Si x_1 seulement varie alors cela engendre une variation de V qui doit s'écrire selon Jellett¹⁰² :

$$\left(\frac{dV}{dx_1} + \frac{dV}{du_1} \cdot \frac{du_1}{dx_1} + \frac{dV}{du_2} \cdot \frac{du_2}{dx_1} + \&c. \right) \cdot dx_1.$$

Maintenant si on fait varier la « forme » de u_1 par exemple, on obtient une variation du type $\frac{dV}{du_1} \cdot \delta u_1$. Pour justifier cette variation Jellett précise la validité du théorème sous-jacent¹⁰³ :

The truth of this theorem depending solely on the fact that V is a function of u_1 , and being altogether independent of the *species*¹⁰⁴ of increment assigned to u_1 .

⁹⁹Jellett écrit : « [...] perfectly analogous to the principle of the superposition of small motions in Mechanics » [Jel50, p.5]

¹⁰⁰[Jel50, p.6]

¹⁰¹[Jel50, p.6]

¹⁰²[Jel50, p.6]

¹⁰³[Jel50, p.8]

¹⁰⁴En italiques dans le texte de Jellett.

L'accroissement total DV s'obtient par somme de ces accroissements et on obtient la « variation proprement dite » en ne retenant que les termes relatifs aux accroissements des « formes »¹⁰⁵ :

$$\delta V = \frac{dV}{du_1} \cdot \delta u_1 + \frac{dV}{du_2} \cdot \delta u_2 + \&c.$$

Ces calculs éclairent la conception de Jellett et en déterminent précisément à la fois l'étendue et la généralité. Nous verrons ensuite les proximités mais aussi les différences avec l'approche de Volterra quelques décennies plus tard.

Une dernière idée sur les « accroissements » va nous permettre de retrouver le texte de Volterra de 1883-84 qui nous a permis de connecter Jellett aux travaux du mathématicien italien.

Les accroissements que Jellett a introduits jusqu'à ce point sont pour lui de type « purement mathématiques ». Si on s'en tient rigoureusement aux idées et aux notations introduites plus haut, le symbole δx ne fait pas sens en tant que conception « purement mathématique ».

Pourtant la mécanique permet d'introduire les concepts de « mouvement » et de « déplacement » qui permettent à Jellett de définir un nouveau type d'accroissement qu'il introduit dans un chapitre intitulé « Application of the calculus of variations to mechanics »¹⁰⁶. C'est précisément à ce type d'accroissement que Volterra fait référence en 1883-84 dans son étude sur l'équilibre des surfaces flexibles et nous y reviendrons un peu plus tard en détail.

Dans ce texte Jellett définit une notion de « fonction » sous la forme très générale de relation de dépendance entre des éléments de natures très diverses : quantités (les x, y, z), « formes » (nature de la relation qui définit la fonction), arcs de courbes¹⁰⁷, etc. La notion de « fonction » définit les rôles des divers éléments dans la dépendance des uns par rapport aux autres, les uns devenant variable indépendante, les autres variable dépendante. Le vocabulaire supplémentaire de « fonction primitive » et « fonction dérivée » permet de bien marquer l'attribution des rôles quand les divers éléments en jeu sont des fonctions.

De plus, grâce aux notions d'accroissements adaptés à chacune des situations, Jellett définit de façon générale la « variation » comme la partie d'ordre 1 relative à un type d'accroissement infinitésimal.

Outre ces notions que nous avons rappelées, Jellett introduit aussi une « variation seconde » ou des « variations successives ». Le texte en allemand de 1860 montre que Jellett conçoit véritablement un calcul différentiel où des fonctions peuvent jouer le rôle de variable¹⁰⁸ :

Die Variation von φ bei den höhern Variationen von $v = F(\varphi)$ als unveränderlich annehmen, heißt offenbar weiter nichts, als $\delta\varphi = i\psi$ für die Variation der unabhängigen Function fortwährend zu nehmen, so dass man hat:

$$\delta v = F(\varphi + \delta\varphi) - F(\varphi) = \frac{dF(\varphi)}{d\varphi} \delta\varphi = F'(\varphi)\delta\varphi,$$

¹⁰⁵[Jel50, p.8]

¹⁰⁶[Jel50, p.311]

¹⁰⁷Dans son ouvrage sur les variations Jellett dira : « Adopting Lagrange's method, we shall take the arc of the curve as the independent variable » [Jel50, p.194]

¹⁰⁸[Jel60, pp.31-32]

$$\delta^2 v = (F'(\varphi + \delta\varphi) - F'(\varphi))\delta\varphi = \frac{d^2 F(\varphi)}{d\varphi^2} \delta\varphi^2 = F''(\varphi)\delta\varphi^2, \text{ etc. etc.}$$

u. f. f. Kurz man braucht nur zu bemerken: dass die unabhängige Function φ hier ganz dieselbe Rolle spielt, wie die unabhängige Veränderliche x in der Differentialrechnung — und dass die eigentlichen Variationen, obgleich sie von Formänderungen der Functionen herrühren, im Calcul doch zuletzt als Wertsänderungen betrachtet und behandelt werden müssen, so dass sich alles wie in den entsprechenden Fällen der Differentialrechnung gestaltet.

Nous avons isolé et commenté quelques éléments du travail de Jellett sur les variations qui permettront de mieux comprendre la progression de Volterra et d'éclairer quelques différences par la suite.

Nous voudrions maintenant saisir l'occasion de cet examen du travail de Jellett pour aborder un thème important dans toute notre étude et qui sera aussi fortement présent dans le travail de Volterra. Il s'agit de la notion de *distributivité* qui apparaît sous des formes multiples qu'il nous faut analyser, d'une part en montrant qu'elle n'est pas toujours entièrement liée à une conception algébrique, et d'autre part dans son rapport avec la notion de *linéarité* qui s'est imposée dans un certain domaine de l'analyse fonctionnelle.

Il est remarquable que Jellett fasse un cas particulier des « fonctions primitives » qui sont « distributives » dès le début de son traité. Il introduit cette notion dès la page 4 comme propriété des « fonctions dérivées » les plus usuelles à son époque, c'est-à-dire celle qui correspondent à la dérivation et à l'intégration des fonctions ordinaires. Après avoir défini sous une forme extrêmement générale le calcul des variations tel que nous l'avons évoqué plus haut, Jellett précise l'étendue que l'on peut effectivement donner à une telle conception dans l'état d'avancement des mathématiques de son époque¹⁰⁹ :

This [as stated above] is the most general problem of the Calculus of Variations.

But as, in the present state of mathematical science, we are concerned with but two species of derived functions, sc., those which are derived by the process of differentiation, denoted by the symbol d , and those which are derived by the process of integration, denoted by the symbol \int , the investigation of so general a problem is quite unnecessary.

Sans restreindre son étude au contexte classique du calcul des variations, Jellett isole une propriété commune de la différentiation ou de l'intégration des fonctions ordinaires pour la convoquer à nouveau dans sa vision générale¹¹⁰ :

We shall, therefore, proceed to consider a particular case, which is, however, sufficiently general for all purposes to which the Calculus of Variations has been hitherto applied.

Let the symbol of derivation, F , be *distributive*, i.e. such as to satisfy the equation

$$F \cdot \varphi + F \cdot \varphi_1 = F \cdot (\varphi + \varphi_1),$$

¹⁰⁹[Jel50, p.4]

¹¹⁰[Jel50, p.4]

where φ and φ_1 are functions of any number of variables.

La « distributivité » donc est pensée à partir des cas particuliers de la dérivation et de l'intégration, mais devient chez Jellett une propriété générale.

Dans ce cadre, et pour n'importe quelle « fonction dérivée » F « distributive », Jellett montre que l'accroissement prend une forme particulière : $F \cdot (\varphi + i\psi) = F \cdot \varphi + F \cdot i\psi = F \cdot \varphi + iF \cdot \psi$ ce qui fait apparaître $iF \cdot \psi$ comme accroissement total (*total increment*) lorsque la fonction qui joue le rôle de variable indépendante φ reçoit un accroissement $i\psi$.

Jellett commente à plusieurs reprises l'égalité $F \cdot i\psi = iF \cdot \psi$ —c'est-à-dire en nos termes l'homogénéité de la fonction— qui pour lui vaut pour n'importe quel réel i . Cette propriété sera présentée plus tard de façon différentes chez divers mathématiciens et nous y reviendrons par la suite.

Dans son livre de 1850 Jellett fait d'abord l'hypothèse de cette égalité sans la commenter dans le corps du texte, puis il renvoie à une note de fin d'ouvrage qui donne quelques précisions sur lesquelles nous revenons maintenant.

Dans un premier temps Jellett montre que la « distributivité » permet de prouver que $F \cdot n\psi = nF \cdot \psi$ pour tout n entier. Puis en remplaçant ψ , qui est une fonction quelconque, par $\frac{1}{n}\psi$ on obtient $F \cdot \frac{\psi}{n} = \frac{1}{n}F \cdot \psi$. Par conséquent $F \cdot r\psi = rF \cdot \psi$ pour tout rationnel r . Dans ce premier texte Jellett invoque alors la « méthode des exhaustions » pour conclure¹¹¹ :

The theorem is, therefore, true for all rational values of i , and may easily be extended to irrational values by the method of exhaustions.

Il reviendra sur cette justification dans son texte en allemand où l'argumentation est donnée dans le corps du texte cette fois-ci. Jellett y montre que grâce à la « distributivité » on a pour tout entier n l'égalité $F \cdot \frac{1}{n}\psi = \frac{1}{n}F \cdot \psi$. Il précise que cette égalité reste vraie pour n aussi grand que l'on désire et « même pour $n = \infty$ ». On a donc $F \cdot i\psi = iF \cdot \psi$ « pour i aussi petit qu'on le désire » (*wo i unendlich klein ist*¹¹²). Plus tard nous verrons comment ce type d'extension d'une égalité pourra être liée à diverses notions de continuité, alors qu'elle n'est pas invoquée ici dans l'approche de Jellett.

Outre ces considérations fines, l'égalité d'homogénéité donnée dès le texte de 1850 permet à Jellett de préciser la variation dans le cas général sous hypothèse de distributivité¹¹³ :

Finally, let it be required to find the complete variation of $U = F \cdot V$, where

$$V = f(x_1, x_2, \dots, u_1, u_2, \dots),$$

and F is a symbol of derivation, which satisfies the condition

$$F \cdot \varphi + F \cdot \varphi_1 = F \cdot (\varphi + \varphi_1).$$

We shall proceed, as before, to determine separately the variations arising from changes in the several variable elements which U contains.

¹¹¹[Jel50, p.355]

¹¹²[Jel60, p. 5]

¹¹³[Jel50, p.4]

Jellett envisage alors les diverses variations qui interviennent du fait des différentes dépendances en jeu.

Il identifie tout d'abord les éléments variables : ce sont les x_i et les formes des fonctions u_i .

En ce qui concerne la variation induite par un accroissement des x_i , Jellett procède en les considérant un par un. Pour x_1 par exemple, il considère que tous les autres éléments sont fixés. Dans ce cas il analyse V , F , et par conséquent U , comme des fonctions « déterminées » (*determinate*) des variables x_i , ce qui lui permet d'écrire l'« évolution »¹¹⁴ (*change*) de U correspondante sous la forme $\frac{dU}{dx_1} \cdot dx_1$.

Jellett fait ensuite remarquer que l'on ne peut pas écrire la variation résultant d'un changement de « forme » de u_1 comme dans le cas classique $\frac{dU}{du_1} \cdot \delta u_1$ car^{115,116} :

it must be remembered that the truth of this theorem of the Differential Calculus depends entirely upon the supposition that U is a *function* of u_1 i.e. a quantity whose *magnitude* depends upon the magnitude of u_1 . But U is not a function of u_1 , inasmuch as the relation between them is a relation of *form*, not of *magnitude*; and it is therefore no longer true that the increment of U is

$$\frac{dU}{du_1} \cdot \delta u_1.$$

Si on adopte la vision de Jellett on voit qu'il y a en fait ici deux dépendances enchaînées qu'il va traiter en deux temps. D'une par V est une « fonction déterminée » de u_1 , et la « variation » due à un changement de « forme » de u_1 peut donc être écrite $\frac{dV}{du_1} \cdot \delta u_1$. Mais quand on veut examiner U il faut penser que V subit une variation de « forme » lorsque u_1 change aussi de « forme ». Donc à ce niveau d'analyse V est conçue comme « fonction indéterminée » dans sa relation avec F , ce qui est noté $F \cdot V$. Comme ici F est supposée « distributive » on a, comme l'a expliqué Jellett précédemment, $\delta F \cdot V = F \cdot \delta V$. En combinant ces éléments et toujours sous l'hypothèse de distributivité Jellett obtient¹¹⁷ :

$$DU = \frac{dU}{dx_1} \cdot x_1 + \frac{dU}{dx_2} \cdot x_2 + \&c. + F \cdot \left(\frac{dV}{du_1} \delta u_1 + \frac{dV}{du_2} \delta u_2 + \&c. \right).$$

Deux conclusions s'imposent après cette analyse du texte de Jellett.

Premièrement la notion de « distributivité » est apparue ici comme un moyen d'évaluer la « variation » dans le cas d'une relation fonctionnelle particulière. Rappelons que la notation $F \cdot V$ indique non seulement une relation de dépendance entre F et V au niveau des « formes » de F et V , mais aussi que la « variation » de F dépend de la « variation » de V . Jellett signale que V est susceptible de changer de « forme » en disant que c'est une fonction

¹¹⁴On prend soin ici de distinguer variation (*variation*) et évolution (*change*).

¹¹⁵[Jel50, p.9]

¹¹⁶Les termes sont écrits en italiques par Jellett.

¹¹⁷[Jel50, p.10]

« indéterminée », et précise les rôles des deux fonctions en nommant F « fonction dérivée » et V « fonction primitive ». Dans ce type de relation entre fonctions ou entre « formes », la distributivité permet d'écrire la variation de $F \cdot V$ en fonction de la variation de V : $\delta F \cdot V = F \cdot \delta V$. Nous retrouverons des idées très similaires chez Volterra un peu plus loin.

Deuxièmement l'approche de Jellett laisse entrevoir des éléments qu'il nous faut précisément positionner dans la scène que nous cherchons à élaborer pour éclairer l'émergence de l'analyse fonctionnelle. Nous avons déjà évoqué la nécessité de se référer tout au long de cette analyse au terme choisi par Jellett de « distributivité » plutôt qu'au terme de « linéarité » qui induit une certaine vision rétrospective. Il nous faut maintenant préciser la nature et les conséquences de cette distinction à la lumière des textes de Jellett.

Nous avons vu que Jellett utilisait la notation $F \cdot \phi$ pour indiquer une relation qu'il nomme « fonction » entre les éléments F et ϕ . Jellett a choisi une notation qui ressemble plus à une notation d'opération qu'à une notation de fonction ordinaire. À l'époque de Jellett, et cette habitude persiste encore, les fonctions sont désignées soit de façon synthétique par une lettre unique, soit en faisant apparaître la variable comme dans $\phi(x)$. L'écriture $F \cdot \phi$ nous ramène donc plutôt à une notation d'opération telle qu'elle est employée en algèbre symbolique et évoque deux aspects. L'un nous rappelle le modèle que constituent les symboles de différentiation ou d'intégration, opérations auxquelles Jellett fait précisément référence comme nous l'avons vu. L'autre, aspect plus algébrique, évoque une certaine similitude avec les opérations d'addition et de multiplication des nombres. Deux aspects qui, comme nous l'avons vu, tendent à recevoir un traitement uniformisé en algèbre symbolique.

Chez Jellett ces notations sont donc choisies d'abord pour mettre en évidence une relation particulière distincte de la notion de fonction ordinaire. D'autre part elles permettent d'exprimer la relation de distributivité sous une forme identique à celle utilisée en algèbre.

Enfin les éléments que nous avons fait ressortir dans l'approche de Jellett montrent une certaine organisation du discours où sont convoqués très tôt dans le texte à la fois des éléments abstraits —les fonctions ou les « formes » dans leurs différents rôles—, une *multiplication* —notée par le point entre deux symboles de fonctions— et une addition. Ces différents éléments sont reliés par une loi principale qui est la distributivité et forment une sorte de tout cohérent sur lequel Jellett appuie sa démarche. Bien que nous ne puissions pas ici parler de *structure*, en particulier parce que Jellett lui-même ne porte pas un regard suffisamment réflexif sur le type d'agencement que nous lisons dans son exposé, il nous semble que certains aspects de ce texte relèvent tout de même d'une organisation d'un certain type et qu'ils font apparaître ce que nous nommerons des *éléments structurels*. Le type de généralité que nous lisons chez Jellett nous semble intimement lié à ces *éléments structurels* qui sont le fruit d'un processus de sélection et d'organisation que nous avons suivi dans notre analyse.

3.3.2 EXTENSION DE LA THÉORIE DE RIEMANN SUR LES FONCTIONS DE LA VARIABLE COMPLEXE

C'est dans une note intitulée « Sopra le funzioni che dipendono da altre funzioni » que Volterra introduit en 1887 pour la première fois le concept de « fonction qui dépend d'autres fonctions ».

Volterra présente cette note comme un préliminaire nécessaire pour prolonger l'étude des fonctions de la variable complexe initiée par Riemann¹¹⁸ :

Mi permetto di accennare in questa Nota ad alcune considerazioni le quali servono a chiarire dei concetti che credo necessari introdurre per una estensione della teoria di RIEMANN sulle funzioni di variabili complesse, e che penso possano tornar giovevoli anche in varie altre ricerche.

En fait Volterra publiera juste après son travail sur les *fonctions de lignes*, entre 1887 et 1888, une série de trois notes intitulées « Sopra una estensione della teoria di Riemann sulle funzioni di variabili complesse ». Ces notes constituent un premier travail sur les fonctions de plusieurs variables complexes où l'on voit à l'œuvre la notion de *fonction de ligne* comme Volterra l'avait annoncé. Les notes dédiées aux fonctions de la variable complexe contiennent quelques éléments qui éclairent les motivations de Volterra et permettent de comprendre la voie générale qu'il cherche à emprunter lorsqu'il développe l'idée de « fonction de ligne ».

En particulier, à la fin de la publication sur la théorie riemannienne, Volterra revient sur les idées qui ont motivé ses recherches et sur la possibilité, grâce au concept de *fonction de ligne*, de concevoir de façon efficace des problèmes concernant les fonctions de plusieurs variables complexes¹¹⁹ :

Ciò prova che la teoria esposta in questa Nota e nelle due precedenti [Sopra una estensione della teoria di Riemann sulle funzioni di variabili complesse] è strettamente legata allo studio delle funzioni di due variabili complesse ed ai loro integrali, onde credo che le idee brevemente accennate potranno mettere in evidenza la utilità di introdurre le funzioni dipendenti da linee nello studio delle funzioni di due variabili complesse. Il sig. POINCARÉ in una importantissima Memoria pubblicata nel volume IX degli «Acta Mathematica» ha esteso il teorema di CAUCHY agli integrali doppi: il teorema enunciato nel §7¹²⁰ coincide colla estensione del teorema di CAUCHY data dal sig. POINCARÉ. Questo teorema è stato il punto di partenza delle mie ricerche.

L'article de Poincaré auquel il est fait allusion s'intitule « Sur les résidus des intégrales doubles »¹²¹ et a été publié cette même année 1887 dans *Acta Mathematica*. Il s'y attaque à une extension de la théorie de Cauchy aux intégrales doubles qui résiste sérieusement aux mathématiciens qui ont tenté

¹¹⁸[Vol54, p.294]

¹¹⁹[Vol54, p.350]

¹²⁰cf. [Vol54, p.347]

¹²¹[Poi87]

cette généralisation. Poincaré commence son étude en précisant qu'une difficulté principale repose sur le mode de représentation d'un couple de variables complexes ($\xi = x + iy; \eta = z + it$), qui, envisagé comme quadruplet (x, y, z, t) de \mathbb{R}^4 nécessite de passer à une conception d'espace à quatre dimensions. Pour Poincaré cette conception n'est pas la bonne car elle ne conserve pas la possibilité d'une interprétation géométrique qui était essentielle dans l'approche classique des intégrales de Cauchy. Le passage à l'« hypergéométrie », c'est-à-dire à la dimension supérieure à trois ne pose pas de problèmes conceptuels, mais ne semble pas bien adapté pour généraliser les idées classiques. L'idée de Poincaré est donc de contourner cet obstacle en n'utilisant que des surfaces dans un espace à trois dimensions¹²² :

Je vais maintenant exposer les artifices à l'aide desquels je compte m'affranchir de la nécessité de considérations hypergéométriques.

Soient λ, μ, ν trois quantités que je regarderai comme les coordonnées d'un point dans l'espace ordinaire, et considérons une surface algébrique ou portion de surface algébrique S sur laquelle se trouve le point λ, μ, ν . Écrivons :

$x = \varphi_1(\lambda, \mu, \nu), y = \varphi_2(\lambda, \mu, \nu), z = \varphi_3(\lambda, \mu, \nu), t = \varphi_4(\lambda, \mu, \nu)$
où $\varphi_1, \varphi_2, \varphi_3, \varphi_4$ sont des fonctions rationnelles de λ, μ, ν dont le dénominateur ne s'annule pour aucune valeur réelle de ces variables. Il est clair que quand le point λ, μ, ν décrira dans l'espace ordinaire la surface ou portion de surface S , le point x, y, z, t décrira dans l'hyperespace une certaine surface ou portion de surface S' ; de telle sorte que la surface S' est définie par la surface S et par les quatre fonctions fondamentales $\varphi_1, \varphi_2, \varphi_3, \varphi_4$.

L'idée de Poincaré est donc de considérer les fonctions de deux variables complexes, par le biais du processus exposée au-dessus, comme des fonctions de points de l'espace à 3 dimensions. Cela lui permettra de donner une version généralisée du théorème de Cauchy dont il énonce la teneur en quelques mots¹²³ :

On peut résumer tout ce qui précède en disant que l'intégrale prise le long d'une surface fermée S ne dépend que des courbes singulières qui sont contenues à l'intérieur de cette surface.

Dans un article écrit en 1889 dans *Acta Mathematica* intitulé « Sur une généralisation de la théorie des fonctions d'une variable imaginaire » Volterra revient sur le théorème classique de Cauchy comme source d'inspiration et en donne une lecture particulière¹²⁴ :

Soit donnée une fonction de la variable $z = x + iy$ continue et uniforme dans une aire limitée par un seul contour. On déduit du théorème de CAUCHY que l'intégrale d'une telle fonction prise entre des limites imaginaires dépend des limites seulement. On obtient donc, par l'intégration d'une fonction des points du plan complexe, une nouvelle fonction des points.

¹²²[Poi87, p.326]

¹²³[Poi87, p.342]

¹²⁴[Vol154, p.365]

Cette nouvelle façon de penser le théorème de Cauchy initial permet à Volterra de relire le théorème de Poincaré pour deux variables complexes en des termes qui tranchent avec la formulation du mathématicien français¹²⁵ :

On peut déduire [du théorème de Poincaré¹²⁶] que, si la surface d'intégration n'est pas fermée, l'intégrale dépend des lignes qui forment le contour de la surface. Donc on voit que l'intégration des fonctions de deux variables conduit aux fonctions des lignes.

Ces quelques commentaires *a posteriori* de Volterra montrent une de ses motivations à concevoir un type de dépendance où la *ligne* est l'élément variable. Dans la suite nous allons montrer que le processus d'élaboration des « fonctions de lignes » repose sur une approche générale qui n'est pas immédiatement reliée aux problèmes de la variable complexe.

3.3.3 LES « FONCTIONS QUI DÉPENDENT D'AUTRES FONCTIONS »

Volterra s'appuie pour commencer sur une notion de fonction qui lui semble propice à une généralisation. Il s'agit pour lui de voir les fonctions suivant la conception proposée par Dirichlet¹²⁷ qui libère la notion de fonction des propriétés d'analyticité qui nuisent à son sens à une vision générale. Volterra se saisit ainsi à son tour de cette idée de fonction¹²⁸ :

Seguendo il ben noto concetto del DIRICHLET si definisce attualmente una funzione nel seguente modo: una variabile è funzione di un'altra se, per ogni valore che questa prende entro certi limiti, la prima assume un dato valore.

Un tal concetto, che non implica nessuna relazione analitica fra l'una variabile e l'altra, discende molto naturalmente dalla considerazione di fenomeni nei quali due grandezze variano simultaneamente in modo que i valori dell'una dipendono da quelli dell'altra.

Pour Volterra cette définition décrit de façon générale une dépendance entre deux éléments, et elle va lui permettre d'étendre le concept de fonction. Il propose de considérer des quantités qui dépendent de « toutes les valeurs

¹²⁵[Vol54, p.365]

¹²⁶Volterra cite le théorème de Poincaré sous la forme suivante : « L'intégrale d'une fonction uniforme de deux variables imaginaires prise sur une surface fermée est nulle, si l'on peut déformer et réduire la surface à un point sans rencontrer de singularités. » (cf. [Vol54, p.365])

¹²⁷Dirichlet définit la notion de fonction par un processus qui relie deux types de grandeurs, la première parcourant un intervalle (ab) donné : « Man denke sich unter a und b zwei feste Werthe und unter x eine veränderliche Grösse, welche nach und nach alle zwischen a und b liegenden Werthe annehmen soll. Entspricht nun jedem x ein einziges, endliches y , und zwar so, dass, während x das Intervall von a bis b stetig durchläuft, $y = f(x)$ sich ebenfalls allmählich verändert, so heisst y eine stetige oder continuirliche Function von x für dieses Intervall. » [Dir37] Bien que le mot « allmählich » (=progressivement) laisse *a priori* penser qu'il s'agit de fonctions continues, une note de bas de page montre que Dirichlet envisage que sa définition recouvre aussi des cas non continus. Il rajoute en effet : « Da im Folgenden nur von stetigen Functionen die Rede sein wird, so kann der Zusatz ohne Nachtheil wegbleiben ». Rappelons que la notion de continuité à son époque est donnée par la propriété qui impose que $f(x + \epsilon) - f(x)$ devienne aussi petit qu'on le désire avec ϵ . (cf. par exemple [Dir29, p.121])

¹²⁸[Vol54, p.294]

que peuvent prendre une ou plusieurs fonctions dans un intervalle donné. »¹²⁹
 Pour illustrer cette conception il citera très souvent des exemples issus de la physique¹³⁰ :

Così per esempio la temperatura in un punto di una lamina conduttrice dipende da tutti i valori che la temperatura ha al contorno; lo spostamento infinitesimo di un punto di una superficie flessibile e inestendibile, dipende da tutte le componenti degli spostamenti dei punti del contorno, parallelamente ad una certa direzione.

Dans un premier temps Volterra définit donc la notion de « fonction qui dépend d'autres fonctions » et éventuellement de certains paramètres supplémentaires¹³¹. Pour indiquer qu'une quantité y dépend d'une fonction φ définie sur un intervalle (AB) Volterra écrira $y = y|[\varphi \underset{A}{\overset{B}{(x)}}]|$, ou encore, lorsque la dépendance met en jeu plusieurs fonctions et des paramètres $y = y|[\varphi_1 \underset{A_1}{\overset{B_1}{(x)}}, \varphi_2 \underset{A_2}{\overset{B_2}{(x)}}, \dots, t_1, t_2, \dots]|$. Volterra envisage aussi pour les fonctions φ — que nous nommerons comme dans d'autres situations déjà évoquées *fonctions-objets*¹³² — des fonctions de plusieurs variables.

Une fois le concept établi comme extension de l'idée de fonction, Volterra dédie sa première note au développement de ce que nous appellerions de nos jours un *calcul différentiel* pour les « fonctions qui dépendent d'autres fonctions » et qu'il entreprend par le biais de l'étude des variations.

La notion de « variation » est le point de départ de cette étude : une fonction (ordinaire¹³³) peut recevoir une variation ψ qui sera continue et en valeur absolue inférieure à un nombre donné ϵ . Le but est alors d'étudier la variation produite au niveau des valeurs de la fonction y lorsque la *fonction-objet* est ainsi modifiée.

Pour étudier le comportement de y en termes de variations, Volterra commence par définir la notion de « continuité »¹³⁴ :

¹²⁹[Vol54, p.294]

¹³⁰[Vol54, p.294]

¹³¹Volterra insiste dès son introduction sur la nécessité de ne pas confondre les « fonctions qui dépendent d'autres fonctions » avec des « fonctions de fonctions » que nous nommerions fonctions composées. Cette distinction qui paraît simple ici est en fait cruciale pour les acteurs qui cherchent à définir un nouveau champ des mathématiques où la notion de fonction joue un rôle particulier. Dans une lettre écrite par Fréchet le 30 juillet 1913 à Volterra — nous reproduisons cette lettre en annexe B —, celui-ci insistera sur la distinction qui est en jeu : « je ne ramène pas la question à la théorie des fonctions d'une variable qui serait un paramètre ϵ dont dépendrait la fonction $f(x)$; je reste dans le Calcul Fonctionnel. » [Fré13]

¹³²Nous employons ici *fonction-objet* et pas *fonction-variable* car la seconde expression nous pousse du côté des variations, et nuit à l'identification d'un processus d'abstraction qui va permettre de redéfinir les éléments en jeu dans le concept de fonction ou d'opération.

¹³³Nous respecterons ici les notations de Volterra qui désigne les fonctions ordinaires que nous avons qualifiées de *fonctions-objets* par des lettres grecques minuscules. Nous désignerons comme Volterra les « fonctions qui dépendent d'autres fonctions » et plus tard les « fonctions de lignes » comme nous l'avons rappelé plus haut en faisant apparaître tous les éléments en jeu, ou par des notations plus synthétiques qu'utilise Volterra, y ou des lettres capitales.

¹³⁴[Vol54, p.295]

Sia

$$y = y|[\varphi \underset{A}{\overset{B}}(x)]|$$

diremo che y è continua se, data a $\varphi(x)$ una variazione $\psi(x)$ tale che in valore assoluto $\psi(x)$ sia sempre inferiore ad ϵ , la variazione corrispondente di y può rendersi inferiore a δ , piccolo ad arbitrio.

Volterra étend cette idée aux fonctions de plusieurs variables et qui dépendent de paramètres en disant qu'une fonction est continue si en perturbant assez faiblement chacune des variables (*fonctions-objets* et paramètres) on obtient une variation plus petite qu'un δ imposé.

Le but central de cette partie est d'obtenir un théorème de représentation intégrale de la variation d'une fonction y continue. Dans le cadre des fonctions ordinaires, le théorème correspondant est un résultat important en analyse qui permet entre autres de voir les deux opérations d'intégration et de dérivation comme inverse l'une de l'autre. Le théorème obtenu par Volterra dans le cas des « fonctions qui dépendent d'autres fonctions » s'écrira avec des notations que nous préciserons par la suite¹³⁵ :

$$\delta y|[\varphi]| = \int_A^B y'|[\varphi, t]| \cdot \delta \varphi(x) \cdot dt.$$

Pour arriver à ce théorème Volterra a besoin de plusieurs ingrédients qu'il va dégager et exposer sous la forme d'une liste de conditions au début de son article. Une partie du travail postérieur à cette première note consistera à relâcher certaines de ces conditions restrictives pour obtenir des résultats aussi généraux que possibles.

Les conditions dégagées par Volterra sont au nombre de quatre. Elles permettent d'une part de concevoir une notion de dérivée y' et d'autre part d'adapter des idées qui mènent dans le cas des fonctions ordinaires à l'écriture de la variation sous la forme d'une intégrale, celle-ci étant conçue en un sens riemannien ici.

Nous rappelons ces quatre conditions qui déterminent le champ de validité des résultats et permettent de préciser les éléments mis en jeu par Volterra¹³⁶ :

Per la y dipendente dalla $\varphi(x)$, oltre alla condizione della continuità, ammetteremo altre condizioni.

Preso un intervallo $h = mn$ entro AB , diamo alla $\varphi(x)$ una variazione continua $\theta(x)$ entro h , tale che $\theta(x)$ sia in valore assoluto inferiore ad ϵ , e denotiamo con δy la variazione corrispondente di y . Ammetteremo:

I. Che il rapporto $\frac{\delta y}{\epsilon h}$ sia sempre inferiore ad un numero finito M .

Suppongasì ora $\theta(x)$ sempre dello stesso segno e si ponga

$$\int_m^n \theta(x) dx = \sigma.$$

¹³⁵[Vol54, p.299]

¹³⁶[Vol54, p.296]

Se rappresentiamo la funzione $\varphi(x)$ mediante una curva $z = \varphi(x)$, avremo che σ sarà l'area compresa fra questa curva e la curva variata. Porremo le condizioni:

II. Che facendo impiccolire indefinitamente ϵ ed h , in modo che questo intervallo contenga sempre nel suo interno un punto G di indice t , esista il limite determinato e finito del rapporto $\frac{\delta y}{\sigma}$.

III. Che il rapporto $\frac{\delta y}{\sigma}$ tenda verso il suo limite uniformemente rispetto a tutte le possibili funzioni $\varphi(x)$ e agli indici t .

Il limite $\frac{\delta y}{\sigma}$ dipenderà dalla $\varphi(x)$ e dall'indice t del punto G ; lo denoteremo con

$$y' | [\varphi(x), t]$$

e lo chiameremo derivata prima di y . Ammetteremo:

IV. Che $y' | [\varphi(x), t]$ sia continua rispetto a $\varphi(x)$ e a t .

La notion de fonction choisie par Volterra induit que la variation δy est une *grandeur* (réelle ou complexe) et elle permet donc de concevoir les rapports différentiels à partir desquels sont construites les conditions que nous venons de rappeler.

On peut distinguer ici deux types de conditions liées à deux types de rapports différentiels.

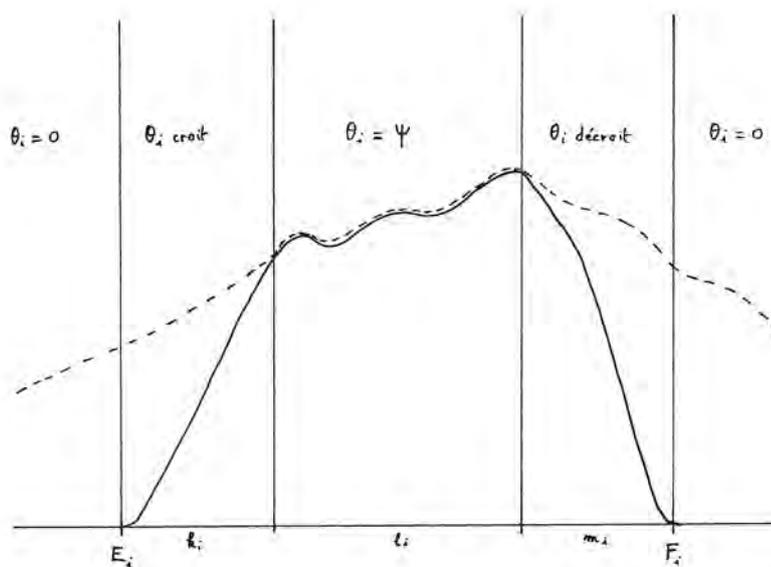
La condition I apparaît comme une nécessité technique qui s'impose à Volterra dans sa démarche pour contrôler la variation de la fonction y lorsqu'il perturbe la fonction φ . Le rapport différentiel associé met en jeu la variation δy ainsi que deux éléments qui déterminent la perturbation que l'on fait subir à la variable φ : la longueur h de l'intervalle sur lequel on effectue la perturbation et la *mesure* ϵ de la perturbation sur cet intervalle. La condition est globale en ce sens qu'elle doit être réalisée pour tout type de perturbation, c'est-à-dire quel que soit l'intervalle sur lequel on effectue la perturbation et quelle que soit la *mesure* ϵ de celle-ci.

Les conditions II à IV définissent la dérivée de la fonction y par la convergence d'un nouveau rapport différentiel. La notion est ici locale : non seulement φ est donnée mais tout se passe aussi autour d'un point G de la courbe représentative de φ repéré par son abscisse t . Volterra considère alors des perturbations de φ qui sont localisées autour de G . Le rapport différentiel est défini pour chacun des couples (φ, t) et dépend d'une *mesure* de la perturbation donnée par l'aire entre la courbe de φ et la courbe obtenue par perturbation. La limite du nouveau rapport différentiel se conçoit désormais lorsque l'aire devient petite. La dérivée existe lorsque le rapport converge uniformément (par rapport à ces deux paramètres) quand l'aire qui mesure la perturbation devient petite. Lorsque la fonction y est dérivable, la dérivée est une fonction qui dépend d'une autre fonction, φ , et d'un paramètre, t . Il y a donc une infinité de dérivées pour une fonction φ donnée. Nous reviendrons sur cet aspect important que Volterra commentera lui-même dans ses publications.

La démarche mise en place par Volterra pour étudier la variation de y consiste à découper habilement un accroissement ψ donné en une somme de

petits accroissements localisés sur des intervalles de plus en plus petits. La variation pour chaque accroissement localisé pourra alors s'interpréter grâce à la dérivée y' et la somme donnera l'intégrale escomptée.

Plus précisément, Volterra découpe l'intervalle de départ AB (en évitant les points où ψ s'annule) en segments $E_i F_i$ eux-mêmes découpés en trois parties k_i, l_i, m_i . On construit alors une fonction θ_i sur chacun des intervalles $E_i F_i$ qui coïncide avec ψ sur la partie médiane comme indiqué sur le schéma ci-dessous¹³⁷.



On peut jouer alors sur la taille des intervalles et le fait que les fonctions θ_i prennent des valeurs non nulles sur des intervalles disjoints pour calculer la variation de y . Nous reprenons rapidement la démarche de Volterra afin de préciser la nature et le rôle des conditions qu'il impose au début de son article.

Dans un premier temps, grâce à la condition I, Volterra montre que $y|[\varphi(x) + \epsilon\psi(x)]|$ peut être approché par $y|[\varphi(x) + \epsilon \sum_{i=1}^n \theta_i(x)]|$. Plus précisément l'écart en valeur absolue est majoré par un terme du type $K\epsilon\delta$ où δ majore la longueur sur laquelle on effectue la perturbation et K est une constante indépendante de ϵ et de δ .

Ensuite on s'appuie sur des sommes qui se télescopent pour écrire :

$$y|[\varphi(x) + \epsilon \sum_{i=1}^n \theta_i(x)]| - y|[\varphi(x)]| = \sum_{r=1}^n \left\{ y|[\varphi(x) + \epsilon \sum_{i=1}^r \theta_i(x)]| - y|[\varphi(x) + \epsilon \sum_{i=1}^{r-1} \theta_i(x)]| \right\}.$$

¹³⁷Volterra ne fournit pas de schéma pour cette démarche.

Mais la variation $y|[\varphi(x) + \epsilon \sum_{i=1}^r \theta_i(x)] - y|[\varphi(x) + \epsilon \sum_{i=1}^{r-1} \theta_i(x)]$ peut s'écrire :

$$\epsilon \sigma_r \left\{ y'|[\varphi(x) + \epsilon \sum_{i=1}^{r-1} \theta_i(x), t_r] + \eta_r \right\}$$

(où $\sigma_r = \int_{E_r}^{F_r} \theta_r(x) dx$) par convergence simple du rapport qui définit la dérivabilité, et même

$$\epsilon \sigma_r \left\{ y'|[\varphi(x) + \epsilon \sum_{i=1}^{r-1} \theta_i(x), t_r] + \eta \right\}$$

avec η aussi petit qu'on le désire lorsque ϵ et h_r deviennent petits grâce à la propriété III qui assure la convergence uniforme par rapport à la courbe φ et au point $G(t)$.

Enfin la continuité de y' imposée par la propriété IV montre que $y'|[\varphi(x) + \epsilon \sum_{i=1}^{r-1} \theta_i(x), t_r]$ peut être rendu aussi proche que l'on désire de $y'|[\varphi(x), t_r]$.

Pour retrouver l'intégrale attendue comme dans le cas des fonctions ordinaires Volterra utilise l'écriture de σ_r en fonction des valeurs de φ aux points t_r et l'oscillation de ψ dans l'intervalle h_r qu'il note D_r : on a $\sigma_r = h_r \psi(t_r) + \tau [h_r D_r + (k_r + m_r) \sup |\psi(x)|]$ où $\tau \in (0; 1)$.

On obtient alors

$$y|[\varphi(x) + \epsilon \psi(x)] - y|[\varphi(x)] = \epsilon \sum_{r=1}^n h_r \psi(t_r) y'|[\varphi(x), t_r] + \epsilon \Theta$$

où Θ devient aussi petit que l'on désire lorsque l'on prend des subdivisions de l'intervalle AB de plus en plus fines. Nous laissons le lecteur vérifier que les paramètres peuvent effectivement être choisis correctement.

Finalement ce raisonnement conduit Volterra à énoncer le résultat sous la forme suivante¹³⁸ :

$$\lim_{\epsilon=0} \frac{y|[\varphi(x) + \epsilon \psi(x)] - y|[\varphi(x)]}{\epsilon} = \lim_{\epsilon=0} \frac{\Delta y}{\epsilon} = \int_A^B \psi(t) \cdot y'|[\varphi(x), t] dt.$$

Volterra reprendra alors des notations que nous avons vues chez Jellett en isolant la partie de premier ordre de Δy , qu'il notera $\delta y|[\varphi(x)]$ ou δy . Il s'agit de la *variation* de y engendrée par une variation de la variable qui est elle-même notée $\epsilon \psi(x) = \delta \varphi(x)$. Ces notations permettent de reformuler le résultat obtenu sous la forme que nous avons indiquée au départ.

Un peu plus loin dans son texte Volterra utilisera la notation empruntée au calcul différentiel des fonctions ordinaires $\left(\frac{dy}{d\epsilon}\right)_{\epsilon=0} = \int_A^B y'|[\varphi(x), t] \psi(t) dt$ qui met en évidence le lien entre la dérivée au sens des « fonctions qui dépendent d'autres fonctions » et la dérivée de la fonction $y|[\varphi(x) + \epsilon \psi(x)]$ vue comme fonction de ϵ . Cette remarque est importante car ce lien permet à Volterra

¹³⁸[Vol154, p.298]

d'adapter plusieurs idées de la théorie des fonctions ordinaires qui s'appuient sur la notion de dérivée.

Comme la dérivée y' peut être vue comme une fonction qui dépend d'une (ou plusieurs) autres fonctions lorsque l'on fixe le paramètre t , et si l'on tombe à nouveau sous le coup des hypothèses énoncées pour la dérivabilité, on peut comme l'indique Volterra définir des dérivées successives qui font apparaître des paramètres multiples : $y^{(n)} = y^{(n)}|[\varphi(x), t, t_1, t_2, \dots, t_{n-1}]$ qui donnent, comme le montrera Volterra, des expressions symétriques dans les différents paramètres t et t_i .

Là encore Volterra précise le lien avec une expression intégrale : $\left(\frac{d^n y}{d\epsilon^n}\right)_{\epsilon=0} = \iint \dots \int \prod_{i=1}^n \psi(t_i) \cdot y^{(n)}|[\varphi(x), t_1, t_2, \dots, t_{n-1}] dt_1 \dots dt_n$ où les intégrales sont prises entre A et B .

Enfin cette première note se clôt sur un paragraphe dédié à une extension de la formule de Taylor. En considérant l'expression $y|[\varphi(x) + \epsilon\psi(x)]$ comme une fonction de la variable ϵ , on peut appliquer la formule de Taylor classique ce qui donne¹³⁹ :

$$y|[\varphi(x) + \psi(x)] = y|[\varphi(x)] + \sum_{n=1}^{\infty} \frac{1}{\pi(n)} \iint \dots \int y^{(n)}|[\varphi(x), t_1, t_2, \dots, t_n] \prod_{r=1}^n \psi(t_r) \cdot dt_1 \dots dt_n.$$

La reprise des démonstrations de cette première note de Volterra sur les « fonctions qui dépendent d'autres fonctions » permet d'éclairer très précisément sa démarche et de saisir la nature de la généralité qui se dégage de son approche.

Tout d'abord nous avons mis en évidence le recours à des *éléments infinitésimaux* qui ont permis à Volterra de définir la dérivée d'une « fonction qui dépend d'autres fonctions ». En fait on peut distinguer plusieurs types d'*éléments infinitésimaux* qui à leur tour permettent de définir plusieurs types de *conceptions localisées*. On trouve dans ce texte de Volterra trois types d'*éléments infinitésimaux* :

1° Un premier *élément infinitésimal* intervient pour perturber la fonction φ qui sert de variable dans la « fonction qui dépend d'autres fonctions » y : il s'agit d'une fonction qui en valeur absolue ne dépasse pas une valeur ϵ que l'on peut choisir aussi petite qu'on le désire. Ceci permet entre autres de définir la continuité de y . La *localisation* opère ici par la considération des fonctions $\varphi + \theta$ où $|\theta(x)|$ est petit sur l'intervalle AB , c'est-à-dire *autour* de φ .

2° Ensuite, la notion de dérivée fait intervenir un point G de la courbe représentative de la fonction φ qui sert de variable. Ce point est repéré par son abscisse t et Volterra considère des intervalles mn aussi petits qu'il le désire autour de t . La *localisation* s'appuie donc ici sur des intervalles réels qui contiennent tous le point t .

3° Enfin un dernier *élément infinitésimal* prend la forme $\sigma = \int_m^n \theta(x) dx$ qui sert à définir le rapport différentiel qui donne la dérivée d'une « fonction qui dépend

¹³⁹[Vol54, p.299]

d'autres fonctions ». Le type de *localisation* qui est alors en jeu fait intervenir à la fois un intervalle mn autour du point t et le contrôle de la valeur absolue ϵ de la perturbation. Volterra fait appel à une vision géométrique de cet *élément infinitésimal* et du type de *localisation* qu'il induit en concevant σ comme l'aire qui sépare les deux courbes de φ et $\varphi + \theta$ respectivement.

Un second aspect que nous voudrions isoler concerne le choix du rapport différentiel qui entre dans la définition de la dérivée d'une « fonction qui dépend d'autres fonctions ». Volterra choisit de rapporter la variation de y à l'aire comprise entre la courbe de la variable φ et de la courbe de la perturbation $\varphi + \theta$, et nous avons vu comment ce choix était fondamental dans sa démonstration afin d'obtenir la forme intégrale qu'il désire.

Il y a là un enjeu que nous avons déjà signalé dans notre détour par le travail de Jellett. Quelques années plus tard Fréchet, en travaillant à son tour sur ces textes de Volterra, commentera les conséquences du choix des éléments qui interviennent dans un rapport différentiel pour définir une théorie générale en analyse fonctionnelle. Nous y reviendrons à nouveau en étudiant l'approche de Fréchet, mais ici Volterra introduit sans faire de commentaires sa définition de dérivée y' et montre qu'elle permet de mener à bien la preuve.

Enfin on peut comparer l'approche de Volterra pour les « fonctions qui dépendent d'autres fonctions » et celle que nous avons étudiée pour les substitutions. Les deux travaux mettent en place un calcul infinitésimal en adaptant des notions initialement définies pour les fonctions ordinaires. Cependant Volterra opère dans les deux situations des choix différents, non seulement quant à la nature du concept qui vient remplacer la notion de fonction ordinaire, mais aussi dans la façon de transposer le modèle initial.

En effet, dans le premier cas nous avons vu que Volterra définissait une différentielle et une intégrale propres aux substitutions. Il s'en sert pour exposer une véritable *théorie* qui doit lui servir comme fondement pour l'étude des équations différentielles linéaires.

Dans le cas des « fonctions qui dépendent d'autres fonctions » Volterra définit seulement une dérivée adaptée au nouveau concept et la notion d'intégrale qui intervient est l'intégrale d'une « fonction qui dépend d'autres fonctions » vue comme *fonction ordinaire* de l'un de ses paramètres, et donc en lui appliquant la théorie riemannienne classique.

Le texte ne cherche pas dès l'abord à établir de parallèle avec le calcul infinitésimal des fonctions ordinaires, il reste focalisé sur la mise en place d'outils généraux pour établir la différentielle d'une fonction (généralisée) sous une unique forme donnée par une intégrale et une notion de dérivée adaptée.

Ces remarques et les analyses qui précèdent montrent que Volterra développe une vision abstraite de « fonction qui dépend d'autres fonctions » à partir de la notion de fonction ordinaire (au sens de Dirichlet) suivant un processus particulier qui fait apparaître plusieurs aspects que nous voudrions préciser.

D'une part Volterra conçoit une notion abstraite de « fonction qui dépend d'autres fonctions » qu'il relie plusieurs fois à des situations concrètes de physique ou d'une *certaine géométrie* elle-même fondée sur une imagerie issue

de la physique mathématique ou de l'étude des fonctions¹⁴⁰. Nous avons par ailleurs établi une parenté entre cette notion et des conceptions antérieures comme celles exposée par Jellett qui pense la fonction comme *fonction-variable* (terme d'observateur), c'est-à-dire comme *variable* d'une autre fonction.

D'autre part le texte de Volterra que nous venons d'étudier isole un rapport différentiel $\frac{\delta y}{\sigma}$ qui fait passer la nature abstraite de la fonction y au second plan, puisque ce rapport est conçu comme rapport de deux quantités qui suivent les règles du calcul infinitésimal ordinaire (on peut faire tendre σ vers 0 et étudier la limite du rapport). Ce rapport sert à définir la notion de dérivabilité, et l'expression « dérivée première » (*derivata prima*) choisie par Volterra indique bien la correspondance qu'il établit avec une théorie des fonctions ordinaires. D'ailleurs, l'organisation de l'exposé de Volterra qui donne ce rapport différentiel dès le début renforce l'idée de transposition d'une théorie préexistante¹⁴¹.

On peut mettre en regard la façon technique et calculatoire qui permet à Volterra d'opérer avec des fonctions dont la variable est désormais d'un type nouveau avec la transposition des idées que Jellett exprimait d'une façon beaucoup plus descriptive et géométrique près de cinquante ans auparavant¹⁴² :

[I]n the Differential Calculus, it is possible to give a geometrical statement of any problem of maxima and minima which does not involve more than two independent variables, so in the Calculus of Variations it is possible to give a geometrical statement of any problem of maxima or minima not involving integrals of an order higher than the second.

In the Differential Calculus it is known that any such problem may be stated as follows: —“To find a point in a given curve or surface at which the ordinate of that curve or surface is a maximum or minimum.” Here the curve or surface is given. But in the Calculus of Variations the principal question is to determine, not a point on a curve or surface, but the curve or surface itself. Geometrically stated the problem is this: —“To determine a curve or surface which may have the property of rendering the value of a given integral a maximum or minimum.” In fact, if the words “nature of the curve or surface” be substituted in any of the preceding chapters for “form

¹⁴⁰Nous reviendrons plusieurs fois sur des pratiques qui s'appuient sur une vision géométrique, c'est-à-dire un type de représentation descriptive, pour élaborer un élément abstrait. Les mathématiciens ont recours à des modes de représentation et font intervenir des éléments dont notre étude illustre une certaine diversité. Nous reviendrons en conclusion de notre travail sur certaines distinctions théoriques que l'on peut élaborer dans le recours à *des visions géométriques*.

¹⁴¹L'exposé de Volterra suit une certaine forme usuelle de l'étude des fonctions ordinaires qui rend difficile l'identification du processus qui l'a mené à concevoir ce rapport différentiel particulier. Ce texte de Volterra ne permet pas de préciser s'il imagine ce rapport différentiel à partir du rapport différentiel que l'on utilise à son époque pour définir la dérivabilité des fonctions ordinaires, s'il le pense à partir du calcul des variations appliqué à la géométrie comme dans Jellett cité juste après ou si celui-ci est conçu à partir de la démarche qui établit le résultat final dont les outils seraient exposés en amont.

¹⁴²[Jel50, p.137]

of the function," we shall have the geometrical statement of the problem¹⁴³.

Enfin l'analyse de la démonstration qui mène à la représentation intégrale montre que tous les calculs sont effectués au niveau des quantités ou au niveau des fonctions ordinaires qui servent ici de *fonctions-objets*. Dans les différentes phases du raisonnement Volterra ne met jamais en jeu le type de relation qui définit la « fonctions qui dépend d'autres fonctions » et seule la notation $y|[\varphi(x)]$ permet de savoir que l'on manipule des fonctions généralisées.

La suite du travail de Volterra, exposée dans une seconde note sur « les fonctions qui dépendent d'autres fonctions », examine les cas où la condition I n'est pas réalisée.

Rappelons que cette condition doit pouvoir être utilisée sur chaque intervalle de longueur h pour remplacer l'accroissement quelconque φ par un accroissement contrôlé θ . Il fallait donc sur tout intervalle pouvoir écrire $y|[\varphi(x) + \epsilon\psi(x)] - y|[\varphi(x) + \epsilon\theta(x)] < \epsilon h M \sup |\psi(x) - \theta(x)|$, où h représente la longueur de l'intervalle dans lequel $|\psi - \theta| \neq 0$. Quand on fait intervenir plusieurs perturbations qui agissent sur des intervalles disjoints de longueur h_i , le rapport différentiel correspondant à l'ensemble des perturbations $\frac{y|[\varphi(x) + \epsilon\psi(x)] - y|[\varphi(x)]}{\epsilon}$ dépend donc des h_i , c'est-à-dire des subdivisions de plus en plus fines de l'intervalle initial AB . Cette première approche de Volterra repose donc sur un double contrôle qu'il exprime en imposant que l'accroissement δy soit un infiniment petit d'ordre au moins ϵh .

Dans la suite Volterra veut donc s'affranchir de la condition 1 et il examine les différentes possibilités lorsqu'il existe un point de la courbe représentative de φ où la condition n'est pas vérifiée. Volterra va ici raisonner par cas en modifiant les idées qui permettaient dans un premier temps d'avoir une approche uniforme pour tous les points du segment AB .

Le premier cas étudié par Volterra envisage l'hypothèse où il existe un point $C(x_1)$ de la courbe de φ tel que la variation de y liée à des accroissements localisés au voisinage de x_1 vérifie : $\lim_{\substack{h \rightarrow 0 \\ \epsilon \rightarrow 0}} \frac{\Delta y}{\epsilon} = 0$ (ici $x_1 \in (mn)$, $h = n - m$ et l'accroissement ne dépasse pas ϵ en valeur absolue). Volterra procède alors en deux temps.

Il établit d'abord que l'intégrale qui décrivait la variation dans le cas I existe toujours (est finie) malgré le point singulier. Pour cela Volterra part d'un accroissement fonctionnel continu quelconque ψ . Puis il remplace, comme nous l'avons déjà vu, la perturbation ψ par une perturbation θ localisée au voisinage du point singulier x_1 comme illustré page 111.

La stratégie consiste à montrer que les hypothèses de travail sont suffisantes pour établir que la variation obtenue pour la perturbation particulière θ est donnée par l'intégrale finie $\int_p^q y'|[\varphi(x), t]|\theta(t)dt$ dont on contrôle la taille uniformément par rapport au voisinage de x_1 . Comme θ et ψ coïncident sur un sous-intervalle de (pq) aussi grand qu'on le désire, on obtient que l'intégrale

¹⁴³Jellet ajoute en note de bas de page: "This was the form in which the problem of maxima and minima was first stated by John and James Bernoulli. —Vid. Acta Erud. 1696-7."

$\int_p^q y' |[\varphi(x), t]| \psi(t) dt$ est elle-même contrôlée, ce qui permet d'établir que l'intégrale $\int_A^B y' |[\varphi(x), t]| \psi(t) dt$ est finie.

Dans un second temps Volterra étudie la variation de y due à une perturbation ψ quelconque. Il s'appuie encore sur une fonction θ particulière qui coïncide avec ψ hors d'un voisinage du point singulier x_1 et qui s'annule au voisinage (plus fin) du point singulier. Volterra décompose $\psi(x)$ en $\psi(x) = \theta(x) + \alpha(x)$ où $\alpha(x)$ peut être non nulle seulement au voisinage du point singulier. Mais la valeur absolue de α est alors contrôlée par $2 \sup\{|\psi(x)| ; x \in (AB)\}$. On a donc à nouveau une stratégie qui consiste à remplacer l'accroissement quelconque par un accroissement particulier qui permet de désenlacer l'étude de la variation locale autour du point singulier de l'étude sur le reste de l'intervalle AB .

Dans ce premier cas où $\lim_{\substack{h \rightarrow 0 \\ \epsilon \rightarrow 0}} \frac{\Delta y}{\epsilon} = 0$ Volterra obtient encore la variation sous forme intégrale¹⁴⁴ :

Dunque

$$\lim_{\eta=0} \frac{y |[\varphi(x) + \eta \psi(x)]| - y |[\varphi(x)]|}{\eta} = \int_A^B y' |[\varphi(x), t]| \psi(t) dt;$$

e quindi anche in questo caso potremo porre

$$\delta y |[\varphi(x)]| = \int_A^B y' |[\varphi(x), t]| \delta \varphi(t) dt.$$

Se la singularità considerata invece di presentarsi nel punto C soltanto si verificasse in più punti entro AB , purché fosse sempre per tutti soddisfatta la condizione $[\lim_{\substack{h \rightarrow 0 \\ \epsilon \rightarrow 0}} \frac{\Delta y}{\epsilon} = 0]$, si giungerebbe pure ai precedenti risultati.

Dans un second cas Volterra envisage une nouvelle forme du rapport différentiel sur lequel il appuie son analyse par cas : cette fois-ci il exprime la condition sous la forme $\lim_{\substack{h \rightarrow 0 \\ \epsilon \rightarrow 0}} \frac{\Delta y}{\epsilon} = a_1 \lim_{\epsilon \rightarrow 0} \frac{\rho}{\epsilon}$ où ρ est la valeur $\psi(x_1)$ de l'accroissement au point singulier (terme d'observateur) et a_1 est un nombre fixé indépendant des autres paramètres.

Volterra va se ramener au cas précédent en considérant la « fonction qui dépend d'autres fonctions » $z |[\varphi(x)]| = y |[\varphi(x)]| - a_1 \varphi(x_1)$. Grâce à l'hypothèse la fonction z vérifie la condition $\lim_{\substack{h \rightarrow 0 \\ \epsilon \rightarrow 0}} \frac{\Delta z}{\epsilon} = 0$ (au voisinage du point singulier) et on a donc d'après ce qui précède $\delta z = \int_A^B z' |[\varphi(x), t]| \delta \varphi(t) dt$. Mais pour tout $t \neq x_1$ on a $z' |[\varphi(x), t]| = y' |[\varphi(x), t]|$ et $\delta z = \delta y - a_1 \delta \varphi(x_1)$. Volterra en déduit la forme de la variation de y qui s'écrit en tenant compte du point singulier :

$$\delta y = \int_A^B y' |[\varphi(x), t]| \delta \varphi(t) dt + a_1 \delta \varphi(x_1).$$

¹⁴⁴[Vol54, p.306]

Les calculs ci-dessus permettent de comprendre la conception de Volterra et le sens des termes qu'il utilise. Les points x_i sont bien des « points exceptionnels » où la fonction y se comporte de façon particulière, cette propriété étant traduite à travers le comportement du rapport $\frac{\Delta y}{\epsilon}$ (qui dépend de ϵ mais aussi de h). Volterra dira que y , en plus de dépendre de φ , dépend de façon « spéciale » de la valeur de φ au point exceptionnel¹⁴⁵. S'il y a plusieurs « points exceptionnels » on a donc $y|[\varphi(x)] = z|[\varphi(x)] + \sum_{i=1}^n a_i \varphi(x_i)$ ce qui permet d'écrire :

$$\delta y = \int_A^B y'|[\varphi(x), t] \delta \varphi(t) dt + \sum_{i=1}^n a_i \delta \varphi(x_i).$$

Enfin dans un troisième cas Volterra utilise une stratégie similaire à celle que nous venons d'examiner dans le cas précédent pour envisager des « points exceptionnels » où le rapport différentiel dépend de façon particulière non seulement de la valeur $\varphi(x_i)$ mais aussi des valeurs des dérivées $\varphi^{(p)}(x_i)$ pour reprendre le vocabulaire de Volterra. Plus précisément Volterra suppose que φ ainsi que les accroissements ψ sont des fonctions m_1 fois dérivables au point singulier. Dans ce cadre on considère $\epsilon > 0$ et un accroissement φ dont les dérivées successives jusqu'à l'ordre m_1 sont en valeurs absolues inférieures à ϵ sur AB . Si $\rho_0, \dots, \rho_{m_1}$ sont les valeurs de ψ et de ses dérivées successives jusqu'à l'ordre m_1 au point x_1 alors Volterra note $k_p = \lim_{\epsilon \rightarrow 0} \frac{\rho_p}{\epsilon}$. La condition généralisée s'écrit alors :

$$\lim_{\substack{h \rightarrow 0 \\ \epsilon \rightarrow 0}} \frac{\Delta y}{\epsilon} = \sum_{p=0}^{m_1} a_p k_p$$

pour certaines valeurs de a_p fixées. Volterra obtient la forme de la variation de y qu'il étend à un nombre fini de « points exceptionnels »¹⁴⁶ :

Se ciò che vale pel punto x_1 , valesse anche analogamente per i punti x_2, \dots, x_n , allora

$$\delta y = \int_A^B y'|[\varphi(x), t] \delta \varphi(t) dt + \sum_{i=1}^n \sum_{p=1}^{m_i} a_{pi} \delta \varphi^{(p)}(x_i),$$

e si scriverebbe

$$y = y|[\varphi_A^B(x)] |(\varphi(x_1), \varphi'(x_1), \dots, \varphi^{(m_1)}(x_1) \cdots \varphi(x_n) \cdots \varphi^{(m_n)}(x_n)),$$

cioè y oltre che da $\varphi(x)$ in tutto AB , dipenderebbe *specialmente* dai valori di $\varphi(x)$ nei punti x_1, \dots, x_n e dalle sue derivate, rispettivamente degli ordini m_1, \dots, m_n , prese nei punti stessi.

Les coefficients a_{pi} dépendent de la fonction φ et seront notés par Volterra $a_{pi} = y'_{\varphi^{(p)}(x_i)}|[\varphi(x)]$.

¹⁴⁵Volterra écrit : « si dirà che y oltre che da $\varphi(x)$ in tutto AB , dipende specialmente dal valore di $\varphi(x)$ in x_1 . » [Vol54, p.307]

¹⁴⁶[Vol54, p.308]

Le cas des points exceptionnels et la notation choisie par Volterra qui met en évidence tous les paramètres qui peuvent varier lorsque l'on étudie la variation de y nous ramène au texte de Jellett. Le résultat énoncé par Volterra ainsi que la démarche qui consiste à écrire $y|[\varphi(x)] = z|[\varphi(x)] + \sum_{i=1}^n \sum_{k=1}^{m_i} a_i \varphi^{(k)}(x_i)$ permettent de concevoir la variation totale en additionnant les variations, principe que nous avons souligné chez Jellett. Nous avons aussi remarqué que Jellett faisait intervenir des éléments du type $\delta \frac{dy}{dx}$ pour une fonction y ordinaire, termes que l'on retrouve chez Volterra sous la forme $\delta \varphi^{(p)}(x_i)$. Il y a cependant des différences importantes entre les deux façons de concevoir les fonctions.

Notons enfin que Volterra utilise une notation qui sépare d'un côté la dépendance générale entre une *fonction-objet* φ et des quantités notées $y|[\varphi(x)]$ et de l'autre une dépendance particulière qui met en jeu la valeur de la fonction φ en des « points exceptionnels ». Ainsi le concept général qui ressort de ce travail de Volterra est celui de « fonction qui dépend d'autres fonctions », qu'il note $y|[\varphi(x)]$. Bien que Volterra pense $\varphi(x_i)$ dans sa variabilité par rapport à φ , cela n'en fait pas pour lui une variable au même sens que celle qu'il note $\varphi(x)$. En effet dans l'approche de Volterra $\varphi(x)$ désigne toutes les valeurs que peut prendre la fonction dans un intervalle donné AB comme il l'indique dans l'introduction de sa première note sur le sujet. Cette différence de nature des éléments variables ici se retrouve dans le choix des notations de Volterra où l'on reconnaît le type d'interprétation qu'il convoque.

La troisième note de Volterra sur les « fonctions qui dépendent d'autres fonctions » s'intéresse à ce qu'il appelle des « questions particulières ».

Dans un premier temps Volterra cherche à caractériser certaines « fonctions qui dépendent d'autres fonctions » lorsque leur dérivée prend une forme particulière. Il montre que si $y'|[\varphi(x), t] = 0$ pour toute fonction φ et tout t alors y est constante. Dans le prolongement de ces idées, si $y'|[\varphi(x), t] = F(\varphi(t))$ où F est une « fonction ordinaire » (*funzione ordinaria*) alors $y = \int_A^B \Psi(\varphi(t)) dt + C$ où C est une constante.

Une seconde approche montre réciproquement que si y s'écrit

$$y = \int_A^B F(\varphi(t), \varphi'(t), \dots, \varphi^{(m)}(t)) dt,$$

que Volterra envisage comme « fonction qui dépend d'autres fonctions », alors

$$\delta y = \int_A^B \left(\frac{\partial F}{\partial \varphi} - \frac{d}{dt} \frac{\partial F}{\partial \varphi'} + \dots \pm \frac{d^m}{dt^m} \frac{\partial F}{\partial \varphi^{(m)}} \right) \delta \varphi(t) dt + \sum_{p=0}^{m-1} N_p \delta \varphi^{(p)}(B) - \sum_{p=0}^{m-1} M_p \delta \varphi^{(p)}(A),$$

et donc dans cette écriture

$$y'|[\varphi(x), t] = \left(\frac{\partial F}{\partial \varphi} - \frac{d}{dt} \frac{\partial F}{\partial \varphi'} + \dots \pm \frac{d^m}{dt^m} \frac{\partial F}{\partial \varphi^{(m)}} \right)_{x=t}.$$

Volterra envisage aussi les intégrales doubles où

$$y|[\varphi(x)] = \int_A^B dt \int_A^B F(\varphi(t), \varphi(t_1)) dt_1$$

dont la variation fait intervenir les deux dérivées partielles de F .

Enfin une dernière partie de cette note montre comment Volterra relit les équations différentielles générales grâce au concept de « fonctions qui dépendent d'autres fonctions ».

Il part d'une équation sous la forme¹⁴⁷ :

$$f\left(y, \frac{dy}{dx}, \frac{d^2y}{dx^2}, \dots, \frac{d^ny}{dx^n}, \varphi(x), \varphi'(x), \dots, \varphi^{(m)}(x)\right) = 0$$

pour laquelle la fonction y ainsi que ses dérivées jusqu'à l'ordre $n - 1$ sont connues en un point A , et en un second point B la valeur Y de y dépend de toutes les valeurs de φ sur l'intervalle AB : on pourra donc écrire avec les notations de Volterra $y(B) = Y[[\varphi(x)]_A^B]$ où Y est une « fonction qui dépend d'autres fonctions ».

Alors que le problème initial semble ne faire intervenir que des fonctions classiques, Volterra réinterprète ce problème en termes de « fonctions qui dépendent d'autres fonctions ». Nous suivons en quelques lignes sa stratégie qui permet préciser le lien établi entre la nouvelle théorie et celle des équations différentielles. Nous ne modifions que très légèrement les notations de l'exposé de Volterra afin de rendre celui-ci plus facile à suivre pour le lecteur.

Le point de départ consiste à analyser l'équation en termes de variations. En notant $\frac{df}{dy^{(i)}} = a_i$ et $\frac{df}{d\varphi^{(i)}} = b_i$, la variation du premier membre de l'équation différentielle doit être nulle, ce qui s'écrit en tenant compte des diverses dépendances :

$$\sum_{i=0}^n a_i \frac{d^i \delta y}{dx^i} + \sum_{i=0}^m a_i \frac{d^i \delta \varphi}{dx^i} = 0.$$

Volterra multiplie cette variation par une fonction quelconque λ et calcule l'intégrale du produit obtenu entre A et B par parties. On obtient :

$$0 = \left[\sum_{i=0}^{n-1} p_i \frac{d^i \delta y}{dx^i} \right]_A^B + \left[\sum_{i=0}^{m-1} q_i \frac{d^i \delta \varphi}{dx^i} \right]_A^B + \int_A^B \left\{ \delta y \sum_{i=0}^n (-1)^i \frac{d^i}{dx^i} (\lambda a_i) + \delta \varphi \sum_{i=0}^m (-1)^i \frac{d^i}{dx^i} (\lambda b_i) \right\} dx$$

où les fonctions p_i et q_i dépendent de la fonction λ et des fonctions a_i . Volterra utilise alors le fait que $\left(\frac{d^i \delta y}{dx^i} \right)_{x=A} = 0$, et $\left(\frac{d^i \delta y}{dx^i} \right)_{x=B} = \frac{d^i \delta Y}{dB^i}$.

De plus on peut choisir la fonction λ de façon à ce qu'elle soit solution de l'équation :

$$(\alpha) \quad \sum_{i=0}^n (-1)^i \frac{d^i}{dx^i} (\lambda a_i) = 0.$$

¹⁴⁷[Vol54, p.311]

En isolant les termes qui contiennent les variations de Y on a donc pour chaque fonction λ solution de (α) :

$$(\beta) \quad \sum_{i=0}^{n-1} P_i \frac{d^i \delta Y}{dB^i} = - \left[\sum_{i=0}^{m-1} q_i \frac{d^i \delta \varphi}{dx^i} \right]_A^B - \int_A^B \delta \varphi \sum_{i=0}^m (-1)^i \frac{d^i}{dx^i} (\lambda b_i) dx$$

où $P_i = p_i(B)$.

Enfin Volterra ayant fait apparaître une équation linéaire (α) , il y a un système fondamental de solutions qui, injectées dans (β) donnent un système de n équations linéaires dont les inconnues sont $\delta Y, \frac{d}{dB} \delta Y, \dots, \frac{d^{n-1}}{dB^{n-1}} \delta Y$. Le déterminant du système fait apparaître le wronskien des λ_i qui est non nul et Volterra peut donc établir que la variation de Y s'écrit sous la forme :

$$\delta Y = \left[\sum_{i=1}^{m-1} \alpha_i \frac{d^i \delta \varphi}{dx^i} \right]_A^B - \int_A^B \beta \delta \varphi \cdot dx$$

où les coefficients sont explicitement donnée en fonction des solutions fondamentales de l'équation (α) et de leurs dérivées jusqu'à l'ordre n .

La variation ainsi obtenue permet à Volterra de montrer que Y « dépend de façon spéciale » des valeurs de $\delta \varphi$ et de ses dérivées jusqu'à l'ordre $n - 1$ aux points A et B . Elle donne aussi la dérivée $Y' | [\varphi(t), x] = \beta$.

Les trois notes de Volterra sur les « fonctions qui dépendent d'autres fonctions » exposent donc une *analyse infinitésimale* adapté à un concept de fonction généralisé.

Nous y avons montré de nombreuses similarités avec l'exposé de Jellett auquel Volterra fait référence dès ses publications de 1883-4. En particulier l'introduction du texte de Jellett fait apparaître les notions de « fonction dérivée » et « fonction primitive » qui permettent de distribuer des rôles particuliers à différentes fonctions. Ces fonctions sont elles-mêmes vues de façon générale comme « correspondance » non seulement entre quantités, mais aussi entre la « forme » d'une fonction et une quantité. Nous avons rapproché ces idées de celle développées ici par Volterra qui définit les « fonctions qui dépendent d'autres fonctions » grâce à une interprétation de fonction comme principe général de mise en relation de certains éléments avec des quantités qu'il puise dans la définition de fonction donnée par Dirichlet.

Volterra développe son *analyse infinitésimale* à partir de la notion de *variation* qu'il interprète de façon abstraite pour les « fonctions qui dépendent d'autres fonctions ». Le schéma d'analyse de la variation est, comme nous l'avons vu, très proche de celui exposé par Jellett tout en laissant apparaître des différences importantes. En particulier si les deux approches identifient que la partie de premier ordre de la variation joue un rôle important, celle-ci va être traitée suivant deux conceptions distinctes. Le texte de Jellett donne une interprétation qui s'appuie sur une vision en termes de développement en série dont il identifie le premier coefficient (cf. page 98). De son côté Volterra choisit d'adapter la notion de « dérivée » qui est associée à un développement d'ordre 1 dans le calcul infinitésimal des fonctions ordinaires.

Enfin le traitement de Volterra met en avant l'intégrale $\int_A^B y' | [\varphi(x), t] \cdot \delta \varphi(t) dt$ comme élément central et général dans la description de la variation première.

On retrouve cette forme dans les exemples particuliers traités dans le texte de Jellett lorsque la « fonction dérivée » est du type intégral, c'est-à-dire pour la variation de $\int_{x_0}^{x_1} V dx$ où V représente différents types de fonctions que l'on a vues un peu plus haut. Volterra s'appuie sur cette « fonction qui dépend d'autres fonctions » particulière, $\int_A^B y' |[\varphi(x), t]| \cdot \delta\varphi(t) dt$, pour décrire uniformément la variation dans le cas général (en y ajoutant les termes liés aux points exceptionnels). Par contraste, Jellett donne aussi une lecture générale mais c'est alors la « fonction dérivée » F qui revient dans l'écriture de la variation lorsque F a la propriété de distributivité. Donc bien que Volterra semble isoler un élément issu du contexte particulier de l'étude de la variations des intégrales, il en fait un élément général grâce au concept de « dérivée » d'une « fonction qui dépend d'autres fonctions ».

Cette notion de dérivée, $y' |[\varphi(x), t]|$, est centrale dans le texte de Volterra car elle fait le lien entre un contexte particulier où l'on étudie la variation des formes intégrales et la notion de « fonction qui dépend d'autres fonctions ». Deux éléments contribuent à établir ce lien et permettent l'élaboration d'une vision générale telle que nous la lisons chez Volterra.

D'une part la notion de dérivée permet une double lecture : elle est à la fois « fonction qui dépend d'autres fonctions » mais aussi fonction ordinaire de son paramètre t . Le premier aspect la relie au concept abstrait de fonction généralisée alors que le second fait le lien avec le calcul différentiel et intégral des fonctions ordinaires.

D'autre part on peut la voir comme un élément *ad hoc* qui apparaît dans une démarche qui part de la forme intégrale de la variation (identifiée comme élément commun à tous les cas) et en donne une analyse en termes d'intégrale de Riemann. Bien que Volterra ne dévoile pas, ni ici ni dans ses cours où il tente de donner une justification d'ensemble à cette théorie, la démarche heuristique qui l'a conduit à sa définition de dérivée, on peut imaginer une démarche fondée sur l'hypothèse que nous avons formulée¹⁴⁸ : c'est la forme intégrale, lue dans des cas particuliers, qui lui sert de point de départ. En effet si la variation s'écrit $\int_A^B \Omega(\varphi, t) \delta\varphi(t) dt$ alors on peut l'écrire comme limite d'une somme de termes du type $\Omega(\varphi, t_k) \delta(t_k) h_k$ dans laquelle on peut interpréter $\delta(t_k) h_k$ comme une aire infinitésimale σ_k . Pour faire le lien avec la variation de y il faut donc que $\Omega(\varphi, t_k) \sigma_k$ représente un morceau de la variation de y localisée sur l'intervalle de longueur h_k .

Du point de vue de ce que nous pourrions classer sous l'appellation *éléments de topologie* nous avons mis en évidence un certain nombre d'éléments locaux qui mettaient en jeu des infiniment petits de diverses natures. Outre ces aspects que nous avons détaillés au cours de notre analyse, l'approche de Volterra met en avant le recours à une vision géométrique particulière qui consiste à considérer les fonctions ordinaires du point de vue de leur courbe représentative

¹⁴⁸Il s'agit de l'hypothèse la plus probable à ce jour. Une étude de grande ampleur des textes de Volterra, Jellett et d'autres auteurs ayant pu nourrir la pensée de Volterra sera nécessaire pour établir avec plus de certitude et de précision les mécanismes de sélection et de réinterprétation auxquels Volterra se livre pour développer la notion de dérivée d'une « fonctions qui dépend d'autres fonctions ».

dans un repère. Cette représentation géométrique qu'il emploie pour manipuler les *fonctions-objets* φ se décline en une version infinitésimale qui permet de concevoir des courbes qui sont infiniment proches l'une de l'autre. La nature de la généralité qui apparaît dans ce texte est largement dépendante du recours à la notion de courbe et au type de localisation géométrique qui en découle. Nous verrons par la suite que Volterra s'appuiera davantage sur cet aspect pour développer et étendre sa théorie.

3.3.4 LES « FONCTIONS DE LIGNES »

Cette même année 1887 Volterra va publier deux nouvelles notes aux *Rendiconti Lincei* qu'il intitule « *Sopra le funzioni dipendenti da linee* ». Ces notes vont nous permettre d'établir une certaine évolution dans la conception des fonctions généralisées que nous venons de présenter.

L'optique adoptée par Volterra dès l'introduction est beaucoup plus géométrique dans ces nouvelles notes. Cette conception qui semble *a priori* plus restrictive va permettre en fait à Volterra de développer des idées qui favorisent une approche générale. En effet, si la conception de « fonction qui dépend d'autres fonctions » permet de penser à des fonctions de plusieurs variables (les *fonctions-objets*), la notion de « ligne » permet de penser cette *fonction-objet* comme une infinité variable. D'autre part nous verrons plus loin que la notion de « ligne » évacue un aspect calculatoire ou mesurable qui était présent dans la notion de « fonction qui dépend d'autres fonctions ». En effet dans les premières notes, la variable était une fonction continue. Volterra s'appuyait donc sur un certain nombre de propriétés ou de conceptions de la variable mettant en jeu des quantités mesurables : maximum de la valeur absolue pour définir des petits accroissements, possibilité de concevoir un accroissement via l'addition $\varphi + \theta$, intégration. Dans le cadre des « lignes » un certain nombre de ces aspects sont à repenser comme nous allons le voir en suivant la démarche de Volterra. Le recours à une version géométrique permet en fait à Volterra de se débarrasser de certaines contraintes et de dégager un cadre qui ouvre le champ à un nouveau type de généralité que nous allons analyser.

Nous passerons ici rapidement sur les détails techniques et nous mettrons principalement en lumière les éléments qui engendrent pour nous une véritable évolution en matière de généralité. En particulier à l'occasion de cette étude Volterra va faire apparaître des éléments qui nous semblent importants dans l'évolution de la conception des *fonctionnelles*.

Les « lignes » considérées par Volterra dans ces deux nouvelles notes sont des lignes¹⁴⁹ d'un espace à 3 dimensions qui sont fermées ou qui se terminent au bord d'une surface donnée. De plus Volterra considère des lignes qui ont une tangente en chaque point sauf éventuellement en un nombre fini de points singuliers. La notation pour les « fonctions de lignes » ressemble à celles des « fonctions qui dépendent d'autres fonctions » et Volterra écrit $\varphi = \varphi|[L]|$ pour une ligne L .

¹⁴⁹Nous emploierons le terme « ligne » qui est utilisé par Volterra et repris en français par ses contemporains. L'usage actuel tend à lui préférer le mot « courbe » mais notre but vise à préciser ce que Volterra convoque comme propriétés ou comment il manipule ce qu'il nomme « ligne ».

Un premier aspect nouveau apparaît dès le début du texte où Volterra définit une sorte de voisinage tubulaire pour les lignes. Étant donnée une ligne L il considère une ligne qui lui est « liée » à L (*concatenata*) et qui décrit une « surface tubulaire » (*superficie tubulare*) σ autour de L . Le « domaine » (*spazio*) S délimité par la surface σ sera appelée par Volterra « voisinage de la ligne L »¹⁵⁰ (*intorno della linea L*). Ce voisinage qui est décrit en termes de domaine, et donc de points de l'espace, permet en fait de penser un voisinage comme une collection de lignes autour d'une ligne donnée. Volterra définit ainsi la notion de « ligne longitudinale » (*linea longitudinale*) qui est une ligne qui traverse le domaine S . La portée de cette définition se conçoit plus précisément dans son utilisation lorsque Volterra définit la continuité d'une « fonction de lignes »¹⁵¹ :

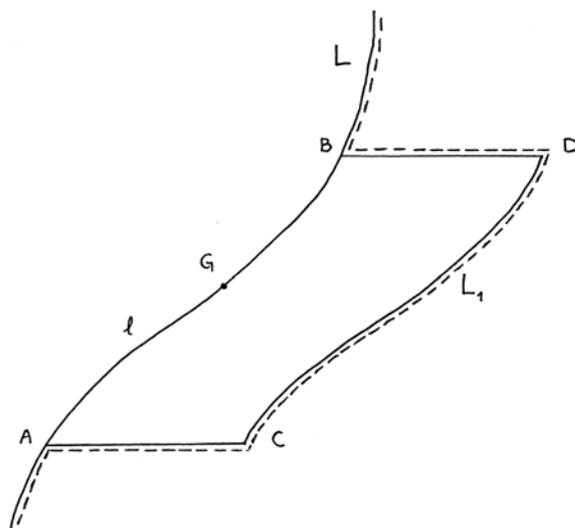
La funzione φ , funzione delle linee L , sarà continua se, preso un numero δ piccolo ad arbitrio, potrà trovarsi un intorno S di L tale che i valori di φ corrispondenti a tutte le linee longitudinali di S differiscano dal valore di φ in L meno di δ .

Quand on compare cette version de la continuité avec celle proposée par Volterra pour les « fonctions qui dépendent d'autres fonctions » on voit précisément l'évolution qu'il met en œuvre ici. La première notion était liée à une notion de mesure de la taille des *fonctions-objets* ou de leur perturbations, qui étaient effectuées par le biais de la valeur absolue des valeurs de la fonction φ sur l'intervalle d'étude. Il s'agissait d'une vision *métrique* en quelques sortes, bien que Volterra n'emploie pas ce type de dénomination. Dans ce cas la notion de continuité est liée à la notion de supremum d'une fonction continue sur un intervalle. Dans la version géométrique la notion de voisinage est dégagée de toute grandeur mesurée, ce qui lui confère un caractère abstrait et engendre une généralité particulière liée à ce mode de représentation géométrique. Nous comparerons plus tard ces éléments à ceux que l'on peut lire dans certains textes de Fréchet par exemple.

La notion de voisinage sans recours à la distance qui permet de définir la continuité n'est pas reprise pour définir les notions de dérivée dont Volterra a besoin. Il s'appuie sur une version *directionnelle* d'accroissement qui réintroduit une mesure de l'écart entre deux lignes. L'espace à 3 dimensions est rapportée à un repère orthonormé qui permet à Volterra de retrouver une notion de déformation localisée de la ligne L autour d'un point dans les trois directions x , y et z . La technique consiste à considérer un point G repéré par la longueur s de l'arc mesuré sur L à partir d'un point origine. Volterra construit alors une déformation de la ligne L issue d'une translation d'une portion l de L autour du point G dans la direction de x qui permet de remplacer localement L par une portion de parallélogramme mixtiligne comme indiqué sur le schéma ci-dessous.

¹⁵⁰[Vol54, p316]

¹⁵¹[Vol54, p316]



La variation de la « fonction de ligne » φ s'écrit alors $\varphi(L) + \Delta_x \varphi$ ce qui permet à Volterra de définir, lorsqu'elle existe, la limite $X = \lim_{\substack{l \rightarrow 0 \\ \epsilon \rightarrow 0}} \frac{\Delta_x \varphi}{\epsilon l}$. On peut raisonner de même pour les trois directions et obtenir de façon similaire deux autres limites. Volterra remarque que ces limites, lorsqu'elles existent, sont des fonctions qui dépendent de la ligne L mais aussi de l'abscisse curviligne de G . Il écrira donc $X = X[[L, s]]$, $Y = Y[[L, s]]$ et $Z = Z[[L, s]]$ respectivement qu'il appellera « dérivées de φ par rapport à x, y ou z » et qu'il notera $X = \varphi'_x$, $Y = \varphi'_y$, et $Z = \varphi'_z$.

Ces notions sont le pendant de la notion de dérivée introduite pour les « fonctions qui dépendent d'autres fonctions », déclinées en termes de dérivées directionnelles suivant les axes x, y ou z . Le rapport différentiel est à nouveau établi par rapport à l'aire entre deux courbes, qui est ici l'aire du parallélogramme mixtiligne $ACDB$. Pour ce type de dérivée la forme de l'accroissement est ainsi particulière et celui-ci pas seulement assujéti à être petit et localisé comme c'était le cas pour les « fonctions qui dépendent d'autres fonctions ».

On retrouve ici des conditions similaires à celles que Volterra avait imposées au début de son étude des « fonctions qui dépendent d'autres fonctions » avec cependant quelques différences que nous précisons. La convergence rappelée ci-dessus, qui définit une notion de dérivée, doit être uniforme relativement au point G et aux courbes L . Volterra suppose aussi que les dérivées directionnelles (termes d'observateur) sont continues par rapport à L et à l'abscisse curviligne s . De plus il impose désormais une continuité supplémentaire relative à la variation des projections des lignes sur les plans de base¹⁵² :

Finalmente supporremo che $\varphi[[K]] - \varphi[[L_1]]$ possa ridursi minore di un numero arbitrariamente piccolo, quando le aree comprese fra le proiezioni delle curve L e L_1 , sui piani coordinati si siano rese inferiori a dati valori.

¹⁵²[Vol54, p.317]

Sous ces conditions Volterra envisage alors à nouveau d'étudier la variation de la « fonction de ligne » φ en s'appuyant cette fois-ci sur un seul paramètre, l'abscisse curviligne s d'un point. Étant données deux lignes L et L_1 , on fait correspondre de façon univoque et continue à chaque point $(x(s); y(s); z(s))$ de L un point $(x_1(s); y_1(s); z_1(s))$ de la ligne L_1 . On a donc trois accroissements notés $\delta x = x_1 - x = \epsilon \zeta$, $\delta y = y_1 - y = \epsilon \eta$, $\delta z = z_1 - z = \epsilon \zeta$. Volterra obtient alors le résultat suivant dont il précise que la démonstration est similaire à celle établie dans les notes précédentes¹⁵³

$$\lim_{\epsilon \rightarrow 0} \frac{\Delta \varphi}{\epsilon} = \int_L (X\zeta + Y\eta + Z\zeta) ds$$

à l'aide d'une intégrale curviligne. Comme dans le cas précédent la « variation première de φ » correspond à la partie de premier ordre de la variation $\Delta \varphi$, c'est-à-dire $\delta \varphi = \int_L (X\zeta + Y\eta + Z\zeta) ds$.

Volterra remarque tout d'abord que dans cette écriture de la variation première les dérivées ne sont pas indépendantes les unes des autres. En effet, lorsque la ligne L est déplacée le long d'elle-même pour donner une ligne L_1 qui n'est qu'une reparamétrisation de L , la variation de φ doit être nulle. Volterra en déduit que les dérivées sont liées par la relation $\alpha X + \beta Y + \gamma Z = 0$ où α , β , et γ sont les cosinus des angles que fait la tangente à la courbe L avec les axes des coordonnées.

Enfin en analysant à nouveau les quantités X , Y , Z comme des « fonctions de lignes » cette fois-ci, Volterra envisage des dérivées partielles secondes qui font intervenir un nouveau paramètre : $X'_x|[L, s, s_1]|$ et les huit autres similaires. Il montre que les dérivées secondes partielles ont alors les propriétés suivantes :

1. $X'_x|[L, s, s_1]|$, $Y'_y|[L, s, s_1]|$ et $Z'_z|[L, s, s_1]|$ sont des fonctions symétriques de s et s_1 .
2. L'ordre de dérivation n'importe pas :

$$\begin{cases} X'_y|[L, s, s_1]| = Y'_x|[L, s, s_1]| \\ X'_z|[L, s, s_1]| = Z'_x|[L, s, s_1]| \\ Y'_z|[L, s, s_1]| = Z'_y|[L, s, s_1]| \end{cases}$$

3. Les dérivées sont encore liées par des relations linéaires :

$$\begin{cases} \alpha_x X'_x + \alpha_y X'_y + \alpha_z X'_z = 0 \\ \beta_x Y'_x + \beta_y Y'_y + \beta_z Y'_z = 0 \\ \gamma_x Z'_x + \gamma_y Z'_y + \gamma_z Z'_z = 0 \end{cases}$$

où les coefficients dépendent de la tangente au point d'abscisse s_1 .

¹⁵³Volterra donne une démonstration dans l'article publié dans *Acta Mathematica* en 1889 (cf. [Vol54, p.366-368])

Cette fois encore Volterra traite les « points exceptionnels » et montre que l'on obtient en plus de l'intégrale une somme finie de termes qui mettent en évidence la dépendance particulière que peut présenter la fonction φ vis-à-vis des coordonnées (et des dérivées des coordonnées) des points de la courbe L .

La seconde note publiée par Volterra sur les « fonctions de lignes » s'appuie sur la version géométrique que nous venons de suivre et montre la portée qu'il donne à cette vision en établissant des résultats extrêmement importants.

Une première étape consiste à montrer que si $\delta\varphi$ est une intégrale curviligne alors pour deux courbes L_1 et L_2 la différence $\varphi|[L_2]| - \varphi|[L_1]|$ peut s'exprimer comme une intégrale de surface. Pour mener à bien cette analyse, Volterra passe de l'idée d'accroissement qui a prévalu jusqu'ici comme nous l'avons vu à l'idée de *déplacement*¹⁵⁴ qui permet de concevoir de proche en proche des positions successives d'une même courbe passant d'une position L_1 à une position L_2 . Pour une portion de courbe ds on peut considérer comme nous l'avons vu précédemment un parallélogramme mixtiligne qui peut être orienté grâce à l'orientation de l'arc ds . Ce parallélogramme infinitésimal permet alors de définir une normale désignée par n et orientée « en prenant pour direction positive celle d'un observateur qui voit le point mobile se déplacer dans le sens des aiguilles d'une montre » comme le précise Volterra. L'aire du parallélogramme infinitésimal est notée $d\sigma$. Enfin Volterra envisage le *déplacement* de la ligne L_1 jusqu'à ce qu'elle vienne coïncider avec la ligne L_2 avec la même orientation que cette dernière. Ce *déplacement* entre les deux lignes engendre une surface Σ ¹⁵⁵, en général annulaire si les courbes sont fermées.

La démonstration se fait alors en deux temps. Au point de vue infinitésimal, c'est-à-dire en considérant le parallélogramme et l'élément de surface $d\sigma$ ainsi que la normale n , Volterra montre qu'il existe A , B et C (d'ailleurs non uniques) tels que :

$$\delta\varphi = \int_{\sigma} (A \cos nx + B \cos ny + C \cos nz) d\sigma.$$

Puis en considérant que le déplacement total résulte de déplacements infiniment petits, Volterra en déduit¹⁵⁶ :

$$\varphi|[L_2]| - \varphi|[L_1]| = \int_{\Sigma} (A \cos nx + B \cos ny + C \cos nz) d\Sigma$$

qui exprime la variation de φ par une intégrale (ordinaire) de surface.

3.3.5 LINÉARITÉ ET SIMPLICITÉ

Nous avons rappelé plus haut que Jellett avait mis en avant la possibilité d'exprimer dans le cas d'un « symbole de dérivation F distributif » la variation de $\delta F \cdot \varphi$ sous la forme $F \cdot \delta\varphi$. Du point de vue de la généralité, nous avons vu que

¹⁵⁴Voir la note (α) de fin de section sur une distinction similaire chez Jellett.

¹⁵⁵On peut préciser avec le texte de Volterra sa conception de la surface Σ . Elle contient pour lui deux « systèmes de courbes » : il y a d'une part les courbes obtenues en observant les positions successives lors du déplacement de la courbe L_1 , et d'autre part l'ensemble des courbes qui sont les trajectoires décrites par chaque point de L_1 jusqu'au « point conjugué » sur L_2 : Volterra dira que l'on a « mené une surface par L_1 et L_2 » (*è condotta una superficie*).

¹⁵⁶[Vol54, p.322]

la *distributivité* était chez Jellett liée aux propriétés des opérations particulières de dérivation ou d'intégration, mais qu'il l'introduisait de façon plus générale. Ici, comme nous l'avons déjà montré, la variation et la « fonction dérivée » s'écrivent avec le même symbole F .

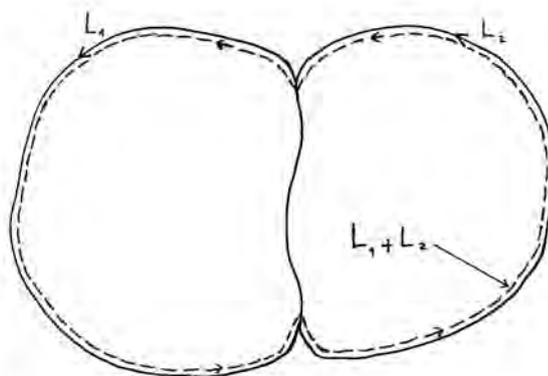
Nous rappelons cette idée sur laquelle Jellett attire très tôt l'attention dans son texte de 1850 car dans la dernière note de Volterra que nous étudions, celui-ci introduit aussi une notion de *distributivité* qui va lui permettre de donner la forme générale de toute « fonction de lignes » distributive en employant une intégrale tout à fait similaire à celle qui décrit la variation $\delta\varphi$. Il y a cependant d'importantes différences entre les deux approches et nous tachons ici d'éclairer les différents types de généralités qui se dégagent des approches des deux mathématiciens.

Volterra propose pour commencer de considérer la concaténation de certaines courbes dont il désigne le résultat par le symbole $+$ ¹⁵⁷ :

Si considerino due linee L_1 e L_2 che hanno un tratto l a comune, e si supponga che le direzioni di L_1 e L_2 siano tali che il tratto l debba venir percorso in senso opposto secondo che si ritiene essere appartenente all'una o all'altra linea. Tolto l le porzioni di L_1 e L_2 formeranno un'unica linea L_3 e ambedue le porzioni verranno percorse in uno stesso senso che si fisserà come direzione della L_3 . Scriveremo:

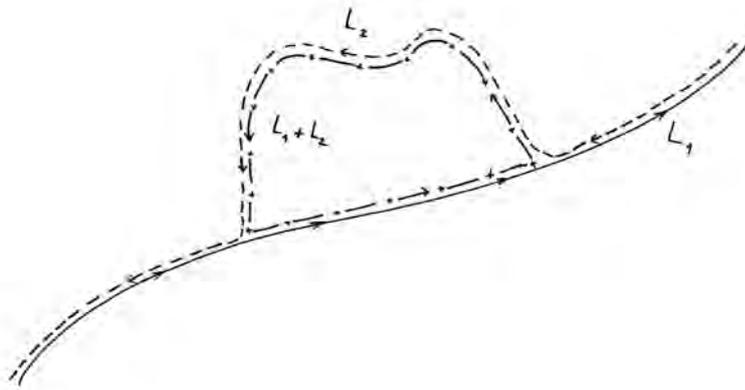
$$L_3 = L_1 + L_2.$$

Les illustrations ci-dessous montrent deux situations de concaténation¹⁵⁸ qui permettront de suivre plus facilement les idées de Volterra dans la suite :



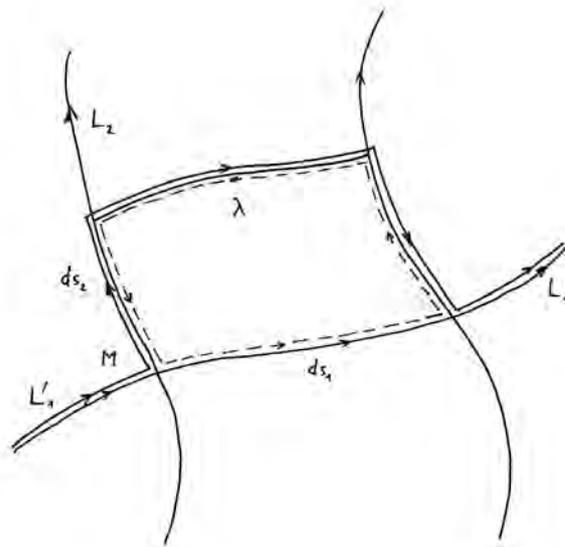
¹⁵⁷[Vol54, p.322]

¹⁵⁸Dans les *Leçons de Stockholm* Volterra fera le rapprochement de cette concaténation avec l'utilisation des courbes dans les intégrales curvilignes.



Bien que Volterra utilise le signe + ici il ne parle pas d'addition et le symbole ne s'applique que pour des couples de courbes particuliers. En fait c'est l'utilisation ultérieure de cette opération sur les courbes qui permet de comprendre la portée générale de la conception de Volterra. Celui-ci ne va par la suite ne considérer que des paires formées d'une courbe donnée et de la même courbe qui aura subi une perturbation locale (ou un déplacement local). Ce qui suit montre le type de situations sur lesquelles Volterra va s'appuyer dans la suite de sa progression.

Dans la continuité des idées qui avaient conduit à la conception d'une surface à partir de différentes lignes et de leurs déplacements, Volterra considère un point M sur une surface et deux lignes qui se coupent en M . Deux portions infiniment petites ds_1 et ds_2 sur L_1 et L_2 respectivement permettent de construire un parallélogramme mixtiligne noté λ . Le schéma ci-dessous montre l'organisation des divers éléments considérés par Volterra dans la suite de son article.



La courbe L'_1 est une perturbation locale de L_1 dans la direction de L_2 . Le couple de courbes à considérer dans la suite est (L'_1, λ) et on a donc $L'_1 + \lambda = L_1$.

Dans cette publication, ou même celle d'*Acta Mathematica* de 1889 Volterra n'indique à aucun moment qu'il considère ce symbole $+$ comme une opération algébrique : les termes d'algèbre comme « addition », « opération », etc. sont absents, ainsi que des notations qui pourraient venir à l'idée de l'observateur, un signe $-$ par exemple pour décrire le parallélogramme à partir des autres courbes. Même dans ce qui va suivre Volterra n'emploiera pas, à propos de cette addition, de vocabulaire de l'algèbre que nous avons déjà vu utilisé en analyse, comme « distributivité » ou « linéarité ». On fera donc ici une distinction entre le choix de Volterra qui n'ancre pas ses idées dans l'algèbre, avec celui par exemple de Jellett qui parle de « symbole distributif »¹⁵⁹ dans un contexte qui se rapproche davantage des idées de l'algèbre symbolique telle que nous l'avons présentée au début de notre thèse.

La concaténation de deux courbes, indiquée par le signe $+$ par Volterra, ne doit donc pas être associée à une idée de structure algébrique telle que nous pourrions la concevoir aujourd'hui ou telle que nous pourrions l'identifier dans des textes qui définissent des groupes par exemple à l'époque de Volterra¹⁶⁰. Pourtant la concaténation des courbes permet à Volterra de définir en ses propres termes une propriété de distributivité (expression d'observateur).

C'est sous l'expression « fonction simple de lignes » (*funzione semplice delle linee*) que Volterra introduit dans un second temps la propriété : φ est dite « simple » si pour chaque paire de lignes $(L_1; L_2)$ qui permettent la concaténation on a :

$$(*) \quad \varphi|[L_1 + L_2]| = \varphi|[L_1]| + \varphi|[L_2]|.$$

Dans les premiers textes sur les « fonctions de lignes » deux choses motivent l'introduction de cette propriété. Elle permet de pousser l'analyse de la variation plus loin et elle est suffisamment générale. En ce qui concerne la généralité, les opérations d'intégration ou de dérivation qui se présentent dans les cas usuels de l'étude des variations, interprétées pour Volterra comme « fonctions qui dépendent d'autres fonctions » ou « fonctions de lignes », possèdent toutes cette propriété. Ces deux points avaient été aussi soulignés par Jellett lorsqu'il introduisait dans son propre contexte une propriété de distributivité. Volterra poursuit alors son étude sous cette hypothèse de distributivité ou de *simplicité*.

Dans la situation indiquée sur le schéma au-dessus on a $L'_1 + \lambda = L_1$, et par la propriété (*) on obtient $\varphi|[L'_1]| - \varphi|[L_1]| = \varphi|[\lambda]|$. Mais Volterra change de point de vue et considère une configuration similaire avec une courbe \tilde{L}_2 qui est L_2 parcourue dans le sens contraire¹⁶¹ comme élément de base, et cette courbe aurait été déplacée dans la direction de L_1 de façon à engendrer le même parallélogramme mixtiligne orienté λ . On a dans ce cas $\tilde{L}'_2 + \lambda = \tilde{L}_2$ et à nouveau $\varphi|[\tilde{L}'_2]| - \varphi|[\tilde{L}_2]| = \varphi|[\lambda]|$. Volterra montre donc que dans le cas des « fonctions

¹⁵⁹[Jel50, p.4]

¹⁶⁰On peut penser à la vision plus *réflexive* que suggère un travail comme celui de Walther Dyck dans « Gruppentheoretische Studien » par exemple. (cf. [Dyc82])

¹⁶¹Nous n'employons pas la notation $-L_2$ afin de ne pas introduire un élément ambigu qui pousserait le symbole $+$ du côté de l'algèbre.

simples » on a $\delta_1\varphi = \delta_2\varphi$, où δ_1 et δ_2 désignent les variations calculées des deux façons, mais associées au même parallélogramme mixtiligne orienté. On voit donc ici que la concaténation des courbes et la propriété (*) permettent à Volterra de décrire une sorte d'invariance géométrique des variations d'une fonction. Dans le cas « simple » la variation dépend du parallélogramme mixtiligne orienté par la normale donnée par deux courbes qui se coupent en M mais ne dépend pas de l'ordre des déplacements qu'on envisage pour engendrer la surface λ .

En prolongeant ces idées pour une « fonction de lignes simple », Volterra prend différents couples de lignes qui se coupent en M pour prouver le théorème suivant¹⁶² :

Se φ è una funzione semplice delle linee di un campo a tre dimensioni, esistono per ogni punto del campo tre valori P, Q, R che possono rispettivamente prendersi come valori di A, B, C ¹⁶³ in quel punto per tutte le linee che vi passano.

Ces considérations permettent à Volterra d'écrire $\varphi|[L_2]| - \varphi|[L_1]| = \int_{\Sigma} (P \cos nx + Q \cos ny + R \cos nz) d\Sigma$. Si maintenant la ligne L_1 est contractée en un point on obtient $\varphi|[L_1]| = 0$ et donc une expression générale de $\varphi|[L]| = \int_{\Sigma} (P \cos nx + Q \cos ny + R \cos nz) d\Sigma$. À nouveau Volterra imagine que la surface Σ peut être réduite indéfiniment autour d'un point M ce qui permet de définir un nouveau rapport différentiel $\frac{d\varphi}{d\Sigma} = \lim_{\Sigma} \frac{\varphi|[L]|}{\Sigma} = P \cos nx + Q \cos ny + R \cos nz$. Lorsque l'on choisit pour Σ des plans normaux aux axes du repère on retrouve les coefficients P, Q et R que Volterra notera donc $\frac{d\varphi}{d(y,z)} = P, \frac{d\varphi}{d(z,x)} = Q$, et $\frac{d\varphi}{d(x,y)} = R$. Volterra note aussi réciproquement que lorsque les coefficients P, Q et R sont donnés alors il existe une « fonction de ligne » φ telle que $\frac{d\varphi}{d(y,z)} = P, \frac{d\varphi}{d(z,x)} = Q$, et $\frac{d\varphi}{d(x,y)} = R$.

Enfin cette troisième note de Volterra sur les « fonctions de lignes » se termine par une section qui donne un théorème de représentation des « fonctions simples de lignes ». Ce résultat sera repris et commenté par Hadamard et Fréchet comme nous le verrons par la suite. Nous donnons la démarche de Volterra que l'on pourra contraster avec celle des deux mathématiciens français.

Étant donnée une « fonction de lignes simple » L , on a $\frac{\partial P}{\partial x} + \frac{\partial Q}{\partial y} + \frac{\partial R}{\partial z} = 0$.

En utilisant la méthode des multiplicateurs de Jacobi¹⁶⁴ on peut trouver deux

¹⁶²[Vol54, p.324]

¹⁶³Les coefficients P, Q, R sont ceux obtenus pour les couples de courbes particulières qui passent par M et qui sont parallèles aux axes du repère. Les coefficients A, B et C sont ceux construits au départ et qui dépendent de la ligne L qui passe par M .

¹⁶⁴Volterra renvoie à [Jac81, p.78].

fonctions λ et μ de x , y , et z qui vérifient le système :

$$\begin{cases} \frac{\partial \lambda}{\partial y} \frac{\partial \mu}{\partial z} - \frac{\partial \lambda}{\partial z} \frac{\partial \mu}{\partial y} = \frac{d(\lambda, \mu)}{d(y, z)} = P \\ \frac{\partial \lambda}{\partial z} \frac{\partial \mu}{\partial x} - \frac{\partial \lambda}{\partial x} \frac{\partial \mu}{\partial z} = \frac{d(\lambda, \mu)}{d(z, x)} = Q \\ \frac{\partial \lambda}{\partial x} \frac{\partial \mu}{\partial y} - \frac{\partial \lambda}{\partial y} \frac{\partial \mu}{\partial x} = \frac{d(\lambda, \mu)}{d(x, y)} = R \end{cases}$$

Enfin en posant $a = \lambda \frac{\partial \mu}{\partial x}$, $b = \lambda \frac{\partial \mu}{\partial y}$ et $c = \lambda \frac{\partial \mu}{\partial z}$ on obtient $\frac{\partial c}{\partial y} - \frac{\partial b}{\partial z} = P$, $\frac{\partial a}{\partial z} - \frac{\partial c}{\partial x} = Q$ et $\frac{\partial b}{\partial x} - \frac{\partial a}{\partial y} = Q$.

Volterra en déduit que toute « fonction simple de lignes » peut s'écrire pour une ligne L qui délimite une surface σ :

$$\varphi|[L]| = \int_{\sigma} \left\{ \left(\frac{\partial c}{\partial y} - \frac{\partial b}{\partial z} \right) \cos nx + \left(\frac{\partial a}{\partial z} - \frac{\partial c}{\partial x} \right) \cos ny + \left(\frac{\partial b}{\partial x} - \frac{\partial a}{\partial y} \right) \cos nz \right\} d\sigma.$$

Le théorème de Stokes permet à Volterra de donner alors la forme générale d'une « fonction simple de lignes »¹⁶⁵ :

$$\varphi|[L]| = \int_L (adx + bdy + cdz) = \int_L \lambda d\mu.$$

Ces résultats seront repris avec quelque nuances dans l'article d'*Acta Mathematica* de 1889 qui renvoie abondamment dans la partie sur les « fonctions de lignes » aux articles initiaux que nous venons d'étudier. Dans cet article de 1889 Volterra nomme les fonctions qu'il avait qualifiées de « simples » précédemment « fonctions du premier degré ». Si la « simplicité » initiale était liée à une propriété de la fonction qui permettait de mettre en place un certain type de raisonnement géométrique, la notion de « premier degré » dégage une vision différente qui s'appuie sur le résultat de représentation des « fonctions de lignes simples » que nous venons de voir. Une « fonction de lignes simple » se présente comme l'intégrale d'une forme différentielle du premier degré. Nous verrons par la suite comment Fréchet repensera et intégrera ces idées dans son approche générale.

3.4 CONCLUSION

L'analyse précise de ces cinq notes dans lesquelles Volterra développe pour la première fois les concepts de « fonctions qui dépendent d'autres fonctions » et sa version géométrique de « fonctions de lignes » a permis de dégager quelques éléments de la pensée de Volterra.

UN PARADIGME FONDÉ SUR UNE NOTION DE FONCTION. Tout d'abord Volterra choisit ici de fonder une vision générale qui permette d'appréhender un certain nombre de problèmes qui ont été jusqu'à présent traités en s'appuyant

¹⁶⁵[Vol154, p.328]

sur une analyse qui conduit à des équations intégrales ou différentielles. Volterra relie alors ces situations diverses en identifiant à chaque fois une dépendance par rapport à une ou plusieurs fonctions ordinaires. Comme le souligne Volterra lui-même, cette notion de dépendance est intimement liée à une conception physique des problèmes envisagés. D'ailleurs les publications du mathématicien italien portent la trace jusque dans leurs titres de ce rapport à la physique, il suffit de consulter sa bibliographie pour s'en convaincre. La physique à ceci de particulier qu'elle fait correspondre des grandeurs à une multitude d'éléments, et cette diversité engendre une certaine généralité que la notion abstraite de fonction permet de saisir.

Volterra choisit d'étendre le concept de fonction en s'appuyant sur une définition de Dirichlet qui propose une vision abstraite fondée sur la notion même de correspondance entre des quantités. En gardant l'idée de correspondance, Volterra conçoit une idée de fonction qui à des éléments divers (fonctions ordinaires, lignes, surfaces, espaces) fait correspondre des nombres réels ou complexes.

La construction du nouveau paradigme est donc fondée sur la notion de fonction étendue et sur l'adaptation du calcul infinitésimal des fonctions ordinaires à ce nouveau concept. C'est la deuxième fois que Volterra cherche à transposer le calcul infinitésimal des fonctions ordinaires à un cadre généralisé. Nous avons éclairé les contrastes qui apparaissent entre l'étude des substitutions et celui des fonctions de lignes. La dernière notion appuie sa généralité sur des éléments spécifiques que nous avons analysé dans les cinq articles fondateurs de Volterra.

La transposition de la théorie infinitésimale des fonctions ordinaires que propose Volterra repose sur la notion de variation qui est apparue comme un outil suffisamment général pour s'adapter à des situations nouvelles. Nous avons vu comment Jellett s'appuyait déjà sur la variation dans un cadre abstrait et Volterra se saisit à son tour de cet outil pour développer une théorie infinitésimale des fonctions de lignes. La notion ingénieuse de dérivée d'une fonction de ligne permet d'établir la forme de la variation :

$$\int_A^B y' |[\varphi(x), t]| \delta\phi(t) dt.$$

Volterra exposera cette théorie générale dans de nombreux cours jusqu'à la fin de sa carrière et il utilisera le concept de fonction de ligne pour étudier divers problèmes.

On peut rattacher l'utilisation des fonctions de lignes chez Volterra à deux types de situations qui font apparaître des aspects généraux différents.

LA DIMENSION $n \geq 2$. D'abord dans un certain nombre de problèmes la notion de fonction de ligne permet à Volterra une généralisation relative à la dimension de l'espace. Cette vision générale contraste avec celle qui tend à concevoir d'abord le point défini par une seule coordonnée, puis le point défini par 2 puis n coordonnées. Pour Volterra la généralisation passe par la notion de point, puis de ligne et enfin d'hyper-espace. Mais la ligne peut aussi bien être considérée dans un plan que dans un espace de dimension supérieure ce qui lui confère une généralité particulière.

L'analyse de l'utilisation des fonctions de lignes dans l'œuvre de Volterra dépasse largement le cadre de notre étude et nous indiquons ici seulement quelques thèmes qui illustrent l'aspect général que nous venons de préciser.

1. Nous avons déjà mentionné l'article « Sur une généralisation de la théorie des fonctions d'une variable imaginaire »¹⁶⁶ publié à *Acta Mathematica* en 1889 qui comporte un premier chapitre dédié aux « fonctions de lignes ».

Cet article vient à la suite d'une série de trois notes publiées aux *Rendiconti Accademia Lincei* sous le titre « Sopra una estensione della teoria di Riemann sulle funzioni di variabili complesse » en 1887.

Nous avons rappelé comment ce travail de Volterra avait en ligne de mire une extension du théorème de Cauchy obtenu par Hadamard dans le cas des fonctions de deux variables complexes. Pour étendre la notion de monogénéité aux espaces à 3 dimensions Volterra introduit la notion d'« isogénéité » qui est une adaptation du concept initial aux « fonctions de lignes »¹⁶⁷ :

Pour généraliser la théorie ordinaire des fonctions monogènes aux espaces à trois dimensions, supposons que nous ayons deux variables imaginaires qui soient des fonctions de 1^{er} degré des lignes de l'espace. Nous allons établir une condition tout à fait semblable à la condition de monogénéité. À cet effet soient F et Φ les valeurs des deux fonctions correspondantes à une ligne L . Déformons un arc AB de L et désignons par ΔF et $\Delta \Phi$ les variations de F et de Φ . Si en diminuant indéfiniment la déformation et la distance entre B et le point fixe A , le rapport $\frac{\Delta F}{\Delta \Phi}$ tend vers une limite qui dépend seulement du point A , on dira que les deux variables ont une liaison d'isogénéité, ou qu'elles sont isogènes.

En s'appuyant sur cette notion et les travaux sur les « fonctions de lignes » développés en 1887 Volterra montre que si f , fonction ordinaire, et F , « fonction de lignes », ont une liaison d'« isogénéité »¹⁶⁸ et si σ est une surface orientée alors la quantité $\int_{\sigma} f dF := \int_{\sigma} f \frac{dF}{d\sigma} d\sigma$ est bien définie.

Volterra peut alors donner une extension du théorème de Cauchy en montrant entre autres que $\int_{\sigma} f dF = 0$ pour une surface σ fermée qui délimite un « champ » S dans lequel f et F n'ont pas de singularités.

2. Les « fonctions de lignes » servent aussi à Volterra pour étendre la théorie de Jacobi-Hamilton. Ici le but est de passer de l'étude de l'annulation de la variation première d'une intégrale simple au résultat pour les intégrales multiples qui sont naturellement présente en physique. La méthode initiale est décrite par Volterra¹⁶⁹ :

Il procedimento JACOBI-HAMILTON si fonda sull'esame dell'integrale semplice (di cui si vuole annullare la variazione) considerato come funzione dei suoi limiti e dei valori assegnati ad arbitrio

¹⁶⁶[Vol54, p.363]

¹⁶⁷[Vol54, p.377]

¹⁶⁸Volterra adapte la notion dans le cas de couples mixtes fonction ordinaire / fonction de lignes : la condition d'isogénéité s'écrit $\frac{dF}{d(y,z)} \frac{df}{dx} + \frac{dF}{d(z,x)} \frac{df}{dy} + \frac{dF}{d(x,y)} \frac{df}{dz} = 0$.

¹⁶⁹[Vol54, p.464]

alle funzioni incognite nei limiti stessi. È una tale funzione (la funzione caratteristica) che soddisfa alle equazioni differenziali a derivate parziali scoperte da HAMILTON e che fornisce gli integrali del problema mediante operazioni di derivazione.

Donc pour étendre cette idée Volterra propose dans le cas des intégrales doubles de les considérer comme fonctions des lignes qui limitent le contour mais aussi de paramètres qui sont les valeurs prises sur ces lignes par les fonctions qui interviennent dans l'intégrale. La généralisation se poursuit alors en envisageant des fonctions d'hyper-espaces et de différents paramètres.

3. Le dernier exemple que nous citons ici concerne l'addition des intégrales elliptiques. Dans le cas initial auquel Volterra se réfère, on considère deux intégrales simples dont les bornes supérieures sont deux variables x et y . Le théorème d'addition dit que la somme des intégrales est constante lorsque le point $(x; y)$ décrit une certaine courbe définie par une relation algébrique entre x et y .

Une première voie de généralisation tentait de prendre des fonctions de plus en plus générales sous le signe d'intégration. Volterra emprunte une autre voie de généralisation qui consiste dans un premier temps à étendre par des méthodes classiques (sans intervention de « fonctions de lignes ») le résultat à des intégrales doubles. Il considère alors des sommes du type $J = \iint_{\sigma_1} \varphi_1(y, z) dy dz + \iint_{\sigma_2} \varphi_2(z, x) dz dx + \iint_{\sigma_3} \varphi_3(x, y) dx dy$ qu'il associe à l'équation différentielle $\varphi_1 \frac{\partial f}{\partial x} + \varphi_2 \frac{\partial f}{\partial y} + \varphi_3 \frac{\partial f}{\partial z}$. Si f est une solution de l'équation différentielle et que σ_1 , σ_2 et σ_3 sont les projections de la surface $f = \text{const}$ sur les plans de base du repère, alors $J = 0$.

Mais dans un second temps, et après avoir cité une remarque de Picard sur le même sujet qui le pousse probablement à regarder le problème de façon plus générale¹⁷⁰, Volterra donne une analyse en termes de « fonctions de lignes ».

En effet le problème fait apparaître des équations de type $\frac{\partial X}{\partial x} + \frac{\partial Y}{\partial y} + \frac{\partial Z}{\partial z} = 0$ où X , Y et Z sont des fonctions ordinaires de 3 variables qui dans la théorie de Volterra expriment qu'il existe une « fonction de lignes » F telle que $X = \frac{dF}{d(y, z)}$, $Y = \frac{dF}{d(z, x)}$ et $Z = \frac{dF}{d(x, y)}$. Soit alors f une solution de l'équation différentielle $X \frac{\partial f}{\partial x} + Y \frac{\partial f}{\partial y} + Z \frac{\partial f}{\partial z}$ et Σ la surface $f = \text{constante}$. Dans ses textes initiaux sur les « fonctions de lignes » Volterra a montré que si L se déplace sur Σ alors $F(L)$ reste constante. Et réciproquement $F(L) = \text{constante}$ peut être interprété comme une équation qui définit l'ensemble des surfaces $f = \text{constante}$ où f est solution de l'équation différentielle $X \frac{\partial f}{\partial x} + Y \frac{\partial f}{\partial y} + Z \frac{\partial f}{\partial z}$.

Ces exemples illustrent une façon d'utiliser les fonctions de lignes et montrent la généralité de ce concept à l'œuvre dans un premier type de situations.

MÉCANIQUE HÉRÉDITAIRE. Mais il y a un autre type de généralisation que Volterra conçoit grâce à l'idée de ligne. Elle est particulièrement bien illustrée

¹⁷⁰[Vol54, p.333-334]

dans un certain type de mécanique dite « héréditaire » que Volterra repense en termes de fonctions de lignes¹⁷¹

Comme M. PICARD a montré dans son intéressant article sur la Mécanique classique et ses approximations successives, il faut distinguer la mécanique en deux branches, celle de l'hérédité et celle de la non-hérédité. Celle-ci se rapporte aux cas où l'avenir d'un système ne dépend à un instant donné que de son état actuel, ou, d'une manière plus générale (si l'on regarde les forces comme pouvant dépendre aussi des vitesses) de l'état actuel et de l'état infiniment voisin qui précède. La mécanique de l'hérédité correspond au cas où chaque action laisse un héritage dans le système, et l'état actuel dépend de toute l'histoire précédente. C'est ainsi que le problème fondamental de l'astronomie appartient à la mécanique de la non-hérédité, tandis que les questions d'*hystérésis*, de l'*élastische Nachwirkung*, du *traînage* rentrent dans la mécanique de l'hérédité, ou, plus général, dans la physique d'hérédité.

Volterra va alors concevoir le type de dépendance particulière qu'engendre l'hérédité en introduisant des fonctions de lignes du type $y = y|[\varphi(t)]|_{-\infty}^x$ où l'on prend en compte toutes les valeurs de φ jusqu'au temps $t = x$.

Nous ne rentrons pas plus avant dans les détails de ces aspects essentiels et passionnants de l'œuvre de Volterra qui illustrent une nouvelle capacité générale du concept de fonction de ligne qui n'est pas associé à un changement de dimension de l'espace ambiant¹⁷².

UNE LECTURE À TRAVERS LE COUPLE LOCAL / GLOBAL. La notion d'*infinitésimal* à laquelle Volterra a recours dans les textes que nous avons étudiés recouvre diverses situations. Nous avons en particulier mis au jour différents types de *localisations* qui permettent de préciser l'extension générale que l'on peut attribuer au travail de Volterra.

Tout d'abord les textes du mathématicien italien font apparaître deux notions qui pour nous relèvent du *global*.

Premièrement la notion initiale de « fonction qui dépend d'autres fonctions » s'appuie sur une dépendance qui met en jeu *toutes les valeurs* d'une fonction sur un intervalle donné AB et nous la qualifierons de notion *globale par quantification (universelle)*. La notation adoptée pour désigner les « fonctions qui dépendent

¹⁷¹[Vol57, pp.487-88]

¹⁷²la question de l'hérédité est à l'époque au centre de vives discussions qualifiées par les mathématiciens de *philosophiques*. Volterra cite la position de Painlevé : « M. PAINLEVÉ dans le chapitre de l'ouvrage *de la méthode dans les sciences* consacré à la mécanique affirme qu'il n'y a pas de vrais problèmes de nature héréditaire. Ceux qui se présentent sous cet aspect ne seraient, à son avis, que des problèmes destinés à disparaître dès que nos connaissances sur la constitution des corps deviendront plus complètes. » La conception mathématique de Volterra a pour effet de donner une existence indiscutable aux fonctions qui dépendent de toutes les valeurs d'une fonction donnée jusqu'au temps t . Il ne cherche d'ailleurs pas à polémiquer et reste loin des considérations extra-mathématiques dans ce domaine : « Je ne discute pas cette opinion, mais je me limite à remarquer qu'à l'état actuel de nos connaissances scientifiques ces problèmes se présentent effectivement et il est nécessaire de les résoudre. » (cf. [Vol57, pp.488-489])

d'autres fonctions » $y|[\varphi(x)]$ laisse apparaître la variable de la *fonction-objet* et la façon dont Volterra conçoit une vision abstraite globale qui embarque toutes les valeurs prises sur un intervalle.

La seconde situation qui pour nous relève du *global* se trouve dans les articles mettant en œuvre les « fonctions de lignes ». Tant que l'on reste au niveau des définitions données en introduction (par contraste avec les emplois ultérieurs dans les démonstrations) la ligne est vue de façon abstraite et sans recours à un paramètre. Celle-ci est conçue comme un élément d'une *géométrie-physique*¹⁷³ qui lui confère un caractère à la fois général et *global*. Pour distinguer ce cas du précédent nous parlerons ici d'une notion *globale-par-abstraction*. La notation choisie par Volterra $\varphi|[L]$ reflète cette vision sans référence à une paramétrisation de la ligne L . De plus cet aspect que nous qualifions de *global-par-abstraction* n'est pas altéré lorsque Volterra définit la concaténation de deux lignes qu'il désigne par le signe $+$. Ce signe, importé de l'algèbre, précise une certaine configuration nécessaire pour rendre possible la concaténation en tenant compte du sens de parcours des lignes. Ces différents aspects développés dans une partie limitée des textes que nous avons étudiés définissent l'extension et la nature particulière de la généralité qui est à l'œuvre ici. Nous rapprocherons ce type de généralité de celle que nous pourrions identifier plus tard chez Fréchet par exemple.

Les deux types de traitement globaux que nous avons fait apparaître vont donner lieu à différentes notions *locales* que nous allons maintenant distinguer.

D'une part à la notion de « ligne » comme support d'une vision *globale-abstraite* permet à Volterra de concevoir un « voisinage » tubulaire d'une ligne que nous avons commenté. Nous qualifierons cette notion de « voisinage » de *locale-abstraite*, où le mot *local* fait à la fois référence au fait que l'idée de voisinage se déploie à partir d'une ligne donnée, mais aussi aux aspects infinitésimaux convoqués par Volterra et dont nous avons analysé la teneur.

Chacun de ces aspects est utilisé par Volterra, encore une fois dans une partie bien délimitée de son texte, en ne faisant référence qu'à la notion générale de ligne dans une conception qui s'appuie sur une *géométrie-physique*. Volterra se sert de cette notion de « voisinage » pour définir dans le même cadre abstrait la continuité d'une « fonction de ligne ». Ici encore c'est un aspect général d'un genre particulier que Volterra met en place dès 1887 et nous le rapprocherons de la conception de Fréchet ou de Banach plus tard.

D'autre part l'aspect *global-par-quantification* qui apparaît dans la grande majorité des textes que nous avons étudiés donne lieu à deux types de notions *locales-quantifiées*.

Le premier type apparaît quand Volterra définit un accroissement de façon générale en écrivant $\varphi(x) + \theta(x)$ où θ est inférieure à ϵ en valeur absolue comme il l'écrit. La fonction θ est donc contrôlée globalement et donne lieu à une notion

¹⁷³Nous distinguons ici le type d'éléments géométriques que convoque Volterra et qu'il décrit très souvent comme issus de situations évoquées par la physique. L'introduction de la première note sur les « fonctions de lignes » montre bien cette inspiration : « Una tale idea è famigliare ai fisici ». Nous soulignons par l'expression *géométrie-physique* qu'il découle de cette conception une vision géométrique particulière qui n'est pas purement mathématique. Cette distinction nous semble importante pour préciser un certain registre du général qui intervient dans les textes que nous étudions ici.

de voisinage (mots d'observateur ici) de φ ou tout du moins d'une collection de fonctions perturbées $\varphi + \theta$ complètement générales où la seule contrainte est que la perturbation soit aussi petite qu'on le désire. C'est ici sur une notion définie pour chaque x de l'intervalle (AB) que s'appuie Volterra pour définir une notion d'accroissement qui a un aspect global et général.

Le second type de notion *locale-quantifiée* joue sur une double localisation, à la fois en référence à une fonction particulière φ et par rapport à un point de la courbe représentative de φ qui est repéré par son abscisse $G(t)$. La perturbation de φ doit rester plus petite qu'un ϵ donné sur l'intervalle (AB) tout entier, mais elle doit aussi être nulle en dehors d'un intervalle de longueur h aussi petite qu'on souhaite l'imposer qui contient l'abscisse t du pont G . Les deux contraintes donnent une perturbation θ qui est petite selon ces deux aspects que l'intégrale $\int_A^B \theta(x)dx$ résume en un seul élément qui mesure la taille de l'accroissement. Ces accroissements sont plus contraints que les précédents et donnent un voisinage (mot d'observateur) de φ particulier.

Dans le cas des fonctions de lignes on retrouve les deux aspects *locaux-quantifiés* qui sont déclinés en ayant recours à une paramétrisation de la courbe.

Nous avons de plus montré qu'il y avait chez Volterra un passage du *local-quantifié* au *global-abstrait* qui s'appuyait sur deux notions.

La première est la dérivée d'une « fonction qui dépend d'autres fonctions » (et ses variantes pour les lignes) qui est au départ construite à partir d'un accroissement particulier θ tel que $\int_A^B \theta(x)dx$ est aussi petit qu'on le désire. Par passage à la limite du rapport différentiel correspondant, Volterra définit la notion de dérivée. Enfin dans une dernière phase, rendue explicite lors de la considération des dérivées successives, il réinterprète cette dérivée en termes de « fonction qui dépend d'autres fonctions ». La notion $y'[[\varphi(x)]]$ laisse encore transparaître un *global quantifié* qui disparaît avec le passage aux « fonctions de lignes » et l'écriture des dérivées $\varphi'_x = X[[L, s]]$.

La seconde est l'intégrale comprise au sens de Riemann, qui par un autre passage à la limite permet de passer d'une somme mettant en jeu des accroissements doublement localisés à une conception en termes de « fonctions de lignes » du type $\varphi[[L]] = \int_L \lambda d\mu$ comme nous avons vu.

Enfin, nous venons de montrer que l'un des aspects importants du travail de Volterra sur les fonctions généralisées et du calcul infinitésimal qu'il adapte consistait à passer d'une conception *locale* à une conception *globale* et abstraite. Un dernier aspect qu'il nous reste à éclairer concerne un mouvement parallèle qui permet de préciser la pensée de Volterra et le type de généralité que nous y lisons.

PASSAGE DU DISCRET AU CONTINU ET DU FINI À L'INFINI. Dans de nombreux textes, et en particulier dans les cours qu'il donnera à plusieurs reprises, Volterra revient sur ses travaux en insistant sur le rôle important qu'a joué pour lui l'idée de « passage du fini à l'infini » ou « du discret au continu ». En fait les textes de Volterra tels que nous les analysons laissent apparaître un tissu complexe d'idées relatives à ces notions.

Dans l'introduction de ses « Leçon sur les fonctions de lignes » professées à la Sorbonne au début de l'année 1912 Volterra met en avant l'idée qu'il faut désormais considérer des lois physiques discontinues. Il cite un travail de Poincaré intitulé « Sur la théorie des quanta »¹⁷⁴ où celui-ci cherche une voie qui permettrait de contourner la *théorie des quanta* développée par Planck dans l'étude du rayonnement des corps noirs. La conception quantique qui y est alors nécessaire semble priver le mathématicien de toute approche fondée sur des équations différentielles¹⁷⁵ :

Il est inutile de faire remarquer combien cette conception s'éloigne des idées habituellement reçues puisque les lois physiques ne seraient plus susceptibles d'être exprimées par des équations différentielles. Il est naturel qu'on cherche à échapper à cette conséquence, sans parler d'une foule de difficultés de détail, et qu'on se demande s'il n'y aurait pas moyen d'expliquer autrement les faits. J'ai donc cherché, si l'on ne pouvait pas rendre compte de la loi de Planck par d'autres hypothèses et je suis arrivé à un résultat négatif.

Volterra répond alors à ce « très vif sentiment de regret »¹⁷⁶ en précisant sa propre vision de l'analyse et du rôle essentiel qu'y jouent pour lui les couples fini/infini et discret/continu.

Pour commencer il met en avant une vision de l'histoire de l'analyse où l'on a ramené les cas de discontinuité à ceux de continuité par un procédé qu'il nomme « statistique »¹⁷⁷ :

La puissance des méthodes des infiniment petits a eu le dessus dans la pratique des calculs sur les conceptions et les hypothèses plus probables se rapportant à la nature même du sujet.

D'après Volterra la rigueur des méthodes permettant de passer du discret au continu s'appuie sur la notion de limite qui s'effectue par « le transport dans le domaine de ce calcul des propriétés connues de l'Algèbre finie et de l'Arithmétique. »

Volterra développe donc une idée de l'analyse où plusieurs points de vue coexistent, où le discret et le continu ne s'opposent pas, la cohérence et la compatibilité du tout étant garantie essentiellement par la notion de limite, qui permet de passer d'une conception à une autre. Notre analyse a montré divers emplois de la notion de limite et a aussi établi que d'autres notions étaient impliquées dans une vision qui réconcilie discret et continu. L'approche de Volterra offre en quelques sortes une réponse à Poincaré en dressant un tableau où le « calcul infinitésimal » garde toute sa place, même dans une analyse qui prend en considération les phénomènes discrets que l'on trouve en physique. En citant plus particulièrement les théories de la chaleur ou de l'élasticité Volterra argumente que le « calcul infinitésimal » permet de traiter ce qui est donné par des lois discontinues de la physique.

Le « calcul infinitésimal » ne semble donc pas mis en danger par les conceptions discontinues et Volterra propose une conception de l'analyse qui

¹⁷⁴[Poi11]

¹⁷⁵[Poi11, p.145]

¹⁷⁶[Vol13, p.2]

¹⁷⁷[Vol13, p.2]

montre comment ce calcul peut être étendu et généralisé afin de s'adapter à des situations nouvelles. Dans la vision de Volterra il y a deux éléments essentiels qu'il isole lui-même : la notion d'*opération* et la notion de *fonction* qui, prises dans un sens particulier, sont deux idées qui lui permettent de concevoir une analyse générale.

Tout d'abord Volterra met en avant l'intégration qui occupe une place de choix dans son approche. Elle représente « le transport de la conception de somme du domaine fini au domaine infini et la dérivation n'est que son opération inverse. Elles sont les bases du calcul infinitésimal. »

Une première voie de généralisation consiste donc pour Volterra à étendre cette idée à d'autres opérations¹⁷⁸ :

Arrivés à ce point on peut se demander : Est-ce que ces opérations sont les seules où ce passage est applicable ? Il est facile de concevoir que le passage du fini à l'infini peut s'appliquer, non seulement à la somme, mais à bien d'autres opérations.

Et l'on peut y arriver, soit moyennant le procédé qu'on appelle des séries, soit par le procédé que nous pouvons comparer à celui du calcul intégral où l'on envisage, d'une manière analogue à ce que faisait Galilée pour étudier la chute des graves, la variation totale d'une quantité comme l'ensemble des variations infiniment petites qui se suivent, obtenues en divisant cette variation totale en intervalles partiels.

Nous avons largement illustré plusieurs situations où Volterra s'appuyait sur des *opérations* élémentaires pour ensuite passer à une conception infinie. Outre les cas de multiplication des substitutions ou de recours à la théorie riemannienne de l'intégration Volterra envisage par exemple la « déformation d'une figure » comme une succession de dilatations qui se combinent par « multiplication de leurs paramètres ».

Dans chacun des cas Volterra propose une idée d'intégration étendue qui traduit le passage d'une succession d'éléments finis et localisés à une conception globale par passage à la limite. Le cas de la multiplication des substitutions que nous avons étudiée permet à Volterra de commenter l'organisation entre le général et le particulier qu'il élabore ici¹⁷⁹ :

Ce passage du fini à l'infini est tout à fait analogue à celui qu'on fait lorsqu'on passe de la somme d'un nombre fini de termes à une intégration, et il est évident que l'opération ordinaire d'intégration n'est qu'un cas particulier de celle que nous venons de définir.

Or, cette suite infinie de transformations infiniment petites correspond à une transformation finie et l'opération ainsi définie, peut s'appeler l'intégration des substitutions linéaires. L'opération inverse sera une dérivation des substitutions et l'on aura un calcul intégral et différentiel des substitutions tout à fait parallèle au calcul intégral et différentiel ordinaire.

Ce calcul correspond à ce qu'on appelle en analyse l'intégration des équations différentielles linéaires. Toute cette théorie peut donc

¹⁷⁸[Vol13, p.7]

¹⁷⁹[Vol13, p.9-10]

s'exposer à ce nouveau point de vue et bien des résultats se coordonnent ainsi dans la partie plus élémentaire comme dans la partie plus élevée.

Ces premières considérations mettent en évidence un principe de généralisation qui est essentiellement fondé sur la notion de limite et la conception riemannienne d'intégrale que Volterra adapte et généralise. On trouve un intérêt pour ce type d'idées dès le début de sa carrière et plus particulièrement dans un article intitulé « Sui principii del calcolo integrale »¹⁸⁰ publié au *Giornale di Matematiche* en 1881. Nous avons vu comment Volterra avait ensuite mis en œuvre et adapté des idées connectées à l'approche de l'intégration riemannienne dans plusieurs domaines.

Un deuxième élément est au cœur de la conception de Volterra et lui permet de développer un nouvel axe de généralisation qui vient se combiner avec celui que nous venons d'évoquer. Il s'agit de la notion de fonction que Volterra convoque, repense et généralise plusieurs fois. Dans ses commentaires *a posteriori* il revient sur différentes conceptions et indique trois perspectives. Pour lui il y a : les fonctions analytiques dont l'idée « est sortie tout naturellement du développement de l'Algèbre » ; une conception géométrique qui « parl[e] aux yeux d'une manière concrète et intuitive » ; et la notion de *dépendance* générale qu'il reprend comme nous l'avons vu des travaux de Dirichlet et qu'il associe à la notion de loi physique, ce qui lui confère dans son optique un caractère général.

Notre analyse des textes nous a permis de voir que la notion de fonction permettaient une progression vers le général qui mettait en jeu des registres multiples. En particulier nous avons vu que Volterra combinait plusieurs fois une approche *discrète* et *locale* pour donner par passage à la limite un concept *continu* et *global* (les termes sont ceux de notre analyse). Nous avons montré que cette façon de penser s'appuyait fortement sur des notions *localisées* (associées à divers types de voisinages en nos termes) qui font le lien entre *discret* et *continu* et sont à la base de ce que Volterra nomme « calcul infinitésimal ».

De plus cette approche, qui est au cœur de la pensée de Volterra dans les textes que nous venons d'étudier, fait apparaître un autre mouvement qui est celui du passage du fini à l'infini qui lui-même s'accompagne d'un type d'abstraction particulier.

Volterra va relire un certain nombre de situations que nous avons qualifiées de *globales-par-quantification* (pour indiquer qu'elles prenaient en considération un ensemble infini de valeurs) comme résultat d'un processus qui permet de passer d'un nombre fini à un nombre infini de variables. De plus Volterra dans ses textes propose une véritable méthode pour opérer ce passage à une infinité de variables : les sommes deviennent des intégrales, les indices de sommation sont remplacés par des variables continues, etc.

Ainsi la « fonction qui dépend d'autres fonctions » est vue comme le résultat d'un processus d'abstraction qui opère depuis une fonction (ordinaire) qui dépend d'une quantité toujours plus grande de variables. De même la dérivée $y' | [\varphi(x), t]$ d'une « fonction qui dépend d'autres fonctions » fait apparaître un paramètre t qui permet de l'interpréter comme dérivée par rapport à une de

¹⁸⁰[Vol54, pp.16–48]

ses infiniment nombreuses variables. Il y a donc une infinité de dérivées, qui ne sont pas exclusivement pensées en chaque point de la courbe représentative de φ , mais par rapport à une infinité de variables que φ représente de façon abstraite.

Volterra précisera une des façons de penser le passage du fini à l'infini dans son cours de la Sorbonne de 1912 en envisageant une ligne polygonale dont la limite est une courbe continue. Volterra propose alors de voir une fonction F de cette ligne polygonale comme une fonction des ordonnées y_i de ses sommets. Il y a donc un nombre fini de dérivées partielles $\frac{\partial F}{\partial y_i} = F'_i(y_1, \dots, y_n)$. La méthode conceptuelle mise en place par Volterra consiste alors à passer à l'infini : les n ordonnées deviennent une courbe continue, l'indice i qui prenait les valeurs $1, \dots, n$ devient un paramètre continu. Dans le même ordre d'idées le procédé qui permettait d'obtenir la dérivée partielle en fixant toutes les variables sauf une est repensé en une version localisée et continue¹⁸¹ :

De même pour obtenir la dérivée partielle d'une fonction de ligne relative au point d'abscisse ξ , il faudra faire varier la ligne dans un domaine h qui renferme le point ξ [...], sans changer les autres parties de la ligne.

Volterra tirera de cette relecture de ses travaux une vision qu'il qualifie d'« analytique » et qui pour nous dévoile le processus d'abstraction à l'œuvre dans le passage à l'infini¹⁸² :

On peut se demander s'il n'y a pas une voie analytique pure qui puisse nous conduire aussi aux mêmes nouvelles conceptions ?

La théorie des fonctions ordinaires renferme l'étude des propriétés de toute quantité obtenue par des opérations algébriques. Comme Lagrange l'a remarqué, l'algèbre n'est qu'une branche de la théorie des fonctions, car les résultats des opérations algébriques sont les fonctions les plus élémentaires envisagées par l'analyse. De même on peut prendre un concept analytique pour point de départ d'une théorie des quantités qui dépendent d'une ligne variable, ou de toutes les valeurs d'une ou plusieurs fonctions de forme variable. La marche à suivre est déjà toute tracée : il suffira de remplacer les opérations de l'Algèbre faites sur un nombre fini de valeurs par des opérations analytiques sur un ensemble continu ou sur toutes les valeurs d'une fonction.

La vision « analytique pure » chez Volterra donne à voir un certain type d'abstraction où la notion de dépendance est centrale. Elle ne fait plus référence au processus de passage à l'infini pour regarder les *fonctions-objets* de façon abstraite et générale. Les mots « toutes les valeurs d'une fonction » ne désignent plus tout à fait un concept *global-par-quantification* mais plutôt un concept *global-par-abstraction* si on reprend le vocabulaire que nous avons introduit. Le processus d'abstraction que nous mettons en évidence ici est supporté par l'idée que se fait Volterra de la définition de fonction proposée par Dirichlet et par le concept général de dépendance qu'il bâtit à partir d'elle.

¹⁸¹[Vol13, p.23]

¹⁸²[Vol13, p.16-17]

Nous avons mis en évidence que la notion de « fonction de ligne » offrait une conception *globale-par-abstraction* de ligne et de voisinage (tubulaire). L'analyse que nous venons de mener montre que la notion de « fonction qui dépend d'autres fonctions » est aussi propice à l'émergence d'une forme abstraite du même type et à la conception d'une « voie analytique pure » comme la nomme Volterra.

NOTE.

^(α) La distinction entre *accroissement* et *déplacement* que Volterra mentionne avait été déjà mise en avant par Jellett. Dans son traité *An elementary treatise on the calculus of variations*¹⁸³ celui-ci dédie un chapitre spécial aux applications des variations à la mécanique, et il isole une partie qu'il intitule « On the application of the calculus of variations to the deduction of equations of equilibrium and motion » dans lequel il introduit un nouveau type d'accroissement.

Jellett qualifie les deux types d'accroissements qu'il a utilisés jusqu'ici d'accroissement de « nature purement mathématique ». Il s'agit des accroissements qui concernent la variation d'une variable de la fonction, notée dx , ou la *forme* d'une fonction pour laquelle Jellett réserve le symbole $\delta\phi$. Il est donc, d'un point de vue « purement mathématique », impossible de considérer une variation δx puisque ce symbole est réservé aux formes¹⁸⁴ :

As long as we confine ourselves to purely mathematical conceptions, an independent variable admits of but one species of increment.

Jellett introduit alors un nouveau type d'accroissement pour la variable qui intervient en mécanique à travers le mouvement¹⁸⁵ :

But the case is materially altered by the introduction of the mechanical conception of *motion*. The symbol δ no longer denotes the increment which is produced by a change in form, but the increment which is produced by a change in position, by the motion of a particle from one point of space to another. In fact the increment denoted by δ is not a variation, but a displacement; and although the science which treats of such increments is, generally speaking, identical in its rules with the Calculus of Variations, it is, nevertheless, in its fundamental conceptions, essentially distinct.

Jellett va alors préciser en deux temps les distinctions qu'il est nécessaire de faire quand on introduit la notion de « déplacement ».

Première distinction. Pour Jellett il est nécessaire de maintenir deux conceptions, l'une notée dx pour la variation de la variable indépendante et δx qui désigne un « déplacement ». Jellett précise encore¹⁸⁶ :

The symbol δ no longer denotes the increment which is produced by a change in form, but the increment which is produced by a change in position, by the motion of a particle from one point of space to another. In fact the increment denoted by δ is not a *variation*, but a *displacement*.

Seconde distinction. Jellett s'appuie sur une conception moléculaire de la matière pour préciser l'emploi du symbole δ et le distinguer du symbole d . Ce dernier est associé pour lui à un changement de forme, c'est-à-dire à un

¹⁸³[Jel50]

¹⁸⁴[Jel50, p.311]

¹⁸⁵[Jel50, p.311]

¹⁸⁶[Jel50, p.311]

changement des distances entre les molécules. Par contraste δ désigne un changement dû au déplacement de la même molécule.

Ainsi dans une intégrale du type $\iiint_a^b V dx dy dz$ le terme $dx dy dz$ désigne pour Jellett un « agrégat d'éléments » (*aggregate of elements*). L'intégrale dépend des bornes et de la forme de V qui peuvent donc tous subir une variation. Par contre l'intégrale ne dépend pas du terme $dx dy dz$ et lui assigner une variation ne fait donc pas sens.

Par contraste Jellett affirme qu'il faut tenir compte de la variation des « éléments » dans les équations d'équilibre ou de mouvement¹⁸⁷ :

[I]n the investigation of the equations of equilibrium or motion of a continuous system, the *variation* of the *element*, that is to say, the change in geometrical magnitude which it undergoes in consequence of the displacement of the molecules which compose it, cannot be neglected.

Il y a pour lui deux types de situations où il faut tenir compte des déplacements des molécules¹⁸⁸ :

We shall find that this species of variation is introduced into mechanical problems in two ways, namely: 1. By the nature of molecular force. 2. By the nature of the equations of condition which subsist in a continuous system.

Enfin après avoir précisé la distinction nécessaire à opérer entre les deux variations Jellett propose de nommer « variation mécanique » (*mechanical variation*) la variation due à un « déplacement ».

L'exemple suivant permet de suivre le traitement de l'infinitésimal $dx dy dz$ et de sa variation. En notant δx , δy et δz les « déplacements virtuels » Jellett obtient¹⁸⁹ :

$$\delta \cdot dx dy dz = \left(\frac{d\delta x}{dx} + \frac{d\delta y}{dy} + \frac{d\delta z}{dz} \right)$$

Finalement l'introduction des « déplacements » permet à Jellett de distinguer la « variation mécanique » de la « variation mathématique » que nous indiquons entre crochets dessous. En tenant compte de tous les types d'accroissements on obtient pour une « molécule » et une fonction $V = f(x, y, z, \xi, \eta, \zeta, \frac{d\xi}{dx}, \frac{d\xi}{dy}, \dots, \frac{d\eta}{dx}, \dots)$:

$$\delta V = \left(\frac{dV}{dx} \right) \delta x + \left(\frac{dV}{dy} \right) \delta y + \left(\frac{dV}{dz} \right) \delta z + \left[\frac{dV}{d\xi} \delta \xi + \frac{dV}{d \cdot \frac{d\xi}{dx}} + \dots \right].$$

Il découle qu'en général la variation mathématique diffère de la variation mécanique, et Jellett utilise cette analyse pour étudier des problèmes d'équilibre et de mouvement grâce à la méthode lagrangienne¹⁹⁰ dans les milieux continus. Nous Laissons le lecteur apprécier les différents exemples et développements que Jellett propose comme application de ces idées.

¹⁸⁷[Jel50, p.311]

¹⁸⁸[Jel50, p.311]

¹⁸⁹[Jel50, p.316]

¹⁹⁰Voir [Lag88]

Vers une analyse générale : Conception abstraite et éléments structurels chez Fréchet.

*The existence of analogies between central features of various theories implies the existence of a general theory which underlies the particular theories and unifies them with respect to those central features.*¹

— E. H. Moore

4.1 INTRODUCTION

Nous avons souligné en introduction de notre thèse l'importance de la vision de structure dans l'élaboration de l'analyse fonctionnelle.

Dans cette partie nous illustrons la progression de Fréchet vers une vision abstraite et générale qui s'appuie sur différentes formes de structures. Nous établirons que certains éléments structurels sont liés à divers types de généralité que nous analyserons. En particulier nous montrerons que l'on peut élaborer une lecture des premiers textes de Fréchet qui met en évidence une réappropriation des travaux de Volterra que nous avons abordés dans la partie précédente.

En effet, comme nous allons le voir, Fréchet travaille rapidement sur les articles de Volterra de 1887 sur les fonctions de lignes et dès 1904 il publie à son tour un article sur le sujet².

Les idées qui ont influencé ou nourri le travail de Fréchet en Analyse fonctionnelle ont fait l'objet de plusieurs études historiques qui se sont particulièrement concentrées sur le concept essentiel qui sous-tendait sa vision,

¹Préface du cours intitulé « Introduction to a form of general analysis » donné par E.H. Moore en 1906 lors du New Haven Colloquium à l'université de Yale. (cf. [Moo10])

²[Fré04b]

la notion d'*Espace abstrait* en Analyse. En particulier on peut consulter la thèse pionnière de Reinhard Siegmund-Schultze intitulée *Die Anfänge der Funktionalanalysis und ihr Platz im Umwälzungsprozeß der Mathematik um 1900*³ et le très ample travail de Angus E. Taylor publié en trois parties sous le titre *A study of Maurice Fréchet* —I: "Fréchet as analyst, 1909–1930"; II: "Mainly about his work on general topology, 1909–1928"; III: "His early work on point set theory and the theory of functionals."⁴

Dans son étude, Taylor qualifie d'*indirecte* l'influence de Volterra sur la conception des espaces abstraits⁵ :

Fréchet's short note of 1904 [on Weierstraß theorem] broke absolutely fresh ground. Although Vito Volterra's work certainly had some influence on Fréchet's work taken as a whole, I think a good deal of it was exerted indirectly, through Hadamard. I see little or no reason for thinking that Volterra contributed directly to the shaping of Fréchet's ideas on L-classes, V-classes, or E-classes.

La note de 1904 à laquelle il est fait référence ici⁶ introduit pour la première fois dans le travail de Fréchet une vision très générale qui lui permet de concevoir des opérations fonctionnelles sur des ensembles constitués de certaines catégories d'éléments quelconques (nombres, surfaces, etc.) en faisant intervenir des éléments de topologie. Fréchet n'utilise naturellement pas ce mot mais il importe, à des fins tout à fait particulières que nous décrivons dans la suite, des idées que nous classerions désormais dans le champ de la *topologie*, comme la compacité ou des propriétés de tel ou tel type de convergence des suites. Taylor considère que cette note marque le point de départ de la progression de Fréchet vers les *espaces abstraits* et relègue en quelques sortes ses études contemporaines sur les fonctions de lignes au rang de travail de jeunesse. Or, il nous semble légitime de réexaminer l'importance de la lecture par Fréchet des travaux de Volterra dans sa conception des espaces abstraits et plus généralement dans son travail en Analyse fonctionnelle.

En particulier nous allons montrer que c'est essentiellement la recherche d'une façon générale d'appréhender les *nouvelles* fonctions mises en avant par Volterra qui va, dès les publications de 1904, donner lieu à deux types de développements qui ne seront réunifiés que petit à petit dans un cadre général.

D'un côté, en saisissant le concept de fonction que met en avant Volterra, Fréchet développe une théorie dans la lignée des idées de Dirichlet⁷ et de Weierstraß qui va prendre une forme très générale obtenue par un processus d'abstraction propre au jeune mathématicien français. La notion d'« espace abstrait » apparaît ici comme véritable préliminaire à une théorie générale des fonctions⁸ ou encore des *opérations fonctionnelles* comme les appelle Fréchet.

³[Sie82]

⁴[Tay82 ; Tay85 ; Tay87]

⁵[Tay82, pp.286-287]

⁶[Fré04a]

⁷Cf. le texte de Dirichlet que nous avons déjà évoqué avec Volterra : [Dir37].

⁸Le mot *fonction* sera employé par Fréchet dans sa thèse pour désigner ce qui deviendra dans son optique un cas particulier d'*opération*, à savoir une fonction classique d'une ou plusieurs variables réelles.

D'un autre côté, en suivant la piste ouverte par Volterra, Fréchet va mettre en avant des outils adaptés aux problèmes de variations issus de la Physique mathématique. Nous analyserons plus loin la démarche de Fréchet qui cherche à réunir dans un même cadre conceptuel les problèmes de ce domaine, notamment à travers la notion d'*opération linéaire* et la recherche de théorèmes de représentation de ces fonctions généralisées.

Ces deux voies, nous allons tenter de le montrer, furent développées en parallèle, et prirent racine dans une relecture des idées de Volterra, avant de se rejoindre progressivement dans une théorie générale des opérations définies sur des espaces topologiques abstraits.

4.2 PREMIERS CONTACTS AVEC LES FONCTIONS DE LIGNES

4.2.1 DEUX PUBLICATIONS D'HADAMARD SUR LES FONCTIONS DE LIGNES

En 1902 et 1903 Hadamard publie deux articles sur les fonctions de lignes qui installent le travail de Volterra dans le paysage mathématique français.

Le premier s'intitule « Sur les dérivées des fonctions de lignes »⁹ et présente sommairement une partie des résultats de Volterra sur la variation d'une fonction de ligne que l'on peut écrire $\delta U = \int_L (U'_x \delta x + U'_y \delta y + U'_z \delta z) ds$ plus des termes de la forme $P_x \delta x + P_y \delta y + P_z \delta z + P_{x'} \delta x' + \dots$ dans le cas de points singuliers. Dans ce court article, Hadamard montre que l'intégrale $\int_L (U''_{xx_1} \delta x_1 + U''_{xy_1} \delta y_1 + U''_{xz_1} \delta z_1) ds_1$ ne peut pas décrire la variation $\delta U'_x$ de la dérivée première en suivant de près les arguments que Volterra avait lui-même indiqués dans ses travaux.

Outre la référence aux travaux de Volterra qu'Hadamard fait alors connaître en France, celui-ci ajoute un commentaire qui éclaire sa propre conception générale¹⁰ :

Il semble d'ailleurs que, depuis les travaux de MM. Pincherle et Bourlet, il soit possible de remplacer utilement ces hypothèses, évidemment particulières, par d'autres plus générales. Il est, en effet, rationnel d'admettre, en tout cas, que δU est ce que M. Pincherle a appelé une *fonction distributive*, M. Bourlet une fonction *additive* et moi-même une fonction *linéaire* des quantités δx , δy , δz , considérées comme fonctions de s .

Hadamard précise que « dans le cas des opérations analytiques » les travaux qu'il cite permettent de préciser la forme générale de δU . Bien sûr la référence au cas analytique détermine grandement le type de généralité qui est en jeu ici, mais pas complètement. Le point de vue a changé par rapport à ce que nous avons analysé chez Volterra. D'un côté les conditions de Volterra définissent un certain type de fonctions de lignes pour lesquelles on peut donner la forme de la variation alors que pour Hadamard c'est une propriété de la variation (la *linéarité*) qui est mise en avant et qui définit une classe de fonctions de lignes suffisamment générale.

⁹[Had02]

¹⁰[Had02, p.41]

La seconde publication d'Hadamard est une note à l'Académie des Sciences présentée par Jordan qui porte le titre « Sur les opérations fonctionnelles »¹¹ dans laquelle il introduit immédiatement la notion de linéarité¹² :

Les opérations fonctionnelles linéaires, c'est-à-dire les lois suivant lesquelles on peut, à toute fonction $f(x)$ définie dans un intervalle $a < x < b$, faire correspondre un nombre U , de telle façon qu'on ait (quels que soient les nombres c_1, c_2 et les fonctions f_1, f_2)

$$U(c_1f_1 + c_2f_2) = c_1U(f_1) + c_2U(f_2)$$

ont été étudiées principalement par MM. Volterra, Pincherle, Bourlet.

Hadamard définit donc une « opération » comme une « correspondance » à son tour et la notation $U(f)$ laisse penser qu'il adopte une notion proche de celle de fonction comme l'avait fait Volterra en faisant référence à la définition de Dirichlet. De plus, bien qu'il fasse à nouveau référence aux même trois mathématiciens que dans sa publication de 1902, il précise en note de bas de page que cette fois-ci c'est du côté des travaux de Volterra qu'il faut trouver la plus grande généralité¹³ :

Je me place ici au point de vue de M. Volterra dans ses études sur les fonctions de lignes, et qui me paraît préférable à celui de MM. Pincherle et Bourlet. C'est, me semble-t-il, introduire un élément parasite dans la question que de considérer U comme une fonction de la même variable x qui figure dans f . Rien n'empêche, lorsqu'on raisonne comme je le fais dans le texte, de faire dépendre U non seulement de la forme de la fonction $f(x)$, mais aussi d'un certain nombre de paramètres, dont l'un sera, le cas échéant, désigné par x , comme la variable qui figure dans f .

La critique d'Hadamard peut sembler un peu exagérée pour un lecteur actuel, non seulement parce que Pincherle comme Bourlet utilisent des notations où ne figure pas la variable de la fonction ($A(\varphi)$ par exemple) mais aussi parce qu'Hadamard désigne lui-même les fonctions en faisant figurer très souvent la variable, comme dans « la fonction $f(x)$ » dans la citation au-dessus. En fait l'enjeu est probablement ce qui est au centre de cette note d'Hadamard, à savoir la nature même des fonctions utilisées qui sont essentiellement analytiques chez Pincherle ou Bourlet, alors que Volterra se dégage de cette contrainte. En effet, le développement en série de la fonction analytique définit la *forme* de la fonction par la relation algébrique exprimée dans la variable x . Dans sa note Hadamard insiste : « il m'a paru nécessaire de traiter la question en faisant complètement abstraction de l'analytisme de $f(x)$ »¹⁴.

Nous reprenons ici rapidement les idées exposées par Hadamard afin de mettre en évidence l'enjeu majeur qui consiste à se départir de cette condition d'analytisme. Ces éléments nous permettront par la suite de mieux appréhender une des voies de progression de Fréchet.

¹¹[Had03]

¹²[Had03, p.351]

¹³[Had03, p.351]

¹⁴[Had03, p.352]

La stratégie générale qu'Hadamard met en œuvre consiste principalement à exprimer une fonction continue comme limite uniforme d'une quantité mettant en jeu une intégrale et la fonction f comme nous allons le préciser. En appliquant l'opération U aux deux membres, puis en utilisant la continuité et la distributivité de l'opération, on peut exprimer l'action de U sur la fonction continue f .

Dans le détail, Hadamard s'appuie sur un résultat qu'il attribue à Weierstraß et Kirchhoff qui permet de trouver une fonction F continue (par exemple $F(x) = \frac{1}{\sqrt{\pi}}e^{-x^2}$) telle que pour toute fonction f continue entre a et b on a :

$$\lim_{\mu \rightarrow \pm\infty} \mu \int_a^b f(x) F[\mu(x - x_0)] dx = f(x_0) \quad a < x_0 < b.$$

Comme la continuité de l'opération U est définie par Hadamard via la convergence uniforme des suites de fonctions continues, cette formulation pourrait sembler insuffisante à un lecteur de notre époque. En fait, Hadamard fait référence aux travaux de Weierstraß où il est spécifié que la convergence est uniforme dans la limite au-dessus. Il est alors possible d'appliquer l'opération linéaire U et d'écrire comme Hadamard : $U[f(x)] = \lim_{\mu \rightarrow \pm\infty} \int_a^b f(x) \Phi(x, \mu) dx$ où $U[\mu F\mu(x - x_0)] = \Phi(x_0, \mu)$.

Enfin Hadamard précise dans quel cas on obtient $U[f(x)] = \int_a^b f(x) \psi(x) dx$ ¹⁵ :

La condition nécessaire et suffisante pour que U ait la forme [indiquée au-dessus], $\psi(x)$ étant fini et continu, est que $\Phi(x_0, \mu)$ tende uniformément vers $\psi(x)$, lorsque μ augmente indéfiniment par valeurs positives ou négatives (mais réelles).

4.2.2 PREMIERS TRAVAUX DE FRÉCHET SUR LES FONCTIONS DE LIGNES

Peu après la fin de sa scolarité à l'École Normale à l'automne 1903 Fréchet écrit une longue lettre à Volterra¹⁶ sur les conseils conjoints de Borel et Hadamard qui lui avaient suggéré qu'il y avait probablement dans les travaux des géomètres italiens matière à engager un travail de thèse. Dans cette lettre Fréchet mentionne qu'il a commencé à travailler sur les fonctions de lignes et les travaux de Volterra qu'il a connus par les cours d'Hadamard deux ans auparavant. Il y fait aussi rapidement référence aux travaux de Pincherle (ce dernier avait publié en 1901 en collaboration avec Amaldi son livre *Le operazioni distributive*).

La lettre n'a pas la forme d'un article, mais elle propose un exposé très détaillé et montre que Fréchet a déjà beaucoup réfléchi aux conceptions de Volterra. Elle contient en fait une discussion autour d'une partie des résultats qui seront publiés rapidement après aux Annales de l'ÉNS dans un article intitulé « Sur les fonctions de lignes fermées »¹⁷.

¹⁵[Had03, p.353]

¹⁶Nous reproduisons cette lettre en annexe -A- ; elle est conservée à l'Académie des Sciences à Paris.

¹⁷[Fré04b]

Un des intérêts de la lettre par rapport à l'article qui suivra vient de l'ordre d'exposition des idées de Fréchet qui y est sensiblement différent. En particulier la lettre commence la présentation mathématique par une discussion qui concerne le type de point de vue général à adopter pour l'étude des fonctions de lignes. Malgré toute la déférence dont fait preuve Fréchet à l'égard du mathématicien italien dans sa lettre, il propose une organisation des idées qui remet en question le point de vue adopté systématiquement par Volterra depuis ses premiers travaux sur les fonctions de lignes. Fréchet suit en fait une ligne proche de celle qu'avait proposée Hadamard en ne partant pas de la propriété de dérivabilité des fonctions de lignes comme l'avait fait Volterra, mais en imposant d'emblée la forme de la « différentielle » δU_L ¹⁸ :

Je ne considère donc dans la suite que des fonctions de lignes fermées qui jouissent de cette propriété que l'on a :

$$\delta U_L = \int_L (U'_x \delta x + U'_y \delta y + U'_z \delta z) ds$$

U'_x, \dots désignant des quantités fonctions de la ligne L et du point (d'arc s).

Fréchet argumente son choix en proposant dans sa lettre un théorème et une analyse que l'on retrouvera dans l'article que nous allons maintenant commenter.

Ce premier article de Fréchet introduit les fonctions de lignes en des termes proches de ceux que nous avons lus chez Volterra quelques années plus tôt¹⁹ :

La physique mathématique conduit à l'étude de fonctions beaucoup plus générales que les fonctions dépendant de la valeur d'une ou de plusieurs variables (fonctions que j'appellerai *ordinaires*). Je veux parler des expressions qui ne sont déterminées que par la connaissance de toutes les valeurs d'une ou de plusieurs *fonctions ordinaires*.

Mais aussitôt cette approche générale rappelée, Fréchet limite son exploration aux cas les plus « simples », et il précise le cadre que couvre son étude²⁰ :

Nous nous bornerons au cas des fonctions U_L dont la valeur varie seulement avec la *forme* d'une ligne L plane ou gauche, continue, *fermée* et dont la tangente varie d'une manière continue, sauf en des points isolés en nombre fini.

Même s'il ne l'explique pas, Fréchet n'envisage que des fonctions U_L à valeurs réelles ou complexes. Le point de vue adopté pour la variable est celui des familles paramétrées de courbes qui permettent de définir des limites (uniformes) et la variation de la fonction U_L ²¹ :

¹⁸[Fré03, p.3-4]

¹⁹[Fré04b, p.557]

²⁰[Fré04b, p.557]

²¹[Fré04b, p.557]

Considérons une famille G de ces lignes, dépendant d'un paramètre α de façon que, si α tend vers α_0 , L tende *uniformément* en tous ses points vers les points correspondants de L_0 . Pour ces lignes L , U_L sera une fonction de α et nous ajoutons l'hypothèse que ce soit une fonction continue et dérivable en α . Dans ces conditions, nous pourrions parler, comme dans le calcul des variations, de la variation première de $U_L : \delta U_L$ ²², laquelle dépend, bien entendu, de la famille G considérée.

À ce moment précis Fréchet élabore son propre contexte général : il convoque de façon particulière la notion de fonction de ligne sous une forme qui lui permet de relire grâce à une notion abstraite unique un ensemble de situations dans lesquelles il identifie des problèmes fonctionnels, c'est-à-dire dont le questionnement porte sur des fonctions²³.

Cette introduction de Fréchet dévoile sa conception et nous permet de contraster son angle de vue avec celui adopté par Volterra en termes de généralité. Fréchet conçoit ici des lignes qui sont définies comme image d'une fonction suffisamment régulière de \mathbb{R} dans l'espace à trois dimensions. De plus il considère une famille G de lignes dépendant d'un paramètre α qui lui permet de définir la convergence des lignes lorsque le paramètre se rapproche d'une certaine valeur. Dans ce cadre l'étude des variations de U_L se ramène donc à une étude de fonctions qui dépendent du paramètre réel α .

Nous avons vu précédemment que Volterra avait lui aussi recours à des idées proches de celles-ci, mais sa visée initiale était de concevoir des fonctions dont la variable était une fonction ou une ligne vue de façon *synthétique* ou *abstraite*, c'est-à-dire sans recours systématique à la variable de la *fonction-objet* ou au paramétrage de la ligne. De même, si souvent Volterra a introduit un paramètre annexe pour mesurer l'*écart* entre deux lignes (ϵ , ou σ avec les notations de la partie précédente), nous l'avons aussi vu convoquer une notion de voisinage tubulaire abstrait qui ne recourrait pas à un paramètre réel de façon explicite.

Ces éléments de comparaison permettent de voir que Fréchet est lui-même en train de repenser le contexte le mieux adapté pour fonder une vision générale et efficace.

Son article s'ouvre sur une critique de la notion de dérivée que Volterra propose comme point de départ et d'appui pour son « calcul infinitésimal » des fonctions généralisées. Fréchet analyse le rapport différentiel qui permet de définir la dérivée dans le cas des fonctions ordinaires et dans le cas des fonctions de lignes pour Volterra. Pour les fonctions ordinaires (de la variable complexe) Fréchet dit que « les fonctions analytiques sont telles que le rapport de leur variation à celle de l'une d'entre elles, z , soit une fonction de la même

²²On peut consulter le cours de Joseph-Alfred Serret (1819-1885) [Ser68] publié pour la première fois en 1868 et réédité à de nombreuses reprises par la suite pour avoir un aperçu des notations et des définitions de *dérivée*, *accroissement*, *différentielle*, etc. en usage à l'époque pour les fonctions ordinaires. Les textes de Volterra, Hadamard ou Fréchet ne redéfinissent pas ces notions qu'il nous faut précisément distinguer et analyser ici. En particulier dans ce texte δU_L désigne la différentielle relative au paramètre α (cf. Ch XII de [Ser68]).

²³On peut relire le cours d'Hadamard du Collège de France (1898-99) publié en 1903 pour voir un panorama des questions qui sont au cœur des recherches que nous évoquons ici.

espèce. »²⁴ Dans la lettre à Volterra que nous avons mentionnée Fréchet présente cette fonction z comme « la fonction analytique la plus simple ». On retrouve là des idées que nous avons évoquées avec Jellett plus avant : on y lisait une singularisation de la fonction $x \mapsto x$ pour concevoir le rapport différentiel d'une fonction ordinaire quelconque par rapport à cette fonction de référence. La question qui se pose à Fréchet dans le mouvement de généralisation qui tente de transposer ces idées aux fonctions de lignes est de savoir si l'on peut identifier une fonction de ligne simple de référence qui jouerait le rôle que jouent $x \mapsto x$ ou $z \mapsto z$ pour les fonctions ordinaires dérivables ou analytiques. Dans la lettre Fréchet précise qu'il pense par exemple comme fonction de référence à la fonction de ligne (fermée) qui donne l'aire enclose par la ligne.

La première partie de son article donne alors un théorème qui va montrer que cette façon de penser ne permet pas d'étendre la conception de dérivée de façon satisfaisante. Le théorème est donné d'emblée sous une forme générale pour une collection finie de fonctions de lignes²⁵ :

Étant données $n + 1$ fonctions de lignes : $U_L, V_L^{(1)}, \dots, V_L^{(n)}$, la condition nécessaire et suffisante pour que U_L soit une FONCTION ORDINAIRE de $V_L^{(1)}, \dots, V_L^{(n)}$ est qu'il existe une relation de la forme

$$\delta U_L \equiv K_L^{(1)} \delta V_L^{(1)} + \dots + K_L^{(n)} \delta V_L^{(n)}$$

(sauf peut-être pour $\delta V_L^{(1)} = \dots = \delta V_L^{(n)} = 0$),

où $K_L^{(1)}, \dots, K_L^{(n)}$ sont des fonctions déterminées de la ligne L indépendantes de la manière dont L est fonction de α .

La démonstration de cette proposition met en évidence la façon de penser en termes de *fonctions ordinaires* d'un paramètre α et le type de généralité qui est à l'œuvre ici.

Le résultat obtenu va amener deux commentaires de Fréchet. Le premier dans sa lettre à Volterra, où il dit après avoir exposé ces mêmes idées : « Ce qui précède m'a fait mieux comprendre combien vous aviez eu raison de prendre votre définition qui consiste au fond à généraliser l'expression de la variation de l'intégrale $\int Pdx + Qdy + Rdz$. » Il reconnaît donc chez Volterra une voie qui permet effectivement la généralisation et évite l'écueil que dévoile le résultat qui ramène à l'étude des fonctions ordinaires. Le second commentaire est déjà esquissé dans la fin de cette dernière phrase écrite à Volterra. Nous avons précisément analysé que la vision développée par le mathématicien italien faisait intervenir un rapport différentiel particulier²⁶ qu'il ne présente pas lui-même comme issu de l'analyse des intégrales du type indiqué par Fréchet, bien qu'un lien avec celles-ci soit établi dans un second temps. Fréchet instille donc

²⁴[Fré04b, p.559]

²⁵[Fré04b, p.558]

²⁶On rappelle que Volterra définit la dérivée comme limite du rapport $\frac{\delta y}{\int_m^n \theta(x) dx'}$, dont le dénominateur n'est pas la variation de ce que Fréchet appelle « une fonction de lignes déterminée une fois pour toutes ». Il y a chez Volterra une *localisation* de la dérivée en un point de la courbe qui donne une dérivée y' qui dépend non seulement de la fonction φ mais aussi de l'abscisse t du point de la courbe de φ .

une première variation de point de vue et commente plus avant le résultat qu'il a obtenu et qu'il a présenté d'emblée dans son article.

Il affirme qu'on ne peut pas développer une théorie générale des fonctions de lignes en prenant un rapport différentiel construit à partir d'une fonction de ligne V_L choisie par avance²⁷ :

Ne pourrait-on obtenir, à partir d'une fonction de ligne très simple, V_L , une classe intéressante de fonctions de lignes U_L , par la condition que $\frac{\delta U_L}{\delta V_L}$ ne dépende que de la ligne L ? La réponse est négative : d'après ce qui précède, ces fonctions U_L seraient tout simplement des *fonctions ordinaires* de V_L ; et, connaissant V_L , l'étude des fonctions U_L ne constituerait plus, à proprement parler, une question de calcul fonctionnel.

Cette remarque importante justifie pour Fréchet le choix d'une autre voie de généralisation qui renverse en quelques sortes le schéma initial de Volterra. Ce dernier avait cherché, comme nous l'avons vu, à adapter la stratégie classique de l'étude des fonctions de la variable réelle : une dépendance, la *fonction*, vue pour lui dans un sens étendu, puis une variation de la variable qui engendre une variation de la valeur de la fonction qui donne à son tour la notion de dérivée. Enfin dans cet enchaînement on obtient une expression de la *variation* dont la partie de premier ordre donne la *différentielle*. L'analyse dans le cas des lignes passe de même chez Volterra, suivant un schéma similaire, par une définition de la dérivabilité d'une fonction de lignes à partir de la déformation d'une ligne L et de la variation engendrée pour la fonction U_L avant de donner la forme de la différentielle.

Le travail de Fréchet propose une approche laissant apparaître une réelle différence de point de vue qu'il justifie en remarquant que la présentation de Volterra ne permet pas de traiter des fonctionnelles classiques issues du Calcul des Variations²⁸. Ce point de vue différent apparaît jusque dans le choix des notations de la fonction de ligne : alors que Volterra utilise la notation $\varphi|[L]$ pour désigner la fonction de ligne, Fréchet reprend une notation déjà utilisée par Hadamard U_L où le lien avec la notion de fonction n'est pas totalement transparent. Ici le symbole U_L sert plutôt pour Fréchet à désigner certains éléments du calcul des variations, éléments qu'il identifie comme fonctions de lignes et dont il cherche à mettre en évidence des propriétés.

La stratégie de généralisation qu'adopte Fréchet consiste à considérer en premier lieu les *fonctionnelles* les plus simples que l'on rencontre dans le Calcul des Variations. Fréchet part donc d'une intégrale du type $I_L = \int_L P(x, y, z)dx + Q(x, y, z)dy + R(x, y, z)dz$ dont la variation peut s'écrire $\delta I_L = \int_L (I'_x \delta x + I'_y \delta y + I'_z \delta z) ds$. Ainsi une façon de généraliser consiste à prendre une

²⁷[Fré04b, p.559]

²⁸En fait Fréchet pendant une longue période ignorera les traitements particuliers que Volterra réserve aux « points exceptionnels » comme nous l'avons vu. Il en découlait une conception générale qui couvre beaucoup plus de cas que Fréchet ne le laisse entendre dans ses travaux lorsqu'il cite Volterra. En fait ce point est important et donnera lieu à une querelle entre les deux mathématiciens et à un échange de lettres où chacun défend sa position à partir de 1913. On peut trouver une analyse de ce différend dans l'article « Lines on the horizon » de Guerraggio, Jaëck et Mazliak [GJM15].

propriété d'une fonctionnelle particulière comme définition générale. Encore faut-il s'assurer que la nouvelle définition recouvre des cas nouveaux, ce que fait Fréchet en donnant des exemples précis.

À partir de cette définition de base Fréchet va décliner plusieurs classes de fonctions de lignes et organiser divers niveaux de généralité. Les fonctions de lignes définies à partir de la forme générale de la variation indiquée au-dessus porteront le nom de Volterra²⁹ :

Nous appellerons donc *fonction de Volterra*, ou *fonction* (\mathcal{V}), toute fonction de ligne fermée, U_L , satisfaisant aux conditions que nous avons posées au n. 1 et telle que l'on ait

$$\delta U_L = \int_L (U'_x \delta x + U'_y \delta y + U'_z \delta z) ds,$$

U'_x, U'_y, U'_z étant des quantités déterminées en chaque point M de toute ligne fermée L .

En outre cette définition permet de définir plusieurs autres classes de fonctions : les *fonctions* (\mathcal{V}) *du premier degré*, c'est-à-dire celle de la forme I_L qui ont servi à forger la définition étendue, et des *fonctions* (\mathcal{V}) *simples* qui se présentent comme fonctions ordinaires d'une seule fonction (\mathcal{V}) du premier degré.

Cette approche amène Fréchet à trouver les propriétés caractéristiques des différentes catégories de fonctions (\mathcal{V}). En particulier, on retrouve ici, dans un second temps, la notion de linéarité que Volterra avait lui aussi mise en avant avec d'autres motivations³⁰ :

Rappelons, pour le généraliser, un théorème de M. Volterra. Désignons avec lui par $L + L'$ un contour fermé (pouvant comprendre plusieurs courbes fermées) et constitué par les contours L et L' , où l'on a supprimé les parties communes (s'il en existe) parcourues en sens contraire (car U_L dépend en général du sens de parcours de L). M. Volterra a démontré que les fonctions du premier degré sont les seules fonctions (\mathcal{V}) qui vérifient l'équation fonctionnelle

$$U_{L+L'} = U_L + U_{L'}.$$

Plus généralement, nous allons montrer que les fonctions simples sont les seules fonctions (\mathcal{V}) qui vérifient l'équation fonctionnelle

$$U_{L+L'} = \varphi(U_L, U_{L'}),$$

φ étant une fonction *ordinaire* (continue et dérivable) de U_L et $U_{L'}$.

Fréchet redéfinit donc le vocabulaire initialement employé par Volterra. Pour ce dernier nous avons vu que les « fonctions simples de lignes » étaient précisément celles qui étaient distributives (mots d'observateur ici). Le critère de simplicité était attaché chez lui à la possibilité de préciser la variation, idée que nous avons aussi repérée chez Jellett.

²⁹[Fré04b, p.560]

³⁰[Fré04b, p.562]

Pour Fréchet ce sont les intégrales $I_L = \int_L Pdx + Qdy + Rdz$ qui forment les modèles les plus « simples » de fonctions de lignes à partir desquelles il élabore des fonctions plus générales. Les « fonctions de Volterra » forment ainsi une classe étendue en prenant comme critère la forme de la variation des « fonctions simples ». D'ailleurs Fréchet précise que les différentes classes qu'il détermine sont bien distinctes en proposant des exemples qui appartiennent à une classe sans appartenir à la classe plus restreinte.

L'article de Fréchet montre une réflexion essentielle sur l'approche générale à adopter et on voit à plusieurs reprises que celui-ci redéfinit les contours et l'organisation de la théorie que nous avons lue chez Volterra. Nous n'étudions pas ici tous les points de l'article qui aborde sur quelque quatorze pages plusieurs aspects précis de la théorie développée par Volterra. Fréchet envisage en particulier les problèmes d'*isogénéité* et d'intégration —c'est-à-dire la détermination de la fonction de ligne par ses dérivées— que Volterra a aussi mis au centre de ses travaux.

Nous souhaitons quand même ici préciser, parce que cela fait écho à un commentaire que nous avons fait en lisant Volterra, la façon particulière de traiter les lignes chez Fréchet. Nous avons mentionné que l'utilisation du signe + pour désigner la concaténation de deux lignes orientées n'impliquait pas chez Volterra une pensée *algébrique* ou un *calcul algébrique* des lignes³¹. Dans son article de 1904 Fréchet utilise aussi un signe – pour indiquer que la ligne qu'il note $-L$ est parcourue dans le sens contraire de L . Pourtant ici aussi il ne faut pas en déduire un calcul sur les lignes ni le recours à une quelconque idée de structure algébrique concernant les lignes. En particulier Fréchet, comme Volterra dans les textes que nous avons étudiés, n'utilise pas la notation $L - L'$ mais désigne la concaténation³² par $-L + L'$. Nous verrons par la suite comment les mathématiciens laisseront tomber la notion de ligne et de concaténation pour les *remplacer*, dans une sorte de rupture en termes d'abstraction, en convoquant diverses notions préexistantes de *structures*.

Ces précisions nous permettent de relire un passage précis de l'article de Fréchet qui éclaire la progression générale des idées qui forment petit à petit l'analyse fonctionnelle.

Tout d'abord quand Fréchet écrit $U_{L+L'} = U_L + U_{L'}$ il dit que la fonction U « vérifie [une] équation fonctionnelle ». Les deux signes + de l'égalité jouent ici des rôles différents. Le premier indique la concaténation des courbes comme nous l'avons dit alors que le second indique un certain type de relation entre U_L et $U_{L'}$. Nous disons *relation* en général et pas *relation algébrique* car la généralisation proposée par Fréchet n'est pas de type *algébrique* mais fait intervenir une *fonction ordinaire* : $U_{L+L'} = \varphi(U_L, U_{L'})$. Il faut donc entendre ici « équation fonctionnelle » et lire les égalités au-dessus en ayant à l'esprit que Fréchet ne cherche pas à adosser sa vision à un *calcul algébrique* dont nous avons vu plusieurs utilisations en rapport avec l'analyse et les fonctions.

³¹Nous ne disons pas que la concaténation de lignes n'a pas été ailleurs vue sous un angle algébrique. Nous précisons seulement les aspects que convoquent les mathématiciens —ici Volterra ou Fréchet— lorsqu'ils utilisent les lignes et les symboles + ou – avec les lignes.

³²Le mot « concaténation » est ici encore un terme d'observateur.

On peut comparer cette approche avec celle de Volterra, en particulier dans son travail sur les substitutions où nous avons montré que l'on pouvait distinguer une *opération* élémentaire (la multiplication) qui engendrait un certain nombre d'opérations par composition. Il est important ici de souligner que Fréchet ne met pas en avant de relation de cette nature en passant de $U_L + U'_L$ à $\varphi(U_L, U'_L)$ et qu'il ne pense pas ici cette égalité comme un moyen de *construire* de nouvelles opérations à partir d'opérations fonctionnelles élémentaires. Chez Fréchet l'écriture fonctionnelle $\varphi(U_L, U'_L)$ ne peut pas désigner simplement une extension constructive à partir de U_L et U'_L .

Dans les deux cas les mathématiciens, par des stratégies différentes, cherchent en particulier à se départir de la notion de fonction analytique, et nous verrons que c'est un point crucial dans la suite.

De plus on peut préciser avec ce texte de 1904 comment s'organisent pour Fréchet les éléments en jeu, l'« équation fonctionnelle » du dessus et la notion de convergence. Dans le cas où l'on a la relation précédente $U_{L+L'} = \varphi(U_L, U'_L)$, Fréchet montre par une considération géométrique s'appuyant sur les lignes que si une ligne Δ se « réduit » en un point, alors U_Δ tend vers une limite déterminée qu'il note U_0 . Plus précisément on peut résumer ses idées de la façon suivante. On considère un point M du plan et une courbe fermée Δ qui va se réduire en M . Pour toute ligne L on a $\Delta = L + (-L + \Delta)$. Lorsque l'on applique U on obtient donc $U_\Delta = \varphi(U_L, U_{-L+\Delta})$. Par le principe de continuité que nous avons rappelé plus haut (convergence des lignes vue à travers un paramètre α) et la continuité de φ on en déduit que U_Δ tend vers une constante $U_0 = \varphi(U_L, U_{-L})$. Cette constante ne dépend ni du point M , ni de la ligne L et par conséquent on peut aussi conclure, comme le fait remarquer Fréchet, que « U_{-L} est une fonction ordinaire de U_L . »

Une des visées de notre travail consiste à montrer que ces idées ne sont pas déconnectées des travaux qui ont permis plus tard à Fréchet de forger ses *espaces abstraits*.

À partir de ce texte que nous venons de commenter, on peut dissocier deux voies dans la progression de Fréchet qui vont être développée en parallèle avant de se rejoindre en une théorie plus aboutie et générale.

La première voie s'appuie sur la notion de fonction dont Volterra lui-même a fait l'élément majeur de sa progression vers une analyse généralisée. Fréchet va repenser à son tour la notion de fonction et essayer d'extraire les éléments essentiels qui permettent de développer une théorie générale, en se débarrassant de la nécessité d'y voir la variable comme un nombre réel avec tous les outils topologiques alors connus sur \mathbb{R} . Cette première approche sera développée dans [Fré04b] et constituera dans son prolongement le cœur de la thèse de doctorat de Fréchet.

D'autre part, Fréchet propose un deuxième point de vue : il décide cette fois-ci de prolonger l'étude de problèmes de la théorie des variations qui ont été relus à travers la notion de fonction de ligne. Dans la continuité des travaux d'Hadamard que nous avons mentionné au début sur les fonctions de lignes, Fréchet va mettre en avant comme objet central d'étude la *fonctionnelle* et petit à petit dégager le rôle de la linéarité dans sa représentation sous forme intégrale. Une série de trois articles publiés dès 1904 va établir une généralisation du

résultat de représentation d'Hadamard que nous avons rappelé plus haut et initier un mouvement d'abstraction distinct de celui que l'on observe dans sa thèse.

Ces deux voies que nous allons maintenant explorer sont clairement intimement liées au travail de relecture des travaux de Volterra que nous venons d'étudier. Leur analyse permet de saisir les divers aspects convoqués par Fréchet pour penser d'abord une analyse fonctionnelle générale et plus tard une analyse générale.

4.3 EXTENSION DE L'ÉTUDE DES FONCTIONS

4.3.1 GÉNÉRALISATION D'UN THÉORÈME DE WEIERSTRASS

Fréchet propose une première direction d'étude générale des fonctions adaptée aux problèmes variationnels dans une courte note publiée à l'Académie des Sciences intitulée « Généralisation d'un théorème de Weierstraß ».

L'introduction pose la visée de l'article en se plaçant dans la lignée d'Hadamard, non seulement pour la notation U d'une fonctionnelle, mais surtout pour les questions de minimum qu'il avait évoquées comme on l'a vu dès 1897³³ :

On sait l'importance qu'il y aurait, dans un grand nombre de problèmes, à savoir si une quantité U dépendant de certains éléments (points, fonctions, etc.) atteint effectivement un minimum dans le champ considéré. Le principe de Dirichlet offre une des justifications les plus frappantes de cette remarque...

Fréchet rappelle aussitôt le résultat qu'il voudrait conserver dans le cadre plus général³⁴ :

La question est résolue dans le cas particulier où U est une simple fonction de x (ou de plusieurs variables indépendantes). Weierstraß a en effet démontré que toute fonction continue dans un intervalle limité y atteint au moins une fois son maximum.

Cette propriété dépend bien d'ingrédients que Volterra avait lui-même convoqués dans son approche, à savoir l'utilisation centrale d'une notion (étendue) de fonction et la continuité. Mais au lieu de concevoir la continuité de diverses manières en fonction de la nature de la variable, Fréchet cherche les ingrédients généraux dans l'approche de Weierstraß.

Il prend d'abord une notion proche de celle de Volterra et que l'on peut donc voir comme une extension de la définition de Dirichlet³⁵ :

Nous supposons donnée une certaine catégorie³⁶ C d'éléments quelconques (nombres, surfaces, etc.), dans laquelle on sache discerner les éléments distincts. Nous pourrions dire que U_A est une

³³[Fré04a, p.848]

³⁴[Fré04a, p.848]

³⁵[Fré04a, p.849]

³⁶Il faut prendre garde de ne pas donner un sens moderne au mot « catégorie ». Ici « catégorie » est employé pour former une certaine typologie. Il s'ensuit que dans cet énoncé une catégorie est un ensemble formé d'éléments d'un certain type déterminé (« nombres », ou « surfaces », etc.)

fonction (ou opération fonctionnelle) uniforme dans un ensemble E d'éléments de C , si à tout élément A de E correspond un nombre bien déterminé U_A .

Si l'on y regarde de près on peut trouver pourtant une différence entre la façon de convoquer une notion de fonction chez Volterra et chez Fréchet. Chez Volterra, comme chez Dirichlet, c'est la notion de *dépendance* qui est mise en avant et qui semble définir son propre champ d'application, tant au niveau de la variable que des valeurs obtenues. Ainsi une fonction qui dépend d'autres fonctions dépend en fait de *toutes les fonctions* définies sur un intervalle donné (AB) et possédant telle ou telle régularité en fonction des situations.

Fréchet introduit une organisation nouvelle en donnant *a priori* une collection d'éléments qui vont intervenir dans la relation qui définit la fonction. Il y a un précédent que nous avons vu dès ses premières interventions sur les fonctions de lignes. Rappelons qu'il avait introduit dès le début de son article une « famille G » de lignes et que la variation d'une fonction de ligne δU_L « dépend, bien entendu, de la famille G considérée. » Ici la mise en forme se cristallise un peu plus et fait apparaître ce que nous avons appelé des *éléments structurels* : une catégorie d'éléments est donnée *a priori*, et la fonction est vue comme une correspondance qui à ces seuls éléments associe des nombres réels ou complexes.

Dans son premier article, comme l'avait fait Volterra avant lui, Fréchet définit la continuité de façon particulière en des termes adaptés au type de variable qu'il manipule. Nous avons vu chez Volterra une adaptation de la définition pour les fonctions, puis pour les lignes et chez Fréchet l'hypothèse de continuité s'énonce via un paramètre réel qui permet de décrire la famille de lignes.

Dans la note qui nous intéresse maintenant Fréchet a une approche beaucoup plus générale et il dégage tout d'abord une notion de convergence séquentielle indépendante de la nature des éléments qui vont servir de variables pour les « opérations fonctionnelles » (ou « fonctions ») qu'il va considérer³⁷ :

Pour arriver à la notion de continuité d'une telle fonction, nous supposerons acquise une définition qui donne un sens précis à cette phrase : *la suite infinie $A_1, A_2, \dots, A_n, \dots$ d'éléments de C a une limite B* . Il nous suffira que cette définition, d'ailleurs quelconque, satisfasse aux deux conditions suivantes : 1° si la suite $A_1, A_2, \dots, A_n, \dots$ a une limite, toute suite $A_{p_1}, A_{p_2}, \dots, A_{p_n}, \dots$ formée d'éléments d'indices croissants de la première suite a aussi une limite qui est la même ; 2° si aucun des éléments $A_1, A_2, \dots, A_n, \dots$ d'une suite quelconque n'est distinct de A , cette suite a une limite qui est A .

Le caractère abstrait de cette notion de convergence va alors permettre de définir les notions de fermé, de continuité et de compact de façon également indépendante de la nature des éléments considérés. Par exemple la notion de compact est définie par Fréchet de la façon suivante³⁸ :

Nous appellerons *ensemble compact* tout ensemble E tel qu'il existe toujours au moins un élément commun à une suite infinie

³⁷[Fré04a, p.849]

³⁸[Fré04a, p.849]

quelconque d'ensembles $E_1, E_2, \dots, E_n, \dots$ contenus dans E , lorsque ceux-ci (possédant au moins un élément chacun) sont fermés et chacun contenu dans le précédent.

Nous verrons par la suite comment ces définitions seront reprises et modifiées dans la thèse de Fréchet dans une approche plus générale.

Dans ce cadre, le théorème de Weierstraß sur les extrema des fonctions (ordinaires) continue trouve une nouvelle formulation dont la généralité garantit l'applicabilité dans de nombreuses situations³⁹ :

THÉORÈME. Toute opération fonctionnelle U_A uniforme⁴⁰ et continue dans un ensemble compact et fermé E : 1° est bornée dans E ; 2° y atteint au moins une fois sa limite supérieure.

Fréchet finit enfin cette note en soulignant l'importance à ses yeux de la notion de compacité dans le théorème de Weierstraß et avance l'idée qu'une étude des « opérations fonctionnelles » doit donc être précédée d'une étude des ensembles compacts⁴¹. Il établit le lien avec la notion de convergence qui permet d'amorcer l'étude systématique des propriétés des ensembles compacts⁴² :

La condition nécessaire et suffisante pour qu'un ensemble E soit compact est que tout ensemble E_1 forme d'une *infinité* d'éléments distincts contenus dans E donne lieu à un élément limite au moins.

Cette propriété lui permet de plus de retrouver, en prenant une « définition ordinaire de la limite d'une suite de points », que les intervalles fermés bornés sont compacts. Il s'ensuit que sa définition est une généralisation qui s'appuie sur une notion étendue de limite et de convergence.

Ces idées seront développées et précisées dans la thèse de Fréchet que nous commenterons plus loin. Mais c'est aussi sa stratégie de recherche d'un cadre général pour traiter les problèmes qui nous intéressera dans la suite de notre étude. Elle illustre en effet de façon spectaculaire combien Fréchet amena les concepts de son prédécesseur italien sur un terrain nouveau.

4.3.2 LA THÈSE DE DOCTORAT DE FRÉCHET

La thèse de doctorat de Fréchet est publiée en 1906 sous le titre *Sur quelques points du calcul fonctionnel*⁴³ aux *Rendiconti del Circolo Matematico di Palermo*. Nous n'en donnons pas une présentation extensive, mais nous isolons quelques éléments qui prolongent les idées que nous venons de mentionner. Sa thèse illustre la voie qui s'appuie essentiellement sur des idées qui se rattachent à l'étude des fonctions et cela la connecte pour nous aux travaux de Volterra comme nous venons de le préciser.

³⁹[Fré04a, p.849]

⁴⁰« Uniforme » signifie ici que U n'associe à chaque A qu'une seule valeur.

⁴¹On peut consulter l'article « Historique de la notion de compacité » de J-P Pier pour avoir un aperçu général du développement de cette notion. Il y reprend en particulier les différentes approches de Fréchet et les met en perspective avec d'autres définitions. On y trouve de nombreuses références pour comprendre le lien avec diverses tentatives de définir la notion de convergence. (cf. [Pie80])

⁴²[Fré04a, p.849-850]

⁴³[Fré06]

La thèse se décompose dans sa structure la plus large en deux parties. La première partie est intitulée « Introduction de la notion de limite dans les ensembles abstraits » et sera celle qui nous servira essentiellement à illustrer notre propos. Elle est décomposée en deux chapitres le premier dédié à l'étude la plus générale où seule une notion de limite est imposée alors que le second chapitre introduit des classes plus restrictives permettant de résoudre de nouveaux problèmes. La seconde partie est intitulée « Applications de la théorie générale » et Fréchet y décline le travail de la première partie dans les cas particuliers (sous-ensembles de \mathbb{R} , ensembles de fonctions, etc.) qui lui semblent les plus importants. Les deux parties sont à peu près de même longueur, soit un peu plus ou un peu moins d'une trentaine de pages.

Nous analysons maintenant quelques aspects de ce texte en indiquant sa structure générale, au point de vue qui nous intéresse ici, lequel permet de mettre en évidence le contraste de ce travail avec les textes antérieurs de Volterra que nous avons étudiés.

LES ENSEMBLES ABSTRAITS POSSÉDANT UNE NOTION DE LIMITE ET LES OPÉRATIONS. Le développement des deux notions —ensembles abstraits et opérations— se fait dans une certaine organisation que nous allons étudier où elles restent indissociables l'une de l'autre pour répondre à un objectif précis : établir une théorie générale des opérations fonctionnelles.

Après avoir rappelé en introduction que dans la décennie qui précède, plusieurs mathématiciens (Le Roux, Volterra, Arzela, Hadamard) ont généralisé la notion de fonction en considérant des cas de plus en plus étendus, Fréchet s'engage dans une stratégie qui consiste alors à concevoir pour la variable le cadre le plus large et indéterminé possible⁴⁴ :

Nous dirons qu'une opération fonctionnelle U est définie dans un ensemble E d'éléments de nature quelconque (nombres, courbes, points, etc.) lorsqu'à tout élément A de E correspond une valeur numérique déterminée de $U : U(A)$. La recherche des propriétés de ces opérations constitue l'objet du Calcul Fonctionnel.

On remarque à nouveau un glissement dans l'utilisation de certains termes chez Fréchet : le « calcul fonctionnel » prend ici une nouvelle dimension et devient sous sa plume une *théorie* dont le but est d'étudier les « opérations ». Nous disons *glissement* car il nous semble qu'il y a une différence entre l'emploi qu'il fait ici de l'expression « calcul fonctionnel » et celui que l'on a cité plus haut dans son article sur les fonctions de lignes. En effet rappelons que sous certaines hypothèses « l'étude des fonctions U_L ne constituerait plus, à proprement parler, une question de calcul fonctionnel⁴⁵. » Ici Fréchet souligne que certaines relations relèvent ou ne relèvent pas d'un calcul portant vraiment sur les fonctions. On retrouve un emploi similaire du mot « fonctionnel » quand il qualifie la relation $U_{L_1} + U_{L_2} = U_{L_1+L_2}$ d'« équation fonctionnelle ». Désormais il emploie l'expression « calcul fonctionnel » comme un substantif, un mot composé qui désigne un champ des mathématiques comme le sont

⁴⁴[Fré04b, p.1]

⁴⁵Je souligne.

l'algèbre ou la géométrie par exemple. Son choix n'est pas sans précédent, mais il montre une évolution par rapport au vocabulaire employé jusqu'ici par Hadamard qui emploie souvent l'expression « opération fonctionnelle » où le mot « fonctionnelle » indique que l'opération agit sur des fonctions et ne fait pas référence à une théorie établie. Si l'on veut trouver une filiation il faut à nouveau la chercher du côté des Italiens. En effet Pincherle a publié en 1897 son texte « Mémoire sur le calcul fonctionnel distributif » aux *Mathematische Annalen* où, malgré la restriction indiquée dans le titre, il définit en introduction le calcul fonctionnel de façon générale⁴⁶

Aux deux premières parties que l'on a aperçues dès l'abord dans la science des nombres, s'en ajoute une troisième, à laquelle on pourrait donner le nom de Calcul fonctionnel : on réunirait sous ce titre les chapitres de l'analyse où l'élément variable n'est plus le nombre, mais la fonction considérée en elle-même.

Ainsi Fréchet se place dans cette lignée dont il découle un certain regard réflexif sur les opérations qui leur confère un statut particulier : elles passent d'outil pour étudier des problèmes (variationnels essentiellement) à objet d'étude en elles-mêmes.

L'introduction de sa thèse montre une vision qui s'appuie essentiellement sur une notion de fonction dont ses prédécesseurs ont cherché à étendre la conception et cherche à développer « un point de vue tout à fait général qui embrasse [les] différents cas ». Elle se termine en citant les divers travaux de Fréchet depuis 1904 sur le sujet, ce qui montre que les deux voies que nous identifions dans notre analyse se nourrissent l'une de l'autre bien que nous y voyions deux angles de vues distincts.

Nous n'évoquerons pas tous les résultats de la thèse mais nous établissons l'organisation générale de la théorie que Fréchet cherche à mettre en place dans cette première voie. Cela nous permettra de préciser la portée du travail de Fréchet en termes de généralité, et plus tard de le contraster avec le travail de Banach qui effectuera des choix sensiblement différents.

LES CLASSES (L). Fréchet fait d'abord référence à la « théorie des ensembles » dont il n'utilisera techniquement que des aspects de base, mais qu'il convoque ici de façon particulière afin d'obtenir une théorie générale⁴⁷ :

Les résultats qu'on obtiendra dans une telle théorie seront d'autant plus étendus qu'on s'adressera à des ensembles plus généraux. Pour obtenir la plus grande généralité possible, il y aurait donc lieu de chercher d'abord des résultats applicable à des ensembles abstraits, c'est-à-dire dont on ne spécifie pas la nature des éléments.

Il s'agit donc ici d'une référence à la théorie des ensembles qui a pour but de se dégager de la nature particulière des champs contenant les variables des opérations et d'éviter ainsi le recours aux « ensembles linéaires » ou à la théorie des fonctions qui seront relus comme des cas particuliers par la suite.

⁴⁶[Pin97, p.325]

⁴⁷[Fré06, p.4]

Ensuite c'est la notion de *limite* qui est revisitée et adaptée à des éléments de nature indéterminée. Nous avons vu une première version proposée pour définir la continuité dès 1904 dans l'optique de généraliser un théorème de Weierstraß. Dans sa thèse Fréchet définit précisément la *méthode* qu'il adopte afin d'obtenir un cadre général. En particulier il va s'appuyer sur une progression qu'il lit dans la théorie des groupes abstraits⁴⁸ :

[D]ans la théorie des groupes abstraits, le mode de composition est supposé défini à l'*avance* dans chaque cas particulier ; mais on ignore volontairement cette définition pour ne retenir que certaines conditions générales qu'elle remplit *mais qui ne la déterminent pas*.

Fréchet dissocie ici deux types de progression en distinguant les définitions « descriptives » du type de définition qui intervient dans la théorie des groupes —il ne leur donne pas de nom. Pour préciser sa pensée Fréchet reprend les termes de Lebesgue qui, dans le chapitre VII de ses « Leçons sur l'intégration » du Collège de France⁴⁹, revient sur différents types de définitions.

En fait il s'agit dans le texte de Lebesgue du chapitre intitulé « Intégrale définie des fonctions sommables » où il isole six propriétés que doit posséder une intégrale définie en un sens général. Les propriétés ont été sélectionnées par Lebesgue à partir de l'étude des intégrales déjà existantes, celle de Riemann ou celle de Duhamel et Serret dans son texte. Pour lui « leurs propriétés doivent nécessairement appartenir à l'intégrale, si l'on veut qu'il y ait une analogie entre cette intégrale et l'intégrale des fonctions continues. »

Lebesgue précise alors sa démarche⁵⁰ :

En énonçant les six conditions du problème d'intégration, nous définissons l'intégrale. Cette définition appartient à la classe de celles que l'on peut appeler *descriptives* ; dans ces définitions, on énonce des propriétés caractéristiques de l'être que l'on veut définir. Dans les définitions *constructives*, on énonce quelles opérations il faut faire pour obtenir l'être que l'on veut définir.

En d'autres termes une définition « descriptive » telle qu'elle est conçue par Lebesgue et plusieurs autres mathématiciens aux yeux de Fréchet (ce dernier cite Drach, Borel, Hadamard⁵¹) permet de penser un nouvel « être » grâce à des propriétés commune à plusieurs situations antérieures. La nouvelle définition embrasse les cas précédents qui, dans une nouvelle organisation des idées, deviennent des cas particuliers. Il y a un mouvement d'abstraction qui permet la sélection des propriétés à conserver et une rupture en termes de généralité : la définition nouvelle est une *extension* des concepts précédents.

La proposition de Fréchet est sensiblement différente et il l'illustre avec l'exemple de la théorie des groupes. En effet, la théorie des groupes abstraits formule que pour chaque groupe particulier la loi (ou, pour reprendre les mots de Fréchet, « le mode de composition ») en jeu possède déjà *toutes* les propriétés

⁴⁸[Fré06, p.5]

⁴⁹[Leb04]

⁵⁰[Leb04, p.106]

⁵¹Il serait intéressant ici d'établir divers aspects qui sont discutés chez ces auteurs. Nous ne citons que Lebesgue afin d'éclairer la pensée de Fréchet et de mieux comprendre ce passage de son texte.

de groupe. Le procédé d'abstraction à l'œuvre s'appuie sur une *ressemblance de famille* entre les différentes situations et isole des propriétés qui ne dépendaient pas de la nature particulière des éléments en jeu dans chaque cas. Il s'ensuit un concept général qui entretient avec les cas particuliers une relation singulière où chaque « groupe de mouvement, de substitution, de transformation etc. » joue le rôle de ce que l'on pourrait nommer un *paradigme axiomatique*⁵². Un groupe donné *concret* (terme d'observateur) devient un archétype de la notion de groupe et une vision abstraite et générale permet de concevoir les différentes occurrences de groupe comme autant de déclinaisons d'un modèle choisi.

Fréchet adopte donc ce que nous nommerons une *définition synthétique*⁵³, par contraste avec les définitions « descriptives » ou « constructives », qu'il élabore à partir de situations choisies et dans une démarche générale⁵⁴ :

Si l'on examine avec soin les diverses définitions classiques de la limite d'une suite de nombres ou d'une suite de points ou d'une suite de fonctions, etc., on remarque immédiatement que toutes ses définitions satisfont à deux conditions I et II que l'on peut énoncer indépendamment de la nature des éléments considérés.

C'est à partir de cette réflexion que Fréchet introduit sa notion de classe (L). Dans un mouvement de sélection et d'abstraction, il isole deux propriétés qui ne dépendent pas de la nature des éléments en jeu dans les suites convergentes et en fait le cœur de sa définition. Les deux conditions retenues ont déjà été données dans l'article sur la généralisation d'un théorème de Weierstraß mais le contexte est étoffé ici et organisé d'une façon particulière qui définit précisément le contexte général convoqué par Fréchet⁵⁵

Dorénavant, nous nous limiterons donc à l'étude des ensembles tirés d'une classe (L) d'éléments de nature quelconque mais satisfaisant aux conditions suivantes : on sait distinguer si deux éléments de la classe (L) sont distincts ou non. De plus, on a pu donner une définition de la limite d'une suite d'éléments de la classe (L). Nous supposons donc qu'étant choisie au hasard une suite infinie d'éléments (distincts ou non) de la classe (L), on puisse dire d'une façon certaine si cette suite $A_1, A_2, \dots, A_n, \dots$ a ou non une limite A (d'ailleurs unique). Le procédé qui permettra de donner la réponse (autrement dit la définition de la limite) est d'ailleurs absolument quelconque, assujéti seulement à satisfaire aux conditions I et II dont nous avons parlé et qui sont les suivantes :

I) Si chacun des éléments de la suite infinie $A_1, A_2, \dots, A_n, \dots$ est identique à un même élément A , la suite a certainement une limite qui est A .

⁵²Nous construisons ce terme sur le modèle du *paradigme de flexion* dans son rôle de modèle des formes d'un certain type de mots.

⁵³Nous avons employé le mot « synthétique » à plusieurs reprises dans notre travail. Ce n'est pas un hasard et nous dégagerons petit à petit en quoi ce terme permet de clarifier un certain nombre de ressemblances que nous aurons pu établir.

⁵⁴[Fré06, p.5]

⁵⁵[Fré06, p.5-6]

II) Si une suite infinie $A_1, A_2, \dots, A_n, \dots$ a une limite A , toute suite d'éléments de la première suite pris dans le même ordre : $A_{n_1}, A_{n_2}, \dots, A_{n_p}, \dots$ (les nombres entiers n_1, n_2, \dots, n_p iront donc en croissant) a une limite qui est aussi A .

Précisément il ne s'agit pas d'une définition de limite que Fréchet choisit de convoquer ici, mais la définition d'un type d'ensemble possédant une notion de limite. La définition *synthétique* désigne ce *type* d'ensemble par l'expression « classe (L) ». Il en résulte pour nous que le rapport de général à particulier ne s'établit pas entre une notion générale de convergence et des définitions particulières (par la distance, la norme, etc.), mais entre une définition abstraite de classe (L) et divers exemples d'ensembles particuliers où existe une notion de limite : ensembles de nombres, de points, de fonctions, chacun avec sa propre notion de limite.

À partir de ce moment Fréchet va considérer des ensembles abstraits qui possèdent une notion de limite qui vérifie les deux propriétés, et donc se placer systématiquement dans une classe (L). En s'appuyant sur la notion de limite qui existe dans une telle classe, Fréchet va établir des notions de topologie (mot d'observateur) adaptées au cadre abstrait dans un paragraphe intitulé « Les définitions usuelles de la théorie des ensembles ponctuels étendues aux ensemble abstraits. » Les principales notions avaient été déjà données par Fréchet dans son article de 1904 sur le théorème de Weierstraß. Nous reprenons l'étude à travers le texte de la thèse qui construit réellement une théorie générale et qui illustre la voie que nous souhaitons dégager dans le travail de Fréchet. On lit dans ce titre comment le *paradigme axiomatique* fait œuvre : il permet à partir d'une situation particulière, ici la définition de la « théorie des ensembles ponctuels » (lus comme cas particulier de classes (L)), de donner des définitions générales. Ainsi Fréchet définit en quelques lignes les notions (que nous qualifierions désormais de séquentielles) d'élément limite⁵⁶, d'ensemble dérivé, fermé, parfait, d'intérieur au sens étroit⁵⁷.

Si ces différentes définitions sont immédiatement transposées au cas général, le cas des « éléments de condensation » montre un aspect du processus d'abstraction qui permet de donner une formulation dans les espaces abstraits. Pour cette notion Fréchet cite une définition initiale de Lindelöf qu'il lit dans le cours de Borel de 1905 « Leçons sur les fonctions de variables réelles et les développements en séries de polynômes »^{58,59} :

Nous appellerons avec [M. Ernst Lindelöf] *point de condensation* d'un ensemble E un point A tel que, dans une sphère de centre A et de rayon aussi petit que l'on veut, il existe une infinité *non dénombrable* de points de E .

⁵⁶Soit E un ensemble d'éléments d'une classe (L) où l'on a pu définir une notion de limite qui vérifie les propriétés I et II au-dessus. Fréchet dit alors que A , élément de (L) est un « élément limite de E » s'il existe une suite d'éléments distincts (nous soulignons) de E qui tendent vers A .

⁵⁷Avec une formulation actuelle $A \in E \subset H$ est un élément « intérieur à E au sens étroit » si A n'est limite d'aucune suite d'éléments distincts de $H \setminus E$.

⁵⁸Cours professé par Borel à l'ÉNS pendant l'année 1903-04, rédigé par Fréchet lui-même et complété par trois longues notes de Painlevé et de Lebesgue.

⁵⁹[Bor05, p.4]

Mais cette définition « ne se prête pas à la généralisation actuelle » d'après Fréchet qui va définir à son tour un « élément de condensation » d'un ensemble E comme « élément limite de E qui est aussi limite de tout ensemble obtenu en supprimant de E une infinité *dénombrable* d'éléments. »⁶⁰ Fréchet remarque que sa définition est « équivalente » à celle de Lindelöf lorsque E est un « ensemble ponctuel ». Dans le contexte des « ensembles abstraits » et des classes (L) la notion de sphère utilisée par Lindelöf semble inappropriée et incompatible avec le type de généralité qu'il met ici en place. Il y a donc un véritable travail de sélection des définitions à prendre en considération dans les cas particuliers car celles-ci doivent suivre un type de critère de compatibilité que nous voyons à l'œuvre ici.

En ce qui concerne les définitions « descriptives » ou « constructives », Lebesgue avait déjà mentionné que les définitions devaient satisfaire à certaines conditions⁶¹ :

Lorsque l'on a énoncé une définition constructive, il faut démontrer que les opérations indiquées dans cette définition sont possibles ; une définition descriptive est aussi assujettie à certaines conditions : il faut que les conditions énoncées soient compatibles.

Pour Fréchet un nouveau critère intervient dans la formation des définitions *synthétiques* : celles-ci ne doivent pas faire intervenir d'élément qui n'ont pas été définis dans le contexte le plus large et le plus général, ici les classes (L). Nous verrons plus tard que Fréchet définira des sous-classes qui permettent d'importer d'autres types de définitions.

Enfin Fréchet introduit dans un paragraphe séparé la notion de « compact » : celle-ci doit jouer « le même rôle dans la théorie des ensembles abstraits que la notion d'ensemble limité dans la théorie des ensembles ponctuels. » C'est Fréchet qui a proposé pour la première fois dans son article de 1904 le mot « compact » comme le souligne Pier dans son article « Historique de la notion de compacité ». Pour Fréchet « un ensemble est *compact* lorsqu'il ne comprend qu'un nombre fini d'éléments ou lorsque toute infinité de ses éléments donne lieu à au moins un élément limite. »⁶² Si le « compact » joue un rôle similaire à celui d'« ensemble limité » dans les ensembles ponctuels, c'est la notion d'« ensemble *extrémal* », un compact fermé, que Fréchet rapproche de celle d'intervalle dans les « ensembles linéaires ».

Fréchet donne alors trois résultats qui montrent pour lui l'analogie qu'il y a avec les « ensembles *limités* ponctuels ». Ils « résultent directement de la définition »⁶³ :

Tout ensemble contenant un ensemble non compact est non compact. Toute partie d'un ensemble compact est compacte. Tout ensemble formé d'un nombre fini d'ensembles compacts est un ensemble compact.

Un seul théorème apparaît dans cette partie avec sa démonstration. Bien sûr les éléments que nous venons de résumer permettent de cerner les différentes

⁶⁰[Fré06, p.6]

⁶¹[Leb04, p.107]

⁶²[Fré06, p.6]

⁶³[Fré06, p.7]

notions *topologiques* (nos termes) que Fréchet met en place pour constituer le contexte général dans lequel il va travailler. Mais la démonstration que nous allons maintenant analyser offre bien plus et permet de préciser le travail d'abstraction de Fréchet dans la conception des « espaces abstraits ». En effet, celui-ci, par une note placée à la fin de sa preuve, renvoie à une autre démonstration antérieure de Baire en précisant que sa lecture permettait de mieux apprécier ce qui était jeu dans le *passage* de la « théorie dans les ensembles linéaires » où la notion d'intervalle constitue un outil de pensée puissant, à la théorie des ensembles abstraits et des classes (L) ⁶⁴ :

Pour se rendre compte de la transformation qu'il faut effectuer sur les démonstrations ordinaires relatives aux ensembles ponctuels, voir la démonstration correspondante de M. Baire [Bai05, p18], où il utilise la division d'un intervalle en intervalles moitiés, ce qui n'est pas généralisable.

L'analyse comparée des deux démonstrations permet d'éclairer précisément le type de vision générale que Fréchet adopte et la véritable transposition qu'il opère à partir des schémas habituels quand la notion de segment est utilisable.

Tout d'abord rappelons le théorème tel qui est donné par Baire⁶⁵ :

Étant donnés, sur un segment fini, des ensembles fermés en nombre infini $P_1, P_2, \dots, P_\nu, \dots$, satisfaisant aux conditions

$$P_1 \supseteq P_2 \supseteq \dots \supseteq P_\nu \supseteq \dots,$$

il existe au moins un point qui appartient à tous les P_ν .

En effet, par hypothèse, le segment AB contient, quel que soit ν des points de P_ν ; on reconnaît que, si l'on divise AB en deux segments AC, CB , la même propriété appartient à l'un au moins de ces segments; en répétant le raisonnement, on forme une suite de segments possédant tous cette propriété, et dont chacun est contenu dans le précédent, et l'on peut faire en sorte que leur longueur tende vers 0; dans ces conditions, il y a un point unique qui est commun à tous ces segments; ce point appartient à chacun des ensembles P_ν , qui sont tous fermés.

Pour bien interpréter cette démonstration il faut préciser avant toute chose quelques points. Les notions de « point limite », d'« ensemble dérivé » et d'« ensemble fermé » sont définies en introduction par Baire pour des ensembles de points P d'un segment de droite AB . La définition sur laquelle reposent les autres est celle de « point limite » $M \in P$ qui vérifie que « tout segment contenant M à son intérieur contient un point de P autre⁶⁶ que M . » Nous avons vu que la version transposée donnée par Fréchet définissait un « élément limite » en imposant qu'une suite d'éléments **distincts** tend vers cet élément.

Si on reprend avec nos notations la preuve de Baire, on a une suite emboîtée d'intervalles $I_{k+1} \subset I_k \subset \dots$ dont la longueur tend vers 0 et qui contiennent **des** points de chacun des ensembles P_ν . Si l'on ne veut pas faire d'ajout à la

⁶⁴[En note de bas de page Fré06, p.7]

⁶⁵[Bai05, p.18-19]

⁶⁶Nous soulignons.

démonstration proposée et utiliser la définition de « point limite » que nous venons de rappeler, il vaut mieux insister sur le fait que I_k contient *plusieurs points* de P_v (c'est le sens du pluriel « contient DES points » au-dessus). Cela permet de penser le point M comme « point limite » de chaque P_v puisque tout intervalle contenant M contient un point de P_v qui n'est pas M .

Mais l'essentiel pour nous est ailleurs, dans l'intervention d'une collection particulière d'ensembles que sont les intervalles (fermés). La généralité de la démonstration pour des ensembles fermés quelconques est obtenue par Baire par une sorte d'extension d'une version préliminaire obtenue pour les intervalles. La technique qui consiste à partitionner l'intervalle de départ en des intervalles de longueurs de plus en plus petites et qui tendent vers 0 est utilisée dans plusieurs situations antérieures dans le texte de Baire. Elle introduit un élément étranger aux notions qui servent à définir les « points limites » ou les « fermés » en faisant intervenir la longueur des intervalles qui permet de passer d'une idée de convergence de nombres (vers 0) à l'existence d'un point commun à des intervalles emboîtés. L'extension aux ensembles quelconques repose alors sur un procédé astucieux de généralisation qui ramène momentanément le cas des ensembles fermés quelconques à une étude qui met en jeu des intervalles.

Il y a pour nous donc deux éléments que Fréchet cherche à éviter. L'un est directement visé par la remarque de bas de page que nous avons citée plus haut. Ce n'est pas l'intervention des intervalles qui pose véritablement problème car ils peuvent être transposés dans la théorie générale en des compacts. Fréchet pense que la démarche qui consiste à « divis[er] un intervalle en intervalles moitiés » ne peut pas se transposer sans ajouter un élément supplémentaire qui *mesure* les intervalles et nuit donc à la généralité qu'il cherche à mettre en place.

Le second élément du texte de Baire qui à nos yeux pose problème et semble conflictuel par rapport à la vision de Fréchet est lié au type de généralité que celui-ci installe dans sa thèse. Nous pouvons la comparer à celle de Baire en étendant le sens donné à des termes que nous avons déjà utilisés. En effet, il nous semble que l'on pourrait qualifier de *généralité constructive* celle qui émerge du texte de Baire. À partir d'un résultat qu'il tient pour un type particulier d'ensembles, les intervalles fermés, il *construit* une démonstration qui s'appuie sur cette première forme et l'étend en un résultat général. Au contraire Fréchet ne veut pas faire reposer ses résultats généraux sur une sorte d'édifice construit par extension et il choisit des démonstrations qui se donnent dans un à-plat qui définit un type de généralité que nous qualifierons de *généralité synthétique*.

On voit dans cette optique que l'enjeu n'est pas seulement de savoir si la théorie de Baire tombe sous le coup de celle de Fréchet qui serait donc en un sens plus générale, mais plus essentiellement de discerner le type de généralité qui est à l'œuvre dans un cas et dans l'autre.

De son côté le théorème donné par Fréchet prend la forme suivante⁶⁷ :

Si on considère une suite d'ensembles $E_1, E_2, \dots, E_n, \dots$ formés d'éléments d'un même ensemble compact E , chacun fermé, contenu dans le précédent et possédant au moins un élément, il y a nécessairement un élément commun à tous ces ensembles.

⁶⁷[Fré06, p.7]

En effet, il y a au moins un élément M_1 dans E_1 , M_2 dans E_2 etc. S'il y a une infinité d'éléments distincts dans la suite M_1, M_2, \dots , soient M_{n_1}, M_{n_2}, \dots ceux-ci. L'ensemble E étant compact, cette dernière suite a au moins un élément limite M ; comme elle est contenue dans E_{n_p} à partir de M_{n_p} , M sera un élément limite de E_{n_p} et par conséquent appartiendra à l'ensemble fermé E_{n_p} . Ainsi M appartient à E_{n_p} et par suite à E_1, E_2, \dots . Comme cela a lieu quel que soit p , il en résulte que M est commun à tous les ensembles E_1, E_2, \dots . Si, au contraire, il n'y a dans la suite M_1, M_2, \dots qu'un nombre fini d'éléments distincts, c'est qu'il y a au moins un élément M qui est répété une infinité de fois dans cette suite, soit $M \equiv M_{n_1} \equiv M_{n_2} \equiv \dots$. Alors M est contenu, quel que soit p , dans E_{n_p} et par suite dans E_1, E_2, \dots, E_{n_p} . Il est donc commun à tous les ensembles $E_1, E_2, \dots, E_{n_p}, \dots$

Ici la démonstration ne ramène pas à une catégorie d'ensembles particuliers qui auraient des propriétés supplémentaires comme l'intervalle par rapport à l'ensemble général dans le cas précédent. Fréchet ne fait intervenir que les propriétés qui sont assignées aux ensembles donnés dans l'énoncé du théorème. Le type et l'extension de la généralité qui en découlent sont alors non seulement donnés par l'énoncé du théorème mais aussi par la démonstration qui ne repose que sur les propriétés des éléments convoqués dans l'énoncé.

Pour préciser ces idées on peut reprendre dans une vue plus large les différents éléments que nous venons d'analyser. Nous avons fait la différence entre deux types de généralité, l'une *constructive* illustrée par le texte de Baire, l'autre *synthétique* dans la thèse de Fréchet. Dans le second cas, l'*à-plat* que nous avons cherché à mettre en valeur consiste à ne pas faire intervenir d'élément parasite qui entretienne une relation de général à particulier comme dans le cas des intervalles par rapport aux ensemble généraux chez Baire. Pourtant *à-plat* ne signifie pas pour nous qu'il n'y a pas une organisation structurée et le texte de Fréchet permet de préciser le type de généralité à l'œuvre chez lui en termes de *structure*.

En effet sa thèse convoque sur un même plan, mais dans une certaine organisation, plusieurs idées. D'abord il convoque des éléments de théorie des ensembles, puis fait l'hypothèse d'une notion de limite en définissant les classes (L) avant de distinguer des ensembles particuliers —fermés, compacts, etc. Pour comprendre la notion d'*à-plat* que nous proposons ici, il faut préciser les relations qu'entretiennent divers types d'éléments. Dans le texte de Baire un intervalle est un ensemble *particulier*. Il se joue entre les deux notions une relation qui donne à l'intervalle un certain nombre de propriétés que l'ensemble général ne possède pas toujours. Pourtant les ensembles généraux sont souvent manipulés en même temps que les intervalles qui sont toujours dans le paysage.

Chez Fréchet le compact est un ensemble que l'on a *distingué* par une propriété, celle donnée par sa définition. Il est distingué dans un second temps et participe d'une organisation des idées qui relève d'une composition plutôt que d'une construction. Au contraire un intervalle est un ensemble particulier dans l'utilisation qu'en fait Baire, non seulement parce qu'il a des qualités qui le distinguent des autres ensembles, mais parce qu'il est consubstantiel aux

ensemble généraux, le contexte convoqué par Baire les donne dans une relation organique où l'ensemble général est pensé alors que l'ensemble particulier — l'intervalle ici— existe déjà. En quelques sortes l'ensemble général chez Baire est une extension de l'idée d'intervalle, d'où l'idée d'une *généralité constructive*, alors que chez Fréchet l'ensemble général est posé d'emblée avant de distinguer le compact.

Enfin, en ce qui concerne le dernier aspect de la notion de *structure* à laquelle nous faisons appel ici, nous remarquons une certaine visée *réflexive* dans l'approche de Fréchet qui détermine un type de structure particulier que nous distinguons par exemple d'*éléments structurels* dans notre étude. Dès l'article de 1904 sur le théorème de Weierstraß, Fréchet a indiqué la nécessité d'étudier certaines propriétés des compacts, et plus tard dès sa thèse, des différentes classes qui interviennent dans sa théorie. Nous citons pour illustrer cette vision *réflexive* le résultat que Fréchet donne dans son article de 1904 sans qu'il y ait une quelconque application autre que préciser une propriété importante des ensembles compacts⁶⁸ :

La condition nécessaire et suffisante pour qu'un ensemble E soit compact est que tout ensemble E , formé d'une infinité d'éléments distincts contenus dans E donne lieu à un élément limite au moins.

Remarque : nous verrons plus loin que la distinction entre *général synthétique* et *général constructif* ne crée pas une ligne de démarcation toujours infranchissable dans les textes que nous étudions. En particulier la notion d'« ensemble continu » montrera que Fréchet à recours aussi à des éléments hors du cadre qu'il semble ici imposer.

On peut avoir l'impression à ce moment-là que le lien que nous avons établi avec le travail de Volterra s'est quelque peu distendu et que le texte de Fréchet se développe désormais suivant une dynamique propre dont nous avons éclairé les principes et l'organisation en termes de généralité et de structure. Pourtant on retrouve parallèlement à cette autonomie grandissante dirigée vers l'établissement d'une théorie générale un certain nombre de thèmes et de références qui nous ramènent inexorablement vers le mathématicien italien. En particulier la section suivante introduite par Fréchet montre que l'un des objectifs est de construire une théorie des « opérations » et pas une *géométrie* des classes (L).

LES OPÉRATIONS SUR LES ENSEMBLES ABSTRAITS (L). En deux parties et environ sept pages Fréchet va étudier ici les propriétés des « opérations » définies sur un ensemble où l'on dispose de la notion de limite. Grâce à l'introduction des ensembles abstraits sous la forme la plus large possible des classes (L) Fréchet va pouvoir développer une étude des « opérations » en faisant ressortir l'importance que l'on donne à l'ensemble « dans » lequel elles sont définies. La première partie sur les ensembles abstraits a fait apparaître un regard réflexif qui donne une certaine indépendance aux développements que l'on a souvent *a posteriori* qualifiés de *topologiques*. L'étude des « opérations » conduit à une deuxième lecture de la thèse de Fréchet et illustre une visée

⁶⁸[Fré04a, p.849-850]

générale qui, dans le prolongement des idées de Volterra, met en place une théorie des opérations en dégageant étape par étape les éléments strictement nécessaires pour chaque conception, pour chaque résultat.

Fréchet aborde l'étude des « opérations fonctionnelle V uniformes dans un ensemble E d'éléments d'une classe (L) » en donnant une définition de la continuité similaire à celle de 1904 que nous avons rappelée p.160. Seule la forme de l'énoncé change et Fréchet utilise la notation $V(A) = \lim_{n=\infty} V(A_n)$ pour indiquer la propriété de continuité.

Pourtant son étude cherche systématiquement à dégager les pourtours de ce que l'on peut énoncer en admettant un minimum de propriétés. Ainsi le premier résultat étend un théorème donné par Weierstraß pour les fonctions de la variable réelle, puis par Arzelà pour des fonctions de lignes⁶⁹ :

Étant donnée une opération U uniforme dans un ensemble extrémal E , il existe au moins un élément A de E tel que la limite supérieure μ (finie ou non) de U dans E soit égale à la limite supérieure de U dans tout ensemble K d'éléments de E auquel A est intérieur au sens étroit [en considérant E comme l'ensemble fondamental]⁷⁰

Nous indiquons ce théorème ici car il représente l'énoncé le plus général donné par Fréchet sur les « opérations ». Il fait intervenir du côté de la variable les propriétés des ensembles abstraits et en particulier des compacts et du côté des valeurs prises par U des propriétés propres à l'ensemble des nombres réels (par exemple si $\mu \geq \mu_1 > \alpha_n$ et $\alpha_n \rightarrow \mu$ alors $\mu_1 = \mu$.) En fait la disparité de traitement des valeurs de U d'un côté et des éléments qui servent de variable nous ramène vers un type de *généralité constructive*. En effet la démonstration, par le biais de la définition d'« opération » que choisit Fréchet, fait intervenir des propriétés des deux types d'éléments comme le faisait Baire en utilisant les intervalles. La différence réside dans l'organisation des deux types d'éléments qui sont ici conçus dans deux identités distinctes qui n'interagissent que par le biais de l'opération. L'aspect *constructif* se décline donc ici de façon différente mais toujours en faisant intervenir deux niveaux de généralité, les ensembles abstraits d'un côté, les réels de l'autre.

On retrouve ce type de *généralité constructive* un peu plus loin dans le texte quand Fréchet définit les « ensembles continus » de la façon suivante⁷¹ :

J'appelle ensemble continu, un ensemble tel qu'étant donnés deux quelconques de ses éléments A, B , on peut extraire de E un ensemble F dont chaque élément correspond à un point de l'intervalle $(0, 1)$ sur un axe $0t$ et inversement. La correspondance est supposée telle que si deux éléments A_1, A_2 de F correspondent à deux points t_1, t_2 , A_1 tend vers A_2 quand t_1 tend vers t_2 .

Cette définition permet de préciser l'aspect *constructif* que nous avons évoqué. En effet Fréchet n'utilise pas ici le vocabulaire des « espaces abstraits »

⁶⁹[Fré06, p.8]

⁷⁰Les crochets font partie du texte de Fréchet. Celui-ci a défini la notion d'« intérieur au sens étroit » un peu avant dans la partie sur les ensembles abstraits : $A \in K$ n'est limite d'aucune suite d'éléments distincts de $E \setminus K$.

⁷¹[Fré06, p.9]

ni celui d'« opération » mais donne une définition qui repose sur des éléments qui sont *rajoutés* au contexte général précédent. On note la différence entre la définition de compact qui *distingue* des ensembles particuliers en s'appuyant seulement sur les définitions des classes (L) et ici une définition qui juxtapose une partie exprimée dans du vocabulaire qui n'appartient pas à celui employé pour définir les classes (L).

SUITES D'OPÉRATIONS. Quand Fréchet envisage les suites d'opérations il utilise plusieurs types de convergence particuliers qui lui permettent d'utiliser les outils des ensembles abstraits sur les ensembles d'opérations. Il définit tour à tour la notion de « convergence », qui correspond à ce que nous appellerions *convergence simple*, la « convergence uniforme » et la « convergence quasi-uniforme » qui a été introduite par Arzelà pour étudier des séries de fonctions et caractériser les sommes continues⁷². Fréchet donne cette dernière définition sous la forme suivante⁷³ :

Nous dirons qu'une suite d'opérations $U_1, U_2, \dots, U_n, \dots$, uniformes dans un ensemble E converge quasi-uniformément vers une opération U , lorsque, étant donné le nombre $\epsilon > 0$ et un entier quelconque N , on peut déterminer, une fois pour toutes, un entier $N' \geq N$ tel que, à tout élément A de E , on puisse faire correspondre un entier n_A tel que

$$N \leq n_A \leq N', \quad |U(A) - U_{n_A}(A)| < \epsilon.$$

Après avoir rappelé qu'il y a des suites d'opérations continues dont la limite est continue sans que la convergence soit uniforme, il montre que c'est la condition de « convergence quasi-uniforme » qui permet d'établir la caractérisation des suites qui ont une limite continue. Il s'agit d'une généralisation du théorème donné par Arzelà pour les séries que nous avons rappelé plus haut.

Ce type de théorème est dédié à l'étude des ensembles d'opérations où l'on a défini un type de convergence et introduit donc une certaine vision réflexive sur les opérations. Bien que Fréchet n'envisage pas le cas le plus général en considérant un type de convergence non défini à l'avance, il propose cependant trois cas et se sert de la quasi-continuité pour donner un résultat propre aux opérations continues.

Finalement c'est la convergence uniforme qui est retenue par Fréchet dans la suite et pour laquelle il va donner un certain nombre de résultats. En particulier la notions « compact » pour un ensemble d'« opérations » est définie relativement à la convergence uniforme⁷⁴ :

⁷²La définition est exposée par Arzelà dans [Arz99, p.153] et l'expression « convergenza a tratti » est employé p.155 puis avec le terme « uniforme » p.160. En fait la définition d'Arzelà est exprimée sur les restes $R_m(x)$ de la série et s'énonce en deux temps. Dans une première partie Arzelà donne une définition qui ne fait pas appel aux segments (*tratti*). La forme est alors très proche de celle donnée par Fréchet dans le cas des espaces abstraits. C'est dans un second temps qu'Arzelà précise la version plus géométrique qui fait intervenir des segments et donne son nom au type de convergence qu'il convoque.

⁷³[Fré06, p.10]

⁷⁴[Fré06, p.11]

Considérons une famille \mathfrak{F} d'opérations continues dans un même ensemble E formé d'éléments d'une classe (L). Dans un grand nombre de questions de l'Analyse, il serait utile de savoir si la famille \mathfrak{F} est telle que de toute infinité d'opérations de \mathfrak{F} distinctes, on peut tirer une suite $U_1, U_2, \dots, U_n, \dots$ qui converge uniformément vers une limite U nécessairement continue dans E . Nous dirons qu'une telle famille \mathfrak{F} est compacte.

Toujours en s'appuyant sur les travaux d'Arzelà dont il cite une dizaine d'articles, Fréchet introduit alors une notion d'« égale continuité »⁷⁵ Nous rappelons la définition car elle intervient de façon centrale ici⁷⁶ :

Nous dirons que des opérations uniformes dans un même ensemble E formé d'éléments d'une classe (L), constituent une famille \mathfrak{F} d'opérations également continues en A dans E , si, étant donné un nombre $\epsilon > 0$ et une suite quelconque d'éléments de E : A_1, A_2, \dots, A_n , ayant pour limite un élément A de E , on peut trouver un entier p tel que l'inégalité $n > p$ entraîne

$$|U(A) - U(A_n)| < \epsilon,$$

quelle que soit l'opération U de la famille \mathfrak{F} .

Fréchet donne à nouveau ici une définition qui est « indépendante de la notion d'intervalle » et donc plus générale que celle d'Ascoli. Il est intéressant de noter un certain lien avec des aspects géométriques incarnés par la notion de ligne ou de courbe. Ces notions ont disparu de la définition que donne Fréchet et elles nuiraient au type de vision générale qu'il met en place. Cependant les références citées par le mathématicien permettent de concevoir qu'il s'appuie sur certaines conceptions antérieures qui mettent en jeu des lignes : Volterra comme nous l'avons vu, mais aussi Arzelà⁷⁷ sur les « fonctions de lignes » et Hilbert sur les géodésiques⁷⁸. C'est principalement à partir des idées développées dans ces travaux que Fréchet trouve des formulations susceptibles d'être transposées au cadre général qu'il construit. On y voit l'importance dans le processus d'abstraction des figurations antérieures qui utilisent la notion de ligne et s'appuient sur une certaine vision géométrique. Comme nous l'avons vu la notion de ligne permet de dissocier plusieurs notions qui vont servir de matériau de base et de se libérer petit à petit d'une organisation *constructive* où les propriétés de \mathbb{R} ou des segments sont non seulement omniprésents mais jouent de surcroît un rôle particulier incompatible avec le type de généralité que Fréchet veut mettre en place.

La manière de procéder de Fréchet est alors la suivante. Dans un premier temps il isole un certain nombre de lemmes puis il démontre un théorème en utilisant les lemmes de façon particulière. Nous reprenons maintenant le

⁷⁵Arzelà convoque cette notion qui a été introduite initialement par Ascoli dans son travail sur les familles de courbes (cf. [Asc83]).

⁷⁶[Fré06, p.11]

⁷⁷[Arz89]

⁷⁸[Hil00]

théorème prouvé par Fréchet en précisant en note de bas de page les lemmes utilisés⁷⁹ :

THÉORÈME - Pour que des opérations continues dans un même ensemble extrémal E formé d'éléments d'une classe (L) , qui appartiennent à un même ensemble dénombrable D ou à son dérivé D' , forment une famille compacte \mathfrak{F} , il faut et il suffit que les opérations de \mathfrak{F} soient également continues et bornées dans leur ensemble en tout élément de E .

La preuve suit l'énoncé :

La condition est nécessaire d'après le dernier Lemme⁸⁰. Pour démontrer qu'elle est suffisante, il suffit de démontrer que, de toute suite infinie d'opérations de $\mathfrak{F} : U_1, U_2, \dots$ on peut extraire une suite qui converge uniformément dans E vers une opération qui sera nécessairement continue dans E .

En effet, appelons A_1, A_2, \dots la suite des éléments de l'ensemble dénombrable D et considérons la suite S_n des nombres $x_1^{(n)} = U_n(A_1), \dots, x_p^{(n)} = U_n(A_p), \dots$. Puisque les opérations de \mathfrak{F} sont bornées dans leur ensemble en tout élément de D , ces nombres $x_p^{(n)}$ sont, pour chaque valeur de p , compris, quel que soit n , entre deux nombres fixes λ_p et μ_p . D'après la dernière proposition que nous avons rappelée⁸¹, on pourra donc extraire de la suite S_1, S_2, \dots une suite S_{n_1}, S_{n_2}, \dots d'indices croissants, telle que $x_p^{(n_q)} = U_{n_q}(A_p)$ converge vers une certaine limite x_p et cela quel que soit p . Cela veut dire qu'on peut extraire de la suite U_1, U_2, \dots une suite d'opérations $U_{n_1}, U_{n_2}, \dots, U_{n_q}, \dots$ convergente en tout élément de D . D'après le deuxième Lemme⁸², cette suite d'opérations sera aussi convergente en tout élément de E ; d'après le premier Lemme⁸³ la convergence sera uniforme, et par suite la limite continue dans E .

L'analyse de ce passage permet de comprendre l'organisation que met en place Fréchet et de préciser le type de généralité que nous cherchons à établir dans ce texte.

Les lemmes ont à nos yeux deux rôles importants.

Premièrement ils permettent à Fréchet de donner une démonstration du théorème très courte qui met en relief un certain enchaînement logique et qui fait ressortir les hypothèses strictement utilisées lors de chaque étape de la progression. Cette utilisation de lemmes confère à la preuve une qualité

⁷⁹[Fré06, p.13-14]

⁸⁰Pour que des opérations continues dans un même ensemble extrémal E forment une famille compacte \mathfrak{F} , il faut qu'elles soient également continues et bornées dans leur ensemble.

⁸¹Même lemme que précédemment.

⁸²Considérons une suite d'opérations U_1, U_2, \dots convergente dans un ensemble D . Si ces opérations sont également continues dans l'ensemble F formé des éléments de D et de son ensemble dérivé D' , la suite considérée est aussi convergente dans F .

⁸³La limite U d'une suite convergente $U_1, U_2, \dots, U_n, \dots$ d'opérations également continues dans un ensemble E , est une opération continue dans E . De plus, si E est extrémal, la convergence est nécessairement uniforme dans E .

particulière où l'on peut comprendre l'enjeu de la modification de telle ou telle hypothèse.

Deuxièmement Fréchet dit que ces lemmes constituent « une série de propositions intéressantes en elles-mêmes ». Ils ont donc une utilité qui dépasse l'établissement du théorème. Nous avons qualifié de *généralité synthétique* un type de généralité que nous observons dans la thèse de Fréchet et qui émerge dans un *à-plat* que Fréchet y organise. Nous avons précisé ce type de généralité par contraste grâce à des situations où l'organisation des idées faisait naître un type de généralité que nous avons qualifié de *généralité constructive*. La remarque de Fréchet permet de préciser l'agencement de ses idées et de mieux appréhender comment *à-plat* et *structure* ne sont pas antagonistes et permettent l'émergence d'un certain type de généralité.

Dans ce texte de Fréchet la notion de *structure* peut s'entendre de deux façons.

D'un côté les classes (L), et plus tard les classes plus restrictives que Fréchet va introduire, déterminent des *structures* en un sens que nous qualifierions désormais de topologique. Nous les nommerons ici *structures ensemblistes*. La première partie de la thèse de Fréchet donne des propriétés des classes (L)⁸⁴ ce qui introduit une vision *réflexive* qui est important dans notre caractérisation de la notion de *structure ensembliste* telle qu'elle apparaît chez Fréchet. Nous verrons à plusieurs endroits que cette notion de point de vue *réflexif* permet de distinguer plusieurs temps et plusieurs aspects dans les textes mathématiques.

D'autre part si l'on regarde la théorie générale que développe Fréchet, théorie des opérations sur un ensemble d'une classe (L), on peut y voir un nouveau type de structure qui ne peut pas être réduit à la notion de *structure ensembliste* que nous venons de mentionner. C'est sur ce plan que la notion de *structure* est un élément fondamental de l'*à-plat* qui nous permet ici de décrire le travail de Fréchet en termes de généralité. L'organisation des idées met en jeu divers aspects que nous avons illustrés et les lemmes jouent ici un rôle particulier. La remarque de Fréchet sur les lemmes introduit en effet une nouvelle vision réflexive, mais cette fois-ci sa visée est la théorie des opérations sur les ensembles d'une classe (L). D'autre part les lemmes sont systématiquement des énoncés qui donnent des implications. Ils apparaissent donc comme des outils attachés à la théorie et qui fondent sa capacité à se développer grâce à des lemmes de base en plus des éléments qui constituent ce que nous avons qualifiés de *structure ensembliste*. C'est l'ensemble de ces éléments et leur organisation qui établit ici la notion de *théorie* et détermine le type de généralité qui est à l'œuvre dans ce texte.

Pour faire écho à la notion d'« activité théorique » introduite par Marion Vorms⁸⁵ nous appellerons *structure théorique* le type d'organisation que nous avons tenté de préciser dans ce texte de Fréchet et qui met en jeu des éléments

⁸⁴Il faudrait parler avec Fréchet d'« ensembles d'éléments d'une classe (L) », ce sont sur eux plutôt que directement sur la classe (L) que portent les énoncés de Fréchet.

⁸⁵« Cette conception repose sur la conviction qu'il y a quelque chose, dans ce que l'on appelle couramment 'théorie', qui résiste à toute entreprise formalisatrice, et à quoi seule une approche prêtant attention à la forme particulière sous laquelle les hypothèses théoriques sont construites et utilisées dans des raisonnements permet de rendre justice. » (cf. [Vor11, p.169])

qui ne peuvent pas se réduire à des *structures ensemblistes* ou qui ne peuvent pas s'appréhender dans une « entreprise formalisatrice » comme le dit Vorms.

Enfin cette partie concernant les classes (L) se termine par une remarque qui montre que Fréchet est à la recherche du meilleur contexte général pour développer une théorie où l'on peut généraliser un certain nombre de résultats établis quand on a à disposition des intervalles de \mathbb{R} ⁸⁶ :

REMARQUE . —Nous avons vu dans ce qui précède comment on peut arriver à généraliser certains théorèmes connus dans les classes si étendues que nous avons appelées classes (L). Néanmoins le petit nombre des hypothèses que nous avons faites sur les classes (L) ne nous permettra pas de poursuivre cette extension. On s'aperçoit bien vite que certaines propositions fondamentales de la théorie des ensembles linéaires et des fonctions continues, dont l'énoncé mis sous forme convenable peut garder un sens pour les classes (L), que ces propositions, dis-je, ne sont plus vraies pour toute classe (L). Par là-même échoue toute tentative de généralisation des propositions ultérieures auxquelles elles servent de base.

L'exemple que cite Fréchet afin d'illustrer l'incapacité à généraliser certaines propositions dans les classes (L) concerne les ensembles dérivés qui ne sont pas nécessairement fermés dans les ensembles issus des classes (L). Il s'appuie sur des suites de fonctions continues et les travaux de Baire qui montre qu'« il existe des fonctions qui sont limites de fonctions continues de classe 1, sans appartenir à cette classe. »⁸⁷

RECHERCHE DE LA « CLASSE » LA MIEUX ADAPTÉE. Dans un second chapitre de sa thèse Fréchet va explorer un certain nombre de « classes » plus restrictives que la classe (L) afin de trouver le contexte le plus adapté à l'élaboration d'une théorie générale. Il ne s'agit pas de laisser de côté les classes (L) ou d'autres classes intermédiaires. Elles jouent un rôle central dans l'élaboration de la théorie et permettent de mettre en évidence comment chaque hypothèse intervient dans le développement de la théorie. En outre l'organisation que propose Fréchet, avec une longue partie réservée à la classe la plus générale (L), offre un paysage théorique *ouvert* où on peut concevoir un certain nombre de points de bifurcation. La théorie développée par Fréchet est aussi générale en ce sens qu'elle permet de s'adapter à des situations où les conditions les plus restrictives ne sont pas réalisées. Nous retrouverons ce type de construction chez Banach un peu plus tard.

Nous ne détaillons pas cette partie de la thèse mais nous faisons ressortir quelques aspects qui nous permettront d'établir certains rapprochements ou parfois des différences avec le travail de Banach.

Le principe adopté par Fréchet dans sa recherche de la classe la mieux adaptée pour concevoir une généralisation des « ensembles linéaires » est qu'elle doit satisfaire à trois conditions⁸⁸ :

⁸⁶[Fré06, p.15]

⁸⁷cf. [Bai05, p.126 N° 81]

⁸⁸[Fré06, p.17]

1) ces restrictions devraient pouvoir s'énoncer indépendamment de la nature des éléments considérés; 2) elles devraient être satisfaites par les classes d'éléments qui interviennent le plus fréquemment dans les applications; 3) elles devraient fournir la généralisation cherchée des théorèmes sur les ensembles linéaires et les fonctions continues.

LES CLASSES (V). Le troisième point est essentiel dans les choix qui suivent et le premier résultat que Fréchet souhaite pouvoir conserver dans le cadre général est celui qu'il a mentionné plus haut; les ensembles dérivés doivent être fermés.

Fréchet remarque que cette condition est restrictive car des ensembles communs de fonctions qui sont des classes (L) ne possèdent pas cette propriété. Pourtant il rajoute « la propriété précédente est compatible avec un grand nombre de définitions de la limite (particulièrement les limites dites uniformes). »

Le choix qui en résulte montre à nouveau comment Fréchet s'éloigne du choix *constructif* pour proposer un schéma d'abstraction dont nous lisons la *généralité synthétique*. En effet, au lieu de rajouter simplement cette propriété pour créer une classe plus restrictive, Fréchet isole dans « les cas connus » les conditions qui « entraînent » la propriété sous sa forme générale et suivant les principes énoncés plus haut.

Fréchet introduit alors la notion de « voisinage » et de « classe (V) » sous la forme suivante⁸⁹ :

Considérons une classe (V) d'éléments de nature quelconque, mais tels qu'on sache discerner si deux d'entre eux sont ou non identiques et tels, de plus, qu'à deux quelconques d'entre eux A, B , on puisse faire correspondre un nombre $(A, B) = (B, A) \geq 0$ qui jouit des deux propriétés suivantes : 1) La condition nécessaire et suffisante pour que (A, B) soit nul est que A et B soient identiques. 2) Il existe une fonction positive bien déterminée $f(\epsilon)$ tendant vers zéro avec ϵ , telle que les inégalités $(A, B) \leq \epsilon, (B, C) \leq \epsilon$ entraînent $(A, C) \leq f(\epsilon)$, quels que soient les éléments A, B, C . Autrement dit, il suffit que (A, B) et (B, C) soient petits pour qu'il en soit de même de (A, C) . Nous appellerons *voisinage* de A et de B le nombre (A, B) .

Fréchet introduit ici la notion de « voisinage », qui, contrairement à l'extension de ce mot que nous avons trouvé chez Volterra lorsqu'il introduisait des « voisinages » tubulaires, s'appuie ici sur une notion de quantité réelle positive.

Si on ne lit que ces lignes on peut penser que les classes (V) sont obtenues par adjonction d'une propriété supplémentaire aux classes (L). Mais Fréchet définit en fait la notion de convergence en la rendant dépendante de la notion de voisinage⁹⁰ :

⁸⁹[Fré06, p.18]

⁹⁰[Fré04b, p.18]

Ceci étant, nous pourrions dire qu'une suite d'éléments de la classe $(V) : A_1, A_2, \dots$ tend vers un élément A , si le voisinage (A_n, A) tend vers zéro avec $\frac{1}{n}$. [...]

De plus, cette définition de la limite satisfait bien aux conditions I et II que nous avons imposées en général à toute définition de la limite [...] et cela grâce aux conditions 1) et 2) imposées à la définition du voisinage.

On voit donc que la restriction imposée pour définir les classes (V) porte sur le choix d'une façon particulière de définir la convergence. Les classes (V) sont des classes (L) particulières car le mode de convergence y est particulier.

Le premier théorème établi dans le cadre des classes (V) est que les ensembles dérivés sont fermés, ce qui a aussi pour conséquence, d'après les remarques antérieures de Fréchet, de démontrer que les classes (E) ne sont pas toutes des classes (V) .

Enfin, un aspect important qui découle de la considération des classes (V) est le retour d'éléments de géométrie dans le texte de Fréchet. Après avoir démontré plusieurs résultats sur les classes (V) celui-ci introduit la notion de « sphéroïde »⁹¹ :

Appelons sphéroïde de centre A et rayon ρ l'ensemble de tous les éléments B tels que l'on ait : $(A, B) < \rho$. Nous dirons qu'un élément C est intérieur à ce sphéroïde, si l'on a : $(A, C) < \rho$.

En s'appuyant sur une démonstration de Zoratti qu'il généralise, Fréchet peut alors démontrer le théorème devenu classique suivant⁹² :

THÉORÈME - Si un ensemble fermé F appartient à un ensemble compact E formé d'éléments d'une classe (V) , on obtiendra F en supprimant de E les éléments INTÉRIEURS à chacun des sphéroïdes d'un certain ensemble dénombrable de sphéroïdes.

La notion de sphéroïde permet aussi à Fréchet de s'approcher d'une caractérisation des compacts en termes d'ensembles « limités », c'est-à-dire tels que le voisinage de deux points quelconques est majoré par un nombre donné.

Enfin, en continuant à construire une classe abstraite par analogie avec les « ensembles linéaires », Fréchet introduit d'un coup les notions de « classe qui admet une généralisation du théorème de Cauchy » (les suites de Cauchy au sens du voisinage sont convergentes), de « classe séparable » (elle coïncide avec l'ensemble dérivé d'une de ses parties dénombrables) et la notion de « classe parfaite » (tout élément est élément limite, et en particulier « une classe séparable est nécessairement parfaite » comme le souligne Fréchet).

Ces définitions permettent à Fréchet d'étudier les classes (V) qu'il nomme « normales » : elles sont « parfaites, séparables et [admettent] une généralisation du théorème de CAUCHY ».

⁹¹[Fré06, p.21]

⁹²[Fré06, p.21]

NOUVELLE DÉFINITION DE LA CONTINUITÉ. La notion de classe (V) permet à Fréchet de donner un théorème qui caractérise les « opérations continues » définies sur un ensemble issu d'une telle classe⁹³ :

THÉORÈME. - Considérons une opération U définie dans un ensemble E formé d'éléments d'une classe (V) . La condition nécessaire et suffisante pour que U soit une opération continue dans E est que, si A est un élément quelconque commun à E et à E' , on puisse faire correspondre à tout nombre $\epsilon > 0$ un nombre η_A tel que l'inégalité $(A, B) < \eta_A$ entraîne $|U(A) - U(B)| < \epsilon$ pour tout élément B de E .

À partir de ce théorème Fréchet élabore une nouvelle définition de la continuité des opérations lorsque l'on est dans une classe (V) . Cette définition amorce l'ébauche d'une théorie qui pourrait se développer à partir des ensembles issus d'une classe (V) . Elle montre que Fréchet n'a pas tout à fait défini dans sa thèse un cadre idéal pour le calcul fonctionnel, mais qu'il cherche le meilleur contexte qui préserve l'analogie avec une théorie des fonctions définies sur des sous-ensembles de \mathbb{R} .

Les classes (V) fournissent un contexte propice pour que l'analogie soit assez étendue. En particulier dans ce cadre Fréchet montre deux théorèmes qui précisent l'étendue de la généralisation.

Le premier s'appuie sur une définition de l'uniforme continuité à partir de la notion de voisinage⁹⁴ :

Nous dirons qu'une opération U est uniformément continue dans un ensemble E formé d'éléments d'une classe (V) si, étant donné un nombre positif ϵ , on peut choisir un nombre positif η tel que, pour deux éléments quelconques de E : A, B , l'inégalité :

$$(A, B) < \eta \quad \text{entraîne} \quad |U(A) - U(B)| < \epsilon.$$

Cette notion revisitée d'uniforme continuité permet alors d'énoncer le théorème suivant⁹⁵ :

THÉORÈME. - Toute opération continue dans un ensemble EXTRÉMAL E formé d'éléments d'une classe (V) est uniformément continue dans E .

Enfin une caractérisation des familles compactes d'opérations est donnée sous la forme suivante⁹⁶ :

Pour que des opérations continues dans un même ensemble extrémal E , formé d'éléments d'une classe (V) séparable, forment une famille compacte \mathfrak{F} , il faut et il suffit que les opérations de \mathfrak{F} soient, en tout élément de E , également continues et bornées.

Enfin Fréchet finit cette première partie en restreignant une dernière fois le contexte dans lequel il travaille afin de pouvoir donner le pendant d'un théorème phare de la théorie des fonctions ordinaires sur \mathbb{R} . Pour cela il faut

⁹³[Fré06, p.28]

⁹⁴[Fré06, p.28]

⁹⁵[Fré06, p.29]

⁹⁶[Fré06, p.30]

se souvenir que la définition de voisinage introduit une condition qui est une sorte de *contrôle triangulaire* : $(A, B) < \epsilon$ et $(B, C) < \epsilon$ implique $(A, C) < f(\epsilon)$ où $\lim_{\epsilon \rightarrow 0} f(\epsilon) = 0$. Mais cette notion l'empêche d'obtenir un résultat important qui est la réciproque du théorème qui, dans une classe (L) , démontre que les opérations continues sur un ensemble extrémal sont bornées et atteignent leurs limites supérieures et inférieures.

En remplaçant ce que nous avons nommé la condition de *contrôle triangulaire* par une inégalité triangulaire Fréchet conçoit des voisinages particuliers qu'il nomme « écarts » (qui correspond à la notion de distance de nos jours). Les classes correspondantes sont notées (E) et Fréchet énonce le théorème qui suit⁹⁷ :

THÉORÈME. - La condition nécessaire et suffisante pour que toute opération continue dans un ensemble E d'éléments d'une classe (E) , 1) soit bornée dans cet ensemble, 2) y atteigne sa limite supérieure, est que cet ensemble E soit extrémal.

Nous n'analysons pas ici la démonstration⁹⁸ qui nous éloignerait trop du type d'illustration que nous souhaitons mener dans notre étude.

Ce dernier résultat montre que les ensembles extrémaux définis dans les classes (E) jouent précisément le rôle des intervalles compacts dans le cas des « ensembles linéaires ». La version abstraite que donne Fréchet lui permet non seulement de concevoir l'analogie avec les fonctions ordinaires de la variable réelle, mais aussi de réécrire des problèmes plus généraux du calcul des variations. En particulier pour lui le théorème précédent reformule ce que Hilbert avait énoncé « d'une façon un peu vague » en ces mots⁹⁹ :

Tout problème du Calcul des variations possède une solution, pourvu que certaines hypothèses restrictives convenablement choisies relatives à la nature des conditions limitatives données soient remplies, et, nécessairement aussi, pourvu que ce que l'on entend par le mot *solution* éprouve une généralisation conforme au sens, à la nature des choses.

APPLICATION DE LA THÉORIE GÉNÉRALE. La deuxième partie de la thèse de Fréchet donne une longue série d'applications de la théorie générale qu'il a développée dans la première partie à des cas particuliers. Les titres distinguent alors à chaque fois un couple formé d'un ensemble d'une certaine classe et un type d'opération. Ainsi Fréchet revisite les couples suivants : (ensembles linéaires/fonctions ordinaires d'une variable réelle) ; (ensembles de fonctions continues sur un intervalle donné/fonctionnelles) ; (espaces de dimension infinie/fonctions) ; (courbes/fonctions de lignes), etc.

Nous ne nous attardons pas sur les détails de cette partie qui nécessiterait une étude précise et complète. Nous mettons cependant en avant ce que

⁹⁷[Fré06, p.31]

⁹⁸Fréchet lui-même « attire l'attention sur le mode de démonstration » de ce théorème. On trouve là une démonstration qui met en œuvre un raisonnement par cas et dont il serait intéressant de faire ressortir comment Fréchet y utilise la notion d'écart et pourquoi la notion de voisinage ne lui suffit pas.

⁹⁹[Hil00, p.340]

Fréchet nomme « méthode [...] pour appliquer la théorie générale aux cas particuliers obtenus en spécifiant la nature des éléments. » Il donne cette méthode en introduction de la seconde partie dédiée aux applications de la théorie générale¹⁰⁰ :

Dans tous les exemples qui suivront, nous commencerons par définir la classe d'éléments sur laquelle nous aurons à opérer ; en particulier nous donnerons une règle permettant de discerner les éléments distincts, ce qui n'est pas toujours aussi simple qu'on pourrait le croire. Puis nous montrerons qu'on peut considérer la définition ordinaire de la limite de ces éléments comme rentrant dans la définition que nous avons donnée à l'aide de l'écart (cas particulier du voisinage). Nous prouverons enfin que la classe (E), ainsi obtenue, est normale. Ceci fait, il n'y aura plus qu'à répéter les théorèmes généraux de la PREMIÈRE PARTIE énoncés dans le langage qui correspond à la classe considérée. Enfin, nous donnerons dans chaque cas une condition nécessaire et suffisante pour qu'un ensemble soit compact. Cela revient à remplacer la définition générale de l'ensemble compact par une définition d'une apparence moins abstraite, mais particulière à chaque cas.

En fait dans chaque cas Fréchet montre que l'on a affaire à une classe (E) en précisant ce qui joue le rôle d'« écart » (il y a parfois plusieurs choix possibles). Il s'ensuit deux types de résultats issus de la théorie générale : des théorèmes sur les éléments et les ensembles de la classe, puis des résultats sur les « opérations » qui agissent sur ces éléments (et à valeurs dans \mathbb{R}).

Une fois les conditions d'applicabilité de la théorie prouvées, Fréchet exprime les résultats généraux dans les termes de la situation particulière qu'il envisage, redisant ainsi les théorèmes de la théorie d'une façon adaptée au différents cas.

Nous pourrions par la suite comparer cette « méthode » avec celle employé par Banach dans sa thèse. L'étude que nous venons de mener de la thèse de Fréchet nous permettra une comparaison précise avec celle de Banach un peu plus d'une décennie plus tard. Nous pourrions observer une grande proximité de méthode entre les deux jeunes mathématiciens qui éclairera *a posteriori* un aspect général qui découle de l'adoption du type de « méthode » que Fréchet choisit ici.

4.4 LES OPÉRATIONS LINÉAIRES

Dans cette section nous suivons une seconde voie établie par Fréchet et qui prend sa source dans une lecture par le mathématicien français de certains aspects du travail de Volterra sur les fonctions de lignes. L'approche est aussi intimement liée à un point de vue promu par Hadamard dont on a vu qu'il assimilait lui-aussi des idées de Volterra pour aborder des questions de variations essentiellement et dans une quête de son côté d'une vision générale.

Bien que l'on rattache souvent la notion de structure à la thèse de Fréchet et en particulier aux *structures ensemblistes* que l'on peut y lire, le travail mené

¹⁰⁰[Fré06, p.34]

en parallèle et suivant une autre ligne directrice entre 1904 et 1907 laisse aussi apparaître des éléments structurels que nous allons maintenant mettre en évidence.

Pendant cette période Fréchet publie une série de trois notes intitulées « Sur les opérations linéaires »¹⁰¹ qui ne font pas référence directement aux travaux sur les classes (L) que nous venons de voir. Afin de saisir la spécificité de ces articles nous les mettrons en perspective avec quelques textes antérieurs d'Hadamard ou de Pincherle.

Cette étude de Fréchet s'ouvre sur une relecture du travail d'Hadamard que celui-ci avait présenté à l'Académie des Sciences à Paris en 1903 et que nous avons déjà considéré au début de ce chapitre.

Dans un premier temps nous montrerons comment Fréchet réinterprète ce travail et en donne un exposé nouveau, enrichi de plusieurs questions. Nous étudierons aussi les éléments essentiels que Fréchet convoque et sélectionne de travaux antérieurs publiés par divers autres auteurs et qu'il adapte à son propre point de vue. Cette première partie permettra d'analyser le regard particulier qu'il porte sur des éléments et des formes qui joueront ensuite un rôle majeur dans sa progression.

Dans un deuxième temps nous étudierons les processus multiples qui ont permis à Fréchet de généraliser et d'étendre le résultat initial d'Hadamard à des situations qui n'avaient pas été envisagées jusque-là. Cette partie mettra en évidence les différentes stratégies de sélection, d'appropriation et d'utilisation de matériaux extraits de la littérature récente de son époque. Nous verrons comment l'approche nouvelle de Fréchet et des éléments que nous aurons mis en évidence dans la première partie jouent un rôle fondamental dans cette première phase de généralisation.

Enfin, nous montrerons comment l'agencement nouveau de certains concepts (ou plus largement l'émergence d'éléments structurels en un sens à préciser à nouveau), est intimement lié à une organisation particulière de l'exposé mathématique : celui-ci met en jeu une véritable dynamique du discours qui participe de la progression vers le général. C'est dans une lecture des mouvements internes du texte que nous trouverons des éléments pour caractériser de façon plus précise la notion de structure dans un champ en passe de se constituer.

4.4.1 UN THÉORÈME D'HADAMARD COMME OBJET CENTRAL D'ÉTUDE

Cette étude de Fréchet dont les trois volets partagent le même titre, « Sur les Opérations Linéaires », introduit cette dernière notion dès le début de la première note par le biais d'une définition attribuée à Hadamard¹⁰² :

Rappelons d'abord quelques définitions. Nous dirons qu'une opération est définie si l'on fait correspondre un nombre réel déterminé et fini U_f à toute fonction $f(x)$ réelle et continue entre deux nombres fixes a, b . Nous appellerons avec M. Hadamard

¹⁰¹[Fré04c; Fré05; Fré07]

¹⁰²[Fré04c, p.493]

opération linéaire toute opération qui jouit des deux propriétés suivantes :

1. elle est distributive, c'est-à-dire que si f_1 et f_2 sont deux fonctions continues entre a et b , on a toujours :

$$U_{f_1+f_2} = U_{f_1} + U_{f_2}.$$

2. elle est continue, c'est-à-dire que U_{f_1} tend vers U_{f_2} lorsque la fonction f_1 tend uniformément vers f_2 entre a et b .

Cette définition appelle quelques commentaires qui permettent de mieux cerner le contexte choisi par Fréchet pour développer sa réflexion.

On peut tout d'abord la contraster avec celle que donnent Amaldi et Pincherle dans leur livre *Le operazioni distributive* de 1901¹⁰³ :

Tra gli elementi di due sistemi \mathcal{S} ed \mathcal{S}' sia stabilita una corrispondenza che coordini uno o più elementi di \mathcal{S} ad ogni elemento di \mathcal{S}' . [...] Se α e β sono elementi qualsivogliano dello spazio lineare \mathcal{S} ed A una operazione applicabile a questi enti, se dice che A gode della proprietà distributiva quando sia

$$(4) \quad A(\alpha + \beta) = A(\alpha) + A(\beta).$$

Risulta da questa che se [...] c è un numero razionale,

$$(5) \quad A(c\alpha) = cA(\alpha).$$

Noi ammetteremo que l'uguaglianza (5) sussista per ogni numero c , reale o complesso, e:

Chiameremo operazioni distributive quelle che godono della proprietà espresse dalle equazioni (4) e (5).

Contrairement à ce texte antérieur, Fréchet n'envisage comme résultat de l'opération que des nombres réels, tout comme l'avait fait Hadamard avant lui. De plus, il ne considère pas dans sa définition ce qui pourrait correspondre au système (*insieme o spazio lineare*) \mathcal{S} de Pincherle et Amaldi : il fait référence à « toute fonction $f(x)$ réelle et continue entre deux nombres fixes a, b », et il se servira lorsqu'il en aura besoin de l'addition de deux fonctions ou de la multiplication par un nombre réel. Pour leur part, Pincherle et Amaldi développent, avant la définition que nous avons citée, un chapitre dédié à la notion de système linéaire et aux opérations (addition et multiplication par un nombre) qui permettent de définir un tel système. Fréchet n'isole dans sa définition initiale aucun objet qui correspondrait à la collection des fonctions qu'il utilise assorties des opérations : il y a là un contexte usuel sur lequel il ne lui semble pas nécessaire d'insister.

Cette remarque est fondamentale ici et montre à nouveau que la présence d'opérations algébriques ne s'accompagne pas forcément de lignes de partage claires qui distingueraient des éléments de nature différentes et permettraient de conclure à l'existence d'une *structure ensembliste* selon le vocabulaire que nous avons introduit plus avant.

¹⁰³[AP01, p.17-18]

D'autre part, et nous verrons par la suite qu'il y a ici un enjeu essentiel, la définition retenue par Fréchet se distingue par sa référence à la continuité (une opération¹⁰⁴ U_f est continue si U_{f_1} converge vers U_{f_2} lorsque la fonction f_1 converge uniformément vers la fonction f_2) et l'absence d'une propriété qui correspondrait au point (5) de la définition de Pincherle et Amaldi (et son extension à tous les nombres réels). Nous verrons par la suite comment ce choix et le lien établi entre les deux notions participe de l'organisation générale des idées développées par Fréchet et contribue à l'émergence de certains aspects structurels. Nous indiquons pour l'instant ce contraste au niveau des définitions dans le but de préciser le contexte précis qui constitue le point de départ de Fréchet dans son étude des opérations linéaires et nous nous concentrerons essentiellement sur la progression des idées dans ces trois notes.

C'est dans ce contexte, tout au début de son premier article, que Fréchet propose un théorème d'Hadamard comme objet central d'investigation et prétexte à son travail¹⁰⁵ :

M. Hadamard a démontré en se basant sur l'étude de l'expression :

$$\lim_{\mu=\infty} \int_a^b f(u) \mu e^{-\mu^2(u-x)^2} du,$$

que toute opération linéaire peut être représentée sous la forme :

$$U_f = \lim_{\mu=\infty} \int_a^b f(x) K_n(x) dx,$$

où $K_n(x)$ est une fonction continue de x entre a et b .

Comme le précise Fréchet en note de bas de page, ce théorème fait référence à la communication d'Hadamard à l'Académie des Sciences en 1903 dont nous avons sommairement rappelé les idées page 150. Nous y avons souligné qu'un des buts d'Hadamard était de se libérer de l'hypothèse contraignante d'analyticité des fonctions en jeu. Cependant, malgré l'avancée importante que représente le résultat précédent, Hadamard souligne que dans le cas des fonctions seulement continues il est « un peu moins aisé d'exprimer que U est de la forme [donnée par Volterra, Pincherle ou Boulet dans le cas des fonctions analytique] », à savoir une représentation en série qui fait intervenir une somme d'opérations linéaires simples.

C'est essentiellement ce résultat d'Hadamard et la difficulté qui subsiste pour faire le lien avec les développements obtenus précédemment pour les fonctions analytiques qui constitue le point de départ du travail de Fréchet : sitôt le théorème d'Hadamard énoncé, il mentionne le cas des fonctions analytiques et le développement qu'elles permettent d'obtenir pour l'opération linéaire dans un paragraphe intitulé « Développements en séries d'opérations linéaires

¹⁰⁴ La notation qui ne distingue pas clairement l'opération de l'objet à laquelle elle s'applique est aussi contraire à la présentation adoptée par quelques prédécesseurs, dont Pincherle et Amaldi cités plus haut. Ce choix éclaire le processus d'abstraction qui diffère dans chacun des cas dans le passage de l'idée de fonction à celle d'opération. Nous avons eu l'occasion d'illustrer à plusieurs reprises cette diversité.

¹⁰⁵[Fré04c, p.493]

canoniques. »¹⁰⁶ Ainsi ces trois notes successives de Fréchet vont-elles tour à tour reprendre, prolonger et généraliser de diverses façons que nous allons étudier, cette avancée initiée par Hadamard.

4.4.2 GÉNÉRALISATION PAR EXTENSION

Après avoir exposé le point de départ de sa réflexion, Fréchet place au premier plan de son étude le désir de généraliser ce résultat initial —le théorème d’Hadamard— et le mouvement va s’opérer suivant plusieurs modalités. La première stratégie mise en place, qui occupe l’essentiel de la première note, consiste à reprendre la technique utilisée par Pincherle dans le cas des fonctions analytiques et à *étendre le domaine de validité* de la preuve du résultat. Fréchet rappelle la technique générale mise en œuvre par ses prédécesseurs avant d’en donner une première analyse et de l’appliquer à des cas nouveaux. Nous esquissons ici l’enchaînement des idées auxquelles Fréchet fait référence en citant le résultat précis de Pincherle¹⁰⁷ :

M. PINCHERLE a déjà démontré que l’on pouvait écrire :

$$U_f = u_0 f(c) + u_1 f'(c) + \dots + u_n f_c^{(n)} + \dots ,$$

c étant un nombre fixe compris entre a et b , en prenant :

$$u_0 = U_1, u_1 = U_{(x-c)}, \dots, u_n = \frac{1}{n!} U_{(x-c)^n}, \dots$$

Cependant, dès ce rappel, Fréchet souligne à son tour le fait que cette méthode dépend fondamentalement non seulement de la nature de la fonction f mais aussi du type de propriété que l’on attribue à l’opération¹⁰⁸ :

[...] un tel développement ne peut s’écrire que si $f(x)$ est dérivable indéfiniment en $x = c$ et encore cela ne suffit pas pour assurer la convergence du développement vers U_f .

Plus précisément, Fréchet fait ici (explicitement dans une note de bas de page) référence au livre d’Amaldi et Pincherle de 1901. Celui-ci contient des idées qui avaient été développées par Pincherle dès 1897 dans son *Mémoire sur le calcul fonctionnel distributif*¹⁰⁹. Dans ce travail Pincherle étudie les opérations linéaires agissant sur les fonctions φ représentées par une série entière avec convergence uniforme. En appliquant terme à terme l’opération linéaire qu’il note A , et en utilisant simplement la distributivité, Pincherle obtient alors *formellement*¹¹⁰ un développement en série du résultat $A(\varphi)$. Pour montrer que cette série formelle permet effectivement de définir l’action de A , Pincherle doit ajouter quelques conditions sur les opérations qu’il considère¹¹¹ :

¹⁰⁶[Fré04c, p.493]

¹⁰⁷[Fré04c, p.494]

¹⁰⁸[Fré04c, p.494]

¹⁰⁹[Pin97]

¹¹⁰Fréchet utilise de nombreuses fois le mot « forme » lorsqu’il fait référence aux résultats de Pincherle ou d’Hadamard. Nous verrons comment ce mot se fait l’écho d’une certaine façon de penser dans les trois notes que nous étudions ici.

¹¹¹[Pin97, p.349]

48. La propriété distributive dont jouit une opération A peut naturellement s'étendre, sous certaines restrictions, au cas où l'on exécute cette opération sur une série infinie de fonctions. Cette extension est admissible, par exemple, si l'on assujettit A aux conditions suivantes :

- a) d'être à détermination unique¹¹²,
- b) de faire correspondre aux fonctions α qui appartiennent à un certain cercle (r) ¹¹³, des fonctions qui appartiennent à un cercle (r') ,
- c) de faire correspondre à un nombre positif arbitraire g un nombre positif h tel qu'à une fonction α appartenant à (r) et dont la valeur absolue maxima en (r) est moindre que h , correspond une fonction $A(\alpha)$ dont la valeur absolue maxima en (r') soit moindre que g ¹¹⁴.

49. Sous ces conditions, on peut en effet démontrer le théorème suivant :

Soit un système de fonctions $\alpha_1, \alpha_2, \dots, \alpha_n, \dots$ appartenant au cercle (r) et telles que la série

$$\varphi = \sum_{n=1}^{\infty} \alpha_n$$

soit convergente uniformément dans ce cercle, la circonférence comprise ; soit A une opération jouissant des propriétés énoncées au § précédent ; on aura :

$$A \left(\sum_1^{\infty} \alpha_n \right) = \sum_1^{\infty} A(\alpha_n)$$

En particulier en appliquant cette idée lorsque la fonction φ est représentée par son développement de Taylor, on obtient formellement la formule citée par Fréchet. Pincherle commente d'ailleurs la validité de cette « série fonctionnelle de Taylor »¹¹⁵ comme il la nomme¹¹⁶ :

62. Le développement trouvé n'a [en général] qu'une valeur formelle. On peut, dans les cas particuliers, discuter de la validité [de la formule] comme on le fait en calcul différentiel pour la série de Taylor ordinaire, en examinant le reste.

Fréchet relit cette stratégie et sélectionne dans ce schéma deux éléments clés que sont : 1) la décomposition de φ en une série de fonctions qui converge uniformément, 2) l'application (justifiée) de l'opération linéaire terme à terme à cette série. Le but est clairement annoncé par l'auteur¹¹⁷ :

¹¹²[NDA] Il faut entendre ici que le résultat de l'opération est unique. Fréchet parle plus volontiers d'opération « uniforme ».

¹¹³[NDA] Pincherle précise que cela signifie que l'on considère des séries de puissances entières de x , ordonnées par puissances croissantes, dont le rayon de convergence est supérieur à r .

¹¹⁴On peut comparer ce point avec plusieurs façons d'écrire la continuité chez différents auteurs. On a ici une formulation qui est adaptée aux fonctions analytiques, mais Pincherle choisit de ne pas avoir recours aux suites de fonctions pour exprimer la continuité.

¹¹⁵[Pin97, p.355]

¹¹⁶[Pin97, p.355]

¹¹⁷[Fré04c, p.494]

Je vais montrer d'abord qu'on peut obtenir un développement de la forme $[U_f = u_0U_f^{(0)} + u_1U_f^{(1)} + \dots + U_f^{(n)} + \dots]$ ayant un champ de validité beaucoup plus étendu que le développement $[U_f = u_0f(c) + u_1f'(c) + \dots]$ puisqu'il n'exige même pas que $f(x)$ soit dérivable.

En fait son étude repose ici sur la considération des fonctions paires obtenues par symétrie des « fonctions $f(x)$ continues dans $(0, \pi)$ et n'ayant qu'un nombre limité de maxima et de minima dans cet intervalle. » La contrainte sur les fonctions retenues est issue des travaux initiés par Dirichlet en 1829¹¹⁸, et il sera ensuite montré que sous ces conditions la convergence de la série de Fourier est en fait uniforme¹¹⁹. C'est cette propriété que Fréchet exploite donc dans un premier temps en utilisant la distributivité et la continuité comme il les a énoncées¹²⁰ :

Si l'on développe en série de Fourier la fonction (satisfaisant aux conditions de Dirichlet) qui est égale à $f(x)$ entre 0 et π et à $f(-x)$ entre 0 et π , on aura entre 0 et π :

$$(*) \quad f(x) = a_0 + a_1 \cos x + a_2 \cos 2x + \dots + a_n \cos nx + \dots,$$

avec

$$a_0 = \frac{1}{\pi} \int_0^\pi f(x) dx, \quad a_n = \frac{2}{\pi} \int_0^\pi f(x) \cos nxdx \quad (n = 1, 2, \dots).$$

Puisque la série (*) converge uniformément dans $(0, \pi)$, on peut lui appliquer terme à terme toute opération linéaire U_f et l'on aura :

$$U_f = \frac{\pi}{2} (2a_0u_0 + a_1u_1 + \dots + a_nu_n + \dots),$$

En posant cette fois :

$$u_0 = \frac{1}{\pi} U_1, \quad u_1 = \frac{2}{\pi} U_{\cos x}, \dots, \quad u_n = \frac{2}{\pi} U_{\cos nx}, \dots$$

De plus, comme l'avait aussi fait Pincherle dans le contexte que nous avons rappelé plus haut, Fréchet relit le résultat obtenu en termes d'analogie avec le développement de Taylor pour les fonctions. Cette interprétation va elle aussi conditionner la stratégie mise en place pour dégager une vision générale. Pour Fréchet il s'agit de mettre en évidence des « opérations linéaires canoniques »¹²¹ qui permettraient de donner une écriture simple et commune à toutes les opérations linéaires¹²² :

¹¹⁸Dans son article de 1829 Dirichlet ne considère que la convergence ponctuelle de la série de Fourier (cf. [Dir29]). Il faudra attendre quelques années pour voir apparaître la nécessité de distinguer plusieurs types de convergence et en particulier la convergence uniforme. On peut consulter le livre de Bottazzini *The higher calculus : a history of real and complex analysis from Euler to Weierstrass* [Bot86] pour un panorama général du développement de ces problèmes.

¹¹⁹Voir l'article de Heine « Ueber trigonometrische Reihen » ([Hei70]) par exemple

¹²⁰[Fré04c, p.494]

¹²¹[Fré04c, p.493]

¹²²[Fré04c, p.494]

Il est assez naturel de chercher à établir pour les opérations linéaires un développement analogue à celui de Taylor pour les fonctions ordinaires¹²³. Il s'agirait de trouver une suite d'opérations linéaires simples : $U_f^{(0)}, \dots, U_f^{(n)}, \dots$ choisies une fois pour toutes et telles que toute opération linéaire U_f puisse s'écrire sous la forme :

$$U_f = u_0 U_f^{(0)} + u_1 U_f^{(1)} + \dots + u_n U_f^{(n)} + \dots,$$

les constantes u_0, u_1, \dots variant seules avec l'opération U_f .

Ainsi Fréchet va-t-il systématiquement interpréter les développements en série de U_f en mettant en évidence des coefficients dépendant seulement de U (donnés par l'action sur les fonctions particulières intervenant dans le développement et qui ne dépendent pas de f) et des « opérations linéaires simples » ne dépendant pas de U (qui sont elles aussi construites à partir des fonctions intervenant dans le développement de f).

Dans l'exemple des fonctions satisfaisant aux conditions de Dirichlet la fonction f est développée suivant les fonctions $\cos nx$ ($n \in \mathbb{N}$) et l'opération U est décrite grâce à des coefficients dépendant de U (qui sont, à un coefficient près, le résultat de l'action de U sur les fonctions $\cos nx$) et d'opérations élémentaires de la forme « $U_f^{(0)} = \int_0^\pi f(x)dx, \dots, U_f^{(n)} = \int_0^\pi f(x) \cos nxdx, \dots$ »¹²⁴ Cette présentation sera encore utilisée par Fréchet dans les différentes extensions du résultat d'Hadamard.

Une nouvelle phase d'*extension* va permettre à Fréchet d'avancer beaucoup plus loin dans la généralisation du théorème d'Hadamard. Il s'agit cette fois encore d'une extension du domaine de validité de la stratégie initiale (développement en série de la fonction f et application de U terme à terme sous hypothèse de convergence uniforme), mais cette fois-ci l'extension se fait en assouplissant les contraintes de la preuve initiale : en particulier le développement représente dans ce nouveau cas la fonction f de façon bien moins évidente « par une série simplement indéterminée »¹²⁵ Cette réflexion, amorcée dès la première note de Fréchet, sera essentielle dans la progression vers une plus grande généralité. Elle est fondée sur les travaux de Fejér¹²⁶ qui utilise la convergence en moyenne (développée par Cesàro) afin d'étudier la convergence du développement en série de certaines fonctions. La méthode est déjà conforme ici à celle qui est désormais devenue classique : la fonction f est définie sur l'intervalle $(-\pi, \pi)$, elle est paire, continue, et Fréchet lui associe sa série de Fourier $\sum_{n=0}^{\infty} a_n \cos nx$ où $a_0 = \frac{1}{\pi} \int_0^\pi f(x)dx, \dots, a_n = \frac{2}{\pi} \int_0^\pi f(x) \cos nxdx$ $n \geq 1$. Enfin il rappelle comment on peut définir la convergence généralisée de la série associée¹²⁷ :

¹²³Nous avons vu que ce « naturel » s'inscrivait dans une longue histoire qui a fait émerger cette question. Le choix de la série de Taylor comme *forme* apte à représenter les opérations est intimement lié à la conception des opérations en étroit lien avec une notion de *fonction* dont on a vu qu'elle était à la source d'un mouvement de généralisation que nous avons en particulier illustré chez Volterra.

¹²⁴[Fré04c, p.495]

¹²⁵[Fré04c, p.495]

¹²⁶Voir [Fej03] par exemple.

¹²⁷[Fré04c, p.495]

Pour une telle série :

$$v_0, v_1, \dots, v_n, \dots,$$

de somme généralisée σ , nous écrivons :

$$\sigma \doteq v_0 + v_1 + \dots + v_n + \dots,$$

ou :

$$\sigma = \lim_{n=\infty} \text{gén} S_n$$

en posant

$$S_n = v_0 + v_1 + \dots + v_n,$$

ou

$$\sigma = \lim_{n=\infty} \sigma_n$$

avec

$$\sigma_n = \frac{S_0 + \dots + S_n}{n} = \frac{nv_0 + (n-1)v_1 + \dots + v_n}{n}.$$

Dans ce cadre-là, « M. FEJER a démontré que toute fonction périodique continue peut être représentée par la somme généralisée de son développement de Fourier, avec une convergence uniforme. »¹²⁸ Ainsi ce résultat permet à Fréchet de lier le développement de f et celui de U_f comme dans les cas précédents¹²⁹ :

[O]n aura ainsi :

$$f(x) \doteq a_0 + a_1 \cos x + \dots + a_n \cos nx + \dots,$$

$$U_f \doteq u_0 U_f^{(0)} + \dots + u_n U_f^{(n)} + \dots$$

On peut remarquer que ces deux expressions sont formellement très voisines de celles obtenues pour les fonctions dont la série de Taylor converge uniformément (au sens classique). Il y a pourtant un écart non négligeable et une réinterprétation subtile qui est opérée à travers le signe \doteq qui désigne ici la convergence de type Cesàro. Malgré l'opacification induite par l'emboîtement des concepts et le jeu entre les notions de limite « lim » et « lim gén », le passage formel à cette nouvelle version suit précisément le schéma initial de la preuve. Fréchet peut à partir de cette extension de la méthode de Pincherle traduire son résultat sous la forme intégrale (que nous avons rappelée plus haut) dans une section de son article intitulée « Le théorème de M. Hadamard »¹³⁰ :

Quelle que soit la fonction continue¹³¹ $f(x)$, [on pourra] écrire :

$$f(x) = \lim_{n=\infty} \int_0^\pi f(y) \sigma_n(x, y) dy,$$

$$U_f = \lim_{n=\infty} \int_0^\pi f(y) K_n(y) dt.$$

¹²⁸[Fré04c, p.495]

¹²⁹[Fré04c, p.496]

¹³⁰[Fré04c, p.496]

¹³¹ Fréchet note que l'on peut, partant d'une fonction définie sur un intervalle (a, b) , se ramener à l'étude sur $(0, \pi)$ par un changement de variable, ce qui rétablit la portée générale des passages cités plus haut.

On remarque ici que Fréchet focalise son intérêt sur cette forme intégrale en associant le résultat au nom d'Hadamard. Pourtant la méthode employée, nous l'avons vu, est à rapprocher des travaux de Pincherle, et le passage à la forme intégrale est immédiat en posant, comme Fréchet le précise¹³² :

$$\sigma_n = \frac{2}{\pi} \left(\frac{\frac{1}{2} \cos y \cos x + \cdots + \cos ny \cos nx}{n} \right)$$

$$K_n(y) = \frac{nu_0 + (n-1)u_1 \cos y + \cdots + u_n \cos ny}{n}.$$

Nous verrons que ce choix de recentrage sur la forme intégrale du résultat jouera un rôle important dans la progression des idées exposées par Fréchet. Avec ce théorème Fréchet semble donc avoir atteint les limites des possibilités d'extension de la démarche initiale qui utilise un développement similaire à celui de Taylor : il s'est affranchi avant tout des contraintes de régularité sur la fonction f tout en adaptant une certaine méthode.

À ce point de la progression, la fonction f est vue comme un élément générique (une fonction continue quelconque), mais jamais dans cette première note Fréchet ne mentionne l'ensemble des fonctions continues, ni même aucune réserve de fonctions sur lesquelles agirait U , contrairement au point de vue qu'avait adopté Pincherle avant lui. On ne peut donc pas exactement parler d'éléments de structure dégagés lors de cette première avancée sur l'étude des opérations linéaires qui semble se développer dans une direction très distincte de celle que nous avons tracée entre quelques écrits de 1904 jusqu'à la thèse en 1906. Pourtant, s'il n'y a pas encore ici de concept mathématique mettant en jeu une structure, nous avons montré qu'apparaissaient dans la démarche et dans le texte trois éléments —le schéma de la preuve, l'analogie avec le développement de Taylor et l'énoncé du résultat sous forme intégrale— qui organisent fondamentalement le point de vue de Fréchet. L'angle d'approche particulier qui en découle va finalement permettre à Fréchet de réinterpréter l'énoncé du théorème d'Hadamard et de dépasser son cadre limite. La recherche de la *bonne distance critique* par rapport à un objet d'étude revisité va amener Fréchet dans la suite de ces notes à adapter son discours, à penser en termes d'éléments clés ou essentiels dont l'assemblage sera propice à l'émergence d'un point de vue général.

Fréchet veut dépasser le cadre limite exploré au début de son texte mais sa stratégie ne consiste pas à convoquer ici les idées les plus générales possibles : on peut voir dans le travail de Pincherle une vision plus générale¹³³ d'opérations agissant sur des espaces linéaires quelconques, alors que Fréchet choisit d'élaborer un cadre particulier pour l'étude du problème d'Hadamard, un cadre adapté aux problèmes posés d'abord en termes de fonctions.

¹³²[Fré04c, p.496]

¹³³Nous employons dans ces quelques lignes le mot « général » au sens de l'étendue des cas couverts par le concept d'opération. Nous avons illustré à maintes reprises l'émergence du *général* sous des formes multiples qui dépassent de loin cette conception étroite.

4.4.3 MODALISATION DU DISCOURS ET ÉMERGENCE D'ÉLÉMENTS STRUCTURELS

Avant d'aller plus loin dans notre étude il convient de préciser le titre de cette section. Nous employons le terme *modalisation* pour faire référence aux éléments dynamiques du discours qui permettent de créer une forme stable et reconnaissable — nous nous rapprochons donc ici d'une acception bien illustrée en théorie de la musique par exemple où le *mode* peut être défini comme une forme reconnaissable. Ainsi, l'étude de la modalisation du discours mathématique consiste à identifier les éléments d'une dynamique du texte nécessaires à l'émergence de théories nouvelles et générales. En d'autres termes nous cherchons à mettre en évidence les conditions d'élaboration de *modes*, c'est-à-dire, pour reprendre une expression wittgensteinienne que nous avons déjà suggérée, les conditions d'élaboration d'une « ressemblance de famille »¹³⁴.

La dynamique que nous convoquons ici rend compte en particulier des processus de sélection de matériaux antérieurs, de mouvements textuels — migrations de certaines parties — de l'ordonnancement ou de la hiérarchisation de divers éléments choisis, de leur assemblage et de la recherche de la meilleure distance interprétative.

Nous l'avons vu dans la partie précédente, la première évolution vers un contexte général est dirigée par la volonté d'étendre le champ d'applicabilité du schéma de la preuve du théorème d'Hadamard. Ce développement est totalement linéaire en ce sens qu'il suit la règle déterminée qui consiste à s'affranchir petit à petit des hypothèses d'abord formulées sur la fonction f , tout en conservant la forme de la preuve : on est passé de fonctions analytiques à des fonctions satisfaisant aux conditions de Dirichlet avant d'obtenir le résultat pour les fonctions seulement continues.

Après avoir mis en valeur le schéma de la preuve, et après en avoir établi les limites pratiques, Fréchet va déplacer son attention et analyser la nature du résultat obtenu, indépendamment de la méthode qui permettait de l'obtenir. Ce changement d'optique, amorcé à la huitième et dernière partie du premier article, va en fait être le déclencheur d'une analyse nouvelle du problème

¹³⁴M. Vorms a montré, en utilisant des exemples de la physique mathématique, les limites d'une « analyse formelle de la structure logico-mathématique des théories scientifiques. » Elle en déduit que « le projet de tracer une frontière nette entre contenu et formulation des théories n'est pas seulement difficile en pratique, mais impossible en principe. La raison de cette impossibilité est qu'une théorie, en tant qu'outil de représentation et d'inférence, est indissociable des esprits qui mettent en œuvre les raisonnements permettant d'atteindre les différentes parties de son contenu, de les développer, et de les appliquer. En conséquence, chercher à dissocier le contenu d'une théorie (structure syntaxique et règles d'interprétation, ou structure mathématique) de la manière dont il est utilisé dans les raisonnements, c'est se condamner à manquer ce contenu ; une théorie scientifique n'a véritablement de contenu qu'en tant qu'elle est comprise et utilisée dans des raisonnements. » (cf. [Vor11, p.168]) Notre travail montre que d'une façon similaire en mathématiques pures on manque une part essentielle en tentant de comparer des formes en les rapportant à ce que nous avons désigné par des *structures ensemblistes*. Une certaine approche historique des textes permet cependant de mettre en évidence des dynamiques que nous avons désignées par le terme *modalisation* et de mettre en évidence des rapprochements ou des contrastes qui échappent à la comparaison formelle.

d'Hadamard mais aussi de son contexte. Cette nouvelle approche est introduite par Fréchet en ces termes¹³⁵ :

Le théorème de M. Hadamard conduit naturellement à étudier le problème suivant : puisque toute opération linéaire peut s'exprimer sous la forme :

$$U_f = \lim_{n \rightarrow \infty} \int_0^\pi f(y) H_n(y) dy,$$

$H_n(y)$ étant une fonction continue, à quelles conditions le second membre pourra-t-il définir une opération linéaire si l'on se donne a priori une suite quelconque de fonctions continues $H_n(y)$?

C'est l'intérêt nouveau porté à la forme du résultat et à l'interprétation de la formule ci-dessus qui va amorcer la progression vers un point de vue plus général et engendrer l'apparition d'éléments structurels que nous allons identifier. Il s'agit d'une progression vers le général d'une nature essentiellement différente de celle que nous avons analysée jusqu'ici : elle se traduit par la convocation et la sélection de certains éléments mathématiques clés et par un réagencement de la structure du texte. L'évolution sensible que l'on peut saisir dans cette suite d'articles s'opère en trois temps, chacun marqué par la migration et l'organisation nouvelle de certains éléments du texte.

La progression est initiée lorsque pour la première fois Fréchet mentionne la nature des fonctions qui apparaissent dans le théorème d'Hadamard. C'est l'objet d'une remarque à la fin de sa première note¹³⁶ :

Ce qui précède montre en outre l'utilité qu'il y a à ne pas rejeter comme trop artificielles des opérations qui seraient définies par une expression telle que :

$$\int_0^\pi f(y) K(y) dy,$$

où K serait une fonction non continue, même pas intégrable au sens de Riemann, mais intégrable au sens plus large de M. Lebesgue [...]

Cette première remarque porte seulement sur la nature du noyau K et pas encore sur la fonction f . Fréchet note d'abord ici qu'on ne peut pas se restreindre à des fonctions K continues « [c]ar on s'exposerait à rejeter en même temps [...] des opérations linéaires très simples et très naturelles. » Ceci l'amène alors à une reformulation du problème portant sur la nature de K ¹³⁷ :

Dans quel cas l'opération $[V_n^{(n)} = \int_0^\pi f(y) H_n(y) dy]$ a-t-elle pour limite une opération de la forme $[\int_0^\pi f(y) K(y) dy]$, où la fonction $K(y)$ est la fonction la plus générale qui puisse donner un sens à cette expression, c'est-à-dire une fonction mesurable¹³⁸ ?

Ce n'est qu'à la fin de son deuxième article que Fréchet va discuter pour la première fois la nature des fonctions f admissibles, dans une section intitulée « [I]mportance du champ fonctionnel dans lequel on définit une opération

¹³⁵[Fré04c, p.497]

¹³⁶[Fré04c, p.498]

¹³⁷[Fré04c, p.499]

¹³⁸Une note de bas de page fait référence aux *Leçons sur l'intégration* de Lebesgue. (cf. [Leb04])

linéaire. »¹³⁹ Dans cette partie Fréchet énonce le problème en ces termes : « Nous avons supposé jusqu'ici que les opérations linéaires sur lesquelles nous raisonnions faisaient correspondre un nombre U_f à toute fonction $f(x)$ continue entre a et b et cela sans savoir si U_f était définie pour d'autres fonctions. »¹⁴⁰ Il est à noter que la notion de « champ fonctionnel » doit être interprétée en tant qu'étendue de candidats ayant des propriétés adéquates sans y voir *a posteriori* les structures que nous y associons désormais usuellement. Il s'agit de considérer une collection de fonctions f la plus large possible et telle que le théorème d'Hadamard puisse faire sens. Fréchet mène alors une analyse du champ fonctionnel le plus général envisageable¹⁴¹ :

Pour donner un exemple plus général, considérons un ensemble dénombrable E de points entre a et b : c_1, c_2, c_3, \dots et soit

$$A_1 + A_2 + \dots + A_n + \dots$$

une série absolument convergente. L'opération

$$U_f = A_1 f(c_1) + \dots + A_n f(c_n) + \dots$$

sera définie et linéaire pour toute fonction $f(x)$ uniforme et bornée entre a et b . Or, on ne peut étendre à de telles opérations le théorème de M. Hadamard. En effet, son énoncé même suppose qu'on peut calculer l'intégrale

$$\int_a^b K_n f dy;$$

il faut pour cela que f soit intégrable. Par conséquent, il ne peut y avoir de sens que dans le champ des fonctions intégrables. Mais cette restriction n'est pas encore suffisante. [...]

C'est essentiellement la forme du théorème d'Hadamard qui dirige alors l'analyse de Fréchet et permet de cerner les contours du champ fonctionnel le plus vaste possible, d'en mettre au jour les caractéristiques nécessaires. D'ailleurs la suite de la note expose quelques cas limites élémentaires : « Si l'on applique la formule $[U_f = \lim_{n \rightarrow \infty} \int_a^b K_n(y) f(y) dy]$ à la fonction sommable $f(x)$ qui est égale à 1 pour $x = a$ et nulle partout ailleurs, on arrive à une impossibilité. »¹⁴² Dans le même ordre d'idées, Fréchet remarque aussi que certaines opérations linéaires peuvent ne pas être définies pour toutes les fonctions continues, ce qui va l'amener à considérer des opérations définies sur un champ plus restreint : « Un autre genre de difficultés se présente quand on considère des opérations linéaires qui ne sont pas définies dans tout le champ des fonctions continues, comme l'opération $U_f = f'(c)$. »¹⁴³

Finalement, après cette analyse précise de Fréchet, l'évolution du point de vue va être complète dans la troisième note et on y observe un réagencement

¹³⁹[Fré05, p.138]

¹⁴⁰[Fré05, p.138]

¹⁴¹[Fré05, p.138]

¹⁴²[Fré05, p.138]

¹⁴³[Fré05, p.139]

essentiel des idées. Le texte s'ouvre sur l'introduction d'un champ fonctionnel quelconque avec lequel on peut définir une opération¹⁴⁴ :

Définition du champ. Considérons un champ de fonctions de la variable x définies dans l'intervalle $(0, 2\pi)$. Je supposerai que si deux fonctions appartiennent au champ, il en est de même de leur somme. À toute fonction du champ, $f(x)$, nous pourrons faire correspondre un nombre bien déterminé U_f . Nous définirons ainsi une opération dans ce champ de fonctions. Nous dirons que cette opération est distributive, si quelles que soient les fonctions $f_1(x), f_2(x)$ du champ, on a :

$$U_{f_1 + f_2} = U_{f_1} + U_{f_2}.$$

On est donc passé d'un discours respectant une forme établie antérieurement —présentation conforme au texte d'Hadamard— à une singularisation de certains éléments qui ont été extraits et ramenés à une place particulière au début de l'exposé. Une analyse du texte semble indiquer ici que c'est le concept d'opération et les deux propriétés principales associées qui déterminent le processus à l'œuvre. La distributivité (de l'opération U) trouve un écho dans le choix de donner une place particulière à l'addition des fonctions (avec la contrainte supplémentaire que le champ fonctionnel doit être additivement fermé). Quant à la propriété $cU_f = U_{cf}$ pour un nombre réel quelconque c , elle est traitée à part comme conséquence d'une propriété plus générale de l'opération (sa continuité) qui sera énoncée dans des termes différents adaptés à chaque champ fonctionnel choisi. Ainsi Fréchet ne place pas sur le même plan l'addition des fonctions et la multiplication par un réel, et il utilise immédiatement cette dernière de façon implicite^{145,146} :

On en conclut en particulier que l'on a pour toute opération distributive :

$$cU_f = U_{cf}$$

quelle que soit la constante c . Pour que cette relation ait lieu même pour c irrationnel, il suffit que U_f satisfasse à une certaine condition complémentaire. Nous allons énoncer plus loin cette condition complémentaire, mais nous l'énoncerons d'une manière particulière pour chacun des champs de fonctions que nous allons examiner.

Les remarques qui précèdent montrent que Fréchet, même à ce point avancé de sa réflexion, ne considère pas les champs de fonctions sous l'angle des *structures ensemblistes*. L'utilisation de ce que nous pourrions appeler *structure d'espace vectoriel* est pourtant bien établie à cette époque depuis les travaux de Peano d'abord et de Pincherle par la suite, dans des domaines connus de Fréchet et spécialement en analyse. Les définitions établies par ces prédécesseurs auraient permis de constater pour chaque champ de fonction étudié dans la suite qu'il y a des opérations algébriques à disposition qui le font rentrer dans la catégorie des espaces (ou systèmes) linéaires —Pincherle

¹⁴⁴[Fré07, p.433]

¹⁴⁵ Ici Fréchet ne propose pas que le champ de fonctions soit fermé pour la multiplication par un réel par exemple.

¹⁴⁶[Fré07, p.433]

a par exemple montré que de nombreux champs fonctionnels tombaient sous sa définition de « spazio lineare ». Mais ici l'enjeu est différent et il y a un véritable mouvement qui s'amorce vers une vision générale qui s'émancipe pour partie des cas particuliers de champs de fonctions qui vont permettre de mener à bout l'étude de Fréchet. Cette démarche accompagne un processus de structuration important dans lequel Fréchet place côte à côte une première catégorie d'objets (« les champs de fonctions de la variable x définies dans l'intervalle $(0, 2\pi)$ ») et une opération (la « somme » de deux fonctions), afin de créer une nouvelle catégorie (que nous nommerions de nos jours des champs additivement fermés¹⁴⁷) que son étude va utiliser. Ce nouvel objet composite n'est pas à proprement parler l'agrégation de deux entités indépendantes ou placées sur le même plan : la notion de champ de fonctions embarque avec elle de multiples concepts ou objets plus ou moins mis en lumière, dont l'addition des fonctions.

Il nous faut cependant préciser le type d'organisation que Fréchet impose ici. La notion d'opération linéaire repose sur la donnée préalable d'une addition qui permet d'en exprimer la propriété (distributivité). Pourtant dans cette définition initiale d'opération linéaire, nous avons vu que le champ des éléments —ici ce que nous avons appelé précédemment des *fonctions-objets*— était convoqué sans y attacher de notion de *structure ensembliste* : ni groupe additif, ni espace vectoriel quelle que soit la déclinaison que l'on choisisse ici de ces locutions. C'est le fait de mettre en avant les opérations linéaires qui selon notre étude permet à Fréchet de changer de point de vue et attribue un rôle nouveau à la fois au champ des *fonctions-objets* et à l'addition.

Le deuxième motif structurant du texte et de la pensée de Fréchet est déterminé par l'introduction de la continuité et le traitement particulier réservé à la deuxième propriété des opérations, à savoir la possibilité d'écrire $cU_f = U_{cf}$ pour tout nombre réel c et toute fonction f du champ considéré (dans la suite de cette section nous noterons cette propriété (\mathfrak{H})). Fréchet va donc réintroduire ici la continuité de l'opération U , laquelle était présente dans la formulation d'Hadarnard, et voir la propriété (\mathfrak{H}) comme conséquence immédiate de celle-ci.

Alors que pour l'addition et la distributivité Fréchet a opté pour un traitement uniforme des champs fonctionnels, et bien que l'approche soit encore uniforme dans l'énoncé de la propriété (\mathfrak{H}), la définition de la continuité sera établie de façon particulière pour les divers champs fonctionnels envisagés. Nous y retrouvons une démarche désormais proche de la « méthode » indiquée dans la thèse de Fréchet. Il distingue dans la dernière partie de son article les opérations définies successivement dans les champs de fonctions suivants :

1°« [L]e champ (M) des fonctions mesurables et bornées. »¹⁴⁸

¹⁴⁷ Fréchet n'emploiera que le mot « champ » tout au long de son étude, et ceci malgré l'évolution de point de vue et la création d'éléments structurels attachés à la collection de fonctions. Le vocabulaire en analyse fonctionnelle ne se stabilisera finalement qu'à partir du livre de Banach [Ban32] où les mots « ensemble », « champ », « espace » permettront de distinguer plusieurs structures et de mieux apprécier la pertinence de l'une par rapport à l'autre.

¹⁴⁸[Fré07, p.433]

Dans ce cas, les fonctions mesurables sont définies au sens de Lebesgue et la convergence adoptée est celle que nous appelons désormais convergence presque partout¹⁴⁹ :

Nous dirons que cette opération est continue dans (M) si l'on a :

$$U_f = \lim_{n \rightarrow \infty} U_{f_n},$$

lorsque $f(x), f_1(x), f_2(x), \dots$ sont des fonctions du champ (M) qui sont bornées dans leur ensemble¹⁵⁰ et que $f_n(x)$ tend vers $f(x)$ sauf peut-être en un ensemble de valeurs de x de mesure nulle.

2° « [L]e champ (Ω) des fonctions sommables et de carré sommable. »¹⁵¹

Dans ce cas Fréchet utilise la « *convergence en moyenne* introduite par M. Riesz et M. Fischer »¹⁵² et cette fois-ci le mode de convergence des fonctions est nettement dissocié dans l'exposé de la continuité de l'application linéaire¹⁵³ :

On dira que $f_n(x)$ converge en moyenne vers $f(x)$ et on écrira :

$$f(x) \sim \lim f_n(x)$$

si $f(x)$ et $f_n(x)$ sont des fonctions de (Ω) telles que l'intégrale :

$$\int_0^{2\pi} [f_n(x) - f(x)]^2 dx$$

tende vers zéro avec $\frac{1}{n}$. Nous dirons alors qu'une opération U_f est continue dans le champ Ω si U_{f_n} tend vers U_f quand $f_n(x)$ converge en moyenne vers $f(x)$.

3° « [L]e champ des fonctions continues. »¹⁵⁴

Dans ce cas Fréchet reprend une définition due à Hadamard pour définir la proximité de deux fonctions et la convergence d'une suite (C, C_n, C_ω, E) désignant respectivement les champs des fonctions continues, n fois dérivables, indéfiniment dérivables ou holomorphes)¹⁵⁵ :

[D]eux fonctions ont un voisinage d'ordre n défini par le nombre ϵ quand, de 0 à 2π , ces fonctions et leurs n premières dérivées diffèrent en valeur absolue de moins de ϵ (n pouvant être nul, entier ou infini). Cela étant, une opération U_f définie dans l'un des champs C, C_n, C_ω ou E s'appellera opération continue si $U_f - U_{f_p}$ est infiniment petit en même temps que le voisinage (d'ordre 0, n, ∞, ∞ respectivement) de $f(x)$ et $f_p(x)$.

¹⁴⁹[Fré07, pp.433-434]

¹⁵⁰ Il faut comprendre ici *qui sont chacune bornée sur $(0, 2\pi)$* .

¹⁵¹[Fré07, p.439]

¹⁵²[Fré07, p.439]

¹⁵³[Fré07, p.439]

¹⁵⁴[Fré07, p.441]

¹⁵⁵[Fré07, p.441]

Ces définitions dont l'unité n'est pas encore vraiment avérée ici (contexte vu de façon non uniforme, fluctuation de vocabulaire, dissociation non systématique de la définition de la convergence des fonctions et de la continuité de l'opération), permettent pourtant à Fréchet de définir la continuité de U_f d'une manière similaire dans les trois situations : U_f est continue si $\lim U_{f_n} = U_f$ lorsque f_n converge vers f ¹⁵⁶. Ainsi une certaine uniformité de vision des champs fonctionnels, qui semblait avoir disparu dans le traitement cas par cas, réapparaît ici dans une formulation (presque) unique de la continuité de l'opération U et dans la formulation isolée au début de la propriété (5). On peut donc penser ici que le mouvement vers le général ne s'effectue pas dans l'élaboration d'une *structure ensembliste*, que l'on qualifierait maintenant de topologique (issue des considérations de convergences des suites de fonctions), mais plutôt dans une formulation unique de la continuité d'une opération, définition uniformément adaptée à chacun des espaces considérés tout au long de ces trois notes. Cette formulation unique est elle-même intimement liée à l'extraction et à la mise en avant de la convergence des suites de fonctions comme élément fondamental et support de la pensée de Fréchet, et ceci pour tous les champs fonctionnels étudiés. Enfin, mentionnons que Fréchet pourra dans chacun de ces cas utiliser sa méthode et le schéma unique de la preuve du théorème d'Hadamard en approchant la fonction f par une série convergente (en un sens adapté à chacun des cas et des types de convergence) et en appliquant U terme à terme grâce à la continuité exprimée de façon adaptée et comme il l'a fait dès la première note.

Nous avons souligné tout au long de cette section comment une certaine *modalisation du discours mathématique* donnait une place particulière à des éléments sélectionnés : des objets mathématiques, mais aussi des formes ou des éléments du discours sur lesquels l'auteur s'appuie pour forger une vision nouvelle. L'angle d'observation que nous cherchons à adopter ici nous permettra de faire par la suite apparaître certaines *ressemblances de famille* avec le travail de Banach mais aussi de mettre au jour de véritables contrastes qui permettront de mieux apprécier l'œuvre du mathématicien polonais. Enfin nous nous appuierons à nouveau sur les éléments que nous avons dégagés ici pour tenter en conclusion de revenir sur la notion de structure telle qu'elle peut être utilement conçue dans le cadre de l'analyse fonctionnelle et de son histoire.

4.5 CONCLUSION : VERS UNE ANALYSE GÉNÉRALE

Dans les sections précédentes nous avons analysé à travers des textes du début de la carrière de Fréchet deux voies qu'il emprunte parallèlement afin d'établir un contexte général pour l'analyse. Celle-ci est tantôt « analyse fonctionnelle » —on y lit encore le lien avec les problèmes de variations tels qu'ils ont été abordés par ses prédécesseurs Hadamard ou Volterra, tantôt elle devient ce que lui-même nommera « analyse générale », terme qu'il reprend de

¹⁵⁶Fréchet ne reprend pas ici le vocabulaire de sa thèse, mais il s'agit à chaque fois de mettre en avant une notion de limite et d'en déduire une formulation particulière de la continuité que nous nommerions aujourd'hui *continuité séquentielle*.

Moore et qui désigne une analyse qui se veut d'une portée plus générale que la précédente.

L'histoire des mathématiques a souvent tendance à dissocier chez Fréchet le travail sur les ensembles abstraits du travail sur les opérations alors que ces dernières sont les éléments centraux et constitutifs de l'analyse telle que Fréchet et d'autres la conçoivent, qu'elle soit analyse fonctionnelle ou générale. C'est une des caractéristiques du type d'organisation que Fréchet met en place et dont nous avons commenté la nature qui contient en elle la possibilité de bifurcation, qualité qui détermine en partie la nature de la généralité à l'œuvre dans le travail du mathématicien. Nous retrouverons cette caractéristique dans le travail de Banach et la même conséquence pour l'analyse de ses travaux où la notion de structure d'espace vectoriel normé complet est souvent étudiée de façon indépendante du travail sur les opérations.

Nous avons établi au-dessus que réduire la notion de *structure* pour en faire un terme qui désigne rétrospectivement une équivalence formelle — ici un espace topologique, ailleurs un espace de fonctions muni de telles ou telles propriétés— laissait de côté un certain nombre de caractéristiques qui ont pourtant aussi à voir avec une forme de *structure*. L'organisation même du discours, le regard réflexif sur certains éléments, l'agencement des propriétés que l'on peut ajouter ou ôter en fonction du degré de généralité que l'on souhaite obtenir, d'autres *éléments structurels* entrent en ligne de compte et organisent le général. Ces assemblages que donnent à voir les textes mathématiques permettent la considération morcelée, la bifurcation, une progression des mathématiques non linéaire depuis de multiples centres, une progression rhizomatique pour reprendre une image deleuzienne.

L'étude des textes entre 1904 et 1907 a mis en évidence un certain type de généralité qui s'appuie justement sur des *éléments structurels* dont on ne peut voir la cohérence que quand on comprend pourquoi Fréchet les convoque et les organise ainsi. Certains textes postérieurs donnent à voir une vision plus large dont l'idée centrale est celle d'*opération*, une vision que Fréchet tente de préciser à plusieurs reprises à un moment où la théorie s'est largement développée et où un certain nombre de chemins de traverse commencent à être empruntés en s'appuyant sur ses travaux et en particulier sa thèse.

Dans plusieurs circonstances Fréchet exprimera précisément cette vision large et tentera de mieux définir ce qu'il nomme « analyse générale ».

En particulier il publie en 1928 dans la *Collection de monographies sur la théorie des fonctions* dirigée Borel un ouvrage intitulé *Les espaces abstraits et leur théorie considérée comme introduction à l'analyse générale*. Le titre principal, imprimé en gros caractères et en gras est « Les espaces abstraits », le reste figurant comme sous-titre. Ceci montre que le sujet principal qui est traité dans ce livre est axé sur ce que nous appellerions désormais les notions de *topologie* dans les ensembles abstraits. C'est même le sujet exclusif développé dans l'ouvrage, et ce qui nous intéresse ici ne se trouve que dans l'introduction écrite par Fréchet. Sa portée dépasse nettement celle de l'exposé du corps principal du livre et replace les idées qui vont être développées à une place précise dans un édifice plus vaste.

Au Congrès international de Bologne Fréchet précise que son livre « a pour objet principal de préparer à l'étude de l'Analyse générale »¹⁵⁷

L'introduction est elle-même structurée en onze points où les « espaces abstraits » apparaissent en cinquième position. Pour commencer Fréchet définit l'*analyse générale* comme le prolongement de plusieurs domaines qui étudient les fonctions. Il cite M. Winter qui a distingué « les idées relatives au développement formel des propriétés opératoires et un autre courant d'idées, celui qui consiste dans la généralisation des notions élémentaires du Calcul infinitésimal. »^{158,159} Les premières idées sont rattachées par Fréchet aux travaux de Pincherle, et par Winter aussi aux travaux de Bourlet. Elles définissent pour Fréchet « le Calcul fonctionnel proprement dit ». Le second courant d'idées est associé par Fréchet aux travaux de Volterra, Hadamard ou Levy. Il constitue ce qu'il nomme l'« Analyse fonctionnelle » en reprenant une expression proposée par Hadamard.

Pour l'exprimer brièvement, le « calcul fonctionnel » a permis de penser les opérations en des termes algébriques. On retrouve ces idées par exemple chez Peano, Pincherle ou Bourlet sous des formes différentes. Dans ce cas les auteurs se sont appuyés sur des propriétés *algébriques* adaptées aux fonctions et aux opérations. Ils ont pu définir des notions de *distributivité* ou *linéarité* sous diverses formes que l'on peut rapprocher, pour certains aspects, du calcul symbolique comme l'avait suggéré Pincherle. Pourtant, dans cette approche *algébrique* que Fréchet rattache au « développement formel des propriétés opératoires » la notion de continuité n'est pas absente. Elle est convoquée par Pincherle ou Bourlet par exemple dans son lien avec la distributivité et la possibilité d'étendre une relation du type $T(cx) = cT(x)$, qui ne dépend que des propriétés de distributivité pour c rationnel, à tout nombre réel c .

Dans l'« analyse fonctionnelle » le point de vue résumé par Fréchet est différent¹⁶⁰ :

A l'origine, il ne s'est agi que d'étudier les fonctions particulières dont l'argument, la variable, est non pas un nombre, mais une ligne ou, ce qui revient au même, une fonction. D'où le nom de *fonctions de lignes*, usité par M. Volterra, ou de *fonctionnelle*, dû à M. Hadamard. Et, pour cette analyse nouvelle, M. Paul Levy, qui y a apporté d'importantes contributions, a proposé le nom d'*Analyse fonctionnelle*.

Notre étude, et en particulier celle portant sur les travaux de Fréchet entre 1904 et 1907, a montré que la distinction entre *Calcul fonctionnel* et *Analyse fonctionnelle* est en fait très subtile. Les idées se développent selon des voies que l'on peut distinguer mais qui dialoguent en permanence à la recherche d'une vision générale et efficace.

¹⁵⁷[Con28, pp. 271-272]

¹⁵⁸[Fré28b, p.3]

¹⁵⁹En fait Winter dans son texte insiste sur les éléments qui permettent « de montrer la dépendance qui existe entre [c]es idées » [Win13, p.477] Fréchet au contraire insiste sur la distinction que Winter a illustrée dans son texte en rappelant les travaux de nombreux auteurs dont Bourlet, Pincherle, Hadamard, Volterra, Fréchet, Levy, Hilbert etc.

¹⁶⁰[Fré28b, p.3]

Dans l'introduction de son mémoire de 1928 Fréchet indique que c'est la notion de *fonction* qui a permis de construire le point de vue général qui définit l'« Analyse générale ». Pour lui il y a deux temps que nous avons aussi cherché à illustrer à travers notre choix de textes. Dans une première étape les mathématiciens ont tenté de transposer une partie de la théorie des fonctions ordinaires, « reprenant ses procédés et les appliquant à un compartiment voisin¹⁶¹ où la variable était d'une nature autre, mais *déterminée*. » Puis dans un second temps un processus d'abstraction a permis de concevoir la variable sous un angle unifié et général¹⁶² :

Nous avons montré qu'il n'est pas toujours nécessaire de faire rentrer en ligne de compte la nature de la variable envisagée, qu'il suffit souvent de retenir les propriétés topologiques de l'espace, du champ de variation. Et nous nous sommes proposé la généralisation des notions élémentaires du calcul infinitésimal au cas où la variable est de nature quelconque.

Enfin Fréchet esquisse l'étendue la plus générale que l'on puisse donner à la notion d'« Analyse générale » et donne une définition très large¹⁶³ :

Si l'on a généralisé la théorie des fonctions en remplaçant la variable numérique par une variable quelconque, pourquoi ne pas généraliser intégralement en remplaçant aussi la fonction numérique par une fonction de nature quelconque.

Et alors *l'Analyse générale apparaît* maintenant comme *l'étude des relations entre deux éléments de nature quelconque* (dont l'un pourra jouer le rôle de variable et l'autre de fonctionnelle).

Nous avons vu dans les textes précédents que la généralisation ne repose pas seulement sur la notion de fonction qui est petit à petit comprise en un sens étendu. Les auteurs comme Volterra ou Fréchet ont conçu une véritable théorie, une *analyse*, dans laquelle on retrouvait sous une forme transposée et généralisée des théorèmes importants de la théorie des fonctions ordinaires. Non seulement les mathématiciens ont généralisé des concepts mais ils ont aussi isolé et sélectionné des éléments de la théorie classique — théorèmes, stratégies, visions géométriques, etc. — qu'ils ont cherché à conserver et à transposer dans le cadre général. Fréchet résume l'objectif visé¹⁶⁴ :

L'édifice à construire exige deux tâches bien différentes : déterminer un système de définitions aussi semblables que possible, — mais sous forme abstraite, — aux définitions adoptées dans l'Analyse classique ; au moyen de ces définitions, généraliser les propositions classiques fondamentales.

Dans ce schéma Fréchet indique que la théorie des ensembles abstraits devient nécessaire comme « une étude préliminaire » qu'il relie essentiellement

¹⁶¹ Il s'agit d'un jugement rétrospectif de Fréchet. Nous avons vu les processus multiples qui ont permis de concevoir ce « compartiment voisin » et de penser des *fonctions* en un sens généralisé.

¹⁶² [Fré28b, p.7]

¹⁶³ [Fré28b, pp.7-8]

¹⁶⁴ [Fré28b, p.15]

à la question des variations des fonctions. Ainsi le sujet principal est de « donner un sens à l'expression : près d'un élément. »¹⁶⁵.

Enfin dans un dernier paragraphe de son introduction Fréchet donne un point de vue plus philosophique¹⁶⁶ sur sa vision de l'Analyse générale. Bien qu'il s'agisse d'une vision rétrospective par rapport aux travaux que nous avons étudiés, elle permet néanmoins de justifier la nécessité d'étudier les espaces abstraits tels qu'ils sont conçus dans l'élaboration de l'Analyse générale.

Dans ce paragraphe Fréchet dissocie « deux points de vue pour dresser le plan d'extension de l'Analyse générale. »¹⁶⁷. Le premier consiste à développer la topologie des ensembles abstraits¹⁶⁸ :

Plusieurs ouvrages [...] sur les ensembles abstraits [...] s'inspireront des tendances générales de l'*Analysis situs* et un de leurs buts les plus importants sera de contribuer à l'étude des fondements de la géométrie.

Mais comme le précise Fréchet « l'étude des fondements de la géométrie n'est pas l'objet principal des travaux de l'auteur. » Il y a pour lui un second point de vue qui diffère par son objet principal d'étude mais aussi par son but¹⁶⁹ :

Notre but est surtout de faire une étude générale des relations entre variables abstraites, étude entreprise, non seulement, pour obtenir des résultats nouveaux, mais aussi pour réaliser l'unification des énoncés classiques de la Théorie des fonctions et de l'Analyse fonctionnelle.

On voit ici réapparaître la distinction que nous avons proposée de marquer entre *généralité synthétique* et *généralité descriptive ou constructive*. Le point de vue *synthétique* que propose Fréchet a été élaboré par un processus de sélection et d'abstraction qui opère à partir des situations connues de la théorie des fonctions ordinaires ou de l'analyse fonctionnelle. Pour Fréchet la théorie qui est ainsi dégagée « apparaîtra mieux comme indépendante des fluctuations de la pensée moderne concernant le modèle mathématique de notre monde physique. »¹⁷⁰ Par contraste la topologie ne jouit pas totalement de cette indépendance puisqu'elle « a en vue l'élaboration de ce modèle »¹⁷¹. Pour Fréchet la topologie est « pour ainsi dire, un des produits dérivés de l'Analyse générale ». Notre étude a montré que l'analyse générale est fondée sur une notion *synthétique* de correspondance sur des ensembles abstraits : c'est donc une théorie des fonctions générale.

Enfin, une remarque nous permettra de faire retour sur un point qui précise aux yeux de Fréchet la position et le rôle de la topologie et des structures

¹⁶⁵[Fré28b, p.11]

¹⁶⁶C'est un des aspects que l'on retrouve plusieurs fois dans les textes de Fréchet où après des considérations historiques ou techniques il dégage souvent des aspects qu'il qualifie lui-même de « philosophiques ».

¹⁶⁷[Fré28b, p.21]

¹⁶⁸[Fré28b, p.21]

¹⁶⁹[Fré28b, p.21]

¹⁷⁰[Fré28b, p.21]

¹⁷¹[Fré28b, p.21]

vectérielles dans l'élaboration de l'analyse générale. Fréchet lors du congrès de Bologne en 1928 revient sur quelques aspects importants de l'analyse générale telle qu'il la voit¹⁷² :

[L]es notions de voisinages, superflues pour définir l'intégration, deviennent, par contre, insuffisantes quand on s'attaque à la différentiation abstraite.

Il ne suffit pas alors qu'on puisse dire si un accroissement de la variable est petit, c'est-à-dire si une des positions de la variable est voisine d'une autre. Il faut encore qu'on puisse retrancher la différentielle de y de l'accroissement de y , il faut savoir ce qu'on entendra en disant que le résultat de la soustraction est infiniment petit par rapport à l'accroissement de la variable abstraite x .

Il semble difficile de pouvoir donner un sens à cette condition sans supposer que les espaces —distincts ou non— où se meuvent x et y , ont non seulement un caractère topologique, mais encore un caractère vectoriel (ces deux caractères devant d'ailleurs être liés l'un à l'autre de façon naturelle).

Les textes que nous avons étudiés, mais aussi ceux de Banach que nous allons analyser, montrent le cheminement et les voies diverses qui ont permis d'élaborer ce programme pour fonder l'analyse générale telle que la conçoit Fréchet vers la fin des années trente. On sait depuis que cette approche sera celle adoptée tout au long du 20^e siècle dans une grande partie de l'analyse fonctionnelle qui se développera dans des espaces vectoriels —ou plus largement dans des structures algébriques diverses— topologiques.

¹⁷²L'intervention de Fréchet s'intitule « L'analyse générale et les Espaces Abstraits » (cf. [Fré28a])

Établissement de l'analyse fonctionnelle moderne : le rôle de Stefan Banach

Der Kreis der Mathematiker, die sich mit der Theorie der Operationen von verschiedenen Standpunkten beschäftigen, wird immer grösser ; [...]

Sie verbinden miteinander in ungewohnter Weise scheinbar voneinander gänzlich getrennte Gebiete, sie steigern die Kraft unserer geometrischen Intuition, sie gestatten komplizierte Beziehungen der Analysis auf einfachere zurückzuführen und dadurch neue Eigenschaften zu erschliessen.¹

— S. Banach

5.1 INTRODUCTION

L'œuvre mathématique de Banach touche à de nombreux domaines et a été d'une influence remarquable tout au long du 20^e siècle. Son travail a été en grande partie rassemblé en deux tomes publiés respectivement en 1967 et 1979 par l'Académie Polonaise des Sciences.

Le premier tome entend rassembler les « travaux sur les fonctions réelles et sur les séries orthogonales », comme le précise son titre, alors que le second tome, qui contient entre autres la version française de la monographie *Théorie des opérations linéaires*, propose de regrouper les « travaux sur l'analyse fonctionnelle. »

¹Conclusion de l'intervention de Banach au congrès International des Mathématiciens à Oslo en 1936. Le titre de son exposé était « Die Theorie der Operationen und ihre Bedeutung für die Analysis. » (cf. [Ban36, p.268])

La thématisation est définie de façon rétrospective par les auteurs et les éditeurs² de ces deux tomes. Banach lui-même ne donnait que très peu de références dans ses articles à des travaux antérieurs ou connectés aux siens qui auraient pu permettre de les classer suivant des champs préexistants des mathématiques.

En fait, de nombreux témoignages de mathématiciens qui ont travaillé avec Banach montrent qu'il était attiré par de nombreux problèmes qui peuvent *a posteriori* être séparés en des domaines distincts : fonctions réelles, séries orthogonales, théorie de la mesure, théorie des opérateurs, analyse fonctionnelle. Et on pourrait poursuivre la liste tant l'œuvre de Banach s'insinue dans de nombreux domaines que toute classification tend à isoler.

Comme nous allons commencer à le montrer le travail de Banach connecte au contraire de nombreux problèmes, qui semblaient dissociés et indépendant les uns des autres avant lui, grâce à une lecture abstraite et générale. Nous proposons donc ici d'entamer une étude systématique des travaux de Banach en éclairant sa façon d'avoir recours à une certaine vision abstraite et en analysant précisément les différents types de généralité qui émergent de cette vision.

Cette étude permettra entre autres de connecter et de mettre en résonance des travaux dans lesquels Banach développe un point de vue général et commun à partir de situations et de problèmes *a priori* disjoints.

Si les travaux de Banach se laissent difficilement classer dans tel ou tel champ des mathématiques sans occulter des aspects essentiels de sa démarche, il n'en reste pas moins qu'il a contribué de façon essentielle à l'élaboration de l'analyse fonctionnelle du 20^e siècle.

Les textes que nous avons sélectionnés (essentiellement sa thèse de doctorat de 1920 et un travail majeur sur les fonctionnelles publié en 1929) permettent de préciser l'approche abstraite et générale de Banach et de la contraster avec les exemples que nous avons déjà étudiés. La littérature secondaire permet de replacer ces travaux dans un cadre plus large et d'en apprécier plusieurs aspects sur lesquels nous nous appuyerons parfois.

Tout d'abord les deux tomes des œuvres de Banach sont assortis d'un appareil critique qui permet d'éclairer certains aspects.

Le premier tome présente une cinquantaine de pages de commentaires qui reprennent un à un les articles de Banach rassemblés dans cette première partie. On y trouve des références à de nombreux travaux qui indiquent des prolongements des articles de Banach. Beaucoup de ces prolongements sont publiés du vivant de Banach et permettent de reconstruire pour des articles postérieurs un faisceau de références que Banach ne donne pas lui-même. Ces commentaires sont aussi l'œuvre de mathématiciens qui ont été des acteurs importants des mathématiques polonaises à l'époque de Banach, et leurs commentaires permettent de se faire une idée des thèmes qui circulaient dans les milieux mathématiques polonais dans les premières décennies du 20^e siècle.

²Le comité de rédaction est composé de A. Alexiewicz, M. Altman, S. Hartman, E. Marczewski, S. Mazur†, W. Orlicz, R. Sikorski et H. Steinhaus† pour le premier tome et C. Bessaga, S. Mazur, W. Orlicz, A. Pełczyński, S. Rolewicz et W. Żelazko pour le second tome. Les noms suivis d'une marque furent des co-auteurs de Banach. Orlicz participa aux réunions du Scottish Café et plusieurs de ses problèmes figurent dans le Scottish Book tel que Ulam l'a publié.

Le second tome ne dispose pas du même appareil critique et seule la monographie de 1932 (la version française) est suivie d'un long article d'Aleksander Pełczyński en collaboration avec Czesław Bassaga. Ils y donnent des aspects de la théorie des espaces de Banach qui leur semblent avoir été initiés dans la monographie *Théorie des opérations linéaires*. Par ailleurs le livre de 1932 contient quelques pages de *remarques* rédigées essentiellement par S. Mazur qui donnent un certain nombre de références bibliographiques que Banach n'a pas fait figurer dans le corps de son texte.

Deux articles publiés au *Colloquium Mathematicae* sont dédiés à l'œuvre de Banach et sont intitulés « Sur l'œuvre scientifique de S. Banach » (I et II). Le premier est écrit par W. Orlicz et porte le sous-titre « Théorie des opérations et théorie des séries orthogonales »³. Le second est écrit par E. Marczewski avec le sous-titre « Théorie des fonctions réelles et théorie de la mesure »⁴. On y trouve à nouveau des références à des travaux que les auteurs connectent à ceux de Banach et qui permettent de mieux cerner la nature des problèmes qu'il a étudiés.

Outre ces études, on peut consulter le livre de A. Pietsch « History of Banach Spaces and Linear Operators »⁵ qui donne un panorama large et très riche en références des idées que nous pouvons désormais rattacher à la théorie des opérateurs. Mentionnons enfin les travaux de R. Duda⁶ et ceux de A. F. Monna⁷ qui permettent de resituer les travaux de Banach dans des contextes plus larges des mathématiques polonaises.

Notre approche est complémentaire à celles mentionnées au-dessus et s'en distingue essentiellement par le type de focalisation que nous mettons en œuvre. Notre objet est de montrer comment Banach s'appuie sur une forme d'abstraction afin de développer une vision générale. Nous montrerons en particulier que se joue un changement important de perspective lié chez Banach au développement d'une *théorie des opérateurs*. Nous mettrons en évidence que la conception de ce que l'on nomme un *espace de Banach*, c'est-à-dire un espace vectoriel normé complet (pour le dire avec des termes actuels), est intimement liée à l'utilisation des opérateurs (applications entre espaces de Banach ici, Banach parlera d'« opérations ») pour relire divers problèmes.

Nous montrerons en outre que Banach développe un certain regard *réflexif* sur les opérateurs qui leur confère un statut général dont nous étudierons les caractéristiques et qui permet d'établir l'émergence d'une véritable *théorie des opérateurs* largement indépendante des problèmes qui motivent leur conception. Nous mettrons au jour les mécanismes de sélection et d'organisation que Banach met en œuvre et qui seront exposés pour la première fois sous la forme d'une théorie nouvelle dans le livre de 1932⁸.

³[Orl48]

⁴[Mar48]

⁵[Pie07]

⁶Les travaux de Duda permettent entre autres de se faire une idée précise des conditions de production des mathématiques en Pologne à l'époque de Banach. (cf. [Dud79 ; Dud96 ; Dud05 ; Dud09 ; Dud13])

⁷On peut lire par exemple [Mon75].

⁸Notre étude s'arrête essentiellement avant la monographie de 1932 intitulée *Théorie des opérations linéaires*. Nous commenterons la cohérence de ce choix en conclusion.

5.2 LA THÈSE DE DOCTORAT DE BANACH

Banach a écrit sa thèse alors qu'il avait déjà plusieurs résultats à son actif, essentiellement des résultats sur les fonctions ordinaires, c'est-à-dire de \mathbb{R} dans \mathbb{R} . En 1920 il devient assistant de Łomnicki⁹ et rédige d'abord sa thèse en polonais avant qu'elle soit publiée en 1922 en français dans un des premiers volumes de la jeune revue *Fundamenta Mathematicae*¹⁰.

Le titre choisi par Banach, « Sur les opérations dans les ensembles abstraits et leur application aux équations intégrales », indique le point de vue qu'il adopte et l'objectif qu'il s'est fixé. Il s'agit de donner des résultats pour des équations intégrales vues sous un angle particulier. D'une part les différents champs fonctionnels qui interviennent dans les équations intégrales seront interprétés à travers un concept d'« ensemble abstrait » qui permettra de les traiter de façon unifiée. D'autre part les équations intégrales seront elles-mêmes considérées en termes d'opérations agissant sur les éléments des « ensembles abstraits ».

L'histoire des mathématiques a souvent identifié dans ce texte la première occurrence de ce que l'on appelle rétrospectivement une *structure d'espace de Banach*, c'est-à-dire un *espace vectoriel normé complet*. Si l'on poursuit le travail critique que nous avons menée tout au long de notre étude il nous faut réinterpréter la manière dont Banach convoque certains éléments, les organise et les utilise afin de déterminer exactement le rôle que jouent les « ensembles abstraits » dans sa thèse. Nous pourrions alors cerner précisément quels types d'organisations et éventuellement quels types de structures sont réellement en jeu dans ce texte, puis dans des travaux postérieurs.

Or l'étude de la thèse de Banach, comme le suggère le titre, a pour objectif central d'élaborer une théorie des « opérations » dans un but précis. Si l'on se concentre sur cet objectif on s'aperçoit que les « ensembles abstraits » jouent un rôle particulier et qu'on ne peut isoler leur conception de celle des « opérations ». On trouve d'ailleurs ici un parallèle frappant avec l'organisation de l'Analyse générale de Fréchet telle que nous l'avons analysée plus haut¹¹, et là encore les contrastes nous permettront de mieux apprécier la démarche de Banach. Le titre « Sur les opérations... » choisi par Banach est suivi d'une introduction dont les premiers mots sont « L'opération c'est... » qui annoncent clairement la volonté de proposer une théorie centrée sur le concept d'opération.

Un de nos objectifs dans cette section est de montrer comment Banach convoque et organise diverses idées autour de la notion d'opération et d'étudier l'articulation entre une certaine notion d'espace (ici espace de Banach si on reprend un qualificatif rétrospectif) et d'opération.

Afin de montrer l'importance de l'articulation entre les notions d'opération et d'espace nous commençons par rappeler deux situations antérieures au texte

⁹Antonin Marian Łomnicki (1881-1941).

¹⁰Le journal fût fondé cette même année 1920 par Zygmunt Janiszewski, Stefan Mazurkiewicz et Waclaw Sierpiński. La thèse de Banach ainsi que l'ensemble de la revue est actuellement accessible en ligne à la *Polish Digital Mathematics Library* à l'adresse <http://matwbn.icm.edu.pl>

¹¹Nous faisons en particulier référence ici à la fin de notre analyse des textes de Fréchet et à ses commentaires sur l'articulation des espaces abstraits avec l'Analyse générale. Il se pourrait que ce soit ici chez Banach que se joue ce que Fréchet n'avait pas totalement réalisé mais qu'il appelait de ses vœux, c'est-à-dire la formulation d'une théorie des opérations sur les ensembles abstraits et la naissance de ce qu'on nommera plus tard la *théorie des opérateurs*.

de Banach. Elles révèlent à la fois des similitudes que l'on peut apprécier dans un certain type de mise en correspondance formelle des textes, mais elles permettront aussi de mettre au jour un véritable contraste entre divers recours à une version abstraite d'opération et d'espace (termes d'observateur ici).

Les exemples que nous citons dessous ont déjà été identifiés par les historiens des mathématiques pour leurs différents rôles dans la conception des espaces modernes tels qu'ils sont en particulier utilisés en analyse. Notre étude s'efforce d'apporter un regard nouveau en montrant que l'on peut contraster ces situations antérieures avec celle mise en place par Banach plusieurs décennies plus tard. En particulière nous ferons apparaître qu'un certain nombre de différences ne peuvent pas être établies lorsque l'on cherche dès l'abord à rapprocher ces textes d'un point de vue purement formel, c'est-à-dire à la lumière d'un *modèle* unique auquel on pourrait rétrospectivement se référer. Des contrastes apparaissent dès lors que l'on cherche à comprendre certains aspects que nous qualifierons de *dynamiques* dans la création et l'utilisation des concepts convoqués par les différents mathématiciens.

5.2.1 PEANO ET LES « TRANSFORMATIONS DES SYSTÈMES LINÉAIRES »

C'est dans le livre *Calcolo geometrico*¹² de Peano, publié en 1888, que l'on trouve ce que les historiens ont identifié comme la première occurrence d'une liste d'axiomes qui définissent un espace vectoriel.

Le livre envisage un certain nombre de situations géométriques dans le plan et dans l'espace et propose à sa toute fin une liste d'axiomes qui définissent ce que Peano nomme un « système linéaire » (*sistema lineare*). Il s'agit en fait du neuvième et dernier chapitre du livre dans lequel Peano développe brièvement une façon de voir les transformations linéaires de façon générale¹³ :

L'ultimo capitolo tratta sommariamente delle trasformazioni dei sistemi lineari in generale e delle formazioni geometriche in particolare.

Le chapitre, intitulé « Transformations des systèmes linéaires » (*Trasformazioni di sistemi lineari*), s'ouvre sur un énoncé qui précise la notion de « système linéaire »¹⁴ :

Esistono dei sistemi di enti sui quali sono date le seguenti definizioni:

1. È definita l'*eguaglianza* di due enti **a** e **b** del sistema, cioè è definita una proposizione, indicata con $\mathbf{a} = \mathbf{b}$, la quale esprime una condizione fra due enti del sistema, soddisfatta da certe coppie di enti, e non da altre, e la quale soddisfa alle equazioni logiche:

$$(\mathbf{a} = \mathbf{b}) = (\mathbf{b} = \mathbf{a}), (\mathbf{a} = \mathbf{b}) \cap (\mathbf{b} = \mathbf{c}) < (\mathbf{a} = \mathbf{c}).$$

2. È definita la *somma* di deux entes **a** et **b**, vale a dire è definito un ente, indicato con $\mathbf{a} + \mathbf{b}$, che appartient pure al sistema dato, e che

¹²Le titre complet est *Calcolo geometrico secondo l'Ausdehnungslehre di H. Grassmann preceduto dalle operazioni della logica deduttiva*. [Pea88]

¹³[Pea88, Introduction, p.vii]

¹⁴[Pea88, p.141]

soddisfa alle condizioni:

$$(\mathbf{a} = \mathbf{b}) < (\mathbf{a} + \mathbf{c} = \mathbf{b} + \mathbf{c}), \mathbf{a} + \mathbf{b} = \mathbf{b} + \mathbf{a}, \mathbf{a} + (\mathbf{b} + \mathbf{c}) = (\mathbf{a} + \mathbf{b}) + \mathbf{c},$$

e il valor comune dei due membri dell'ultima eguaglianza si indicherà con $\mathbf{a} + \mathbf{b} + \mathbf{c}$.

3. Essendo \mathbf{a} un ente del sistema, ed m un numero intero e positivo, colla scrittura $m\mathbf{a}$ intenderemo la somma di m enti eguali ad \mathbf{a} . È facile riconoscere, essendo $\mathbf{a}, \mathbf{b}, \dots$ enti del sistema, m, n, \dots numeri interi e positivi, che

$$(\mathbf{a} = \mathbf{b}) < (m\mathbf{a} = m\mathbf{b}); m(\mathbf{a} + \mathbf{b}) = m\mathbf{a} + m\mathbf{b}; (m + n)\mathbf{a} = m\mathbf{a} + n\mathbf{a}; \\ m(n\mathbf{a}) = (mn)\mathbf{a}; 1\mathbf{a} = \mathbf{a}.$$

Noi supporremo che sia attribuito un significato alla scrittura $m\mathbf{a}$, qualunque sia il numero reale m , in guisa che siano ancora soddisfatte le equazioni precedenti. L'ente $m\mathbf{a}$ si dirà *prodotto* del numero (reale) m per l'ente \mathbf{a} .

4. Infine supporremo che esista un ente del sistema, che diremo *ente nullo*, e che indicheremo con 0 , tale che, qualunque sia l'ente \mathbf{a} , il prodotto del numero 0 per l'ente \mathbf{a} dia sempre l'ente 0 , ossia

$$0\mathbf{a} = 0.$$

Se alla scrittura $\mathbf{a} - \mathbf{b}$ si attribuisce il significato $\mathbf{a} + (-1)\mathbf{b}$, si deduce:

$$\mathbf{a} - \mathbf{a} = 0, \mathbf{a} + 0 = \mathbf{a}.$$

DEF. I sistemi di enti per cui sono date le definizioni 1, 2, 3, 4, in guisa da soddisfare alle condizioni imposte, diconsi sistemi lineari.

Ce passage est suivi de quelques exemples issus essentiellement de la géométrie du plan ou de l'espace. On y trouve un exposé qui cherche à embrasser sous une forme abstraite un certain nombre de situations qui ont été étudiées en détails dans les chapitres précédents. La sélection et l'organisation des éléments qui ont été explorés dans les situations de géométrie se fait dans le but de dégager un contexte abstrait pour définir des opérations (*operazioni*) sous une forme indépendante des éléments particuliers rencontrés dans les chapitres précédents. Ceci n'est pas sans conséquence sur la nature de la généralité que l'on peut lire dans le dernier chapitre du livre de Peano comme nous allons le voir.

De plus dans son texte Peano envisage le cas de la dimension infinie¹⁵ : « Un sistema lineare può anche avere infinite dimensioni. »¹⁶ Ceci montre que les « systèmes linéaires » tels qu'il les définit jouent un rôle qui dépasse la simple unification des cas qu'il a étudiés auparavant. Il y a une volonté de proposer un contexte indépendant des éléments particuliers qui ont été évoqués dans les

¹⁵Peano ne donne aucun exemple en dimension infinie mais mentionne tout de même cette possibilité.

¹⁶[Pea88, p.143]

chapitres antérieurs afin de penser la notion d'« opérations » de la façon la plus large et la plus générale possible.

La généralisation s'opère ici en abandonnant la nature spécifique des éléments de géométrie rencontrés précédemment —points, droites, volumes, « formations géométriques » (*formazioni geometriche*¹⁷) de diverses dimensions— qui sont remplacés par des « éléments » abstraits (*enti*) tout en conservant la notion de « dimension » (*dimensione*) qui peut désormais prendre toutes les valeurs entières ou être infinie.

Une fois les définitions posées, le texte de Peano se poursuit en mettant en avant deux aspects qui déterminent la nature de la généralité de ce texte.

Premièrement les définitions que nous venons de rappeler sont suivies de quelques propriétés des « ensembles linéaires ». Un premier théorème précise avec le vocabulaire abstrait juste introduit que l'on peut considérer des « systèmes » engendrés par n éléments « indépendants » (*independenti*) $\mathbf{a}_1, \dots, \mathbf{a}_n$ et que tout élément \mathbf{a} du « système » s'écrit alors de manière unique comme combinaison de ces éléments.

Deuxièmement, après avoir développé des considérations de cet ordre Peano envisage des « systèmes linéaires » dont les éléments dépendent d'un paramètre. Il fait, dans cette partie, référence aux résultats qui ont été établis dans le cas des petites dimensions (de 1 à 6) dans les chapitres précédents. Peano mentionne que les différentes notions de « dérivée, dérivées successives, intégrale indéfinie et définie » qui ont été introduites précédemment peuvent simplement se repenser ici en remplaçant les notions de « formations géométriques » par des « éléments d'un système linéaire » (*ente d'un sistema lineare*) quelconque. De plus il remarque que l'on peut conserver un certain nombre de théorèmes qui ne mettent en jeu que des sommes, des multiplications par un nombre ou les coordonnées¹⁸ :

Un ente variabile a d'un sistema lineare può essere funzione d'una variabile numerica t ; su questa funzione si possono dare le definizioni di derivata, derivate successive, integrale indefinito e definito che si ricavano dai N^o 66, 67 e 70, leggendo 'ente d'un sistema lineare' invece di 'formazione geometrica'; sono pure applicabili i teoremi ivi trovati, che si riferiscono puramente alle somme, moltiplicazioni per numeri, e coordinate.

Dans cette partie du livre de Peano on a donc un exposé qui synthétise en des termes abstraits un certain nombre de propriétés qui ont été étudiées dans les chapitres précédents pour des éléments de géométrie en dimension réduite. Peano a pris soin de faire reposer la notion de « système linéaire » sur des opérations algébriques (internes ou externes) : addition des éléments ou multiplication par un nombre.

¹⁷On peut penser dans un premier temps à ces « formations géométriques » comme des éléments de géométrie —points, figures planes, volumes, etc. assortis de certaines propriétés que nous ne redonnons pas ici. En fait cette idée développée par Peano est complexe et nécessiterait une étude séparée. Il développe pour ces éléments de nombreuses notions où interviennent à la fois des idées de convergence, des possibilités de mener des calculs algébriques, des façons de mesurer les éléments, etc.

¹⁸[Pea88, p.145]

Après avoir défini les « systèmes linéaires » Peano introduit les « opérations » en deux temps.

Il définit d'abord une notion de *dépendance continue* (termes d'observateur) entre des éléments d'un « système linéaire »¹⁹ :

Un ente y , funzione d'un ente x , dicesi *funzione continua*, se il limite della funzione è la funzione del limite. Si suppose naturalmente definito il limite dell'ente x e dell'ente y .

Le type de dépendance en jeu ici repose sur une notion de *fonction* que Peano ne précise pas mais sur laquelle il semble faire reposer la conception des « opérations » qui suit. L'énoncé précédent permet en outre de définir la notion de continuité quelle que soit la définition choisie de limite²⁰ en ne faisant référence qu'aux « éléments » du « système linéaire ».

Les « opérations distributives » sont tout d'abord introduites sans faire appel à la continuité²¹ :

DEF. Un'operazione \mathbf{R} , a eseguirsi su ogni ente \mathbf{a} d'un sistema lineare \mathbf{A} , dicesi distributiva, se il risultato dell'operazione \mathbf{R} sull'ente \mathbf{a} , che indicheremo con \mathbf{Ra} , è pure un ente d'un sistema lineare, e sono verificate le identità

$$\mathbf{R}(\mathbf{a} + \mathbf{a}') = \mathbf{Ra} + \mathbf{Ra}', \quad \mathbf{R}(m\mathbf{a}) = m(\mathbf{Ra}),$$

ove \mathbf{a} e \mathbf{a}' sono enti qualunque del sistema \mathbf{A} , ed m un numero reale qualunque.

Les « systèmes linéaires » permettent donc à Peano de définir la distributivité d'une opération en des termes abstraits et indépendants de la nature particulière des éléments sur lesquels l'opération agit.

Puis une liste de vocabulaire va préciser la notion d'opération et la relier à la notion de « fonction » et donc à la notion de dépendance à laquelle nous faisons allusion plus haut²² :

L'ente \mathbf{Ra} , cioè il risultato dell'operazione distributiva \mathbf{R} sull'ente \mathbf{a} , dicesi *funzione distributiva* di \mathbf{a} . Un'operazione distributiva si chiamerà anche *trasformazione lineare*, o *trasformazione senz'altro*. Se l'ente \mathbf{Ra} appartiene allo stesso sistema lineare \mathbf{A} cui appartengono gli enti \mathbf{a} , l'operazione \mathbf{R} si dirà una *sostituzione*. L'ente \mathbf{Ra} dicesi anche prodotto della trasformazione \mathbf{R} per l'ente \mathbf{a} .

Le contexte mis en place par Peano est construit à partir de situations géométriques où l'on considère des « transformations » entre espaces. Ceci se traduit ici par une conception d'« opérations » qui *transforment* des éléments d'un système en des éléments d'un autre système. Bien que l'étendue du concept soit en un sens plus large que celle que nous avons repérée chez Volterra ou Fréchet par exemple, pour lesquels le domaine d'arrivée est systématiquement

¹⁹[Pea88, p.145]

²⁰Dans un chapitre précédent Peano a introduit la notion de « limite » pour un volume variable, puis pour des « formations variables ».

²¹[Pea88, p.145]

²²[Pea88, p.145]

\mathbb{R} ou \mathbb{C} , il serait évidemment inadapté de penser que l'on a ici un énoncé d'une plus grande généralité.

Enfin Peano observe que la relation $\mathbf{Rma} = m\mathbf{Ra}$ lorsqu'elle est valable pour des nombres rationnels m (m è *commensurabile*) peut s'étendre à tous les nombres réels sous certaines conditions, et en particulier lorsque l'« opération » \mathbf{R} est continue.

Peano a donc dissocié la notion de « transformation », dont les propriétés sont données en des termes algébriques pour définir un concept général, et celle de « continuité », qui n'est qu'une façon parmi d'autres d'obtenir l'homogénéité ($\mathbf{Rma} = m\mathbf{Ra}$) étendue à tous les nombres (réels).

La fin du chapitre du livre de Peano fait évoluer le point de vue en mettant en place l'addition et la multiplication des opérations, leur inverse éventuel²³, les règles des puissances et l'exponentielle d'une opération définie par une série avec convergence automatique en dimension finie. Immédiatement la notion d'exponentielle est associée à une équation différentielle²⁴ :

Essendo t una variabile numerica, \mathbf{a}_0 un ente costante del sistema A , e fatto

$$\mathbf{a} = e^{\mathbf{R}t} \mathbf{a}_0,$$

sarà a funzione di t , che per $t = 0$ si riduce ad \mathbf{a}_0 , e che soddisfa all'equazione differenziale

$$\frac{d\mathbf{a}}{dt} = \mathbf{R}\mathbf{a}.$$

Le lien avec l'analyse ne s'arrête pas là dans ce texte de Peano et il va définir une notion de dérivée pour une fonction définie entre deux « systèmes linéaires »²⁵ :

Sia x un ente d'un sistema lineare A , e sia $y = f(x)$ una sua funzione, il cui risultato appartenga pure ad un sistema lineare B . Essendo x' un ente qualunque del sistema A , porremo per definizione del primo membro

$$\left(d \begin{matrix} x' \\ x \end{matrix} \right) f(x) = \lim \frac{1}{h} [f(x + hx') - f(x)]$$

il limite essendo ottenuto col far tendere il numero h a zero. L'ente $\left(d \begin{matrix} x' \\ x \end{matrix} \right) f(x)$ appartiene al sistema lineare B , ed è funzione di x ed x' . All'ente x' si dà il nome di differenziale della variabile indipendente x , e all'ente $\left(d \begin{matrix} x' \\ x \end{matrix} \right) f(x)$ quello di differenziale della funzione $f(x)$.

Peano montre entre autres que $\left(d \begin{matrix} x' \\ x \end{matrix} \right) f(x)$ est une fonction distributive de l'élément x' qu'il nomme « dérivée de l'élément y par rapport à l'élément x . »²⁶

²³Chez Peano l'inverse d'une opération est envisagé seulement dans le cas des transformations d'un espace à n dimensions qui transforment un système de vecteurs indépendant en un autre.

²⁴[Pea88, p.150]

²⁵[Pea88, p.151]

²⁶[Pea88, p.152]

Si l'inspiration a été puisée dans une certaine géométrie grassmannienne, on voit ici l'émergence d'idées qui s'en échappent. Dans le dernier chapitre intitulé « Applicazioni » Peano donne des propriétés de la différentielle qui sont énoncées dans le cadre abstrait des opérations sur les ensembles linéaires tels qu'il les a définis.

Cette analyse montre donc que chez Peano les « systèmes linéaires » ont été conçus dans un mouvement d'abstraction qui repense des situations géométriques ordinaires. Ils servent pour Peano à définir de façon abstraite une notion d'« opération distributive » très générale qu'il fonde en convoquant à la fois une notion de transformation et des propriétés algébriques qui sont pensées sous forme de « système linéaire ». Finalement les « opérations distributives » apparaissent comme un outil abstrait qui permet d'aborder des problèmes de géométrie, de substitutions, de systèmes d'équations ou des problèmes différentiels grâce à un concept unique.

On note de plus que les seules propriétés des « opérations distributives » qui sont étudiées par Peano sont des propriétés calculatoires (somme, produit, inverse). Ces propriétés algébriques sont bien sûr nécessaires pour former un outil général d'interprétation des situations multiples que nous avons rappelées plus haut. Mais, contrairement à ce que nous verrons plus tard chez Banach, Peano ne développe pas de théorème qui énoncerait les propriétés des opérations autres que leur capacité à se combiner selon les lois de l'algèbre.

5.2.2 PINCHERLE ET LES « OPÉRATIONS DISTRIBUTIVES »

Le deuxième exemple qui nous permettra de mieux comprendre la spécificité du travail de Banach concerne à nouveau une focalisation sur des opérations et l'articulation à une certaine notion d'espace.

Un peu plus d'une décennie après le texte de Peano, en 1901, Salvatore Pincherle écrit en collaboration avec Ugo Amaldi un livre intitulé *Le Operazioni Distributive e le loro Applicazioni All'Analisi*.

Le premier chapitre s'intitule « L'insieme lineare generale ad n dimensioni » et définit la notion d'espace linéaire (nous allons préciser le vocabulaire). Alors que Peano avait placé au début de son livre des éléments de logique générale et des relations permettant de manipuler les ensembles, ce qui fait que les notions de « système linéaire » ou d'« opération distributive » sont tout à fait détachées de cette introduction, Pincherle inclut la notion d'égalité à la liste des relations algébriques qui permettent de définir sa notion d'espace²⁷ :

1. Indichiamo con S una classe od *insieme* di enti per i quali ci basterà, in questi primi capitoli, di ammettere le proprietà enumerate in ciò che segue, senza che sia necessario, per ora, di definirli in modo più determinato. Designiamo gli enti di S col nome di *elementi* e li rappresentiamo con lettere greche minuscole: con lettere latine minuscole rappresentiamo, invece, numeri (reali o complessi).

2. Ammetteremo dapprima che, dati due elementi α e β di S , si possa riconoscere se fra essi passa o no una relazione, che sarà detta uguaglianza, caratterizzata dalle seguenti proprietà:

²⁷[AP01, pp.1-2]

- I. Se α è uguale a β anche β è uguale ad α .
 II. Se α è uguale a β e β è uguale a γ , α è uguale a γ .

Per esprimere che α è uguale a β si scriverà $\alpha = \beta$ e le proprietà sopra enunciate si scriveranno perciò:

- I. Se $\alpha = \beta$, è $\beta = \alpha$;
 II. Se $\alpha = \beta$ e $\beta = \gamma$, è $\alpha = \gamma$.

Risulta da queste definizioni che:

Un ente è uguale a sè stesso.

Dans un second temps mais dans un processus de définition continu sont introduites les notions relatives aux opérations algébriques qui permettent de définir les « espaces (ou ensembles) linéaires ». Tout d'abord l'addition est donnée sous forme de « loi » (*legge*) ou de « propriété » (*proprietà*). On y retrouve l'associativité, la commutativité et l'existence d'un élément neutre désigné par un symbole particulier²⁸ :

3. Ammetteremo poi che, dati duo elementi qualsivogliano α , β di S , si possa sempre dedurre da essi un terzo elemento determinato, il quale verrà indicato con $\alpha + \beta$. L'operazione con cui da α e β si deduce il terzo elemento $\alpha + \beta$ è caratterizzata dalle seguenti uguaglianze:

$$\begin{aligned} \text{[I.]} & \quad \alpha + \beta = \beta + \alpha, \\ \text{[II.]} & \quad (\alpha + \beta) + \gamma = \alpha + (\beta + \gamma), \end{aligned}$$

dove $\alpha + (\beta + \gamma)$, $(\alpha + \beta) + \gamma$ rappresentano gli elementi dedotti da α e $\beta + \gamma$, da $\alpha + \beta$ e γ rispettivamente, come $\alpha + \beta$ è dedotto da α e β . Gli elementi uguali $\alpha + (\beta + \gamma)$ e $(\alpha + \beta) + \gamma$ si indicano indifferentemente con $\alpha + \beta + \gamma$. L'operazione così definita dicesi *addizione*, ed $\alpha + \beta$ è detto *somma* di α e β .

Le uguaglianze I e II si dicono *leggi* o *proprietà* dell'addizione; la prima è detta *legge commutativa*, la seconda *legge associativa* dell'addizione. In una somma di quanti addendi si vogliono sarà lecito, in forza delle leggi I e II, aggruppare gli addendi e permutarne l'ordine come più piacerà.

4. Nell'insieme S ammetteremo l'esistenza di un elemento ω , che noi diremo *elemento zero*, o semplicemente *zero*, il quale, per qualsivoglia elemento α , soddisfa all'eguaglianza

$$\alpha + \omega = \alpha.$$

Designeremo senz'altro l'elemento ω col simbolo 0; ma sarà conveniente non confondere questo elemento con lo *zero* dell'insieme dei numeri.

Enfin la multiplication par un nombre est introduite en deux temps. D'abord à partir de l'addition répétée n fois qui permet de définir la multiplication par un

²⁸[AP01, p.2]

entier. Puis Pincherle admet que la relation obtenue s'étend à tous les nombres réels. En fait une remarque que l'on trouve dans cette citation montre que Pincherle conçoit de façon abstraite la multiplication par un nombre quelconque qui peut être rationnel, réel ou complexe et cette multiplication doit coïncider avec celle définie précédemment pour les entiers²⁹ :

5. Le somme

$$\alpha + \alpha, \alpha + \alpha + \alpha, \dots$$

saranno indicate rispettivamente con $2\alpha, 3\alpha\dots$; così, se m è un numero intero positivo, $m\alpha$ rappresenterà la somma di m addendi uguali ad α . Dalle leggi dell'addizione seguono immediatamente le seguenti uguaglianze:

- I. $m(\alpha + \beta) = m\alpha + m\beta,$
- II. $m\alpha + n\alpha = (m + n)\alpha,$
- III. $m(n\alpha) = (mn)\alpha = (nm)\alpha.$

Generalizzando ora queste proprietà, ammetteremo che dato un numero a ed un qualsivoglia elemento α , esista nell'insieme S un elemento, che noi designeremo con $a\alpha$ e diremo *prodotto* dell'elemento α pel numero a ; l'operazione con cui da α si deduce $a\alpha$ dicesi moltiplicazione dell'elemento α per il numero a .

Questa operazione è caratterizzata dalle seguenti proprietà. Per a numero intero e positivo, $a\alpha$ sia la somma $\alpha + \alpha + \dots + \alpha$ di a elementi uguali ad α ; per a e b numeri qualsivogliano, rimangono valide le proprietà precedenti, cioè:

- I. $a(\alpha + \beta) = a\alpha + a\beta,$
- II. $(a + b)\alpha = a\alpha + b\alpha,$
- III. $a(b\alpha) = (ab)\alpha,$

alle quali aggiungiamo l'altra:

- IV. $0\alpha = 0.$

Si noti che nell'ultima uguaglianza il simbolo 0 rappresenta al primo membro il *numero zero*, al secondo l'*elemento zero*.

A seconda dei casi i numeri, di cui si è parlato in questo § si potranno intendere sia razionali, sia affatto liberi (reali complessi). Nel seguito, quando non sia avvertito il contrario, parlando di numeri ci riferiremo ai numeri complessi.

6. Ad ogni elemento α corrisponde l'elemento $(-1)\alpha$ che, addizionato col primo, dà per somma lo zero in forza delle leggi II e IV del § precedente. Questo elemento si rappresenta con $-\alpha$ e si ha:

$$\alpha + (-\alpha) = 0.$$

Dati gli elementi α e β , l'elemento $\beta + (-\alpha)$ ha la proprietà, che addizionato con α dà per somma β , come segue immediatamente

²⁹[AP01, pp.2-3]

dalle leggi dell'addizione. L'elemento $\beta + (-\alpha)$ si indica più semplicemente con $\beta - \alpha$ e si dice *differenza* fra β ed α ; si dice poi *sottrazione* l'operazione con cui $\beta - \alpha$ si deduce da α e β .

Un insieme S che abbia le proprietà ammesse nei §§2, 3, 4 e 5 di così *insieme* o *spazio lineare*.

On peut bien sûr voir des similarités entre ce texte et l'introduction des « systèmes linéaires » de Peano que nous avons rappelées. Pourtant une différence majeure subsiste et provient du fait que les éléments convoqués dans les deux cas par les mathématiciens le sont à des fins différentes.

Chez Peano nous avons vu que le recours aux « systèmes linéaires » avait lieu à la fin de son livre et avait pour but de dépasser un certain nombre de situations qui étaient elles-mêmes présentées sous une forme hiérarchisée : géométrie de la ligne, puis du plan, de l'espace et évocation de situations jusqu'à la dimension 6. Les « systèmes linéaires » et les opérations sont un moyen de concevoir ces différentes situations sous une forme unique et abstraite, mais ils ne constituent pas le cadre d'une géométrie générale qui rendrait caducs les situations et les modèles antérieurs. Ils permettent de choisir le meilleur niveau de généralité et de proposer un vocabulaire commun aux différentes situations. Pourtant aucun des niveaux de généralité ne disparaît dans un cadre plus large : des façons de travailler adaptées pour chaque dimension restent valables mais les notions générales exposées à la fin de l'ouvrage mettent en évidence certaines similarités, certains outils communs

La notion de dérivation qui apparaît dans l'avant-dernier chapitre et qui est reprise de façon généralisée dans le dernier chapitre ne dépasse pas ce cadre. La dérivée est d'abord introduite dans les petites dimensions (formations de 1^{re}, 2^e ou 3^e espèce avec le vocabulaire de Peano) avec une légère différence par rapport à ce qui précède. En effet ici, même au niveau des petites dimensions, Peano propose un traitement unifié de la dérivée qui intervient dans un chapitre dédié à cette notion intitulé « Derivate » (cf. Ch. VIII p.128) Les formations des différentes espèces (1, 2 ou 3) y sont traitées de façon uniforme et simultanément³⁰ :

66. Se $A(t)$ è una formazione geometrica di specie qualunque, funzione d'una variabile numerica t , porreno:

$$A'(t) = \lim_{h} \frac{1}{h} [A(t+h) - A(t)].$$

La formazione $A'(t)$, che è della stessa specie di $A(t)$, si dice la *derivata* di $A(t)$.

Peano adopte donc ici une vision où les éléments (les formations géométriques) sont variables et dépendent d'un paramètre, ce qui permet d'adapter la notion de dérivée des fonctions ordinaires.

Bien que la notion de dérivée semble au départ un élément annexe de la conception de l'espace mis en place par les mathématiciens, elle est en fait un ingrédient qui permet d'éclairer de façon détournée le type de généralité à l'œuvre dans les différents textes.

³⁰[AP01, p130]

L'approche de Pincherle est différente et la position des définitions qui introduisent la notion d'« espace linéaire », en ouverture du livre, n'est pas fortuite. En effet, Pincherle va puiser à la fois dans la géométrie et dans l'algèbre symbolique des outils pour relire un certain nombre de situations d'analyse et plus précisément du calcul fonctionnel³¹.

Le point de vue adopté par Pincherle est rappelé en introduction ou il fait apparaître deux aspects. Tout d'abord on peut considérer des « classes de fonctions » qui sont constituées de toutes les « combinaisons linéaires » de certaines fonctions. On obtient ainsi ce que Pincherle nomme des « espaces à un nombre généralement infini de dimensions »³². Les « opérations distributives » permettent de généraliser et d'étendre la notion d'opération que l'on connaît en géométrie³³ :

Le operazioni distributive applicabili alle funzioni di una simile classe si presentano allora come una generalizzazione naturale di ciò che sono le omografie negli spazi lineari ad un numero finito di dimensioni; e questo concetto, tanto più se sussidiato da una notazione semplice ed espressiva, permetterà di intuire in modo sintetico, e colla guida di continuate analogie colla geometria, molteplici relazioni di composizione, di scomposizione, di classificazione in gruppi, di trasformazioni di siffatte operazioni.

Puis dans un second temps Pincherle introduit lui-aussi la dérivation, non pas comme résultat limite d'un rapport comme pour les fonctions ordinaires, mais comme une opération fondamentale symbolique qui va permettre d'engendrer d'autres opérations. Il s'agit donc d'une vision beaucoup plus proche de celle du calcul symbolique que nous avons rappelée en tout début de notre étude et que Pincherle avait explicitement convoquée lorsqu'il cherchait, quelques années plus tôt, à définir l'analyse fonctionnelle³⁴ :

Ma nei problemi d'indole funzionale, in cui l'ente incognito o variabile non è più un numero, bensì una funzione, la derivazione si presenta non più come una ricerca di limite, ma invece come un'operazione che si viene ad aggiungere, come elemento fondamentale di calcolo, a quelli già menzionati. Questi, in unione alla derivazione applicata un numero finito od infinito numerabile di volte, bastano alla costruzione di tutte le operazioni distributive che, applicate a funzioni analitiche, generano funzioni del pari analitiche: operazioni studiate già, ma ordinariamente sotto la forma di integrali definiti curvilinei. Forma, a dir vero, sommamente efficace e suggestiva, ma quasi divinata, non discendente dall'organismo del calcolo e non in tutto logicamente dipendente dai soli postulati fondamentali della scienza dei numeri: ragioni che forse indussero il WEIERSTRASS a

³¹Cette approche n'est pas nouvelle dans l'œuvre de Pincherle : voir son mémoire « Sur le calcul fonctionnel distributif » par exemple que nous avons mentionné plusieurs fois et qui propose un large panorama des idées qu'il a développées. (cf. [Pin97])

³²[AP01, Introduction p.iii]

³³[AP01, p.iii]

³⁴[AP01, p.iv]

privarsi sistematicamente del suo potente ausilio nelle sue lezioni sulla teoria delle funzioni analitiche.

Pincherle fait donc reposer sa conception sur les deux notions d'espace (de fonctions analytiques) et d'opération (sous une forme symbolique et abstraite). Les notions de limite et de représentation des opérations par des intégrales sont donc écartées au profit d'une vision plus algébrique et symbolique.

On a là un choix radicalement différent de celui que nous avons étudié chez Volterra et une façon de penser qui s'oppose aussi à celle que développe Fréchet dans la « première voie » que nous avons analysée et qui conduit à sa thèse. Malgré des différences notables que nous avons mises en valeur, c'est l'étude locale des fonctions, qui s'appuie sur des conceptions diverses de voisinages (terme d'observateur), que les mathématiciens convoquent comme modèle pour penser en des termes abstraits et généraux. Pincherle de son côté voit dans le développement analytique des fonctions la double possibilité d'une part de considérer les éléments comme des combinaisons linéaires d'un espace linéaire et d'autre part d'envisager les opérations dans leur interaction particulière avec les combinaisons linéaires des éléments sur lesquelles elles agissent. Le calcul symbolique et la distributivité sont les éléments qui permettent de penser les espaces de fonctions et les opérations générales que l'on y fait agir.

Nous contrasterons cette approche avec celle de Banach qui, bien qu'il développe une vision très détachée de la géométrie et même de la conception des « opérations » que l'on trouve ici, préférera lui-aussi une approche qui ne reprend pas l'aspect *différentiel* de la théorie des fonctions.

Le second chapitre du livre de Pincherle s'intitule « Generalità sulle operazioni » et nous permet de préciser sa façon de penser les opérations. La présentation adoptée par Pincherle se décompose en deux parties, la première traitant des opérations générales alors que la seconde se concentre sur les opérations distributives.

La première partie sur les opérations, intitulée « Le operazioni in generale », définit la notion d'opération entre deux systèmes S et S' comme une « correspondance » (*corrispondenza*) qui à un ou plusieurs éléments du premier système S fait correspondre un ou plusieurs éléments de S' (Pincherle précise en fait que chaque élément de S' est mis en correspondance). La notion d'« opération » est vue par Pincherle comme le mécanisme de production d'un élément de S' à partir de ceux de S , mécanisme qui installe une transformation de S en S' : « l'operazione così considerata *trasforma* S in S' . »³⁵ Dans le cas où la correspondance associe à chaque élément de S un unique élément de S' Pincherle parle de correspondance à « détermination unique » (*a determinazione unica*).

La notation adoptée par Pincherle est similaire à celle employée pour les fonctions³⁶ :

Se all'elemento α di S l'operazione A fa corrispondere l'elemento unico α' di S' , scriveremo

$$A(\alpha) = \alpha'$$

³⁵[AP01, p.17]

³⁶[AP01, p.17]

On remarque que Pincherle désigne par une lettre unique A l'opération en jeu. À son époque les fonctions étaient souvent désignées en faisant apparaître la variable en notant la fonction par exemple $f(x)$. En s'appuyant sur le calcul symbolique Pincherle utilise la notation $A(\alpha) = \alpha'$ qui lui permet de distinguer les rôles joués par les différents éléments A , α et α' . Cette manière d'appréhender les opérations comme dans le calcul symbolique permet d'écrire les calculs relatifs aux différents éléments de façon dissociée qui ne se joue pas simplement sur la différence de nature des éléments (fonctions-objets ou opérations agissant sur les fonctions).

Un des aspects importants de l'exposé de Pincherle se trouve dans son choix d'exposer prioritairement certaines propriétés des opérations vues comme « symboles d'opérations ». En effet Pincherle précise d'emblée comment les opérations doivent être manipulées du point de vue de leur calcul algébrique. En reprenant un schéma proche de celui qui a permis de définir les « espaces linéaires » il définit d'abord l'égalité de deux opérations qu'il note $A = B$, puis l'opération nulle notée O (on a donc $O(\alpha) = 0$), et l'« opération identique » ou « unité » (*operazione identica od operazione unità*) notée I .

Puis viennent la définition de l'addition de deux opérations, les règles d'associativité présentées de façon synthétique et sans la variable, le produit par un nombre réel ou complexe avec les diverses règles d'interactions entre les opérations algébriques³⁷ :

$$\begin{aligned} A + B &= B + A, \\ (A + B) + C &= A + (B + C), [\dots] \\ A + 0 &= A[\dots] \\ a(A + B) &= aA + aB, \\ (a + b)A &= aA + bA, \\ a(bA) &= (ab)A. \end{aligned}$$

Pincherle définit aussi la différence de deux opérations en s'appuyant sur la différence qu'il a définie dans les « systèmes linéaires ».

Ces opérations algébriques effectuées sur les opérations A , B , etc. permettent à Pincherle de considérer l'ensemble des opérations univoques comme un « nouvel ensemble linéaire ».

Dans une nouvelle étape Pincherle introduit le « produit » des opérations qui correspond à ce que nous appellerions désormais la composition³⁸ :

Siano A e B due operazioni, la prima delle quali trasformi lo spazio lineare S in S' , la seconda lo spazio lineare S' nello spazio lineare S'' . Si può allora porre

$$A(\alpha) = \alpha', \quad B(\alpha') = \alpha''$$

e considerare la trasformazione di S in S'' tale che ad α corrisponda α'' . Questa trasformazione è un'operazione che si dice prodotto di A per B ; essa viene designata con BA .

³⁷[AP01, p.19]

³⁸[AP01, p.20]

Pincherle définit alors l'associativité de la multiplication (il admettra que cette propriété est toujours valable dans les cas qu'il considère) et la commutativité dont il précise qu'elle ne sera pas toujours vérifiée. La multiplication des opérations permet alors de définir les puissances et d'écrire la « propriété d'itération » (*proprietà iterativa*) sous la forme $A^m A^n = A^{m+n}$ pour des entiers positifs m et n .

Enfin Pincherle met en évidence les propriétés de l'opération identique I par rapport à la multiplication ($AI = IA = A$) et note $A^0 = I$. Il introduit la notion d'inverse en disant que A' est l'inverse de A (*inverso di A*) si $AA' = 1$. Il remarque alors que dans le cas des transformations à détermination unique inversibles on a $A'A = I$, c'est-à-dire que A est l'inverse de A' . La notation finalement adoptée sera A^{-1} qui permet, comme le montre Pincherle, de combiner cette écriture avec celle précédemment adoptée de puissance d'une opération.

On pourrait préciser les détails de ces propriétés ici mais notre but est principalement de montrer la conception des opérations chez Pincherle et comment la notion d'« ensemble linéaire » est lié à celle d'opération. Ces éléments nous permettront de mieux appréhender les choix opérés par Banach un peu moins de vingt ans plus tard.

Pincherle isole dans une seconde partie l'étude des « opérations distributives » (*Le operazioni distributive*). Elles sont définies par la propriété de « distributivité » $A(\alpha + \beta) = A(\alpha) + A(\beta)$. Il s'agit ici d'un point important que nous avons déjà évoqué chez Fréchet et qui sera capital chez Banach.

Pincherle montre que pour une opération distributive A et pour tout entier naturel n on a :

$$A(n\alpha) = nA(\alpha).$$

Cette propriété repose dans son texte sur une itération finie de la propriété de distributivité. En remplaçant α par $\frac{\alpha}{n}$ dans l'égalité précédente, Pincherle obtient $A(\frac{1}{n}\alpha) = \frac{1}{n}A(\alpha)$ et donc pour tout nombre rationnel c :

$$A(c\alpha) = cA(\alpha).$$

Pincherle définit alors les opérations distributives à l'aide de deux propriétés³⁹ :

Chiameremo operazioni distributive quelle che godono delle proprietà espresse dalle equazioni [$A(\alpha + \beta) = A(\alpha) + A(\beta)$] e [$A(c\alpha) = cA(\alpha)$ per ogni numero e, reale o complesso].

Les opérations étudiées ensuite par Pincherle seront toutes distributives au sens de cette définition. Pourtant une note de bas de page précise que le passage, dans l'égalité $A(c\alpha) = cA(\alpha)$, de c rationnel à c quelconque peut être obtenu en faisant sur A des « restrictions appropriées, analogues à la continuité des fonctions. »⁴⁰

Le but poursuivi par Pincherle est de former des polynômes d'opérations distributives et d'utiliser cette vision pour montrer en particulier qu'une grande

³⁹[AP01, p.26]

⁴⁰[AP01, p.25]

partie des résultats du calcul fonctionnel peuvent être exprimés rigoureusement en concevant la dérivation D d'une façon abstraite et selon les modalités que nous venons de rappeler. En particulier il semble à Pincherle que cette approche, qui laisse de côté la notion de dérivée associée à la limite d'un quotient pour en faire un élément abstrait d'un calcul algébrique, permet de concevoir le calcul fonctionnel (avec des fonctions analytiques) sous une forme générale et efficace.

Nous pensons que les travaux exposés dans le livre de Pincherle en 1901, qui ont été parfois qualifiés d'assez pauvres en termes de résultats mathématiques, développent un point de vue important et nouveau. Nous avons vu que l'approche de Pincherle intégrait des idées qui n'étaient pas présentes dans le calcul symbolique tel que nous l'avons examiné au début de notre travail. Le livre de 1901 propose une approche qui éclaire le lien entre « opération » et conception des « ensembles linéaires ». Le modèle de la dérivation D , qui avait déjà servi dans l'élaboration du calcul symbolique tel que nous l'avons vu avec Gregory, est repris à nouveau dans une vision abstraite. La nouveauté vient du fait que Pincherle ne s'intéresse pas simplement aux possibilités de mener des calculs algébriques avec les opérations A, B, \dots . Il définit les opérations distributives en s'appuyant sur les opérations algébriques $(+, -)$ effectuées sur les objets sur lesquelles elles opèrent. La conception des « ensembles linéaires » permet de définir et d'établir les propriétés des opérations distributives. Les éléments des « ensembles linéaires », leurs opérations $(+, -)$ sont convoqués et assemblés par Pincherle pour former ce que nous appellerions désormais un *espace* et dans le but de permettre de concevoir les opérations de façon abstraite et générale.

La propriété de distributivité est donc essentielle dans la conception des « ensembles linéaires ». Celle-ci était déjà mise en lumière dans les textes sur l'algèbre symbolique et même dans des textes de Servois que nous avons rappelés. Elle est systématiquement utilisée dans la manipulation de la dérivation D .

On peut préciser encore un peu plus les choses ici. L'algèbre symbolique, telle que nous l'avons illustrée avec les textes de Gregory, tend à concevoir tous les éléments en jeu comme des symboles d'opération de façon uniforme et indifférenciée. Ici au contraire l'accent est mis sur une différence qui se conçoit au niveau des rôles joués par les opérations (distributives) d'un côté et par les opérations algébriques de l'autre. Il y a bien à la base une conception commune qui relève de l'algèbre symbolique et qui met en jeu les lois de ses opérations. Cependant la conception des « espaces linéaires » est développée à partir de la distinction des rôles attribués à ce que nous avons appelé des *symboles-objets* ou des *symboles-agissants*. Ici ce que nous pourrions rétrospectivement nommer *structure d'espace* se joue non seulement à travers les lois de l'algèbre qui sont convoquées (l'addition et la multiplication) mais aussi dans l'attribution des rôles aux différents éléments, c'est-à-dire de façon indissociable de la notion d'opération distributive, ou pour reprendre le vocabulaire utilisé dans notre analyse de l'algèbre symbolique, la notion de structure se joue dans la dépendance entre les symboles-objets et les symboles-agissants.

En fait Pincherle va encore un peu plus loin pour établir l'opération distributive comme concept central en considérant ce qu'il nomme des « formes

linéaires » (*forma lineare*) : ce sont des polynômes en A qu'il note $c_0 + c_1A + c_2A^2 + \dots + c_mA^m$. Ces derniers forment un groupe, commutent entre eux et peuvent se multiplier entre eux pour donner à nouveau une « forme linéaire ». C'est pour Pincherle le véritable concept général qui sert de base à toute sa théorie.

Ces exemples de constructions d'espaces que nous avons illustrés chez Peano ou Pincherle se conçoivent donc dans leur lien avec des opérations, transformations de l'espace en géométrie ou dérivation en analyse par exemple, toujours conçues de façon abstraite et générale. Nous allons pouvoir contraster ces approches avec celle de Banach en mettant à nouveau en lumière le rôle joué par les opérations linéaires.

5.2.3 LES ESPACES DE BANACH EN QUESTION

Le premier chapitre de la thèse de Banach est dédié à un certain nombre d'axiomes et de définitions qui définissent ce que l'on nomme rétrospectivement un *Espace de Banach*. En fait l'introduction qui précède cette première partie annonce une sorte de méthodologie qui cherche à établir une approche générale⁴¹ :

L'ouvrage présent a pour but d'établir quelques théorèmes valables pour différents champs fonctionnels, que je spécifie dans la suite. Toutefois, afin de ne pas être obligé à les démontrer isolément pour chaque champ particulier, ce qui serait bien pénible, j'ai choisi une voie différente que voici : je considère d'une façon générale les ensembles d'éléments dont je postule certaines propriétés, j'en déduis des théorèmes et je démontre ensuite de chaque champ fonctionnel particulier que les postulats adoptés sont vrais pour lui.

Cette introduction est immédiatement suivie d'une liste de « champs fonctionnels »⁴² :

J'ai introduit pour plus de simplicité les notations suivantes de quelques champs fonctionnels :

L'ensemble de fonctions continues (\mathcal{C})

L'ensemble de fonctions sommables (intégrables au sens de Lebesgue) (\mathcal{S})

L'ensemble de fonctions intégrables (L) avec la $r^{\text{ième}}$ puissance (\mathcal{S}^r)

L'ensemble de fonctions mesurables bornées (\mathcal{M})

L'ensemble de fonctions duhameliennes bornées (\mathcal{D})

L'ensemble de fonctions ayant la $(p - 1)^{\text{ième}}$ dérivée absolument continue et la $p^{\text{ième}}$ dérivée

continue ($\mathcal{C}\mathcal{C}$)

intégrable (L) ($\mathcal{C}\mathcal{S}$)

intégrable (L) avec la $r^{\text{ième}}$ puissance ($\mathcal{C}\mathcal{S}^r$)

⁴¹[Ban22, p.134]

⁴²[Ban22, p.134]

bornée ($\mathcal{C}\mathcal{M}$)
duhamelienne ($\mathcal{C}\mathcal{D}$)

Ces ensembles de fonctions que Banach nomme « champs fonctionnels » joueront plusieurs rôles que nous déterminerons au cours de notre analyse. C'est à partir d'eux que Banach entame un processus d'abstraction et qu'il isole des caractéristiques communes aux collections de fonctions qu'il vient d'énumérer. La démarche n'est pas sans précédent comme nous l'avons vu, et Banach lui-même cite en introduction divers mathématiciens auxquels il connecte ses travaux, dont Volterra et Fréchet que l'on retrouve de façon majeure dans notre étude, mais aussi Hadamard, Riesz, Pincherle, Steinhaus, Weyl ou Lebesgue.

Nous verrons que le matériau extrait par Banach sera organisé d'une façon particulière avec une visée double : d'abord constituer un cadre permettant de traiter de façon efficace et unifiée les ensembles de fonctions qu'il vient de mentionner et ensuite permettre un traitement général des opérations qui agissent sur des ensembles de fonctions. Comme nous l'avons déjà mentionné, ce deuxième aspect relatif aux opérations intervient de façon essentielle dans les choix opérés dès le début de la thèse dans la constitution des espaces de Banach. Un des objectifs de notre étude est de préciser cet aspect.

LES ESPACES VECTORIELS. La thèse de Banach s'ouvre dès le premier chapitre sur une liste d'axiomes qui sera complétée en trois temps. La première série d'axiomes définit ce que nous appelons de nos jours un *espace vectoriel*.

Banach n'utilise pas cette expression dans sa thèse. Les collections d'éléments qu'il convoque seront désignées par divers termes comme « classe », « système » ou « champ » qui serviront plus tard, essentiellement dans son livre de 1932⁴³, à différencier des situations mettant en jeu divers aspects que nous pourrions analyser en termes de structures.

Les axiomes que nous nommerons « axiomes d'espace vectoriel » pour simplifier les choses sont introduits dès le début de la première partie de la thèse sous la forme suivante^{44,45} :

§1. AXIOMES ET DÉFINITIONS FONDAMENTALES. Soit E une classe composée tout au moins de deux éléments, d'ailleurs arbitraires, que nous désignerons p. ex. par X, Y, Z, \dots

a, b, c désignant les nombres réels quelconques, nous définissons pour E deux opérations suivantes :

1) l'addition des éléments de E

$$X + Y, X + Z, \dots$$

⁴³Dans le livre *Théorie des opérations linéaires* de 1932 Banach utilise l'expression « espace vectoriel général » et nous montrerons comment ce vocabulaire accompagne l'aboutissement d'un changement de perspective et l'émergence d'une *structure ensembliste* pour reprendre une terminologie que nous avons introduite précédemment.

⁴⁴Remarque : les éléments de numérotation des paragraphes ou des théorèmes cités sont ceux du texte de Banach. Nous les avons incorporés aux citations afin de faciliter le retour aux sources. Nous utiliserons dans notre propre texte une numérotation des théorèmes du type *Théorème I.2.3* pour indiquer l'emplacement du théorème : [partie].[paragraphe].[numéro du théorème].

⁴⁵[Ban22, pp.134-135]

2) la multiplication des éléments de E par un nombre réel

$$aX, bY, \dots$$

Admettons que les propriétés suivantes sont réalisées :

I_1 $X + Y$ est un élément bien déterminé de la classe E ,

I_2 $X + Y = Y + X$,

I_3 $X + (Y + Z) = (X + Y) + Z$,

I_4 $X + Y = X + Z$ entraîne $Y = Z$,

I_5 Il existe un élément de la classe E déterminé θ et tel qu'on ait toujours
 $X + \theta = X$,

I_6 $a \cdot X$ est un élément bien déterminé de la classe E ,

I_7 $a \cdot X = \theta$ équivaut à $X = \theta$ ou $a = 0$,

I_8 $a \neq 0$ et $a \cdot X = a \cdot Y$ entraînent $X = Y$,

I_9 $X \neq \theta$ et $a \cdot X = b \cdot X$ entraînent $a = b$,

I_{10} $a \cdot (X + Y) = a \cdot X + a \cdot Y$,

I_{11} $(a + b) \cdot X = a \cdot X + b \cdot X$,

I_{12} $1 \cdot X = X$,

I_{13} $a \cdot (b \cdot X) = (a \cdot b) \cdot X$.

Nous introduisons en même temps les définitions suivantes :

(a)
$$-X = (-1) \cdot X,$$

(b)
$$X - Y = X + (-1) \cdot Y.$$

Il s'agit ici pour Banach d'introduire un « ensemble abstrait » en sélectionnant les éléments nécessaires pour concevoir certaines propriétés des champs de fonctions qu'il a énumérés et pour introduire la notion d'opération linéaires un peu plus tard.

Banach fait une brève coupure dans son exposé pour indiquer la portée générale que revêt une telle conception abstraite⁴⁶ :

Toutes les règles algébriques des signes $+$, $-$ et de la multiplication (d'un élément de E par un nombre) restent valables pour un système E . Comme exemples d'un tel système peuvent servir : les vecteurs, les formes de Grassmann, les quaternions, les nombres complexes etc.

On peut bien sûr, comme nous l'avons évoqué, trouver des définitions antérieures qui peuvent être rapproché d'un point de vue formel de cette liste d'axiomes. L'article de Jean-Luc Dorier intitulé « A general outline of the genesis of vector space theory »⁴⁷ permet de situer un certain nombre d'idées qui ont permis de dégager la notion d'espace vectoriel, essentiellement depuis des travaux de la moitié du 18^e siècle sur les systèmes d'équations jusqu'aux textes d'algèbre moderne de la première moitié du 20^e siècle.

⁴⁶[Ban22, p.135]

⁴⁷[Dor95]

Les exemples cités au-dessus avec les textes de Peano ou Pincherle montrent une coïncidence formelle qui pourrait laisser penser qu'il n'y a rien de bien nouveau ici. L'objectif de cette partie est de montrer que c'est dans la façon d'utiliser certains assemblages formels que se joue aussi la nature des mathématiques développées par un auteur. Pour apprécier précisément ce que fait un mathématicien avec une telle notion il faut l'appréhender dans une organisation plus large d'un article, d'une série d'articles ou d'un livre et dans son rapport avec d'autres éléments. De même l'éclairage que nous amènerons en contrastant le texte de Banach avec des textes antérieurs doit se faire non pas terme à terme, ni dans leur équivalence ou leurs différences formelles reconstruites à la lumière d'un concept pensé rétrospectivement, mais en contrastant les rôles joués par les notions dans un texte ou un autre.

En particulier les axiomes d'espace vectoriel que nous venons de rappeler dans le texte de Banach sont à considérer dans leur lien avec d'autres axiomes qui les suivent immédiatement, dans leur lien avec les « opérations » que le titre de la thèse met au centre de cette publication, et en analysant plus largement le rôle que jouent ces axiomes pour le mathématicien.

Le texte de Banach précise l'organisation mise en place en proposant dans une énumération continue de nouveaux axiomes qui définissent une norme⁴⁸ :

II. Il existe une opération⁴⁹ appelée norme (nous la désignerons par le symbole $\|X\|$), définie dans le champ E , ayant pour contre-domaine l'ensemble de nombres réels et satisfaisant aux conditions suivantes :

$$\text{II}_1 \quad \|X\| \geq 0,$$

$$\text{II}_2 \quad \|X\| = 0 \text{ équivaut à } X = \theta,$$

$$\text{II}_3 \quad \|a \cdot X\| = |a| \cdot \|X\|,$$

$$\text{II}_4 \quad \|X + Y\| \leq \|X\| + \|Y\|$$

$$\text{III Si } 1^\circ \{X_n\} \text{ est une suite d'éléments de } E, 2^\circ \lim_{\substack{r=\infty \\ p=\infty}} \|X_r - X_p\| = 0,$$

il existe un élément X tel que

$$\lim_{n=\infty} \|X - X_n\| = 0.$$

On voit donc ici que Banach convoque plusieurs éléments qu'il adjoint à la première liste d'axiomes pour constituer un assemblage qui embarque toutes les propriétés dont il aura besoin et qui sont communes aux ensembles de fonctions qu'il a énumérés. Les différentes propriétés sont classées en trois catégories qui montrent la structure de l'assemblage et introduisent un type de généralité particulier. En effet l'adjonction par lots de certaines propriétés permet d'envisager des situations intermédiaires et confère donc à la théorie une qualité générale que nous qualifierons de *généralité ouverte*. Bien que

⁴⁸[Ban22, pp.135-136]

⁴⁹La notion d'opération est définie par Banach dès les premières lignes. Elle est présentée de la façon la plus générale possible comme une relation binaire (termes d'observateur) entre les éléments de deux ensembles. Nous y reviendrons dans une section dédiée à cet aspect. Pour l'instant on peut penser aux opérations dans les termes modernes d'application de E dans \mathbb{R} .

Banach n'explore pas toutes les possibilités ainsi potentiellement offertes dans sa thèse, le livre de 1932 ainsi que d'autres travaux activeront ces possibilités de bifurcations. L'organisation des axiomes dans ce texte de Banach lui permettra en outre, comme nous allons le voir, de spécifier dans chaque démonstration très précisément les hypothèses qui sont nécessaires et sur lesquelles il s'appuie.

NORME, LIMITE ET COMPLÉTUDE. La notion de norme est très souvent reliée à l'idée de topologie. Le texte de Banach permet de préciser certains aspects qu'il nous semble essentiel de distinguer.

Dans la thèse de Banach la norme va intervenir à différents moments et sera utilisée à des fins tout à fait différentes. Dans ce premier paragraphe de la première partie, ainsi que dans le paragraphe qui suit, la norme ne sert qu'à définir la notion de limite et à préciser certaines propriétés.

Cette façon d'envisager de façon abstraite la notion de convergence n'est là encore pas nouvelle. On retrouve une formulation proche par exemple dans un texte de F. Riesz⁵⁰ :

Definitionen und Hilfssätze.

Den folgenden Betrachtungen legen wir die Gesamtheit der auf der Strecke $a \leq x \leq b$ erklärten, daselbst überall stetigen Funktionen $f(x)$ zu Grunde. Die Veränderliche x wird also für reell vorausgesetzt, dagegen sind als Funktionswerte auch komplexe gestattet. Doch will ich sofort betonen, dass unsere Entwicklungen auch für die engere Gesamtheit der reellen Funktionen ohne Weiteres gelten.

Die zu Grunde gelegte Gesamtheit werden wir der Kürze halber als *Funktionalraum* bezeichnen. Ferner nennen wir Norm von $f(x)$ und bezeichnen mit $\|f\|$ den Maximalwert von $|f(x)|$; die Grösse $\|f\|$ ist danach im Allgemeinen positiv und verschwindet nur dann, wenn $f(x)$ identisch verschwindet. Ferner bestehen für sie die Beziehungen

$$\|cf(x)\| = |c|\|f(x)\|; \|f_1 + f_2\| \leq \|f_1\| + \|f_2\|.$$

Unter Distanz der Funktionen f_1, f_2 verstehen wir die Norm $\|f_1 - f_2\| = \|f_2 - f_1\|$ ihrer Differenz. Danach ist die gleichmässige Konvergenz einer Funktionenfolge $\{f_n\}$ gegen die Grenzfunktion f gleichbedeutend damit, dass die Distanz $\|f - f_n\|$ gegen Null konvergiert. Eine notwendige und hinreichende Bedingung für die gleichmässige Konvergenz einer Folge $\{f_n\}$ besteht nach dem sogenannten allgemeinen Konvergenzprinzip in der Beziehung $\|f_m - f_n\| \rightarrow 0$ für $m \rightarrow \infty, n \rightarrow \infty$. Speziell wird also eine Folge $\{f_n\}$, für welche sämtliche Distanzen $\|f_m - f_n\|$ ($m \neq n$) eine von Null verschiedene, also wesentlich positive untere Schranke besitzen, keinesfalls gleichmässig konvergieren.

Malgré les ressemblances, un certain nombre de différences majeures apparaissent ici. Tout d'abord cet exemple met en évidence le procédé d'abstraction

⁵⁰[Rie16, p.72]

qui est à l'œuvre chez Banach quand il dégage l'idée de norme pour concevoir de façon uniforme les différents ensembles de fonctions qu'il a mentionnés.

Pour Riesz la norme est celle du supremum sur un intervalle compact $[a, b]$ qui semble naturelle lorsqu'on ne considère que des fonctions continues sur $[a, b]$. Il s'agit pour lui d'isoler certaines caractéristiques de l'ensemble $\mathcal{C}([a, b])$ pour faire apparaître précisément les propriétés qui permettent de définir et d'étudier des « transformations » (*Transformation*) linéaires et continues⁵¹. Il y a bien un travail d'abstraction et la mise en place de notations synthétiques qui confèrent au texte de Riesz un caractère général. Il y a d'une part la généralité qui découle du fait qu'il traite de façon uniforme tous les « espaces fonctionnels » (*Funktionalraum*) du type $\mathcal{C}([a, b])$, quel que soit l'intervalle où sont définies les fonctions. Mais il y a encore un aspect qui relève aussi du général ici, et qui consiste en quelques sortes à concevoir *a posteriori* —il n'y a aucune preuve que Riesz lui-même pense dans cette direction— l'espace $\mathcal{C}([a, b])$ comme archétype des espaces vectoriels normés complets. On ne peut pas négliger cet aspect dans une analyse de la portée générale du texte de Riesz. Pourtant nous allons montrer que l'on ne peut pas comparer des textes seulement en termes d'étendue des situations couvertes, que l'on considère cette étendue donnée directement et explicitement par les auteurs eux-mêmes ou qu'elle soit conçue rétrospectivement par un observateur. Il nous faut à nouveau considérer précisément ce que les auteurs font des idées qu'ils convoquent, c'est-à-dire introduire en quelques sortes un élément dynamique dans notre analyse qui permette de distinguer différentes natures d'expression du général, différents types de généralité.

La différence essentielle entre les deux situations et qui singularise aussi le travail de Banach dans sa thèse par rapport aux situations antérieures que nous avons étudiées, c'est le fait de mélanger à la fois une approche *synthétique* et une approche *constructive*. Le point de vue synthétique est clair dans ce texte et la longue liste des espaces fonctionnels du début de sa thèse montre qu'un premier travail de sélection et d'abstraction s'effectue à partir des collections de fonctions qu'il énumère.

Nous avons analysé une démarche similaire chez Fréchet dans sa série d'articles sur les opérations linéaires. Pourtant ce que nous donne à voir Banach ne relève pas exclusivement d'une synthèse, ni comme dans l'exemple de Riesz de l'identification de certaines caractéristiques communes à diverses situations.

La liste d'axiomes donnée par Banach reflète un travail d'abstraction différent : des propriétés énoncées de façon abstraite sont organisées de façon particulière comme nous l'avons vu. Là encore on peut contraster le texte de Banach avec celui de Riesz. Pour le mathématicien hongrois les différentes propriétés qu'il sélectionne sont données sur un même plan. Il énumère quelques-unes des propriétés de la norme : la multiplicativité, l'inégalité triangulaire ou la complétude sont *coexistantes*, ce sont des qualités de la norme du supremum sur $\mathcal{C}([a, b])$.

Banach de son côté élabore une nouvelle *construction* en organisant plusieurs propriétés bien déterminées dans la liste d'axiomes. Il s'agit de ce que nous nommons un procédé *constructif* et les éléments en jeu sont conçus par Banach

⁵¹Nous reviendrons un peu plus loin sur la façon d'écrire ces propriétés chez Riesz ou Banach.

d'une façon qui se veut indépendante des ensembles de fonctions du début du texte. Le vocabulaire employé par Banach montre ce choix de dépasser une synthèse aussi abstraite soit-elle : « Soit E une classe [...] composée [d']éléments [...] arbitraires », « admettons que les propriétés suivantes sont réalisées », « Admettons [qu'] il existe une opération appelé norme » etc.

On a donc une liste de propriétés, qui sont de plus en plus restrictives, et qu'un espace donné peut posséder en totalité ou seulement en partie. D'ailleurs, une longue dernière partie de la thèse de Banach sera dédié à l'examen des ensembles de fonctions cités au départ et il y montrera que chacun vérifie l'ensemble des propriétés établies en introduction.

Banach cherche donc non seulement à établir une conception qui recouvre beaucoup de situations, celles données par la liste des ensembles de fonctions, auxquelles on peut rajouter celles rappelées par le biais de référence à divers auteurs dont Hadamard, Fréchet, Pincherle, Volterra, Riesz, Steinhaus, Weyl, Lebesgue ou Hilbert, mais il cherche aussi à dépasser ces situations en établissant ce qu'il nomme des « ensembles abstraits » —expression tirée de son titre⁵²— par le biais d'une liste d'axiomes.

Les deux aspects, synthétique et constructif, sont donc présents dans ce texte et ils participent du type de généralité que Banach met en place dans sa thèse.

Enfin la norme permet à Banach de définir la convergence d'une suite d'éléments de l'« ensemble abstrait »⁵³ :

DÉFINITION 1. Nous dirons que la suite $\{X_n\}$ tend vers X suivant la norme, ce que nous écrirons

$$\lim_{n \rightarrow \infty} X_n = X,$$

lorsque

$$\lim_{n \rightarrow \infty} \|X - X_n\| = 0.$$

Cette définition, ainsi que l'axiome III (sur la complétude) présuppose une autre notion de convergence, celle des réels. Ceci montre que Banach n'est pas à la recherche de ce que nous pourrions appeler un *exposé général de topologie*. Il convoque et organise un certain nombre d'éléments qui vont lui permettre d'étudier de façon abstraite des opérations. Les nombres réels ou complexes ont donc leurs propres propriétés qui ne sont pas repensées avec les nouvelles idées de norme ou de limite. Banach propose donc un assemblage où les différentes entités convoquées ne sont pas appréhendées selon le même procédé d'abstraction. Si l'ensemble des réels ou des complexes peuvent être repensés en termes d'« ensembles abstraits », il n'en reste pas moins que ces conceptions

⁵²Il faut être ici prudent avec le vocabulaire employé. En fait Banach ne désigne pas l'assemblage résultant de la liste des axiomes par une expression précise. Nous employons l'expression « ensemble abstrait » parce qu'elle figure dans le titre, même si elle ne désigne pas précisément chez Banach tout ensemble qui posséderait les propriétés énumérées dans la liste d'axiomes. Nous employons aussi cette expression pour éviter l'expression « espace de Banach », non seulement parce qu'elle est conçue rétrospectivement, mais plus essentiellement parce que nous souhaitons discuter par la suite de certains aspects qui seraient occultés par son emploi. Pour être précis il faudrait écrire « ensemble abstrait qui vérifie l'ensemble des propriétés... » que nous abrégons donc ici.

⁵³[Ban22, p.136]

reposent toujours sur une vision *à-côté* et préalable de \mathbb{R} ou \mathbb{C} avec une notion de convergence qui n'est pas redéfinie mais qui est utilisée⁵⁴.

Banach définit ensuite la convergence des séries car elles constituent un outil central de la démarche qu'il va mettre en œuvre pour étudier les opérations sur un « ensemble abstrait ».

Le second paragraphe est consacré à établir certaines propriétés de la norme et des limites de suites ou de séries d'éléments d'un « ensemble abstrait ». Ce paragraphe contient deux types de résultats. La grande majorité d'entre eux établit des propriétés qui seront utilisées dans les différentes démonstrations pour alléger la démarche, créer des automatismes et éviter certaines répétitions. Il s'agit de propriétés de la norme comme des inégalités déduites de l'inégalité triangulaire ($\|X\| - \|Y\| \leq \|X - Y\| \leq \|X\| + \|Y\|$) ou adaptées aux séries ($\|\sum_{n=1}^{\infty} a_n \cdot X_n\| \leq \sum_{n=1}^{\infty} |a_n| \cdot \|X_n\|$), ainsi que des propriétés de la limite (par exemple l'unicité de la limite quand elle existe, la linéarité, etc.)

Un théorème de ce paragraphe doit être distingué car il a une utilité différente des autres. Il s'agit du théorème 8 de cette première partie que Banach énonce ainsi⁵⁵ :

THÉORÈME 8. Lorsque 1° $\{X_n\}$ est une suite d'éléments,
2° $\sum_{n=1}^{\infty} \|X_n\|$ existe, la série $\sum_{n=1}^{\infty} X_n$ est convergente suivant la norme.

Nous donnons la preuve de ce théorème du fait de son importance et aussi parce qu'il illustre bien le style de Banach qui précise à chaque fois les axiomes qui sont en jeu :

DÉMONSTRATION. Posons $S_r = \sum_{n=1}^r X_n$. En admettant que $p > q$,
on a

$$(4) \quad \|S_p - S_q\| = \left\| \sum_{n=q+1}^p X_n \right\| \leq \sum_{n=q+1}^p \|X_n\|.$$

mais par hypothèse 2° :

$$\lim_{\substack{p \rightarrow \infty \\ q \rightarrow \infty}} \sum_{n=q+1}^p \|X_n\| = 0,$$

il résulte donc de l'inégalité (4) que

$$\lim_{\substack{p \rightarrow \infty \\ q \rightarrow \infty}} \|S_p - S_q\| = 0,$$

ce qui prouve, en vertu de l'axiome III et selon la définition 1, la convergence de la suite $\{S_p\}$. Or, la convergence de la suite $\{S_p\}$ est,

selon la définition 2, équivalente à celle de la série $\sum_{n=1}^{\infty} X_n$.

⁵⁴D'ailleurs on peut remarquer que même dans le livre de Banach de 1932 il n'y a aucune partie préliminaire sur ce que nous appellerions la topologie du corps de base, ici \mathbb{R} ou \mathbb{C} . La première phrase du livre est « Nous admettons que le lecteur connaît la théorie de la mesure et de l'intégrale de Lebesgue. »

⁵⁵[Ban22, pp.138-139]

Comme on le lit dans la démonstration, ce théorème utilise l'inégalité triangulaire, un résultat sur les séries de nombres réels positifs et l'hypothèse que les suites de Cauchy (termes d'observateur) sont convergentes, c'est-à-dire l'axiome III. Ce qui le distingue des autres résultats du paragraphe c'est essentiellement le rôle qu'il va jouer dans la suite comme nous allons le voir.

D'autre part ce théorème indique que la convergence d'une série de nombres (les normes ici) implique la convergence d'une série d'éléments de l'ensemble abstraits (les X_n). Dans la suite nous retrouverons diverses façons de relier des éléments abstraits d'un ensemble à des grandeurs, c'est-à-dire, en quelques sortes, diverses façons de faire de la *géométrie dans les ensembles abstraits*. En effet il s'agit, ici par le biais des normes, plus tard à travers la notion de « sphère » (des boules fermées pour la norme), de *mesurer* l'espace abstrait. Mais le premier but n'est pas, comme nous allons le voir dans la suite, de déterminer des propriétés de l'« ensemble abstrait » par le biais de cette mesure, mais de déterminer des propriétés de la façon de mesurer, donc ici pour commencer, de déterminer des propriétés de la norme. Il s'agit bien sûr d'une perspective d'observateur que nous proposons d'explorer et d'étayer tout au long de l'analyse que nous menons.

Si on adopte cette optique on s'aperçoit que dès le départ le texte de Banach donne une seule collection d'axiomes fixée une fois pour toutes pour définir l'« ensemble abstrait » (un espace vectoriel) et cinq axiomes (ou collection d'axiomes) pour définir différentes propriétés des normes et donc différents types de normes. Notre lecture cherche donc à distinguer les rôles que vont jouer les notions d'« ensemble abstrait » et de norme dans le travail de Banach. Nous verrons que l'utilisation que fait Banach de ces deux notions permet de montrer qu'il développe une certaine géométrie où les enjeux liés à la norme ou à la structure algébrique de l'espace sont différents. Cette distinction permettra en outre de constater que l'étude des opérations, dont nous montrerons qu'elle est centrale ici, s'organise aussi essentiellement à partir des propriétés de la norme. Nous reviendrons sur ces aspects tout au long de notre étude.

Pour le moment, notons l'importance du théorème 8 au-dessus qui va être utilisé par Banach à de nombreuses occasions pour déduire d'une convergence en termes de rayons, donc de nombres réels, une convergence en termes de points de l'ensemble abstrait.

SPHÈRES ET TOPOLOGIE. Un troisième paragraphe de la première partie de la thèse de Banach met en place la notion de « sphère » qui va jouer un rôle fondamental dans l'étude des « opérations » définies sur un « ensemble abstrait ». Elles sont introduites sous forme de définition⁵⁶ :

DÉFINITION 3. X_1 désignant un élément de E et r un nombre positif, l'ensemble de tous les éléments (X) qui satisfont à l'égalité $\|X - X_1\| \leq r$ est dit *sphère*⁵⁷. L'élément X_1 sera appelé *centre* et r – *rayon* de la sphère. Nous la désignerons par le symbole $K(X_1, r)$.

⁵⁶[Ban22, p.140]

⁵⁷Banach désigne par « sphère » ce que nous appelons de nos jours une *boule fermée*. Nous n'utiliserons le mot *sphère* nous-même que dans ce sens.

Banach convoque ici un concept de géométrie *ordinaire* dont il propose une version adaptée au contexte abstrait qu'il développe. La notion de « sphère » doit garder son efficacité géométrique et parler à l'esprit. Banach propose d'ailleurs d'employer ici « l'expression „point” au lieu d'„élément” » et il donne quelques résultats qui établissent des notions qui sont usuelles dans le cas de la géométrie ordinaire⁵⁸ :

THÉORÈME 10. Étant donné deux sphères : $K_1(X_1, r_1)$ et $K_2(X_2, r_2)$, l'inégalité

$$\|X_1 - X_2\| \leq r_1 - r_2$$

est une condition nécessaire et suffisante pour que la sphère K_2 soit située dans la sphère K_1 .

[...] LEMME 7. Lorsque la sphère $K_1(X_1, r_1)$ est contenue dans la sphère $K_2(X_2, r_2)$, on a l'égalité $r_1 \leq r_2$.

LEMME 8. Lorsque la sphère $K_1(X_1, r_1)$ est contenue dans la sphère $K_2(X_2, r_2)$ et la sphère K_2 est contenue dans la sphère K_1 , on a

$$X_1 = X_2 \text{ et } r_1 = r_2.$$

Ces énoncés établissent donc de façon rigoureuse un certain nombre de propriétés que l'on attend d'une notion de « sphère » transposée de la géométrie ordinaire au contexte des « ensembles abstraits ».

Le paragraphe propose de plus un certain nombre de propriétés topologiques (Banach n'emploie pas ce terme) énoncées à l'aide de « sphères » ou de « suites de sphères ». D'abord un premier théorème pour des suites de sphères emboîtées⁵⁹ :

THÉORÈME 11. 1° $\{K_n(X_n, r_n)\}$ étant une suite de sphères, 2° la sphère K_n contenant K_{n+1} pour chaque n , il existe une limite de la suite $\{X_n\}$ et elle appartient à toutes les sphères simultanément.

La démonstration de ce résultat utilise le THÉORÈME 8 que nous avons rappelé plus haut et donc l'axiome de complétude III.

Enfin le paragraphe se termine sur une série de définitions et deux théorèmes qui seront utilisés pour démontrer un premier résultat important sur les opérations définies sur un « ensemble abstrait » —ayant les propriétés décrites par les axiomes. Banach donne la définition de « point d'accumulation », « ensemble dérivé », « ensemble fermé », « ensemble parfait », « ensemble dense » dans un autre ensemble en utilisant les « sphères » introduites au-dessus. Deux théorèmes sur lesquels nous ne nous attardons pas montrent qu'une intersection non vide et dénombrable de fermés est fermée d'une part et que tout point hors d'un fermé est le centre d'une sphère qui ne rencontre pas ce fermé.

Cette partie laisse donc apparaître des différences importantes avec la précédente qui donnait des propriétés de la limite. Ici Banach introduit à

⁵⁸[Ban22, p.141]

⁵⁹[Ban22, p.143]

nouveau des définitions qui donnent aux sphère et aux outils que nous qualifions de topologiques une place à part que nous voudrions préciser.

D'un certain point de vue rétrospectif et avec des termes d'observateur actuels, ce paragraphe semble donner des propriétés topologiques d'un espace de Banach, et de ce fait pourrait laisser penser que Banach introduit dans ce passage un point de vue *réflexif* sur les « ensembles abstraits » qu'il a définis grâce aux axiomes, c'est-à-dire sur les *espaces de Banach*. Mais cette idée doit être revue à la lumière d'une étude plus précise : en effet chacun des théorèmes que Banach prouve dans ce paragraphe sera utilisé pour démontrer un résultat concernant les « opérations ». Le fait que ces résultats aient été ramenés dans une partie isolée du texte n'est bien sûr pas sans conséquences, et ce mouvement structure le discours de Banach. Dans son livre de 1932 l'organisation du discours qui vise à mettre en évidence le rôle des différents éléments qui sont convoqués sera encore accentuée. Ce passage introduit probablement un embryon de regard *réflexif*, et tout du moins une pause dans le discours qui clôt une première partie et annonce l'étude des opérations. Mais il nous semble que dans ce texte il est encore trop tôt pour affirmer que Banach commence une étude des « ensembles abstraits » pour eux-mêmes. Tous les résultats seront en effet réutilisés plus tard pour simplifier des preuves et mettre en évidence certaines progressions logiques concernant les opérations.

C'est donc l'articulation au reste du texte et à l'étude des opérations qui suit qui permet de préciser le type de généralité que l'on peut lire dans cette première partie. L'identification dans ce texte d'espaces de Banach occulte en outre le rôle que Banach attribue à ses axiomes et aux propriétés qu'il établit en première partie de son texte. Nous contrasterons ce point de vue avec celui que l'on peut lire dans le livre de 1932 à plusieurs reprises dans notre analyse.

5.2.4 LES OPÉRATIONS SUR DES ENSEMBLES ABSTRAITS

Le travail sur les « opérations » est central dans la thèse de Banach et nous allons montrer que leur traitement conditionne chez lui l'organisation des éléments qui constituent et définissent les « ensembles abstraits ».

Plus précisément les « opérations » apparaissent dès la première partie en deux temps qui viennent enserrer l'exposé sur les ensembles abstraits, les limites et les normes. Cette présence des « opérations » dans la partie qui développe les « ensembles abstraits » établit le lien qui existe entre ces deux pans du travail de Banach qui ne peuvent donc pas être analysés indépendamment l'un de l'autre.

La thèse de Banach s'ouvre sur une première définition d'« opération » à laquelle nous avons déjà fait référence et que nous précisons ici⁶⁰ :

L'opération c'est une relation univoque $y\mathfrak{R}x$ c'est-à-dire, telle que

$$y\mathfrak{R}x \text{ et } z\mathfrak{R}x \text{ entraîne } y = z$$

pour tout x, y, z .

Chaque relation yRx comporte un *domaine* (c'est la réserve des x)⁶¹ est un *contre-domaine* (la réserve des y) ou *champ*.

⁶⁰[Ban22, p.133]

⁶¹Nous soulignons ici car le texte de Banach indique en fait que le « domaine » constitue la réserve des y alors que le « contre-domaine » est celle des x . Nous pensons qu'il s'agit là d'une

Banach, contrairement à ce que nous avons vu chez Fréchet ou Volterra par exemple, ne rattache pas explicitement la notion d'« opération » à celle de fonction. Le choix de désigner la mise en relation des éléments par $x\mathfrak{R}y$ permet de penser au « domaine » et au « contre-domaine » de la même façon, c'est-à-dire comme deux ensembles abstraits.

En fait Banach n'utilisera pas cette notation ailleurs dans son texte et le seul moment où l'on ne dispose que de cette définition pour concevoir la notion d'« opération » se trouve dans la deuxième liste d'axiomes qui définissent la norme. Celle-ci est conçue comme une « opération [...] définie dans le champ E , ayant pour contre-domaine l'ensemble des nombres réel »⁶².

Le mot « opération » est aussi utilisé dans les premiers axiomes pour définir l'addition et la multiplication par un nombre réel mais aucune indication ne permet d'affirmer que Banach pense ces opérations grâce au concept de « relation » \mathfrak{R} qu'il a introduit dès le départ.

Enfin un paragraphe dédié aux opérations ⁶³ vient clore la première partie du travail de Banach. Ainsi les « ensembles abstraits » ne sont pas détachés de la notion d'opération et ne sont conçus principalement que dans leur dépendance à la notion d'opération. L'agencement du texte et la présence des opérations dans la continuité de cette première partie éclaire le rôle joué par les « ensembles abstraits » et la liste des axiomes. Ce n'est que dans l'analyse du travail sur les opérations que l'on peut saisir les mécanismes de sélection et d'abstraction que Banach a mis en œuvre pour produire cette liste d'axiomes et que l'on pourra en dégager un certain type de généralité.

Dès le début du paragraphe introduisant les opérations dans la première partie, Banach précise que celles-ci mettent en jeu un « domaine » et un « contre-domaine » qui satisfont aux différents axiomes du départ. Ces deux « domaines » sont donc considérés comme nous le soulignons plus haut de la même manière comme deux « ensembles abstraits ». Banach signale à ce sujet qu'il gardera les mêmes notations d'addition et de multiplication pour les deux ensembles, mais qu'il faut les considérer comme deux ensembles bien distincts, chacun avec leurs propres opérations qui vérifient les propriétés énoncées par les axiomes.

Dès ce moment Banach abandonne la notation \mathfrak{R} qui avait servi à définir une « opération » comme « relation » pour adopter une notation qui rappelle la notation des fonctions avec la variable attachée au nom de la fonction. L'étude des opérations va alors occuper la majeure partie de la suite du travail de thèse de Banach et va être organisé en plusieurs temps que nous allons analyser.

coquille dans le texte de Banach. La lecture de la thèse, des autres articles et du livre de 1932 ne laisse pas de doute sur le vocabulaire : le domaine est bien l'ensemble de départ et le contre-domaine celui de l'arrivée. Dans ce cas l'interprétation stricte du texte de Banach conduit à définir des opérations injectives, ce qui ne correspond pas encore au mot « uniforme » employé par Banach. Le plus probable est qu'un élément x n'est en relation qu'avec un seul y , ce qui correspond à l'emploi du mot « uniforme » au début du 20^e siècle.

⁶²[Ban22, p.135]

⁶³Nous utilisons le mot *opération* de différentes façons qu'il nous faut distinguer tout au long de notre travail. Les opérations $+$, $-$, \times , etc. seront toujours nommées *opérations algébriques*. Le terme *opération* employé dans les autres cas désigne dans cette partie sur les travaux de Banach une « opération » définie sur un « ensemble abstrait », c'est-à-dire ce que nous appelons de nos jours un *opérateur* sur un espace de Banach.

CONTINUITÉ. Nous avons déjà vu que la notion de continuité était importante dans de nombreux textes d'analyse fonctionnelle. Non seulement elle permet de développer des propriétés des opérations, mais elle conditionne aussi la façon de concevoir les ensembles sur lesquels opèrent les opérations.

Ici la continuité est donnée par Banach à l'aide des notions topologiques qu'il a introduites plus haut⁶⁴ :

DÉFINITION. L'opération $F(X)$ est *continue pour un point* X_0 *relativement à un ensemble* A , lorsque

- 1° $F(X)$ est défini pour chaque point de l'ensemble A ,
- 2° X_0 appartient à A et est son point d'accumulation⁶⁵,
- 3° $\lim_{n \rightarrow \infty} F(X_n) = F(X_0)$ toutes les fois que la suite $\{X_n\}$ est contenue dans A et $\lim X_n = X_0$.

On peut contraster cette définition avec celle, introduite un peu plus loin, d'« opération uniformément continue »⁶⁶ :

DÉFINITION. L'opération $F(X)$ sera dite *uniformément continue* *relativement à l'ensemble* A , lorsque

- 1° $F(X)$ est défini pour chaque point de cet ensemble.
- 2° Pour tout nombre positif ϵ on peut trouver un $m > 0$ tel que, X et X' étant deux éléments quelconques de A l'inégalité, $\|X - X'\| \leq m$ entraîne

$$\|F(X) - F(X')\| \leq \epsilon.$$

Les deux définitions mettent en jeu des idées différentes et il émane de ces énoncés deux types distincts de généralité. La première définition utilise le symbole *lim* et fait donc appel à la définition de limite que Banach a introduite au départ. Il faut se souvenir que cette notion de limite repose elle-même sur une définition de limite de suites réelles qui n'est pas redéfinie par Banach : $X_n \rightarrow X$ si $\|X_n - X\| \rightarrow 0$. Nous avons vu à plusieurs reprises que ce type de définition s'adaptait dès que l'on pouvait *mesurer* l'écart entre deux éléments d'un ensemble abstrait par un nombre réel (bien sûr avec quelques propriétés supplémentaires suivant les cas). Ici c'est la norme qui permet d'effectuer cette *mesure* et dont Banach se sert pour définir la continuité d'une opération quelconque entre deux « ensembles abstraits ». Banach fait un choix particulier en définissant ce que nous qualifierions de *continuité séquentielle* qui sera adaptée aux espaces de fonctions qu'il traitera à la toute fin de son travail. Il s'agira, en des termes modernes, d'espaces vectoriels normés complets de fonctions pour lesquels Banach précisera ou reprécisera la norme.

Nous avons vu d'autres choix, opérés par exemple par Volterra concernant la continuité grâce à la notion abstraite de voisinage tubulaire, que ne reprend pas ici Banach. Celui-ci est à la recherche du cadre général le mieux adapté

⁶⁴[Ban22, p.145]

⁶⁵Banach a défini la notion de « point d'accumulation » grâce aux « sphères » : quelle que soit la sphère de centre X_0 elle rencontre A en au moins un point distinct de X_0 .

⁶⁶[Ban22, p.148]

pour traiter les ensembles de fonctions dont il a établi une longue liste en introduction. Il développe dans ce travail une théorie des opérations pour des espaces vectoriels normés, et l'ensemble des définitions qu'il introduit reposent donc sur la notion de norme, et par leur biais sur la notion de limite dans les conditions que nous venons de rappeler. Il ne s'agit pas de définir la vision la plus étendue possible, mais de définir une vision qui en quelques sortes se suffit à elle-même, en s'appuyant sur des axiomes ou des définitions posées dès le départ ou en utilisant des principes mathématiques courants (propriétés des réels, des suites de réels, etc.)

La définition de l'uniforme continuité n'utilise pas le symbole *lim* et fait apparaître un traitement différent. Alors que la présence des ϵ pourrait faire penser qu'elle met en œuvre une vision plus abstraite, elle repose en fait de façon essentielle sur la notion de norme. La notion de limite et le symbole *lim*, employés dans la définition de la continuité, introduisent en fait une certaine distanciation en désignant une propriété qui peut, dans un second temps et par le biais de la définition de limite, être exprimée grâce à la norme. Cette distanciation n'est plus présente dans la définition de l'*uniforme continuité* qui est donnée cette fois-ci explicitement en termes de norme.

Bien que nous soulignons cette différence subtile qui apparaît entre les deux définitions, le texte de Banach installe une cohérence et définit un contexte bien précis en faisant reposer ses définitions essentiellement sur la notion de norme.

LE STYLE DE BANACH. C'est réellement dans le développement de l'étude des « opérations » que l'on peut voir à l'œuvre la façon d'organiser certaines idées mathématiques chez Banach. En particulier, les théorèmes qui ont été développés dans les parties précédentes vont lui permettre d'exposer des preuves qui font ressortir clairement les moments où interviennent les différents axiomes et qui mettent en avant la progression et l'organisation des idées.

Le théorème suivant qui clôt la première partie de la thèse est une illustration parfaite du style de Banach⁶⁷ :

THÉORÈME 17. Lorsqu'une opération $F(X)$ définie dans un ensemble E est de première classe de Baire⁶⁸ relativement à cet ensemble, $F(X)$ est pantachiquement⁶⁹ continu relativement à chaque ensemble parfait.

La preuve de ce théorème non trivial est extrêmement bien structurée. À chaque étape Banach renvoie précisément aux théorèmes qu'il a démontrés précédemment. Dans ce résultat on trouve pas moins de sept renvois à des définitions ou des résultats précis en deux pages pour justifier la progression de la preuve. Ceci confère à la démonstration un aspect clair que l'on retrouve dans la majorité des textes de Banach. Certains aspects techniques sont retirés

⁶⁷[Ban22, p.149]

⁶⁸En nos termes, une opération est de *première classe de Baire* si elle est limite simple d'opérations continues. La définition précise est donnée par Banach quelques lignes au-dessus du théorème que nous citons ici.

⁶⁹Une opération est *pantachiquement continue relativement à A* si l'ensemble de ses points de continuité est dense dans A . La définition précise est donnée par Banach quelques lignes au-dessus du théorème 17.

de l'énoncé de la preuve laissant mieux apparaître l'enchaînement des idées et certaines stratégies.

En particulier les aspects que nous qualifierions désormais de *topologiques* sont traités par Banach principalement grâce à la notion de sphère qu'il a introduite précédemment, et en faisant intervenir des suites de sphères dont les centres convergent.

Une autre idée mise en avant par Banach est la considération de *tous les éléments qui satisfont à une condition donnée*. Cette stratégie permet entre autres de réinterpréter des inégalités en termes de *géométrie de l'espace*. Par exemple dans le théorème que nous évoquons, Banach part d'une suite d'opérations $\{F_n\}$ qui converge simplement (termes d'observateur⁷⁰) vers F . Il fixe un certain $\epsilon > 0$ et considère alors « l'ensemble des points qui satisfont à l'inégalité $\|F_n(X) - F_{n+p}(X)\| < \epsilon$ » qu'il désigne par $A_{n,p}$. On a donc une infinité dénombrable d'ensembles fermés pour lesquels Banach peut appliquer les résultats sur les ensembles qu'il a établis préalablement. Des inégalités de l'analyse sont donc, en employant à nouveau un vocabulaire d'observateur, réinterprétées en termes de notions topologiques.

Cette technique sera utilisée par Banach à différents endroits dans sa thèse et se révèle être un outil puissant au centre dans son approche.

THÉORIE DES OPÉRATION ADDITIVES. La seconde partie de la thèse de Banach se concentre sur les « opérations additives » et constitue le cœur de la théorie qu'il élabore ici. Alors que les *axiomes* ont été réservés pour définir un « ensemble abstrait » et ses diverses propriétés, les opérations voient leurs propriétés sélectionnées et précisées par le biais de *définitions*.

Les opérations algébriques de l'« ensemble abstrait », addition et multiplication par un réel, ont été très souvent utilisées dans la première partie, et en particulier pour étudier les propriétés des opérations les plus générales. Banach a bien précisé qu'il y avait en toute généralité deux « ensembles abstraits » en jeu, l'un jouant le rôle de « domaine » alors que l'autre est le « contre-domaine ».

Si l'on regarde précisément comment sont intervenues les opérations algébriques des « ensembles abstraits » dans la partie précédente lorsqu'il s'agissait de traiter des « opérations », on identifie deux situations. Premièrement les opérations algébriques du « contre-domaine » ont permis de définir des combinaisons linéaires (termes d'observateur) d'opérations continues, dont Banach démontre que ce sont encore des opérations continues⁷¹. Deuxièmement, les opérations algébriques interviennent pour *mesurer* l'écart entre deux éléments, comme dans les expressions du type $\|X_1 - X_2\|$ ou $\|F(X_1) - F(X_2)\|$. C'est alors la *continuité* qui permet de relier une information donnée en termes d'opérations algébriques et de normes du côté du « domaine » à une information utilisant les opérations algébriques et la norme du côté du « contre-domaine ». Tous les

⁷⁰ Les divers types de convergence ne seront nommés par Banach que plus tard. Ici il utilise ce que nous appelons la convergence simple ou ponctuelle : pour chaque x on a $F_n(x) \rightarrow F(x)$. Nous verrons par la suite d'autres définitions de la notion de limite, et Banach nommera de façon cohérente les divers modes de convergence dans son livre de 1932. Il y examinera en outre les rapports qu'entretiennent ces diverses notions de limites et les topologies qui en découlent.

⁷¹ Voir le théorème 14 de la première partie [Ban22, p.146]

théorèmes de cette fin de première partie démontrent une propriété équivalente à la continuité ou font l'hypothèse de la continuité des opérations en jeu.

Dans la seconde partie de la thèse la notion d'*additivité* (d'une opération) va permettre de faire le lien entre des calculs algébriques dans le « domaine » et des calculs correspondants dans le « contre-domaine » et de lier la continuité avec une maîtrise de la norme des éléments-images⁷².

Banach introduit la notion en ces termes⁷³ :

DÉFINITION. On dit qu'une opération $F(X)$ est additive, lorsqu'on a pour tout X et Y :

$$F(X + Y) = F(X) + F(Y).$$

Il remarque que cette définition implique $F(\theta) = \theta$ et que pour tout couple d'entiers p, q avec $q \neq 0$ on a $F(\frac{p}{q} \cdot X) = \frac{p}{q} \cdot F(X)$.

Le premier paragraphe de cette seconde partie est alors consacré à établir des propriétés pour les opérations additives. On retrouve ici les résultats qui sont devenus des grands classiques de la théorie des opérateurs : les opérations additives bornées dans une sphère sont continues (Théorème 1), une opération additive continue en un point est continue en chaque point de l'espace (Lemme 1), réciproquement toute opération additive continue est bornée dans chaque sphère (Lemmes 3 et 4), toute opération additive de première classe de Baire est continue (Théorème 2) pour ne citer que quelques exemples.

On peut remarquer ici que Banach ne précise pas dans ses démonstrations les axiomes utilisés, et en particulier quand la complétude (c.-à-d. l'axiome III) est en jeu. En fait le théorème 2 sur la continuité des limites simples d'opérations continues fait appel à un théorème antérieur (le théorème I.17, qui énonce que la limite est continue dans un ensemble dense) qui lui-même fait intervenir dans sa preuve le théorème I.9 (équivalence entre propriété de Cauchy et convergence) et en particulier l'implication non triviale de ce théorème. Ce n'est que dans la preuve de ce dernier théorème que l'on retrouve explicitement la nécessité d'utiliser l'axiome III sur la complétude. On voit donc que le lien précis à la complétude s'est totalement évaporé de l'énoncé du théorème II.2. Ceci montre que Banach cherche dans cette partie à développer une théorie des opérations sur des « ensembles abstraits » qui vérifient tous les axiomes exposés au départ. Ainsi, s'il n'y a pas à notre sens une vision *réflexive* sur un type d'espace que nous qualifions d'espace de Banach, il y a cependant un cadre unique formulé à l'aide des axiomes qui permet de développer une théorie des opérations sur les « ensembles abstraits » ainsi définis.

L'extension de la propriété de multiplicativité (ou homogénéité) aux nombres réels sous hypothèse de continuité est ici donnée comme un lemme⁷⁴ :

⁷²Le lien que nous faisons ressortir et qui consiste à mettre en relation par différents moyens des éléments d'un espace abstrait avec des nombres réels ou d'autres éléments d'un espace abstrait n'est pas établi tel quel par Banach. Il s'agit pour nous de mettre en évidence certains aspects et d'établir une façon de lire le texte de Banach.

⁷³[Ban22, p.151]

⁷⁴[Ban22, p.153]

LEMME 2. $F(X)$ étant une opération additive continue et a désignant un nombre réel, d'ailleurs quelconque, on a :

$$F(a \cdot X) = a \cdot F(X).$$

Cette propriété sera utilisée par la suite par Banach sans qu'il fasse explicitement référence au lemme du dessus. On notera simplement ici que seule la continuité intervient et cette propriété sera donc vérifiée pour toutes les opérations additives (continues) qui seront considérées par la suite. On peut contraster cette présentation et le traitement unifié que Banach propose avec, par exemple, celle de Fréchet en 1907, dans sa troisième note sur les opérations linéaires, où il démontre cette propriété dans chaque espace de fonction séparément, en utilisant la définition de convergence particulière donnée dans chaque espace.

Le second paragraphe de la seconde partie de la thèse donne deux résultats fondamentaux de la théorie des opérations car ils répondent de façon extrêmement générale à des problèmes d'équations intégrales. Le titre de la thèse de Banach faisait mention de l'« application aux équations intégrales » et celles-ci sont traitées suivant deux aspects. L'un consistera à montrer dans la dernière partie que les ensembles de fonctions que Banach a cités en introductions et qui interviennent dans les problèmes d'équations intégrales satisfont bien aux axiomes qui en font des « ensembles abstraits ». L'autre est un traitement général qui s'effectue en obtenant deux résultats sur les opérations continues. Banach n'indique pas leur lien avec les équations intégrales, mais leur formulation ne laisse aucun doute sur le rôle des deux théorèmes.

Nous rappelons ces théorèmes qui restent très importants même dans une théorie des opérateurs actuelle qui ne serait pas systématiquement reliée à une théorie des équations intégrales.

Le premier est ce que l'on nomme de nos jours un *théorème de point fixe* (E et E_1 sont toujours, en nos termes, des espaces vectoriels normés complets)⁷⁵ :

THÉORÈME 6. Si 1° $U(X)$ est une opération continue dans E , le contre-domaine de $U(X)$ étant contenu dans E_1 .

2° Il existe un nombre $0 < M < 1$ qui pour tout X' et X'' remplit l'inégalité

$$\|U(X') - U(X'')\| \leq M \cdot \|X' - X''\|.$$

— il existe un élément X tel que $X = U(X)$.

Bien sûr, bien qu'elle ne soit pas explicitement mentionnée ici, la complétude est essentielle et Banach l'utilise pour passer d'une série du type $\sum_{n=1}^{\infty} \|X_{n+1} - X_n\|$ qui converge à la convergence de la série $\sum_{n=1}^{\infty} (X_{n+1} - X_n)$, propriété qui a été démontrée dans la première partie.

Le résultat repose sur la construction d'une suite d'itérés à partir d'un élément quelconque $Y : X_1 = Y$ puis $X_{n+1} = U(X_n)$. Cette technique sera celle qui sera utilisée aussi dans le théorème suivant, conjointement à l'utilisation

⁷⁵[Ban22, p.160]

de séries qui permettent de construire un élément limite sur lequel s'appuie la démonstration. Nous y reviendrons un peu plus loin.

Dans ce paragraphe on aperçoit le retour d'un vocabulaire non-géométrique où les « points » ont à nouveau laissé la place aux « éléments ». Il émane de ce choix un certain type de généralité lié à l'organisation des idées. On a tout d'abord un contexte énoncé en termes d'axiomes qui définissent des « ensembles abstraits » en convoquant des opérations algébriques (addition et multiplication par un réel) et une norme dont la définition s'appuie sur celle plus générale d'« opération ». Puis des résultats techniques ont été établis en ayant recours à une vision plus géométrique, et en faisant intervenir en particulier des sphères. Comme nous l'avons vu les énoncés des résultats que nous venons d'examiner n'utilisent que le vocabulaire abstrait et général du début. Cette organisation permet à Banach d'établir ici des résultats où les formulations géométriques n'interviennent plus et où l'énoncé ainsi que la preuve reposent exclusivement sur des concepts énoncés dans les termes du contexte abstrait.

Le théorème de point fixe est en fait très important en ce qui concerne l'analyse de la thèse de Banach en termes de généralité. En effet, comme nous venons de le voir, il est exprimé sous une forme abstraite qui lui confère un caractère général. Mais il y a davantage ici et un certain type de généralité provient du rôle que joue ce théorème dans le travail de Banach. Nous en dégageons deux aspects.

D'une part ce théorème, analysé en termes d'« application », pour reprendre le vocabulaire du titre de la thèse de Banach, expose un résultat directement lié à un problème classique d'équation différentielle, que l'on nomme *problème de Picard-Lindelöf*. Il s'agit de l'équation différentielle $\frac{dy}{dx} = f(x, y)$ où f est k -lipschitzienne en y (uniformément en x). Picard et Lindelöf ont traité ce problème en utilisant la *technique des itérés*⁷⁶ dans le cadre où les éléments en jeu sont des fonctions ordinaires dérivables.

D'autre part on voit que ce théorème de Banach dépasse beaucoup le cadre des fonctions ordinaires et donne un résultat général dans les « ensembles abstraits ». Ce théorème a par ailleurs une place particulière dans la thèse de Banach car il est un des rares théorèmes qui n'est pas utilisé ailleurs dans le texte. Il constitue donc un achèvement en soi, une vision *réflexive* qui participe de la construction d'une véritable théorie générale : celle des opérations continues définies sur un « ensemble abstrait ». Cette précision non seulement positionne le développement de l'étude des opérations comme le cœur de la thèse de Banach et en détermine la portée générale, mais permet de porter un regard précis sur le rôle joué par les axiomes de départ et comment leur généralité découle de ce rôle en interaction avec les opérations.

Le second théorème établi par Banach amène à nouveau une double lecture qui permet de préciser le type de généralité qui s'en dégage. Ce théorème est énoncé sous la forme suivante⁷⁷ :

⁷⁶Cette technique est d'abord mise en avant par Picard puis reprise par Lindelöf. On peut lire la courte note de Lindelöf à l'Académie des Sciences en 1893, intitulée « Sur l'application de la méthode des approximations successives aux équations différentielles ordinaires du premier ordre » pour voir la stratégie initiale développée par Picard ou reprise ici par Lindelöf. (cf. [Lin94])

⁷⁷[Ban22, p.161]

THÉORÈME 7. $X + \alpha \cdot F(X) = Y$ étant une équation où Y désigne un élément donné et X l'élément inconnu, soit :

1° $F(X)$ une opération additive continue dans le champ E à contre-domaine contenu dans E ,

2° M — le plus petit des nombres qui satisfont à l'inégalité

$$\|F(X)\| \leq M\|X\|$$

3° α — un nombre réel quelconque ;

pour tout Y et pour tout h satisfaisant à l'inégalité⁷⁸ $|h \cdot M| < 1$ il existe une solution de cette équation et on peut la mettre sous la forme

$$X = Y + \sum_{n=1}^{\infty} (-1)^n \cdot h^n \cdot F^{(n)}(Y)$$

où les opérations $F^{(n)}(Y)$ sont déterminées pour chaque n par les relations :

$$F^{(1)}(Y)F(Y) \text{ et } F^{(n)}(Y) = F(F^{(n-1)}(Y)).$$

Ce théorème prend donc la forme d'une équation qui la relie directement aux « équations intégrales de Fredholm » (expression qui constituera le titre d'un paragraphe du livre de 1932), et plus précisément des équations du second type. Ces équations⁷⁹ se présentent en général sous la forme $f(s) - h \int_0^1 K(s,t)f(t)dt = y(s)$.

Là encore le principe de base qui est à l'œuvre est l'itération à partir cette fois-ci de l'élément Y donné. Cependant ce théorème n'est pas présenté de la même façon que le précédent. Le théorème II.2.6 est énoncé sous la forme d'une propriété (le point fixe) d'une opération dès que celle-ci est contractante (lipschitzienne avec une constante strictement entre 0 et 1). Dans le théorème II.2.7 l'objet d'investigation est donné dès le départ : il s'agit, comme nous l'avons souligné, d'une équation qui reprend une formulation classique de la théorie des équations intégrales pour en donner une expression abstraite et générale dans le vocabulaire des « ensembles abstraits ».

Pourtant, malgré leur rapport plus ou moins affirmé avec les équations intégrales, ces énoncés sont donnés sous une forme abstraite et générale qui tend à faire naître un point de vue réflexif. Les deux résultats peuvent en effet, du fait de leur expression générale, être lus comme des théorèmes qui énoncent des résultats concernant les opérations linéaires (continues) sans référence à un quelconque contexte extérieur ou à des équations préétablies. Il y a ici une

⁷⁸La publication de *Fundamenta Mathematicae* fait intervenir un élément h qui n'est pas très clair. Il faut relire le théorème avec $h = \alpha$.

⁷⁹Ce type de problèmes a une longue histoire que plusieurs études contribuent à établir et à analyser. On peut consulter l'article récent d'Archibald et Tazzioli intitulé « Integral equations between theory and practice: the cases of Italy and France to 1920 » [AT14] qui donne une analyse qui éclaire les aspects essentiels de ces problèmes et le rôle des acteurs principaux que nous avons cités à diverses occasions. L'article donne en outre une idée des enjeux généraux des problèmes liés aux équations intégrales au début du 20^e siècle. On y trouve de plus les références essentielles de la littérature primaire et secondaire sur le sujet (bien au-delà du cadre Franco-Italien).

sorte d'émancipation du discours qui est soutenue par le fait que Banach ne fait aucune mention des équations intégrales ni de quelconques applications liées à ces énoncés. En d'autres termes il nous semble que ces deux théorèmes participent aussi de la constructions d'une *théorie des opérations linéaires sur des espaces de Banach*.

Ces diverses considérations montrent que le traitement des opérations fait émerger de ce texte un type de généralité que nous avons caractérisé par plusieurs éléments : étendue des résultats qui couvrent de nombreux cas désormais vus d'une manière uniforme, point de vue *réflexif*, titre de la thèse « *Sur les opérations ...* » qui présente un traité dédié aux opérations, théorèmes qui sont établis pour eux-mêmes et ne seront plus utilisés par la suite⁸⁰. Un point de vue rétrospectif permet donc de penser que ce texte constitue l'un des moments clés dans l'élaboration de la théorie des opérateurs telle qu'elle s'est développée au 20^e siècle.

UTILISATION DES SÉRIES. La seconde partie de la thèse de Banach dont nous venons d'analyser les principaux aspects en termes d'organisation met en œuvre une technique qui justifie les différents théorèmes sur les séries que nous avons rappelés plus haut. Ceci montre que ces résultats dédiés aux séries n'étaient pas seulement destinés à donner des résultats concernant les « ensembles abstraits », mais qu'ils étaient façonnés afin de servir précisément dans l'étude des opérations.

Les séries interviennent dans l'étude des opérations de deux façons. Dans le premier exemple (Théorème II.1.4 qui établit la continuité d'une opération sous certaines hypothèses) Banach raisonne par l'absurde. Si l'opération F n'est pas continue, alors il existe une suite $\{X_n\}$ de la sphère de rayon 1 telle que $\lim_{n \rightarrow \infty} \|F(X_n)\| = \infty$. Banach construit alors de proche en proche une suite de combinaisons linéaires des X_n (en laissant tomber autant de X_n qu'il le désire) : $\sum_{i=1}^n a_i Y_i$ où Y_i est un certain X_n . Sans rentrer dans les détails techniques, on voit que la convergence de la série des normes $\sum_{i=1}^n \|a_i Y_i\|$ est assurée par la convergence de la série des réels $\sum_{i=1}^{\infty} |a_n|$. Dans ce cas le Théorème I.2.8 du chapitre 1 (convergence de la série des normes implique convergence de la série) ainsi qu'un lemme technique (Lemme I.2.5 : $\{X_n\}$ bornée et $\sum |a_n|$ converge implique $\sum a_n X_n$ converge) sont utilisés et on voit leur véritable rôle qui n'est pas (essentiellement) de préciser des propriétés des « ensembles abstraits ». Banach a isolé ces théorèmes pour mieux faire ressortir certains principes et l'organisation de la preuve de son théorème.

La seconde utilisation des séries apparaît couplée à l'itération dont nous avons parlé plus haut. Il s'agit des théorèmes II.2.6 et II.2.7. Dans les deux cas on a une série de normes qui est majorée par une série convergente ($\sum \|X_{n+1} - X_n\|$

⁸⁰La partie sur les opérations (fin de la première partie et seconde partie de la thèse) fait apparaître plusieurs résultats qui ne seront pas réutilisés ensuite. Ils déterminent des propriétés générales des opérations sur les ensembles abstraits. Outre ce que nous étudions dans cette section on peut y ajouter les théorèmes 14 et 16 de la première partie ou les théorèmes 4, 5, 6 de la seconde partie.

dans le cas où U est M -contractante, donc contrôlée par la série $\sum M^n$ et $\sum \|(-1)^n h^n F^{(n)}(Y)\|$, pour l'équation de Fredholm généralisée où F est additive continue, qui est contrôlée par $\sum |h \cdot M|^n$ qui converge pour $|h| < \frac{1}{M}$.) Dans ces cas c'est encore le théorème I.2.8 de la première partie qui permet de conclure que la série des éléments de E converge.

Nous avons vu précédemment que plusieurs mathématiciens utilisaient des séries pour définir des éléments qui donnent des solutions à des problèmes d'analyse fonctionnelle. Banach utilise cette idée pour en faire un outil général qui permet de construire des éléments dans un « ensemble abstrait ». La série apparaît ici comme un outil puissant qui permet de concevoir de nombreux nouveaux éléments qui ont des propriétés particulières lorsque l'on leur applique une opération continue. Nous avons ici illustré le recours aux séries dans deux théorèmes importants de la thèse, mais on retrouve leur utilisation dans de nombreux théorèmes que Banach développera par la suite dans diverses études sur les « fonctionnelles linéaires » dont nous parlerons un peu plus bas.

Ce sont les notions de continuité et souvent d'additivité qui sont centrales dans l'utilisation des séries. Banach a isolé ces caractéristiques dans une vision abstraite et a fait des séries un outil de la théorie qu'il développe. Leur présence et les théorèmes développés dans la première partie participent de l'organisation de la théorie des opérations et du type de généralité qui en émane.

5.2.5 TRAITEMENT DES ENSEMBLES DE FONCTIONS

La troisième et dernière partie de la thèse de Banach revient sur les ensembles de fonctions qu'il avait listés tout au début de son exposé.

On y distingue deux paragraphes dont les rôles sont différents. Le premier expose des résultats pour les opérations particulières du type $X \rightarrow Y$ où X et Y sont des fonctions et Y est définie par $Y(s) = \int_a^b K(s, t) \cdot X(t) dt$. Les fonctions X et Y seront considérées comme des éléments d'ensembles de fonctions réinterprétés comme des « ensembles abstraits » (avec une norme qui vérifie les axiomes). Banach ajoutera l'hypothèse importante que Y est bien définie et finie presque partout. Le second paragraphe montre que chaque ensemble de fonctions identifié au début, muni d'une norme que Banach précise pour chacun d'eux, vérifie bien sa liste d'axiomes et peut donc être vu comme un « ensemble abstrait ». Tous les théorèmes établis de façon générale seront donc valides dans le cas des espaces fonctionnels munis d'une norme adéquate.

Le premier paragraphe est orienté vers la résolution des équations intégrales et va donner lieu à l'introduction de nouveaux axiomes. En fait il s'agit désormais de traiter de problèmes concernant des fonctions *mesurables* sur un intervalle. Or contrairement au cadre précédent où la notion de limite découlait de la norme, ici une autre notion de limite préexiste dans la théorie des fonctions mesurables au sens de Lebesgue. Banach doit donc s'assurer d'une certaine compatibilité des deux limites.

La définition de limite asymptotique est donnée dès le départ par Banach⁸¹ :

⁸¹[Ban22, p.163]

DÉFINITION. Soit 1° $\{X_n(t)\}$ une suite des fonctions mesurables, définies dans (a, b) sauf, peut-être, un ensemble de mesure nulle.

2° $X(t)$ une fonction mesurable définie également dans (a, b) sauf un ensemble de mesure nulle.

Nous dirons que la suite $\{X_n(t)\}$ converge *asymptotiquement* vers $X(t)$, lorsque pour tout $\epsilon > 0$ la mesure L de l'ensemble des valeurs de t qui vérifient l'inégalité $|X_n(t) - X(t)| \geq \epsilon$ tend vers 0 avec $\frac{1}{n}$.

Par écrit :

$$\lim_{n=\infty} \text{asym} X_n(t) = X(t).$$

Cette notion de limite précède ici le choix de la norme sur les ensembles de fonctions et il est donc nécessaire pour Banach d'introduire trois nouveaux axiomes portant sur la norme qui garantissent une certaine interopérabilité entre les deux notions de limite. Banach précise que cette définition a été donnée par Hardy et Landau et il s'appuiera sur plusieurs propriétés qui ont été étudiées par Lebesgue ou Riesz par exemple pour relier cette convergence en mesure à la convergence simple presque partout ou à la convergence de certaines intégrales⁸².

La liste d'axiomes est donc allongée de trois nouvelles propositions que nous rappelons ici⁸³ :

Les axiomes que nous allons introduire sont les suivants :

IV] Si 1° $\{X_n(t)\}$ est une suite des fonctions appartenant à E

2° $X(t)$ est une fonction appartenant à E

3° $\lim_{n=\infty} X_n(t) = X(t)$,

on a

$$\lim_{n=\infty} \text{asym} X_n(t) = X(t).$$

V] Si 1° $\{X_n(t)\}$ est une suite des fonctions appartenant à E

2° $\lim_{n=\infty} \|X_n(t)\| = 0$,

on peut extraire de la suite $\{X_n(t)\}$ une suite $\{\overline{X}_n(t)\}$ qui possède la propriété suivante : il existe une fonction $X(t)$ appartenant à E et satisfaisant à l'inégalité

$$|\overline{X}_n(t)| \leq |X(t)|$$

pour tout n et pour tout t sauf, peut-être, un ensemble de mesure lebesgienne nulle.

VI] Si 1° $\{X_n(t)\}$ est une suite des fonctions appartenant à E

2° $X(t)$ est une fonction appartenant à E

⁸²Le lecteur peut ici penser en termes de convergence simple (ou ponctuelle) presque partout au lieu de convergence asymptotique. Les deux notions sont en effet liées par le théorème suivant : $\{f_n\}$ converge asymptotiquement vers f (nous dirions qu'elle converge *en mesure*) si et seulement si de toute sous-suite de $\{f_n\}$ on peut extraire une sous-suite qui converge presque partout vers f . Ce théorème reste d'ailleurs vrai pour les mesures σ -finies mais la convergence en mesure n'a lieu que sur les ensembles mesurables de mesure finie. Dans ce texte et dans son livre de 1932 Banach utilise la convergence en mesure. Dans les traductions plus tardives du livre elle est parfois remplacée par la convergence presque partout (cf. [Ban87, p.53]

⁸³[Ban22, pp.163-164]

3° $\lim_{n \rightarrow \infty} \text{asym} X_n(t) = X(t)$,
on a

$$\liminf_{n \rightarrow \infty} \|X(t)\| \geq \|X(t)\|.$$

Ces axiomes ont plusieurs utilités et sont façonnés par Banach pour répondre à des besoins spécifiques liés aux opérations particulières (intégrales à noyau) qu'il veut étudier dans des ensembles de fonctions intégrables et sont particulièrement utiles pour établir la continuité (séquentielle).

Pour cela la démarche de Banach consistera à prendre une suite d'éléments de l'espace des fonctions $\{X_n\}$ qui tend en norme vers 0 et à démontrer que la suite des images $\int_a^b K(s,t) \cdot X_n(t) dt$ converge vers 0, asymptotiquement dans un premier temps, puis en norme lorsque l'on admet une condition supplémentaire.

L'axiome IV permet donc à partir de la convergence en norme de $\{X_n\}$ de déduire la convergence asymptotique de la suite $\{t \rightarrow K(s,t) \cdot X_n(t)\}$.

L'axiome V va permettre d'utiliser la convergence dominée. En effet $|K(s,t) \cdot X_n(t)| \leq |K(s,t) \cdot X(t)|$ et une des hypothèses que fait Banach est que $\int_a^b K(s,t) \cdot X(t) dt$ existe pour tout X de l'« ensemble abstrait » sur lequel l'opération à noyau est définie. Le théorème que nous qualifions de *convergence dominée de Lebesgue* permet d'affirmer, en passant à une sous-suite indiquée par le symbole \overline{X}_n , que l'on a $\lim_{n \rightarrow \infty} \int_a^b K(s,t) \cdot \overline{X}_n(t) dt = 0$ pour presque tout s . Banach évoque alors un nouveau théorème⁸⁴ qui permet de conclure que la suite de fonctions $\Phi_n(s) = \int_a^b K(s,t) \overline{X}_n(t) dt$ définie presque partout converge asymptotiquement vers 0.

Ces deux premiers axiomes sont donc liés par le schéma que nous venons d'indiquer à la preuve d'un premier théorème que Banach énonce ainsi⁸⁵ :

THÉORÈMES 1. Si 1° $K(s,t)$ est une fonction mesurable à deux variables définie pour tout couple des nombres (s,t) dont chacun appartient à (a,b)

2° Pour tout $X(t)$ appartenant à E l'ensemble des valeurs de s , pour lesquelles le symbole $\int_a^b K(s,t) \cdot X(t) dt$ n'existe pas est de mesure lebesgienne nulle

3° $\{X_n(t)\}$ est une suite des fonctions appartenant à E et $\overline{\lim}_{n \rightarrow \infty} X_n(t) = 0$, on a

$$\lim_{n \rightarrow \infty} \text{asym} \int_a^b K(s,t) \cdot X_n(t) dt = 0.$$

⁸⁴Ce théorème ainsi que celui de la convergence dominée sont donnés par Banach en note de bas de page. On y lit :

Si $\lim_{n \rightarrow \infty} \text{asym} f_n(t) = 0$, où $\{f_n(t)\}$ est une suite de fonctions mesurables dans (a,b) , 2° $f(t)$ est une fonction intégrable positive vérifiant l'inégalité $|f_n(t)| \leq f(t)$ pour chaque n et t sauf un ensemble de mesure nulle, on a $\lim_{n \rightarrow \infty} \int_a^b f_n(t) dt = 0$.

ainsi que

Si 1° $\{f_n(t)\}$ est une suite des fonctions mesurables définies dans (a,b) , 2° presque partout $\lim_{n \rightarrow \infty} f_n(s) = 0$, on a $\lim_{n \rightarrow \infty} \text{asym} f_n(s) = 0$. (cf. [Ban22, p.165])

⁸⁵[Ban22, p.164]

Enfin l'axiome VI permet de se ramener aux théorèmes généraux établis pour les opérations additives et pour toute norme qui satisfait les premiers axiomes. Pour cela il faut bien sûr admettre que le résultat de l'opération appartient lui-même à un certain « ensemble abstrait » muni de sa propre norme (Banach précise que celle-ci peu être totalement différente de la première). Si on part d'une suite quelconque $\{X_n\}$ qui tend en norme vers X le théorème précédent indique que $\lim_{n \rightarrow \infty} \text{asym} \int_a^b K(s, t) \cdot X_n(t) dt = \int_a^b K(s, t) \cdot X(t) dt$. Mais alors l'axiome VI permet d'obtenir $\| \liminf_{n \rightarrow \infty} \int_a^b K(s, t) \cdot X_n(t) dt \| \geq \| \int_a^b K(s, t) \cdot X(t) dt \|$ pour toute suite convergente $\{X_n\}$ de l'espace. Le théorème général II.1.3 s'applique alors et garantit que l'opération $X \rightarrow \int_a^b K(s, t) \cdot X(t) dt$ est continue. Banach énonce ce théorème en indiquant que l'opération linéaire est bornée sur chaque « sphère » : $\| \int_a^b K(s, t) \cdot X(t) dt \| \leq M \cdot \| X(t) \|$.

Cette analyse ne dépend pas du noyau $K(s, t)$ qui est une fonction mesurable tout à fait quelconque, mais de l'existence d'une norme adéquate qui vérifie désormais six axiomes. Les axiomes IV et V définissent une notion de norme qui dépend de la notion de convergence asymptotique spécifique aux ensembles de fonctions. Ils sont bien sûr adaptés aux ensembles de fonctions donnés en introduction par Banach, mais ils sont donnés sous une forme générale sans référence à un espace particulier de fonctions. Banach conçoit donc des espaces de fonctions normés où la norme est complète et, en un sens donné par les derniers axiomes, compatible avec la convergence asymptotique.

En fait, la limite asymptotique qui semble apparaître d'un coup et de façon incontournable vient d'emblée avec la théorie de l'intégration telle que l'a développée Lebesgue⁸⁶. Cette notion de convergence est donc intimement liée aux équations intégrales que Banach se propose de traiter et pour lesquelles c'est la théorie de l'intégrale de Lebesgue qui est adoptée.

On observe là une différence de traitement essentielle par rapport à ce que nous avons analysé par exemple chez Fréchet. En considérant des fonctionnelles $X \xrightarrow{\varphi} \int_a^b K(s, t) \cdot X(t) dt$ Fréchet avait pensé que la nature des fonctions K et X était déterminante pour déduire les propriétés de φ . Il en avait conclu que la nature du « champ fonctionnel » (et sa norme) était centrale pour établir la continuité de ce genre d'opération.

Pour Banach, une fois les fonctions X et $\varphi(X)$ installées dans des « ensembles abstraits », seules les natures des normes du domaine et du contre-domaine sont en jeu et permettent de garantir la continuité de l'opération φ .

⁸⁶Rappelons que Lebesgue développe ses idées en trois temps. D'abord il donne une définition « descriptive » de l'intégrale d'une fonction bornée en citant six propriétés que doit satisfaire l'intégrale, la sixième énonçant le principe de la convergence monotone. Puis il donne une définition constructive en introduisant d'abord la notion de mesure d'un ensemble, puis en définissant l'intégrale comme limite d'intégrales de fonctions en escalier. Pour démontrer alors que la condition 6 (convergence monotone) est réalisée, il s'appuie sur une décomposition de l'intervalle d'intégration en deux ensembles. Si on note $\{f_n\}$ et f les fonctions en jeu, pour $\epsilon > 0$ donné, l'un des ensembles est l'ensemble E_n des points où l'écart $|f_n - f|$ est plus grand que ϵ et l'autre son complémentaire $C(E_n)$ avec ses notations. La preuve repose alors sur le fait que « $m[C(E_n)]$ tend vers zéro avec $\frac{1}{n}$ ». (cf. [Leb04])

Les théorèmes que nous venons de mentionner et d'analyser seront les seuls deux théorèmes où seront explicitement mentionnées des opérations du type $\int_a^b K(s,t)X(t)dt$ qui font pour nous directement écho aux préoccupations que nous avons lues chez plusieurs auteurs et plus particulièrement chez Hadamard ou Fréchet. Le dernier paragraphe de la thèse ne fera plus mention d'opérations mais passera en revue chacun des ensembles de fonctions donnés dans l'introduction : tour à tour Banach y précise une norme et montre qu'elle vérifie les six axiomes. Ainsi pour chaque « champ de fonctions » devenu « ensemble abstrait » particulier et concret, Banach montre que le théorème de continuité de l'opération intégrale à noyau $\int_a^b K(s,t) \cdot X(t)dt$ s'applique.

Nous ne nous attardons pas sur les détails techniques mis en œuvre pour prouver que les différentes normes (qui sont les normes usuelles des ensembles considérés) satisfont les axiomes. Il reste dans cette dernière partie des démonstrations non triviales qui nécessitent parfois plusieurs lemmes ou théorèmes d'approche.

Il est cependant important de noter que Banach démontre en une seule fois dans cette dernière partie que l'ensemble des six axiomes sont vérifiés. Bien qu'il ait découpé l'introduction des axiomes en deux temps « afin de faire mieux ressortir les conséquences qui résultent des axiomes I-III »⁸⁷, Banach ne considère jamais explicitement les ensembles de fonctions *seulement* du point de vue des propriétés qui découlent des trois premiers axiomes. Il n'y a pas non plus de théorème *pratique* ne faisant intervenir que ces trois axiomes et qui serait adapté aux ensembles de fonctions car les formulations sont toujours dans la seconde partie très abstraites. Même le théorème II.2.7 que nous avons vu pour l'équation $X + \alpha \cdot F(X) = Y$ est donné pour F continue et ne s'applique donc à l'équation intégrale classique de Fredholm (second type) que pour $F(X) = \int_a^b K(s,t) \cdot X(t)dt$ dont la continuité dépend d'une norme qui satisfait les six axiomes. Nous verrons une évolution dans l'exposé du livre de 1932 à cet égard, mais ici les ensembles de fonctions avec leurs normes ne peuvent pas être simplement nommés *espace de Banach* car on occulterait alors que Banach les considère comme des cas particuliers décrits avec un modèle général qui comporte six axiomes.

5.2.6 CONCLUSION

L'étude de la thèse de Banach nous a permis de mettre en évidence un certain nombre de points qui éclairent le type de généralité que l'on peut lire dans ce texte. Un des éléments centraux qui ressort de cette étude est que Banach entame ici la construction d'une *théorie des opérations linéaire continues sur des espaces abstraits*. Cette théorie comporte plusieurs aspects qui déterminent le type de généralité à l'œuvre.

Certains résultats développés par Banach sont en effet obtenus sous une forme qui est très générale qui ne fait référence à aucun problème particulier prédéterminé. Nous avons mis en évidence dans ce cas un regard *réflexif* sur la

⁸⁷[Ban22, p.164]

notion d'opération qui nous permet de soutenir l'idée qu'il y a dans ce texte le point de départ de la construction d'une théorie des opérations.

D'autres résultats de la théorie sont formatés de façon plus utilitaire pour pouvoir être rapprochés de l'étude des équations intégrales (de Fredholm par exemple). Ces résultats auront une double portée générale.

D'une part ils permettent de traiter d'équations intégrales dans des ensembles de fonctions qui étaient vus de façon séparée avant Banach. De nombreux résultats partiels obtenus par Fréchet, Riesz ou d'autres établissent certains théorèmes similaires dans des champs de fonctions particuliers alors que Banach adopte un traitement général unique. Pourtant nous avons vu que cette généralité était aussi façonnée par la lecture de certains problèmes intégraux posés dans des champs fonctionnels particuliers. Il est difficile de lire précisément quel théorème est mis sous une forme qui correspond pour Banach à un problème préexistant d'équation intégrale. Le titre de sa thèse indique cet objectif qui n'est donc pas une reconstruction rétrospective qui introduirait un aspect téléologique ici. Dans la seconde partie, le théorème II.2.7, comme nous l'avons montré, peut se rapprocher d'une équation de Fredholm. Le théorème II.2.6 sur le point fixe peut aussi être pensé dans son lien avec divers problèmes d'équations intégrales. D'autres théorèmes peuvent encore être repensés en liaisons avec des problèmes de représentation intégrale. Par exemple les divers théorèmes sur la convergence des suites d'opérations font pour nous, à la suite des éléments que nous avons étudiés dans cette étude, directement écho au problème d'Hadamard tel qu'il est traité par Fréchet (étude de $\int_0^\pi H_n(y)f(y)dy$). Fréchet avait évoqué dans son article « Sur les opérations linéaires » de 1904 des noyaux généraux (fonctions seulement mesurables) pour lesquels $\int_0^\pi H_n(y)f(y)dy$ pouvait converger vers une opération de la forme $\int_0^\pi K(y)f(y)dy$. Les théorèmes II.2 et II.5 de la seconde partie de la thèse de Banach en particulier donnent des résultats sur les suites d'opérations additives continues qui permettent de directement réinterpréter les problèmes que nous avons évoqués avec Fréchet.

D'autre part les mêmes théorèmes que nous venons d'évoquer sont donnés sous une forme abstraite et sans mentionner explicitement les problèmes intégraux auxquels nous avons pu les rattacher. Ceci leur confère une qualité générale qui dépasse les situations envisagées précédemment. Le théorème II.2.7 sur l'équation $X + \alpha \cdot F(X) = Y$ est ici une parfaite illustration de cette double nature générale. Sa forme générale sans référence à des espaces fonctionnels mais seulement à des « ensembles abstraits » en fait un des théorèmes importants qui justifie l'efficacité de la théorie des opérations additives continues. *A posteriori* nous y voyons un théorème central de la théorie des opérateurs qui met en œuvre des notions de valeur propre, de vecteur propre⁸⁸ ou de spectre d'un opérateur sur un espace de Banach.

⁸⁸Banach n'emploie pas le vocabulaire de valeur ou vecteur propre ici. On retrouve les expressions « Eigenwerte » et « Eigenfunktion » dans l'exposé de Hilbert sur les équations intégrales de 1904 (cf. [Hil04, p.51]) Banach voit peut-être dans ce vocabulaire un lien trop fort à un cadre qui entrave la vision générale alors qu'il voudrait évoluer ici sans référence aux fonctions ou aux équations intégrales.

Enfin la notion d'« ensemble abstrait » construite au début du texte permet à Banach de développer une théorie des opérations additives continues qui est générale et qui répond à un certain nombre de questions identifiées dans des situations particulières que nous venons de rappeler. Les trois premiers axiomes (espace vectoriel, norme et complétude avec nos termes) forment un tout et définissent un type d'espace abstrait qui est en jeu dans les deux premières parties de la thèse, c'est-à-dire pour développer l'aspect le plus général et abstrait de la théorie des opérations. Malgré le style rigoureux des démonstrations de Banach il ne cherche pas à mettre en évidence pour chaque théorème s'il est obtenu dans le cadre d'une norme complète ou pas. Nous avons souligné cet aspect plus haut et nous avons montré qu'il contribue à donner une cohérence à la notion d'espace normé complet (mots d'observateur).

On peut contraster ce traitement global relatif aux axiomes I à III avec la précision des énoncés relativement à l'additivité des opérations. Pour chaque théorème Banach précise lorsque l'opération doit être additive (ici pour le point fixe l'opération n'a pas besoin d'être additive et cette hypothèse est absente de l'énoncé du théorème). Cette façon différente de traiter les opérations leur confère à nouveau un rôle particulier dans la vision que développe Banach.

Finalement cette étude de la thèse de Banach nous a permis de dégager deux modalités dans la conception de Banach. D'un côté certaines propriétés découlent du type de norme que l'on impose sur l'espace. Les diverses qualités de la norme ont été énoncées sous formes d'une liste de six axiomes. Il ne s'agit pas de combiner librement ces axiomes afin d'explorer toutes les géométries possibles, mais de donner en deux temps (axiomes I à II, puis axiomes IV à VI) les conditions minimales pour construire une géométrie générale adaptée à l'étude des opérations. D'un autre côté les opérations sont explorées en variant les hypothèses des théorèmes et en mettant en avant leur nécessité. Différents types d'opérations sont envisagés : d'abord des opérations générales (non forcément additives), qui peuvent être continues ou pas, puis des opérations additives, des opérations additives et continues, des opérations Lipschitziennes non forcément additives, et des opérations qui ont des propriétés spéciales issues de leur forme intégrale.

Cette double modalité permet à Banach de dégager un contexte extrêmement général pour les opérations tout en conservant une forme adaptée à l'appréhension de problèmes d'équations intégrales sur des champs de fonctions donné par avance. Nous allons examiner dans la suite une évolution substantielle de ce point de vue et montrer comment Banach prolonge ces idées dans plusieurs travaux qui seront finalement exposées sous une forme synthétique dans son livre de 1932.

5.3 LES FONCTIONNELLES LINÉAIRES

Après sa thèse Banach consacrera une partie de ses travaux à développer des propriétés des opérations additives définies dans des « ensembles abstraits ». Nous n'étudions pas en détail tous les articles qui relèvent des opérations additives ici et nous nous concentrons sur deux publications qui nous semblent essentielles pour établir certains aspects des mathématiques que Banach élabore.

En 1929 Banach publie en deux parties insérées dans le même volume de *Studia Mathematica* un travail intitulé « Sur les fonctionnelles linéaires »⁸⁹. Ces publications sont importantes non seulement parce qu'elles établissent des résultats qui répondent à des questions bien identifiées à l'époque de Banach, mais aussi parce que Banach continue ici à développer la vision que nous avons qualifiée de *réflexive* sur les opérations additives, vision qui détermine en partie la portée générale et le type de généralité que nous pouvons lire dans son œuvre.

Bien que les deux parties de la publication que nous étudions ici soient liées et forment un tout dont la cohérence est assumée dans le titre choisi par Banach, nous allons distinguer les rôles différents que jouent ces deux parties dans la progression construite par le mathématicien.

5.3.1 LE PRINCIPE D'EXTENSION DES FONCTIONNELLES.

La première partie du travail de Banach définit la notion de fonctionnelle sur un « ensemble abstrait » en s'appuyant sur les idées développées dans la thèse. Il y a cependant une évolution ici et l'introduction de certains termes qui n'apparaissaient pas dans la thèse.

En introduction Banach considère ce qu'il nomme un « ensemble vectoriel normé ». Une note de bas de page renvoie à sa thèse et précise encore le vocabulaire⁹⁰ :

L'ensemble E est dit *vectoriel* lorsque pour ses éléments sont définies les opérations d'addition et de multiplication par un nombre réel, conformément aux règles d'algèbre. Un ensemble vectoriel est dit *normé* lorsqu'à tout son élément x est attribué un nombre réel, désigné par $\|x\|$ —la *norme* de cet élément—, de manière que :

1° $\|x\| > 0$ pour tout $x \neq \theta$; $\|\theta\| = 0$, le symbole θ désignant l'élément *nul* (module d'addition),

2° $\|\alpha x\| = |\alpha| \|x\|$ pour tout α réel,

3° $\|x_1 + x_2\| \leq \|x_1\| + \|x_2\|$. La suite $\{x_n\}$ des éléments d'un ensemble vectoriel normé est, par définition, *convergente* vers l'élément x lorsque $\lim \|x_n - x\| = 0$.

Plusieurs niveaux de *structures* (mot d'observateur) reçoivent donc des noms qui permettent de les différencier en fonction des axiomes qui sont utilisés. Banach fait référence à un « module d'addition », qui lui permet de désigner le zéro de l'opération $+$, et distingue bien les deux opérations en jeu dans la notion d'« ensemble vectoriel ». La référence à l'« algèbre » pour les lois qui définissent l'addition et la multiplication par un réel est elle aussi nouvelle : elle participe du type de généralité que Banach met en place depuis sa thèse. La notion de « norme » est exposée ici sans la complétude, c'est-à-dire sans l'axiome III de la thèse. En particulier Banach insistera un peu plus loin en note de bas de page à la suite d'un théorème important : « Notre démonstration est très simple et ne suppose sur l'ensemble E que ce qui est dit dans l'introduction. Notamment nous ne supposons pas que l'ensemble E soit complet ou séparable. »⁹¹ Nous

⁸⁹[Ban29a ; Ban29b]

⁹⁰[Ban29a, p.211]

⁹¹[Ban29a, p.213]

reviendrons un peu plus tard sur cette remarque qui est importante à plusieurs égards.

Le vocabulaire employé par Banach est donc enrichi par rapport à la thèse et il désigne désormais divers éléments qui sont en jeu et qu'il agence dans une vision générale. Les termes choisis permettent de préciser des *structures* ou *sous-structures* et d'identifier les diverses propriétés qui découlent des axiomes, mais aussi d'articuler les concepts à d'autres idées mathématiques qui prennent à cette époque une place importante dans les mathématiques auxquelles Banach s'intéresse (référence à l'« algèbre », utilisation de la notion d'« ensemble »⁹² auquel il ajoute des qualificatifs —vectoriel, linéaire.)

Banach introduit ensuite un « ensemble linéaire » qui, en nos termes, est un sous-ensemble de E stable par combinaisons linéaires (réelles). Les fonctionnelles seront alors définies sur un « ensemble linéaire » G comme des fonctions^{93,94} de G dans \mathbb{R} . La notion d'« ensemble linéaire » permet alors de définir la notion de « fonctionnelle linéaire » qui est une « fonctionnelle » « additive » et « continue », la continuité étant définie à l'aide des suites convergentes de G comme dans la thèse.

Banach signale que l'« on prouve aisément » qu'une « fonctionnelle linéaire » définie sur un « ensemble linéaire » G est bornée sur la boule unité, résultat qui est une adaptation simple du théorème correspondant énoncé dans sa thèse pour les opérations additives sur l'« ensemble abstrait » E qui vérifie les axiomes d'espace vectoriel normé de la thèse (la preuve n'utilisant pas la complétude comme nous l'avons signalé). Il en déduit la définition de la norme $\|f\|_G$ de la fonctionnelle sur G comme « le plus petit des nombres M [tels que] $|f(x)| \leq M\|x\|$ ».

Les résultats que l'on trouve dans ce texte sont souvent reliés aux problèmes que l'on nomme désormais *problèmes de prolongement d'une fonctionnelle*. On peut donc relire le travail présenté dans les articles de Banach que nous étudions maintenant avec cette idée de *prolongement* en tête. Ce n'est pas inutile pour se repérer dans les travaux de Banach et dégager une séquence de textes qui s'occupent d'un thème cohérent. Cependant les mots *prolongement* ou *extension* (que Banach a utilisés dans d'autres articles⁹⁵ mais pas dans celui-ci) doivent être ici précisés et il nous faut examiner précisément ce qui est en jeu ici de façon interne mais aussi en le contrastant avec d'autres travaux qui donnent aussi des *théorèmes de prolongement de fonctionnelles*.

Nous proposons donc une lecture qui nous permet de déterminer une certaine dynamique d'élaboration des idées dans cet article en deux parties de Banach et qui fait apparaître précisément un type de généralité à l'œuvre dans le texte.

⁹²On peut contraster cet emploi unique du mot « ensemble » avec la diversité de termes que nous avons vus dans la thèse : « champ », « ensemble », « classe », « système » par exemple.

⁹³Terme qui n'apparaît pas ici chez Banach qui dit que la fonctionnelle fait « correspondre » à chaque élément x de G un nombre réel $\zeta = f(x)$. En particulier la « fonctionnelle » n'est pas *a priori* linéaire (ou distributive, le mot « linéaire » étant employé par Banach pour désigner le cas distributif et continu.)

⁹⁴Dans ces textes Banach fait le choix de travailler avec des fonctions à valeurs dans \mathbb{R} et ses espaces sont des \mathbb{R} -espaces-vectoriels.

⁹⁵Cf. par exemple l'article de 1925 intitulé « Sur le prolongement de certaines fonctionnelles » [Ban25].

Le premier paragraphe de l'article de Banach s'ouvre sur un théorème que nous qualifierons *d'extension élémentaire*. Banach y démontre qu'une fonctionnelle linéaire (donc continue) f définie sur un ensemble linéaire G peut s'étendre à l'ensemble linéaire $G + \mathbb{R}y_0$ (nos notations) où $y_0 \notin G$, en une fonctionnelle φ qui a la même norme que f . La démonstration de Banach consiste à trouver une valeur adéquate qu'il note λ pour $\varphi(y_0)$. On aura alors $\varphi(x + \alpha y_0) = f(x) + \alpha\lambda$ puisque φ doit être linéaire. Si l'on veut que φ soit continue (on peut supposer $\|f\|_G = 1$) il faut donc choisir λ tel que $|f(x) + \alpha\lambda| \leq \|x + \alpha y_0\|$, inégalité que Banach écrit :

$$-\| \frac{x}{\alpha} + y_0 \| - f(\frac{x}{\alpha}) \leq \lambda \leq \| \frac{x}{\alpha} + y_0 \| - f(\frac{x}{\alpha}).$$

L'existence de λ est obtenue par une technique en deux temps. D'abord pour tout couple $(x_1; x_2)$ d'éléments de G l'inégalité triangulaire donne $f(x_1 - x_2) \leq \|x_2 + y_0\| + \|x_1 + y_0\|$ ou encore :

$$-\|x_1 + y_0\| - f(x_1) \leq \|x_2 + y_0\| - f(x_2).$$

Puis Banach utilise une technique à laquelle il avait déjà pensé dans son article « Sur le problème de la mesure »⁹⁶ de 1923 et qui consiste, pour y_0 fixé, à prendre la borne supérieure et inférieure des quantités de chaque côté du signe d'inégalité. On obtient alors l'existence de λ tel que :

$$m = \sup_{x_1} \{-\|x_1 + y_0\| - f(x_1)\} \leq \lambda \leq \sup_{x_2} \{\|x_2 + y_0\| - f(x_2)\} = M.$$

Pietsch indique que les inégalités obtenues sont présentes dès 1912 chez Helly⁹⁷ et chez Hahn en 1927 dans un travail sur lequel nous reviendrons plus bas.

Le théorème que nous rappelons ici est donc une extension qui s'effectue d'un ensemble linéaire à l'ensemble linéaire engendré par l'adjonction d'un seul vecteur. C'est, comme nous l'avons déjà vu dans la thèse de Banach, une façon d'isoler certains aspects techniques dans un résultat préliminaire sur lequel vont pouvoir s'appuyer des théorèmes dont les démonstrations seront alors extrêmement claires et simplifiées. D'autre part le théorème et sa preuve sont mis en avant comme des outils généraux qui participent de l'élaboration d'une *théorie* des fonctionnelles (le mot « théorie » sera repris dans le titre du livre de 1932) et plus largement des opérations linéaires.

Le second théorème du premier paragraphe illustre ce que nous venons de mentionner et introduit une technique importante dans l'approche de Banach et de plusieurs de ses contemporains. Il est donné sous une forme qui est adaptée aux utilisations que Banach va développer par la suite, une forme simple d'extension des fonctionnelles dont la démonstration très courte s'appuie sur le résultat technique précédent⁹⁸ :

G étant un ensemble linéaire et $f(x)$ une fonctionnelle linéaire définie dans G , il existe une fonctionnelle linéaire $\varphi(x)$, définie dans E , telle que

$$\varphi(x) = f(x) \quad (x \in G),$$

⁹⁶[Ban23]

⁹⁷cf. [Hel12]

⁹⁸[Ban29a, p.213]

$$\|\varphi\| = \|f\|_G.$$

La preuve de ce résultat utilise le théorème précédent et une « induction transfinie » appliquée « aux éléments de l'ensemble $E - G$ (supposé bien ordonné). » Ce n'est que dans une parenthèse et dans la démonstration que Banach indique qu'il travaille sous l'hypothèse que certains ensembles sont bien ordonnés et il peut donc mettre en œuvre une récurrence transfinie.

De nombreux autres théorèmes s'appuieront par la suite sur ce théorème I.2 sans mentionner l'hypothèse qui permet de mener la récurrence transfinie⁹⁹.

Le second paragraphe de ce premier volet de l'étude des fonctionnelles permet de mettre au jour une certaine progression de la théorie élaborée par Banach tout autant qu'il permet de se faire une idée du contexte dans lequel il a élaboré les idées que nous venons de rappeler. Nous allons donc pouvoir préciser la portée générale des théorèmes I.1 et I.2 en analysant le théorème qui suit qui traite d'équations que nous noterons (*) par la suite¹⁰⁰ :

THÉORÈME 3. Soient $\{x_n\}$ une suite des éléments de E , $\{c_n\}$ une suite numérique et M un nombre positif. La condition nécessaire et suffisante pour l'existence d'une fonctionnelle linéaire $f(x)$, remplissant les relations

$$1^\circ f(x_n) = c_n \quad (n = 1, 2, \dots)$$

2° $\|f\| \leq M$, est que l'on ait l'inégalité

$$\left| \sum_{i=1}^r \lambda_i c_i \right| \leq \left\| \sum_{i=1}^r \lambda_i x_i \right\|$$

pour tout système fini des nombres réels λ_i .

En note de bas de page Banach signale que ce théorème avait « été démontré pour certains ensembles particuliers par M. F. Riesz » en 1910 et par Helly « dans quelques cas plus généraux » en 1912¹⁰¹. Le texte de Riesz auquel Banach fait allusion permet de préciser les choses. Il s'agit d'un essai publié aux *Mathematische Annalen* sous le titre « Untersuchungen über Systeme integrierbarer Funktionen ». Riesz y considère le système d'équations :

$$\int_a^b f_i(x) \zeta(x) dx = c_i$$

où les fonctions f_i sont dans L^q et ζ est dans L^p avec p et q conjugués ($\frac{1}{p} + \frac{1}{q} = 1$).

Il démontre¹⁰² que le système admet une solution $\zeta \in L^p$ telle que $\|\zeta\|_p \leq M^p$ si et seulement si :

$$\left| \sum_{i=1}^n \mu_i c_i \right| \leq M \cdot \left\| \sum_{i=1}^n \mu_i f_i \right\|_p$$

⁹⁹Nous ne nous étendrons pas davantage sur les divers positionnements des mathématiciens par rapport à l'axiome du choix ou aux conceptions cantorielles telles qu'elles sont perçues et utilisées à l'époque. C'est un sujet de première importance qui nécessitera une étude précise qui dépasse le cadre de cette thèse.

¹⁰⁰[Ban29a, p.213]

¹⁰¹Pour une vision plus large des problèmes connexes on peut consulter [Pie07, p.36-37] qui donne de nombreux cas particuliers relatifs à des ensembles de fonctions ou de suites chez Schmidt, Riesz, Helly, Hahn ou Banach.

¹⁰²cf. [Rie10, p.474]

pour toutes les valeurs de n et toutes les suites $(\mu_i)_{i \geq 1}$.

Banach donne donc une formulation plus générale de ce théorème en termes d'« ensembles vectoriels normés » et de « fonctionnelles » (continues). Mais cette forme abstraite n'est pas seulement le fruit d'une retranscription dans un langage synthétique, elle participe d'un mouvement plus large vers une conception générale dont nous pouvons apprécier ici plusieurs aspects.

Riesz dans le mémoire que nous venons de citer avait lui-même envisagé de traiter le cas de certaines opérations de L^p en faisant intervenir une notion abstraite de « lineare Funktionaltransformation » qu'il note $T[f(x)]$ et dont il donne les propriétés sous la forme $T[c_1 f_1(x) + c_2 f_2(x)] = c_1 T[f_1(x)] + c_2 T[f_2(x)]$. Une différence importante avec la façon d'opérer de Banach se joue ici dans la manière d'utiliser la notation abstraite et le rôle que lui font jouer les deux mathématiciens. Dans le texte de Riesz le symbole T est intimement lié à des opérateurs intégraux qui sont repensés de façon abstraite et générale en termes de « transformations » (*Transformationen*). On retrouve une utilisation similaire des notations abstraites qui désignent des « transformations » dans d'autres articles de Riesz et en particulier dans celui de 1916 intitulé « Über lineare Funktionalgleichungen » publié à *Acta Mathematica*¹⁰³. Là encore les propriétés de T sont celles que Riesz isole à partir de l'étude des équations intégrales pour en faire une « théorie générale des équations intégrales » (*Theorie der Integralgleichungen*). Dans ces articles les résultats principaux sont obtenus pour les équations intégrales dans les espaces L^p et le recours à une notion abstraite d'opération est dédié à l'obtention de ces résultats ou à la conceptualisation des équations intégrales sous une forme efficace.

Chez Banach la vision abstraite est aussi liée à l'analyse des équations (*) que nous avons rappelées plus haut page 253. La démonstration montre précisément que les théorèmes qui précèdent, ainsi que le contexte abstrait dans lequel ils ont été énoncés et prouvés, peuvent être conçu à partir de l'analyse du problème particulier des équations du type (*). L'organisation même de cette première partie d'article, qui ne propose que cette application (le théorème 3), donne du crédit à la thèse que c'est bien à partir de ce type de problèmes que Banach opère le passage à une vision abstraite. Mais le but n'est pas simplement de produire quelques lemmes qui simplifieraient la démonstration du théorème 3. Les théorèmes 1 et 2 sont formulés d'une façon générale qui dépasse largement le cadre du problème résolu dans le théorème 3 et le cadre des équations (*). Ils ont été choisis par Banach comme outils importants d'une théorie des fonctionnelles et plus largement comme nous allons le voir d'une théorie des opérations linéaires. Dans cette optique le théorème 3 n'est plus un but unique et constitue désormais une illustration de la théorie générale, en assure l'efficacité à démontrer des théorèmes et à répondre à des problèmes connus. Cette approche où la théorie abstraite est d'abord développée de façon très générale, puis appliquée à des exemples choisis (ils sont tous des repères importants dans le développement de l'analyse fonctionnelle à l'époque de Banach) sera reproduite systématiquement comme schéma de présentation dans le livre de 1932 qui couvrira de nombreuses questions dans une approche parfaitement structurée : d'abord un contexte abstrait organisé du plus général

¹⁰³[Rie16]

au plus restreint, puis à chaque fois des applications qui illustrent l'efficacité de la conception abstraite.

Par ailleurs le résultat même du théorème 3 est réinterprété dans l'optique des résultats précédents, et ce nouvel éclairage fait ressortir la lecture générale qu'en fait Banach. Celui-ci ajoute en effet juste après la démonstration une remarque qui montre que le problème n'est pas essentiellement lié à la donnée de deux suites $\{x_n\}$ et $\{c_n\}$ mais à la nature « linéaire » de l'espace et des possibilités que ceci implique en matière de fonctionnelles linéaires¹⁰⁴ :

Remarque. On peut formuler ce théorème d'une manière plus générale que voici :

$\varphi(x)$ étant une fonctionnelle définie dans l'ensemble W (linéaire ou non), la condition nécessaire et suffisante pour l'existence d'une fonctionnelle linéaire, définie dans E , et remplissant les relations

$$1^\circ f(x) = 0 \quad (x \subset G,$$

$$2^\circ \|f\| \leq M, M \text{ étant un nombre positif donné, est que l'on ait}$$

$$\left| \sum_{i=1}^r \lambda_i \varphi(x_i) \right| \leq \left\| \sum_{i=1}^r \lambda_i x_i \right\|,$$

quel que soit le système des r éléments x_i de W , le système des r nombres réels λ_i et l'entier positif r .

Enfin Banach, en s'appuyant sur le théorème 2 d'extension des fonctionnelles, va introduire un nouveau lien entre des propriétés que nous pourrions qualifier de *géométriques* de l'espace et les fonctionnelles. Ce rapprochement des propriétés des fonctionnelles et de celles de l'espace définit en partie le type de généralité que Banach cherche à mettre en place dans cette étude¹⁰⁵ :

THÉORÈME 4. G étant un ensemble linéaire et y_0 un élément, dont la distance à G est $d > 0$, il existe une fonctionnelle linéaire $f(x)$, définie dans E , remplissant les équations

$$1^\circ f(x) = 0 (x \subset G)$$

$$2^\circ f(y_0) = 1$$

$$3^\circ \|f\| = \frac{1}{d}.$$

La démonstration ne fait intervenir que des arguments similaires à ceux qui ont déjà été utilisés : considération de l'ensemble linéaire $G' = G + \mathbb{R}y_0$ (nos notations), construction d'une fonctionnelles de norme inférieure à $\frac{1}{d}$ sur G' et utilisation du théorème 2 pour étendre cette fonctionnelle à tout l'espace sans modifier sa norme.

Ce théorème, s'il est lu comme résultat sur les fonctionnelles seules, paraît être une variante technique qui n'apporte pas grand-chose à la théorie. Pourtant il est donné par Banach car il introduit une optique nouvelle en faisant intervenir la distance de y_0 à G et des premiers éléments que nous qualifierions de *topologiques* dont il a montré l'importance dès sa thèse. C'est en observant les théorèmes qui suivent que l'on se rend compte du rôle que le théorème 4 cité au-dessus joue dans l'édifice que Banach met en place.

¹⁰⁴[Ban29a, p.214]

¹⁰⁵[Ban29a, p.215]

Le théorème 5 qui suit montre précisément le nouveau point de vue introduit¹⁰⁶ :

Si G est un ensemble linéaire fermé tel que toute fonctionnelle linéaire $f(x)$ égale à 0 pour tout x de G est égale à 0 pour tout x , alors

$$G \equiv E.$$

C'est une application directe du théorème précédent et Banach ne donne pas de démonstration ni ne fait de commentaire au sujet de ce théorème. Mais nous voyons que l'optique a changé ici : le théorème 5 est un résultat sur les sous-ensembles linéaires fermés d'un « ensemble vectoriel normé » qui est établi par la considération des fonctionnelles. De plus ici on voit apparaître, outre la linéarité de l'ensemble, outre sa qualité topologique d'être fermé (au sens de la norme, et donc des suites convergentes pour la norme), une nouvelle propriété qui va venir jouer un rôle dans l'étude des opérations sur les espaces linéaires abstraits : c'est l'observation de toutes les fonctionnelles linéaires qui sont nulles sur G qui donne des informations sur G lui-même. Banach ne donne pas de nom à ces fonctionnelles et l'expression « fonctionnelle orthogonale à x »¹⁰⁷ apparaîtra seulement dans le livre de 1932.

Le dernier théorème de cette première partie dédiée aux fonctionnelles permet à Banach de considérer des sous-ensembles quelconques W de E (non forcément linéaires). Dans ce cas il construit un ensemble linéaire dont les éléments sont de la forme $\sum \alpha_i x_i$ où les x_i sont dans W et applique le théorème 4 sur les équations (*) pour obtenir le résultat suivant où l'hypothèse de fermeture est remplacée par le choix d'un élément adhérent¹⁰⁸ :

THÉORÈME 6. Soit W un ensemble arbitraire $\subset E$ et Y_0 un élément qui n'y appartient pas. La condition nécessaire et suffisante pour l'existence d'une suite $\{w_n\} \subset E$, remplissant les relations

$$1^\circ w_n = \sum_{i=1}^{r_n} \alpha_i^{(n)} x_i,$$

2^o $\lim_{n \rightarrow \infty} w_n = y_0$, est que l'on ait $f(y_0) = 0$, pour toute fonctionnelle linéaire $f(x)$ définie dans E et identiquement nulle dans W .

Ce théorème établit sous une forme générale le lien subtil qu'il existe entre l'ensemble des fonctionnelles nulles sur un ensemble et l'ensemble linéaire fermé engendré par les points cet ensemble. Ici l'espace linéaire engendré par un sous-ensemble et sa description à l'aide de limites de combinaisons linéaires apparaissent dans un théorème qui illustre un point de vue général que Banach a exposé en amont. Nous verrons dans la suite que cette conception de l'espace sera le modèle choisi par Hahn pour développer de son côté une approche générale légèrement antérieure à celle de Banach et nous contrasterons leurs deux façons de progresser.

¹⁰⁶[Ban29a, p.216]

¹⁰⁷[Ban32, p.59]

¹⁰⁸[Ban29a, p.216]

5.3.2 THÉORÈMES D'EXTENSION DE HAHN ET DE BANACH

Les théorèmes d'extension des fonctionnelles que nous venons de mentionner portent de nos jours, sous des formes plus ou moins généralisées, le nom de théorème de Hahn-Banach analytique. Cette double filiation est fondée sur une lecture de travaux de Hahn qui propose aussi dès 1927 des théorèmes d'extension des fonctionnelles dont la forme est proche de celle que nous venons de voir chez Banach.

Pietsch dans son livre¹⁰⁹ a signalé une partie des ingrédients communs à diverses approches que l'on retrouve chez Banach comme nous l'avons vu mais aussi chez Hahn et d'autres mathématiciens :

1. L'utilisation d'inégalités que l'on retrouve dans plusieurs travaux et que nous avons rappelée au-dessus. Pietsch précise qu'elles sont dues à Helly en 1912¹¹⁰.
2. L'utilisation de l'hypothèse que l'espace est bien ordonné pour pouvoir faire une récurrence transfinie et répéter un processus d'extension élémentaire.

Hahn dans son article de 1927 établit le théorème suivant¹¹¹ :

SATZ III. Sei \mathfrak{R}_0 ein vollständiger linearer Teilraum von \mathfrak{R} und $f_0(x)$ eine Linearform in \mathfrak{R}_0 der Steigung M . Dann gibt es eine Linearform $f(x)$ in \mathfrak{R} der Steigung M , die auf \mathfrak{R}_0 mit $f_0(x)$ übereinstimmt.

Pietsch émet l'idée qu'une différence entre la version de Hahn et celle de Banach réside dans le fait que Hahn utilise une norme alors que Banach, dans une version plus générale que nous verrons plus loin, fait intervenir une fonctionnelle sous-additive. Cette différence doit être précisée car Hahn propose de faire intervenir une norme en prenant soin de la donner sous une forme abstraite¹¹² :

Sei in \mathfrak{R} eine konvexe Maßbestimmung gegeben; d. h. es sei in \mathfrak{R} eine Funktion $D(x)$ definiert mit folgenden Eigenschaften:

1. Es ist $D(x) > 0$, und zwar $= 0$ dann und nur dann, wenn $x = 0$.
2. $D(\lambda x) = |\lambda|D(x)$.
3. $D(x_1 + x_2) \leq D(x_1) + D(x_2)$.

Le fait que Hahn décide de faire intervenir une norme sous la forme indiquée, très abstraite, et en redonnant la liste des propriétés qui la définissent, confère à ce résultat une valeur générale obtenue par une approche différente de celle que nous avons suivie chez Banach. Il y a donc une différence notable mais qui ne permet pas de hiérarchiser clairement du point de vue du général les deux approches du simple fait que l'une consiste à prendre une fonction sous-additive quelconque alors que l'autre prend une fonction sous-additive nulle seulement en 0.

¹⁰⁹ cf. [Pie07, p.39]

¹¹⁰ cf. [Hel12]

¹¹¹ [Hah27, p.217]

¹¹² [Hah27, p.215]

Pour nous, bien que les énoncés laissent apparaître quelques différences qui dénotent des contextes plus ou moins restrictifs (essentiellement la complétude chez Hahn alors qu'elle est rejetée par Banach, ou le choix d'une norme plutôt qu'une fonction sous-additive), c'est principalement dans l'organisation des idées et des textes que l'on voit la différence des rôles qu'ils jouent pour l'un et l'autre mathématicien. Il nous faut donc examiner de plus près l'enchaînement des idées dans les deux textes auxquels nous faisons référence.

Chez Hahn trois théorèmes se suivent. Le premier met en place la suite de sous-espaces qui permettent d'étendre de proche en proche la fonctionnelle à l'espace tout entier¹¹³ :

SATZ I. Ist \mathfrak{R}_0 ein (vollständiger linearer Teilraum von \mathfrak{R} , so gibt es eine mit \mathfrak{R}_0 beginnende wohlgeordnete Menge linearer Räume $\{\mathfrak{R}_\xi\} (\xi \leq a)$, die mit $\mathfrak{R}_a = \mathfrak{R}$ endigt, und in der stets \mathfrak{R}_ξ echter Teil aller folgenden $\mathfrak{R}_{\xi'}$ ($\xi' > \xi$) ist.

Le second donne une condition pour étendre une fonctionnelle linéaire donnée sur un sous-ensemble à l'espace linéaire complet qu'il engendre¹¹⁴ :

SATZ II. Sei \mathfrak{A} eine Punktmenge des linearen Raumes \mathfrak{R} , und sei \mathfrak{R}_0 der von \mathfrak{A} aufgespannte lineare Raum. Damit es zu der auf \mathfrak{A} definierten Funktion $f_0(x)$ eine Linearform $f(x)$ in \mathfrak{R}_0 gebe, die auf \mathfrak{A} mit $f_0(x)$ übereinstimmt und deren Steigung $\leq M$ ist, ist notwendig und hinreichend, daß für jede endliche Linearkombination $\lambda_1 x_1 + \lambda_2 x_2 + \dots + \lambda_n x_n$ aus Punkten von \mathfrak{A} die Ungleichung bestehe:

$$|\lambda_1 f_0(x_1) + \lambda_2 f_0(x_2) + \dots + \lambda_n f_0(x_n)| \leq MD(\lambda_1 x_1 + \lambda_2 x_2 + \dots + \lambda_n x_n).$$

Enfin le troisième théorème s'appuie sur le second pour étendre la fonctionnelle à tout l'espace en utilisant des arguments proches de ceux employés par Banach tels que nous les avons isolés plus haut.

La conception de Hahn repose donc de façon importante sur une vision de l'espace où les éléments sont décrits comme des limites de combinaisons linéaires. Cette vision est liée à la nature du problème central qu'aborde Hahn qui est la résolution des équations du type (*) comme l'indique le titre « Lineare Gleichungssysteme in linearen Räumen ». Une première condition nécessaire et suffisante d'extension est alors donnée relativement aux combinaisons linéaires finies des éléments d'une partie de l'espace.

Chez Banach la conception de l'espace est liée à la notion d'opération dès le début et ne se construit pas à partir d'un ensemble *générateur*. L'espace est donné dès l'abord comme un ensemble auquel Banach adjoint ce qu'il nomme de façon générale des « opérations », terme qui lui sert aussi bien à désigner les opérations algébriques que la norme. De ce fait les théorèmes d'extension, qui d'un point de vue rétrospectif peuvent sembler très proches, sont en fait complètement dépendants de deux conceptions différentes de l'espace et du rôle que lui font jouer les deux mathématiciens dans leur démarche.

¹¹³[Hah27, p.215]

¹¹⁴[Hah27, p.216]

5.3.3 UN ESPACE DE SUITES NUMÉRIQUES COMME MODÈLE

La seconde partie de l'étude des fonctionnelles linéaires est reçue par la rédaction à peine un mois après la première et contient de nombreux résultats qui s'appuient sur ceux que nous venons d'évoquer. Cette publication donne une dimension nouvelle à la théorie des fonctionnelles qu'elle développe dans plusieurs directions en introduisant un point de vue emboîté : les fonctionnelles forment un ensemble linéaire qui permet à son tour de définir des fonctionnelles qui agissent désormais sur des fonctionnelles. Pour aborder cette nouvelle partie Banach va utiliser une stratégie qui lui permet de s'appuyer sur des propriétés de l'ensemble des suites bornées qu'il considère comme un espace vectoriel normé avec une norme particulière.

Le premier paragraphe introduit pour les fonctionnelles une nouvelle norme qui est associée à la donnée de deux suites¹¹⁵ :

Soient $\{\sigma_n\}$ ($\sigma_1 = 1$) une suite numérique à termes positifs croissant indéfiniment et $\{a_n\}$ une suite des éléments telle que $\|a_n\| \leq 1$ ($n = 1, 2, \dots$). $f(x)$ étant une fonctionnelle linéaire, nous désignerons, dans ce qui va suivre, par $\|f\|^*$, le plus petit nombre positif K pour lequel

$$|f(a_n)| \leq K\sigma_n \quad (n = 1, 2, \dots).$$

Nous dirons que $\|f\|^*$ est la norme de f relative aux suites $\{\sigma_n\}$, $\{a_n\}$.

Banach montre que l'on a bien une norme et qu'elle est moins fine (termes d'observateur) que $\|f\|$, le supremum sur la boule unité. Il définit aussi la norme $\|\{\zeta_n\}\|^* = \inf\{K / |\zeta_n| \leq K\sigma_n\}$ pour une suite. On peut donc repenser la norme $\|\cdot\|^*$ d'une fonctionnelle comme la norme $\|\cdot\|^*$ de la suite $\{f(a_n)\}$.

Le premier paragraphe établit deux lemmes qui donnent des résultats pour des ensembles de fonctionnelles L qui jouent désormais le rôle que jouaient les ensembles G de points de l'espace vectoriel dans la première partie. Bien que Banach puisse considérer depuis sa thèse l'ensemble des fonctionnelles avec leur norme usuelle comme un espace vectoriel, ce point de vue n'est pas directement transposé ici. Il modifie l'approche de la première partie pour l'adapter à l'un de ses objectifs essentiels qui est d'identifier la forme des fonctionnelles qui agissent sur les fonctionnelles. Les questions abordées dans cette partie et qui mènent au développement d'une théorie des fonctionnelles sont aussi intimement liées aux problèmes de la mesure ou de l'intégrale ainsi qu'aux problèmes de représentations des fonctionnelles comme nous allons le voir.

Le premier lemme introduit les *fonctionnelles agissant sur des fonctionnelles* et montre qu'elles ont une forme particulière¹¹⁶ :

LEMME 1. Soit L un ensemble linéaire des fonctionnelles linéaires.

Supposons qu'à chaque fonctionnelle de L correspond un nombre $A(f)$ remplissant les relations :

$$(5.1) \quad A(f_1 + f_2) = A(f_1) + A(f_2),$$

$$(5.2) \quad |A(f)| \leq M\|f\|^*,$$

¹¹⁵[Ban29b, p.381]

¹¹⁶[Ban29b, p.381]

M étant une constante positive. Il existe alors un élément x_0 tel que

$$\|x_0\| \leq M, \quad A(f) = f(x_0) \quad (f \in L).$$

La démonstration de ce théorème met en place grâce à la norme qui dépend des deux suites une véritable stratégie de *représentation* pour étudier les nouvelles fonctionnelles. Nous entendons ici *représentation* au sens moderne du terme, c'est-à-dire l'étude d'un problème en passant à un nouvel espace sur lequel on possède davantage d'information.

Précisément la démonstration va à nouveau faire jouer, comme dans la thèse, un rôle particulier aux suites d'éléments de l'espace E et aux séries qui vont permettre de *construire* l'élément x_0 indiqué dans l'énoncé du théorème, mais ces éléments vont tout d'abord être construits dans un espace de suites numériques.

Pour mettre en place la *représentation* Banach va définir un premier « ensemble linéaire » W qui est l'ensemble des suites que l'on obtient comme image des termes de la suite $\{a_n\}$ par toutes les fonctionnelles de L . Sur ce nouvel ensemble linéaire, que Banach équipe de la norme $\|\cdot\|^*$, il construit une fonctionnelle F avec $w = \{f(a_n)\}$. Banach ne commente pas cette stratégie qui n'apparaît qu'au cœur d'une démonstration mais ses idées permettent néanmoins de dégager le schéma suivant :

$$\begin{array}{ccc} (L, \|\cdot\|^*) & \xrightarrow{\Theta} & (W, \|\cdot\|^*) \\ A \downarrow & & \downarrow F \\ \mathbb{R} & & \mathbb{R} \end{array}$$

où on a :

$$f \in L, \quad \Theta(f) = \{f(a_n)\}_n \in W, \quad F(\Theta(f)) = A(f).$$

D'une part la fonctionnelle F peut être étendue grâce au théorème I.2 de la partie précédente à l'ensemble des suites bornées en une fonctionnelle Φ . En effet, Banach démontre que $|f(w)| \leq M\|w\|^*$, c'est-à-dire que la fonctionnelle F est continue, et on peut donc appliquer le théorème.

Mais du côté des suites bornées Banach a à sa disposition des propriétés qui ne sont pas visibles au niveau des fonctionnelles d'un espace quelconque. En effet toute suite bornée $z = \{\zeta_n\}$ peut s'écrire comme limite en norme $\|\cdot\|^*$ d'une série de suites particulières $z_i = \{\zeta_n^{(i)}\}$ où $\zeta_n^{(i)} = 0$ si $n \neq i$ et $\zeta_i^{(i)} = 1$:

$$\lim_{n \rightarrow \infty} \left\| z - \sum_{i=1}^n \zeta_i z_i \right\|^* = 0.$$

La continuité de Φ relativement à la norme $\|\cdot\|^*$ permet d'établir l'expression de $\Phi(z)$ pour toutes les suites bornées z : $\Phi(z) = \sum_{i=1}^{\infty} \Phi(z_i)\zeta_i$. Une suite z bien choisie permet de montrer que $\sum_{i=1}^{\infty} |\Phi(z_i)\zeta_i| \leq M$. Enfin $x_0 = \sum_{i=1}^{\infty} \Phi(z_i)a_i$ est lui aussi majoré en norme par M et a les bonnes propriétés : $A(f) = \sum_{i=1}^{\infty} \Phi(z_i)f(a_i) = f\left(\sum_{i=1}^{\infty} \Phi(z_i)a_i\right) = f(x_0) \quad \forall f \in L$.

Le résultat obtenu pour les suites est donc transféré à la situation générale par le biais de l'association de F , fonctionnelle agissant sur les suites, et de A , fonctionnelle agissant sur des fonctionnelles¹¹⁷.

Le lemme 2 utilise alors une stratégie que l'on a déjà vue dans la première partie de l'étude des fonctionnelles de Banach. Il est proposé sous une forme qui en fait un outil général sur lequel Banach s'appuiera par la suite pour démontrer d'autres résultats. Son rôle à nouveau dépasse celui de simple outil technique et il vient étoffer la théorie des fonctionnelles¹¹⁸ :

LEMME 2. Soit L un ensemble linéaire des fonctionnelles linéaires et φ une fonctionnelle linéaire qui n'y appartient pas. Supposons qu'il existe un nombre $M > 0$ tel que

$$\|f - \varphi\|^* \geq M \quad (f \in L).$$

Il existe alors un élément x_0 tel que

$$f(x_0) = 0 \quad (f \in L),$$

$$\varphi(x_0) = 1,$$

$$\|x_0\| \leq \frac{1}{M}.$$

On voit ici que Banach fait à nouveau jouer un rôle particulier à la norme et par son biais à la distance entre une fonctionnelle φ et un ensemble de fonctionnelles. La stratégie est celle que nous avons déjà commentée : Banach considère l'ensemble que nous notons $L' = L + \mathbb{R}\varphi$ engendré par L et φ . La fonctionnelle A est alors définie sur L' par $A(f + \lambda\varphi) = \lambda$ dont on maîtrise cette fois-ci la norme $\|\cdot\|^*$. Le lemme précédent de représentation s'applique alors et $A(f + \lambda\varphi) = f(x_0) + \lambda\varphi(x_0)$ pour un certain x_0 , ce qui donne le résultat annoncé.

Le second paragraphe s'ouvre sur un théorème auquel on fait *a posteriori* référence comme forme générale du théorème de Hahn-Banach analytique¹¹⁹ :

THÉORÈME 1. $p(x)$ étant une fonctionnelle remplissant les relations

$$1^\circ p(x + y) \leq p(x) + p(y),$$

$$2^\circ p(\lambda x) = \lambda p(x) \quad (\lambda \geq 0),$$

il existe une fonctionnelle additive $f(x)$ remplissant l'inégalité

$$-p(-x) \leq f(x) \leq p(x) \quad (x \in E).$$

Ce théorème requiert une analyse précise. Il est en effet une version généralisée des énoncés que l'on a lus dans la première partie chez Banach ou encore chez Hahn. Cependant si l'on analyse les deux publications sur les

¹¹⁷Les stratégies de représentation isométriques d'espaces d'opérateurs se développeront considérablement au 20^e siècle. L'étude des fonctionnelles des espaces mis en correspondances et muni de diverses topologies donnera en particulier naissance à la théorie des algèbres duales qui est un exemple de la puissance d'un tel mode de pensée.

¹¹⁸[Ban29b, p.383]

¹¹⁹[Ban29b, p.384]

fonctionnelles dans une certaine unité de temps et comme un exposé cohérent d'une théorie, la hiérarchisation des énoncés du fait qu'ils recouvrent des situations plus ou moins étendues ne suffit pas. C'est à nouveau le rôle que jouent deux versions à leur place dans le texte qui détermine véritablement leur portée générale.

Nous allons mettre au jour plusieurs caractéristiques de ce théorème qui en déterminent la portée et le rôle que Banach lui attribue ici.

Tout d'abord le résultat n'est pas donné sous une forme d'extension d'une fonctionnelle. Il garantit seulement l'existence d'une fonctionnelle contrôlée par p . Pourtant la démonstration s'appuie sur une stratégie proche de celle employée dans la première partie pour démontrer le théorème d'extension. Banach part cette fois-ci d'une fonctionnelle quelconque sur un ensemble linéaire G , majorée par p , et il montre que l'on peut l'étendre à un ensemble $G + \mathbb{R}y_0$. Comme précédemment la technique consiste à établir une inégalité dont la forme ressemble à celle employée précédemment. Cette fois-ci, pour deux éléments quelconques, que Banach note x et \bar{x} on a¹²⁰ :

$$f(x + \bar{x}) \leq p(x + y_0) + p(\bar{x} - y_0),$$

ce qui donne

$$-p(\bar{x} - y_0) + f(\bar{x}) \leq p(x + y_0) - f(x).$$

Encore une fois le supremum de l'ensemble des valeurs prises par les termes du membre de gauche est inférieur à l'infimum des valeurs prises les termes du membre de droite et on peut donc intercaler, indépendamment de x et \bar{x} , un nombre α entre les deux. Ceci permet de définir une fonctionnelle $\varphi(x + \lambda y_0) = f(x) + \lambda\alpha$ sur $G_1 = G + \mathbb{R}y_0$ qui reste majorée par p .

Ensuite, en partant de $G = \{0\}$ et de la fonctionnelle nulle définie sur G Banach l'étend à tout l'espace par une récurrence transfinie. Il y a donc bien dans la démonstration un résultat d'extension qui a une portée plus générale que le précédent, mais Banach préfère ici énoncer un théorème d'existence —existence d'une fonctionnelle contrôlée par une fonctionnelle sous-additive quelconque. On observe donc que le rôle qu'attribue Banach à ce résultat est différent de celui qu'ont joués les théorèmes d'extension dans la première partie.

Au passage nous voyons que l'interprétation que nous donnons de ce théorème permet de distinguer à nouveau le travail de Hahn de celui de Banach, non pas seulement du point de vue de l'étendue de leurs théorèmes dans une lecture rétrospective, mais dans le rôle de ces résultats dans l'organisation de la pensée de chacun des mathématiciens.

Ici le résultat d'existence d'une fonctionnelle qui vérifie $-p(-x) \leq f(x) \leq p(x)$ a une forme adaptée à certains objectifs que s'est fixé Banach et qui vont être développés dans la suite de l'article. En fait ce résultat est une version abstraite de l'inégalité qui intervient dans divers travaux que nous avons mentionnés et dans divers domaines.

Sa forme est adaptée au travail avec les limites supérieures ou inférieures comme nous l'avons montré et le choix opéré par Banach nous permet de rapprocher ce résultat d'une réflexion qu'avait déjà entamée Banach sur la

¹²⁰[Ban29b, p.385]

définition de la mesure ou de l'intégrale¹²¹. Pourtant Banach ne fait aucune allusion ici à un quelconque objectif préétabli et le théorème a une portée générale qui participe de l'élaboration d'une théorie elle-aussi générale. La suite de notre analyse permettra d'éclairer le rôle de ce théorème dans la progression de Banach.

La forme abstraite du théorème répond aussi à un critère de *simplicité*, notion que Banach a évoquée à plusieurs reprises et qui confère aux résultats un caractère général qui découle de leur capacité à être mis en œuvre dans des situations diverses sans procéder à de longues adaptations. Ce résultat est *simple* dans la forme de son énoncé très court mais aussi du fait de nombre très restreint des hypothèses. Banach, qui ne précise pas le contexte de son énoncé, dépasse en fait les conditions qu'il avait indiquées dès le début de son article dans la première partie. Ici on a un « ensemble vectoriel » et aucune norme n'entre en jeu. D'ailleurs ce résultat se retrouvera dans le livre de 1932 dans un chapitre intitulé « Espaces vectoriels généraux », fruit d'une réorganisation générale des idées, et dans lequel toute notion de norme est absente.

Plus précisément, dans le livre de 1932 Banach ne donnera pas le théorème sous la forme que nous venons de rappeler. L'idée principale qu'il met en avant dans le livre est à nouveau la possibilité d'étendre les fonctionnelles. Nous donnons les deux résultats qui font apparaître cette évolution de point de vue dans le livre et plus précisément dans le chapitre dédié aux « espaces vectoriels généraux »¹²² :

THÉORÈME 1. Étant données

1° une fonctionnelle $p(x)$ définie dans E et telle que l'on ait pour tous x et y de E

$$p(x + y) \leq p(x) + p(y) \text{ et } p(tx) = tp(x) \text{ pour } t \geq 0,$$

2° une fonctionnelle additive et homogène¹²³ $f(x)$ définie dans un espace vectoriel $G \subset E$ (bien entendu, avec les mêmes définitions des opérations fondamentales) et telle que l'on ait pour tout $x \in G$

$$f(x) \leq p(x),$$

il existe toujours une fonctionnelle additive et homogène $F(x)$ définie dans E et telle que l'on a

$$F(x) \leq p(x) \text{ pour tout } x \in E \text{ et } F(x) = f(x) \text{ pour tout } x \in G.$$

Puis son corollaire¹²⁴ :

COROLLAIRE. Étant donnée une fonctionnelle $p(x)$ définie dans E et telle que l'on ait pour tout $x \in E$ et $y \in E$

$$p(x + y) \leq p(x) + p(y) \text{ et } p(tx) = tp(x) \text{ pour } t > 0,$$

¹²¹Nous avons déjà mentionné que la technique mise en place avait été utilisée précédemment par Banach dans son approche du « problème de la mesure » sous une forme légèrement différente.

¹²²[Ban32, pp.27-28]

¹²³NDA « homogène » signifie ici $f(tx) = tf(x)$.

¹²⁴[Ban32, p.29]

il existe une fonctionnelle additive et homogène $F(x)$ définie dans E et telle que l'on a pour tout $x \in E$

$$F(x) \leq p(x).$$

C'est donc bien la forme « extension » qui est ici retenue par Banach dans un second temps. Le théorème énoncé dans l'article de 1929 joue donc un rôle différent dans la démarche que Banach y développe. En particulier nous verrons un peu plus loin comment la forme de ce théorème est liée à des considérations que nous qualifierons de *topologiques* et comment il permettra à Banach d'utiliser à nouveau des suites.

D'autre part, avant de poursuivre l'étude de l'article sur les fonctionnelles, précisons encore que le livre *Théorie des opérations linéaires* de Banach éclaire le lien qu'il établit entre le résultat que nous étudions et le problème de la mesure que nous avons mentionné.

Une des applications importantes y est donnée en termes de mesure sur le tore, qui est dans les termes de Banach « une circonférence de longueur 1 ». Il note K la « classe de tous les ensembles situés sur la circonférence en question », A un élément de la « classe » K et A_0 la circonférence elle-même. Enfin E désigne « l'ensemble des fonctions réelles bornées $x(s)$ définies sur $[A_0]$ et où s désigne des arcs comptés à partir d'un point fixe dans un même sens de parcours. »¹²⁵ Banach considère E comme un « espace vectoriel » avec les opérations usuelles. L'application du théorème précédent se fait avec alors avec :

$$p(x) = \inf_{\{\alpha_n\}} \left\{ \sup_s \frac{1}{n} \sum_{k=1}^n x(s + \alpha_k) \right\}$$

où $\{\alpha_n\}$ est « une suite arbitraire de nombres » et donne l'existence d'une fonctionnelle F majorée par p sur E .

Banach obtient alors le résultat suivant¹²⁶ :

[E]n désignant par le symbole $\int x(s)ds$ la fonctionnelle $\frac{1}{2}\{F[x(s)] + F[x(1-s)]\}$, on a le théorème :

À toute fonction $x(s)$ de la classe E on peut faire correspondre un nombre $\int x(s)ds$ de façon que les conditions suivantes (où $x(s)$ et $y(s)$ sont des fonctions arbitraires de la classe E et a, b, s_0 des nombres) soient remplies :

- 1) $\int [ax(s) + by(s)]ds = a \int x(s)ds + b \int y(s)ds,$
- 2) $\int x(s)ds > 0,$ lorsque $x(s) > 0,$
- 3) $\int x(s + s_0)ds = \int x(s)ds,$
- 4) $\int x(1-s)ds = \int x(s)ds,$

¹²⁵[Ban32, p.]

¹²⁶[Ban32, p.31]

$$5) \int 1 ds = 1.$$

L'article sur les fonctionnelles linéaires lui ne fait pas référence directement au problème de la mesure alors que Banach avait déjà publié sur ce problème en 1923 (et à nouveau en 1929 avec Kuratowski). Si on ne peut donc pas éliminer de la scène cette référence à la mesure, il n'en reste pas moins que Banach cherche ici à mettre en avant un théorème indépendant et qui a une utilité qu'il nous faut établir.

5.3.4 TOPOLOGIE DES FONCTIONNELLES

La suite de l'article de Banach s'attache à mettre en place des suites (généralisées) bornées (indexées par un ordinal limite que Banach note ϑ) qui vont lui permettre d'introduire des notions de limite, fermé, densité, séparabilité, etc. La notion de suite choisie par Banach permet d'introduire les limites supérieures et inférieures. L'intérêt est que ces notions mènent directement à des fonctionnelles sous-additives. En effet, comme le remarque Banach dans un lemme, si $\{f_\zeta\}_{1 \leq \zeta \leq \vartheta}$ est une suite (généralisée) bornée par M alors $p(x) = \limsup_{\zeta \rightarrow \vartheta} f_\zeta(x)$ est une fonctionnelle sous-additive qui permet d'appliquer le théorème précédent et d'obtenir une fonctionnelle additive (et homogène) f qui vérifie $-p(-x) \leq f(x) \leq p(x)$. Comme la suite est bornée on obtient $|f(x)| \leq M\|x\|$ ce qui donne une fonctionnelle linéaire (continue donc, cette fois-ci, avec le vocabulaire de Banach).

À partir de la notion de suite bornée on peut donc toujours trouver une (non unique) fonctionnelle linéaire (continue) qui vérifie :

$$\liminf_{\zeta \rightarrow \vartheta} f_\zeta(x) \leq f(x) \leq \limsup_{\zeta \rightarrow \vartheta} f_\zeta(x).$$

Banach nomme les fonctionnelles continues obtenues des « fonctionnelles-limites »¹²⁷. Un ensemble L de fonctionnelles sera dit « faiblement fermé » si toute suite bornée de L admet une « fonctionnelle-limite » dans L . Cette notion de fermeture faible va permettre à Banach de proposer un théorème général qui correspond au lemme 2 qu'il a énoncé dans le premier paragraphe¹²⁸ :

THÉORÈME 2. Soit L un ensemble linéaire des fonctionnelles linéaires faiblement fermé et φ une fonctionnelle linéaire qui n'y appartient pas. Désignons par M la borne inférieure des nombres $\|f - \varphi\|$ ($f \in L$) et par M_1 un nombre positif $< M$, d'ailleurs quelconque. Il existe un élément x_0 tel que

$$f(x_0) = 0 \ (f \in L), \quad \varphi(x_0) = 1, \quad \|x_0\| \leq \frac{1}{M_1}.$$

Ce théorème ne fait plus apparaître la norme liée aux suites que Banach notait $\|\cdot\|^*$ dans le lemme 2 mais la démonstration va s'appuyer sur ce lemme à l'aide de suites construites de façon appropriée. La démonstration laisse apparaître une stratégie que nous avons soulignée dès la thèse de Banach

¹²⁷[Ban29b, p.386]

¹²⁸[Ban29b, pp.386-387]

et qui consiste à considérer l'ensemble des éléments de l'espace qui jouissent d'une certaine propriété. Ici, pour un nombre fixé M_2 Banach considère les ensembles G tels qu'il existe une fonctionnelle f de L qui vérifie $\|f\| \leq M_2$ et $\|f(x) - \varphi(x)\| \leq M_1 \|x\|$ pour $x \in G$. Banach montre alors que tous les ensembles G ont un cardinal majoré par un nombre fini $\aleph < \aleph_0$.

Il s'ensuit l'existence d'un sous-ensemble fini de E , noté K_1 pour lequel l'ensemble des fonctionnelles $\{f : \|f\| \leq M_2, \|f - \varphi\|_{K_1}\}$ ne rencontre pas L . En reproduisant le raisonnement en remplaçant M_2 par M_3 plus petit et ainsi de suite on obtient une suite $\{a_n\}$ de points de l'espace, et une suite $\{\sigma_n\}$ de nombres tels que $\{f : |f(a_n) - \varphi(a_n)| \leq M_1 \sigma_n \quad n = 1, 2, \dots\} \cap L = \emptyset$, ce qui implique que pour la norme donnée par les suites a_n et σ_n on a :

$$\|f - \varphi\|^* \geq M_1 \quad \forall f \in L.$$

La conclusion et l'existence du point x_0 du théorème 2 au-dessus résulte alors du lemme 2. On voit encore ici l'importance que joue la norme $\|\cdot\|^*$ et la stratégie de passage aux espaces de suites que nous avons précisée plus haut. Ces arguments seront réordonnés dans le livre de 1932 et ces idées, qui seront mieux isolées et étudiées séparément garderont une place importante.

Les suites de fonctionnelles qui sont utilisées ici permettent de penser plusieurs types de limites associées dans le texte de Banach au mot « faible ». Le recours à la notion de « fonctionnelle-limite » —non unique et construite à partir de la fonctionnelle sous-additive, est aussi l'occasion pour Banach de faire entrer en jeu des éléments de *topologie* qu'il avait introduits dans sa thèse¹²⁹.

Nous reprenons les notions introduites par Banach :

1° Toute fonctionnelle linéaire (continue) qui vérifie $\liminf_{\zeta \rightarrow \emptyset} f_\zeta(x) \leq f(x) \leq \limsup_{\zeta \rightarrow \emptyset} f_\zeta(x)$

est appelée « fonctionnelle-limite » de la suite (transfinie) $\{f_\zeta\}_{1 \leq \zeta \leq \emptyset}$.

2° Un ensemble linéaire L de fonctionnelles linéaires (continues) est « faiblement fermé » si pour toute suite (transfinie) bornée de L il existe une « fonctionnelle-limite » qui appartient à L .

3° Une suite (ordinaire, indexée par \mathbb{N}) de fonctionnelles linéaires est « faiblement convergente » si elle a en tout point de l'espace E une limite simple.

4° Un ensemble L de fonctionnelles linéaires (continues) est « faiblement séparable » s'il contient une suite dénombrable de fonctionnelles $\{f_n\}_{n \in \mathbb{N}}$ telle que tout élément f de L est limite d'une sous-suite de $\{f_n\}$.

Les notions que nous rappelons dans cette liste permettent de voir que la notion d'ensemble « faiblement-fermé » peut être définie de deux façons différentes par la « fonctionnelle-limite » ou par la limite d'une suite de fonctionnelles au sens de la convergence ponctuelle. Ces notions seront analysées plus en détail et comparées dans le livre de 1932¹³⁰ :

1. Un ensemble vectoriel Γ de fonctionnelles linéaires définies dans un espace E du type (B) s'appelle *régulièrement fermé*, lorsqu'il

¹²⁹En particulier rappelons que ce qu'il nomme ici « convergence faible » d'une suite (au sens ordinaire, indexée par \mathbb{N}) de fonctionnelles est défini par : $\lim_{n \rightarrow \infty} f_n(x) = f(x)$ ($x \in E$), la limite étant conçue pour la norme de l'espace E . (cf. [Ban29b, p.388])

¹³⁰[Ban32, Chap. VIII]

existe pour toute fonctionnelle linéaire définie dans E , mais n'appartenant pas à Γ , un élément $x_0 \in E$ qui remplisse les conditions [$f(x_0) = 1$ et $f(x) = 0$ pour tout $x \in \Gamma$].

2. [N]ous appellerons une fonctionnelle $f(x)$ remplissant les conditions [$\liminf_{\zeta \rightarrow \emptyset} f_\zeta(x) \leq f(x) \leq \limsup_{\zeta \rightarrow \emptyset} f_\zeta(x) \leq M$ pour tout $x \in E$] *limite-transfinie de la suite* $\{f_\zeta(x)\}$.

3. La *convergence faible* sera la convergence ponctuelle que nous avons rappelée plus haut.

Banach remarquera alors que dans un espace de Banach séparable ces notions de fermeture coïncident¹³¹.

Concernant l'étude menée dans les articles de 1929, alors que jusqu'ici Banach a insisté sur le fait que l'espace E de départ était un espace vectoriel normé quelconque (on se souvient que Banach avait mentionné la portée générale de ses résultats qui ne font pas intervenir la complétude), il réintroduit dans les deux derniers paragraphes des hypothèses d'abord de séparabilité puis de complétude.

Le troisième paragraphe introduit la notion d'ensemble « faiblement séparable » que nous avons donné dans la liste de définitions au-dessus et il le relie à la séparabilité de l'espace donnée par la norme¹³² :

LEMME 5. Si l'ensemble E est séparable, alors toute suite des fonctionnelles dont les normes forment une suite bornée, contient une suite faiblement convergente.

[...]

LEMME 6. Si l'ensemble E est séparable, tout ensemble L des fonctionnelles linéaires est faiblement séparable.

Sous cette condition de séparabilité de l'espace E Banach peut alors comparer les notions de convergence des fonctionnelles¹³³ :

THÉORÈME 3. Si l'ensemble E est séparable, la condition nécessaire et suffisante pour qu'un ensemble L des fonctionnelles linéaires soit faiblement fermé, est que la limite de toute suite faiblement convergente de fonctionnelles de L soit aussi une fonctionnelle de L .

Pour comprendre la portée générale de ces énoncés il faut à nouveau examiner le rôle que leur fait jouer Banach dans son texte. En effet on pourrait être tenté de voir ici un début d'étude systématique des différentes topologies que l'on peut définir sur un espace de fonctionnelles qui agissent elles-mêmes sur un espace de Banach donné (nos termes). En fait l'organisation du texte de Banach suit scrupuleusement un plan qui mène à l'étude des fonctionnelles et en particulier à l'identification des *fonctionnelles de fonctionnelles* comme nous les avons appelées.

La première étape que nous avons rappelée a montré les ingrédients en jeu, ingrédients que Banach organise en Théorèmes ou en Lemmes selon le rôle général qu'il souhaite leur faire jouer dans la théorie des fonctionnelles.

¹³¹Voir le Lemme 2 et le théorème 5 du chapitre VIII dans [Ban32].

¹³²[Ban29b, p.389]

¹³³[Ban29b, p.390]

Pour mettre en évidence la sélection et l'organisation de divers éléments nous récapitulons sous une forme synthétique avec nos notations quelques résultats de Banach de la première partie de l'article de 1929 que nous avons commentée. Nous conservons notre numérotation (qui inclut celle de Banach) qui permet éventuellement de revenir aux commentaires antérieurs ou de se référer directement au texte de Banach :

1. Théorème I.1.1 : Soient f une fonctionnelle linéaire (continue donc) définie sur un sous-espace vectoriel G et $y_0 \notin G$. Il existe alors une fonctionnelle φ définie sur l'espace $G' = G + \mathbb{R}y_0$ qui prolonge f et telle que $\|\varphi\|_{G'} = \|f\|_G$.
2. Théorème I.1.2 : Soient f et G *ut supra*. Il existe une fonctionnelle linéaire qui prolonge f à tout l'espace E et $\|\varphi\| = \|f\|_G$.
3. Théorème I.2.4 : Soient G un sous-espace vectoriel et y_0 tel que $\text{dist}(y_0, G) = d > 0$. Alors il existe une fonctionnelle f qui vérifie $f(x) = 0 \forall x \in G, f(y_0) = 1$ et $\|f\| = \frac{1}{d}$.

Si l'on relit les théorèmes donnés dans la première partie de l'article (Th. 1, 2 et 4) on peut concevoir qu'ils établissent un lien entre certaines propriétés de l'espace —le fait qu'il puisse se construire de proche en proche grâce à ses opérations algébriques ou certaines propriétés géométriques lues à travers la norme— et l'existence de fonctionnelles qui ont des propriétés particulières —que l'on peut à son tour lire comme une propriété de l'ensemble des fonctionnelles.

La seconde partie du travail que nous analysons change de point de vue et étudie des propriétés qui mettent en jeu l'ensemble des fonctionnelles (reçu comme espace vectoriel avec diverses topologies, via la norme ou des suites) et des fonctionnelles agissant sur des fonctionnelles.

Le premier résultat est présenté sous forme de lemme (Lemme II.1.1). Il s'appuie comme nous l'avons vu sur un passage à un espace de suites dans lequel nous pouvons désormais lire la présence de tous les ingrédients nécessaires à une théorie générale —avec une norme adaptée ($\|\cdot\|^*$). Banach y utilise le théorème d'extension général présenté dans la première partie et la séparabilité de l'espace des suites bornées avec la norme $\|\cdot\|^*$ qui lui permettent de donner la forme générale d'une fonctionnelle agissant sur des fonctionnelles. Mais ce résultat ne constitue pour Banach ni un exemple important (qui jouerait un rôle d'archétype) ni l'un des piliers de sa théorie des fonctionnelles. Il est désigné comme *lemme* par Banach qui, par contraste, va choisir de désigner comme théorème certains autres résultats. Ce lemme est assez important pour intervenir dans la preuve de certains théorèmes que Banach juge centraux dans sa théorie. Il permet d'en simplifier la preuve et montre une idée importante en jeu. Mais probablement sa forme n'est pas assez générale aux yeux de Banach car son énoncé fait apparaître une norme particulière, la norme $\|\cdot\|^*$, qui disparaîtra de l'énoncé des théorèmes principaux.

De même le lemme II.1.2 peut être vu comme le pendant du théorème I.4 de la première partie lorsque l'on passe aux fonctionnelles agissant sur un espace de fonctionnelles. Cependant, encore une fois, Banach le désigne comme lemme et là encore l'énoncé fait intervenir la norme $\|\cdot\|^*$.

Les théorèmes qui suivent permettent de lire un autre type de généralité. Non seulement ils sont exprimés sans recours à la norme $\|\cdot\|^*$ mais ils correspondent aussi aux outils les plus importants et les plus généraux que nous avons mis en évidence dans la première publication. Banach a lui-même identifié, isolé et repensé certains outils qu'il propose sous forme de théorème dans des versions très générales.

Le théorème II.1 est une version de la fameuse inégalité que nous avons rencontrée plusieurs fois qui est désormais donnée sous une forme générale avec une fonctionnelle sous-additive. C'est ici autant le rôle que Banach fait jouer à ce résultat que sa formulation plus étendue que les précédentes qui détermine la nature du général qu'il véhicule. D'un côté le résultat est donné sous une forme générale qui couvre aussi bien les normes que les limites supérieures des suites ou des suites généralisées en une formulation unique. Mais Banach en fait aussi un outil général de la théorie des fonctionnelles dont il a identifié l'utilité à plusieurs reprises.

Dans la même lignée, le Théorème II.2 est le pendant du théorème I.4. Sa forme générale permet de voir l'exacte transposition opérée entre les deux situations et laisse *a posteriori* comprendre la généralité que revêt en fait le théorème I.4 autant que le théorème II.2. La forme similaire des deux énoncés montre qu'un certain rapport entre espace et fonctionnelle peut se lire à double sens : d'une partie de l'espace on construit une fonctionnelle et d'un ensemble de fonctionnelles on construit un élément de l'espace. Les deux preuves sont très différentes et néanmoins Banach choisit ici d'exprimer les deux résultats sous une forme complètement parallèle. Cette forme unique de l'énoncé dans deux situations différentes participe aussi du type de généralité qui est ici à l'œuvre.

Enfin, sans revenir sur tous les résultats *topologiques*, le théorème II.4 identifie les fonctionnelles faiblement continues¹³⁴ qui agissent sur un espace faiblement fermé de fonctionnelles linéaires (continues pour la norme de l'espace), elles-mêmes agissant sur un espace complet. Banach montre que ce sont des évaluations en des points de l'espace dont on contrôle la norme (ils sont dans une boule centrée en 0 et de rayon plus grand que $\|A\|$ mais aussi proche que l'on désire de $\|A\|$).

Dans ce dernier résultat comme dans de nombreux autres on s'aperçoit que l'approche de Banach repose de façon essentielle sur le théorème II.2 qui apparaît donc comme très important dans son optique. On le retrouvera dans le livre de 1932 sous la même forme.

Enfin la dernière partie de cette étude des fonctionnelles introduit la notion d'« opération adjointe » et montre comment les fonctionnelles peuvent être importantes dans l'étude des morphismes entre espaces de Banach (formulation en termes d'observateur). On retrouve dans cette partie des idées que nous avons rencontrées en particulier dans l'utilisation des suites comme espace modèle. Ici une « opération » est à nouveau une opération additive¹³⁵ entre

¹³⁴On peut consulter [Pie07, p.30 et seq.] pour apprécier diverses approches de l'identification de ce que l'on appelle de nos jours le bi-dual. Il serait intéressant, concernant ces sujets, de mener une analyse comparée des travaux de Hahn et de Banach qui prolonge les remarques que nous avons faites plus haut.

¹³⁵Ici Banach dit « additive » pour indiquer qu'elle est « additive » et « homogène ». « Linéaire » signifie qu'elle est en outre continue pour les normes.

deux « ensembles vectoriels normés et complets » E et E' . L'opération est « linéaire » lorsqu'elle est continue. Banach définit alors l'opération adjointe \bar{U} de l'opération linéaire U par :

$$X = \bar{U}(Y) \text{ si } X = Y[U(x)]$$

pour toute fonctionnelle linéaire Y définie dans le second espace vectoriel E' . Il montre que cette opération qui établit une correspondance entre des fonctionnelles définies sur E' et des fonctionnelles définies sur E est une opération linéaire dont la norme n'excède pas celle de U (il montrera dans son livre de 1932 que les normes de U et \bar{U} sont en fait égales).

Une des motivations essentielles de cette partie est de donner des éléments d'analyse des équations du type $y = U(x)$ ou $X = \bar{U}(Y)$. Le théorème principal est précédé de quelques lemmes dont nous résumons les résultats. Ici U est une opération additive (et homogène, et nous préciserons quand elle est continue), \bar{U} est l'opération conjuguée (pas forcément continue) : 1° U est inversible si et seulement si $U(x) = 0$ implique $x = 0$. 2° Une opération additive (non nécessairement continue) possède un inverse continu si et seulement si il existe m tel que $\|U(x)\| \geq m\|x\| \forall x \in E$. 3° Si U est linéaire (continue), inversible et que U^{-1} est continue alors $U(E)$ est fermé. 4° Si \bar{U} possède un inverse continu alors l'image par \bar{U} de tout ensemble linéaire faiblement fermé (dans \bar{E}') est faiblement fermé (dans \bar{E}).

Ces propriétés qui sont ici énoncées sous forme de lemmes seront données comme remarques en introduction du chapitre intitulé « Équations fonctionnelles linéaires » dans le livre de 1932. Elles permettent d'établir des théorèmes qui sont adaptés à la résolution de certaines équations fonctionnelles¹³⁶. Les deux théorèmes que Banach établit dans cette fin d'étude gardent une forme très générale et abstraite, sans qu'aucune application aux équations fonctionnelles ne soit donnée¹³⁷ :

THÉORÈME 5. I. Si l'inverse de l'opération adjointe $X = \bar{U}(Y)$ existe et est continue, alors l'équation $y = U(x)$ est soluble pour tout y donné.

II. Si l'équation $X = \bar{U}(Y)$ est soluble pour tout X donné, alors

a) l'inverse de l'opération linéaire $U(x)$ existe et est continue,

b) le contre-domaine de l'opération $U(x)$ est l'ensemble des y remplissant la condition $Y(y) = 0$ toute fois que $\bar{U}(Y) = 0$.

La preuve s'appuie sur les propriétés énoncées dans les lemmes et sur le théorème II.2 que nous avons rappelé plus haut. Banach énonce alors le théorème correspondant pour l'opération \bar{U} et la démonstration montre la portée des conceptions générales et abstraites qu'il a mises en avant et qui permettent de transposer facilement un énoncé pour U en un énoncé valable pour \bar{U} ¹³⁸ :

¹³⁶Le chapitre sur les équations fonctionnelles dans le livre de 1932 donnera une série d'applications importantes du théorème qui suit. Couplé aux notions de valeur propre ou régulière il permet de réinterpréter de façon abstraite et uniforme des équations intégrales de Fredholm, Volterra etc.

¹³⁷[Ban29b, p.393]

¹³⁸[Ban29b, p.394]

DÉMONSTRATION. Elle est tout à fait analogue à celle du théorème précédent. On remplacera dans celle-ci les lettres x, y, X, Y, U, \bar{U} resp. par Y, X, y, x, \bar{U}, U et on s'appuiera sur le théorème [I.4] au lieu du théorème [II.2] et réciproquement.

Cette preuve permet d'apprécier la nature de la transposition sur laquelle Banach s'appuie et la correspondance subtile qu'il y a entre les théorèmes I.4 et II.2. Ces deux théorèmes qui sont exposés sous des formes très proches et qui semblent être issus d'une simple retranscription de l'espace E vers l'espace des fonctionnelles, reposent en fait comme nous l'avons vu sur des preuves très différentes. Le premier construit une fonctionnelle selon un schéma déjà employé lors de l'extension : $\varphi(x + \alpha y_0) = \alpha$ pour $y_0 \notin G$. Le second a nécessité une approche beaucoup plus longue chez Banach qui fait intervenir plusieurs théorèmes et donne une fonctionnelle qui dépend de deux suites que l'on ne connaît pas explicitement.

Enfin Banach donne une application à l'équivalence des normes comparables sur un Banach¹³⁹ :

Soit E un espace vectoriel admettant deux normes différentes : $\|x\|$ et $\|x\|_1$, et complet pour chacune d'elles. Supposons encore que $\lim_{n \rightarrow \infty} \|x\|_1 = 0$ entraîne toujours $\lim_{n \rightarrow \infty} \|x\| = 0$. Alors aussi $\lim_{n \rightarrow \infty} \|x\| = 0$ entraîne toujours $\lim_{n \rightarrow \infty} \|x\|_1 = 0$, c'est-à-dire : les deux normes donnent lieu à la même définition de convergence, ou encore : il existent deux nombres m, M ($0 < m < M$) tels que

$$m \leq \frac{\|x\|_1}{\|x\|} \leq M$$

Ce théorème repose donc de façon immédiate sur les résultats antérieurs. Sa forme abstraite et le fait qu'il ne fasse plus référence aux fonctionnelles en font un résultat à part dont nous pensons qu'il est l'un des premiers résultats de *géométrie des espaces de Banach*.

¹³⁹[Ban29b, p.395]

Conclusion

6.1 L'ANALYSE FONCTIONNELLE : SITE DE L'ÉLABORATION D'UNE APPROCHE GÉNÉRALE

L'idée d'*opération*, en tant qu'elle désigne un mode de mise en relation de certains éléments, nous a permis de construire ce qui apparaît désormais comme une scène où l'on peut analyser diverses manières de progresser et de se saisir de problèmes que nous classerions de nos jours dans le champ de l'analyse fonctionnelle. Cette approche nous a en particulier permis de dégager notre propre analyse de certains éléments —fonctions, équations intégrales, équations différentielles, espaces de fonctions, espaces vectoriels normés etc.— qui sont au centre du travail d'abstraction et de généralisation que nous cherchions à définir et à comprendre. Il est maintenant temps de préciser le rôle de certaines de ces notions qui sont intimement liées à notre vision de l'analyse fonctionnelle.

6.1.1 UNE CONCEPTION ABSTRAITE DE FONCTION

La notion de fonction est omniprésente lorsque l'on examine les travaux connectés à l'analyse fonctionnelle. Nous avons pu voir qu'elle était au cœur de l'élaboration de plusieurs approches générales et qu'elle était convoquée par les mathématiciens suivant diverses modalités afin de développer des conceptions abstraites.

LA FONCTION SOUMISE AUX LOIS DE L'ALGÈBRE. Nous avons illustré à plusieurs reprises l'utilisation de notions issues de l'algèbre pour manipuler les fonctions. Les opérations algébriques permettent de définir des calculs sur les fonctions quel que soit le rôle qui leur est attribué : fonction-agissante, fonction-objet, fonction-variable d'une autre fonction ou fonction-inconnue d'une équation. De plus la notion de distributivité, qui est associée à une loi en algèbre symbolique, est reprise pour définir des fonctions ou des opérations particulières qui sont à leur tour nommées distributives, additives ou linéaires. L'utilisation des lois et des opérations de l'algèbre pour manipuler les fonctions n'est cependant pas uniforme. Nous avons observé leur présence dans différents contextes où nous avons précisé leur rôle.

Tout d'abord avec Gregory nous avons pu voir une conception des fonctions comme symbole indifférencié du calcul symbolique. La fonction, tout comme les autres symboles, y est soumise à des lois qui permettent d'effectuer des

calculs et de mener des raisonnements. Les résultats obtenus doivent alors subir une phase d'interprétation qui leur donne éventuellement un sens, et il arrive parfois que l'expression *calculée* reste impossible à interpréter. Le travail de Gregory permet de montrer que la rigueur des résultats obtenus ne dépend pas uniquement de la précision avec laquelle on a appliqué les différentes lois pour manipuler les symboles, ni encore seulement d'une certaine logique qui valide les différentes étapes du raisonnement. La phase d'interprétation des résultats formels obtenus en suivant les règles fait elle-même partie de la procédure de validation des théorèmes. Ainsi la validité de la chaîne inférentielle qui établit un résultat repose-t-elle essentiellement sur une double acceptation de la part du lecteur. Il faut en effet s'accorder d'une part sur ce que l'on entend par suivre une loi calculatoire correctement et d'autre part sur la possibilité ou l'impossibilité de donner telle ou telle autre interprétation. L'exemple de l'étude des logarithmes a montré la subtilité de ce dernier point qui est au centre des préoccupations de Gregory.

De plus l'examen de l'algèbre symbolique chez le mathématicien écossais nous a permis de mettre au jour une façon particulière d'utiliser les fonctions. Comme *forme* algébrique donnée *a priori*, la fonction permet de reconnaître un certain nombre d'expressions qui définissent à leur tour de nouveaux symboles. Les développements en séries de e^x ou $\cos x$ par exemple, ou certaines formes polynômiales que l'on sait factoriser, permettent à Gregory de reconnaître ou de construire de nouvelles expressions symboliques. Nous parlons ici de *forme* de la fonction car nous avons montré que les questions concernant la définition du concept de fonction, la convergence des séries ou certaines propriétés comme la continuité ou la dérivabilité n'entraient pas en ligne de compte dans l'utilisation qu'en faisait Gregory dans le cas que nous évoquons. La *forme* des fonctions lui permet ainsi d'envisager, sans que leur sens soit donné *a priori*, des expressions du type $e^{\frac{d}{dx}}$, qui est simplement la somme infinie $1 + \frac{d}{dx} + \frac{1}{2} \frac{d^2}{dx^2} \dots$, ou de réécrire $\frac{d^2}{dx^2} - 1$ sous une forme factorisée $(\frac{d}{dx} - 1) \cdot (\frac{d}{dx} + 1)$. Toute formule algébrique qui définit une fonction est donc susceptible à son tour de donner une expression où la variable x est remplacée par n'importe quel symbole tant que celui-ci peut être soumis aux opérations algébriques en jeu dans l'écriture de la fonction.

Dans cette approche symbolique nous avons pu mettre en évidence une première forme générale de dissociation des rôles attribués à différents éléments. Les fonctions étant conçues comme des symboles au même titre que les autres éléments en jeu, on retrouve dans leur utilisation la distribution différenciée des rôles sur laquelle s'appuie le mathématicien pour progresser. Nous avons ainsi pu dégager ce que nous avons nommé des *symboles-agissants* et des *symboles-objets*, leurs rôles précis étant marqués dans les textes de différentes façons — parenthèses, écriture organisée de gauche à droite, etc.

Chez Volterra nous avons aussi trouvé l'utilisation d'opérations algébriques pour manipuler les fonctions et les lignes sous différentes formes. Les opérations algébriques y apparaissent principalement dans deux situations que Volterra combine dans une approche générale.

Lorsque Volterra considère des fonctions dont la variable est elle-même une fonction, l'addition des *fonctions-variables* (ordinaires) lui permet de concevoir et de manipuler divers types d'accroissements qu'il note $\varphi + \psi$, où φ et ψ sont deux fonctions ordinaires. Ces accroissements servent entre autres à créer un rapport différentiel grâce auquel le mathématicien italien développe une théorie locale des fonctions dont la variable est une fonction ou une ligne. Nous avons vu qu'il était difficile de dire si ce rapport était lui-même considéré par Volterra comme un rapport de fonctions ordinaires, un rapport de fonctions qui dépendent d'autres fonctions ou simplement un rapport de nombres réels construit à partir des valeurs prises par les fonctions. On peut donc donner au rapport différentiel des interprétations multiples qui nourriront la réflexion de Fréchet quelques années plus tard.

Une fois établis les outils différentiels pour les fonctions généralisées, Volterra choisit d'étudier principalement des fonctions additives pour lesquelles on peut préciser la variation. Nous avons montré que ce qui motive ce choix apparaissait déjà dans un texte de Jellett de 1850. Dans le cas additif le calcul de $F(\varphi + \delta\varphi) - F(\varphi)$ fait apparaître l'accroissement total sous la forme $F(\delta\varphi)$. Ici l'utilisation de l'addition des fonctions-variables est intimement liée à la conception des fonctionnelles additives que Volterra qualifie de « fonctions simples de lignes ».

Chez Fréchet nous avons établi deux lignes parallèles. L'une, reprenant un formalisme proche de celui d'Hadamard, s'intéresse à des champs de fonctions où les opérations algébriques sont disponibles dès l'abord. Nous avons examiné un processus d'abstraction qui permettait à Fréchet de dégager peu à peu une notion d'opération linéaire qui s'applique de façon uniforme à tous les champs de fonctions. Fréchet introduit alors la notion de champ général de fonction qui est stable par addition afin de pouvoir donner une définition d'opération distributive qui est commune à tous les champs.

Dans un second mouvement parallèle Fréchet tente d'établir une théorie des fonctions générales, nommées « opérations fonctionnelles ». Celles-ci agissent sur des ensembles abstraits auxquels il attribue des propriétés minimales choisies de proche en proche. Les opérations algébriques n'ont donc plus immédiatement leur place dans la définition des espaces abstraits où les opérations sont définies, mais elles apparaissent parfois dans la recherche du cadre général le mieux adapté pour l'analyse fonctionnelle.

Les notions topologiques avec lesquelles il fonde d'abord sa théorie ne font appel qu'à une notion de convergence des suites de points suivant un critère qui peut être choisi de manière quelconque pourvu qu'il respecte quelques principes généraux dont nous avons étudié la portée. Fréchet établit alors que les espaces ainsi formés sont trop généraux pour permettre de démontrer suffisamment de résultats pour les opérations fonctionnelles qui y sont définies. Il introduit donc une nouvelle classe d'espaces en définissant une notion de « voisinage » qui associe à deux points de l'espace un nombre réel (avec certaines propriétés sur lesquelles nous reviendrons à nouveau un peu plus bas). Un effet immédiat est la réintroduction des nombres réels —les valeurs prises par le « voisinage »— et donc potentiellement de leurs opérations algébriques. Cette remarque qui paraît anodine permet de distinguer deux situations qui ne mettent pas tout à fait en jeu le même type de généralité. La notion de « voisinage », qui détermine

les classes (V) comme les appelle Fréchet, s'appuie sur une notion subtile de relation triangulaire qui ne fait pas intervenir l'addition de \mathbb{R} . Fréchet résume la propriété ainsi : « il suffit que (A, B) et (B, C) soient petits pour qu'il en soit de même de (A, C) »¹. Par contraste la notion d'« écart » —lequel est un « voisinage » particulier qui correspond à ce que nous nommons une distance— introduite ensuite pour définir les classes plus restrictives (E) , vérifie l'inégalité triangulaire classique qui fait appel à l'addition des réels : $(A, C) \leq (A, B) + (B, C)$.

Il y aura une évolution et un effort pour s'affranchir de cette référence aux réels plus tard chez Fréchet. Celui-ci mettra en particulier en place une notion d'« écart abstrait » à valeurs prises dans une suite ordonnée d'éléments de nature quelconque sans premier élément, à laquelle Fréchet adjoint un élément O . À chaque point A une telle suite ordonnée définit une « échelle des écarts » pour laquelle $(A, A) = O$ et telle que pour tout $B \neq A$ on a $(A, B) > 0$. La notion d'« écart abstrait » permet ainsi de concevoir des « sphéroïdes » autour de A qui permettent à Fréchet de définir la convergence des suites et donc la topologie de l'espace abstrait².

Il y a donc un effort remarquable chez Fréchet de se départir de toute conception mettant en jeu une structure algébrique préexistante. Il démontre ainsi que l'on peut bâtir des espaces abstraits topologiques tout en conservant certaines idées géométriques —comme les « sphéroïdes »— et donc retenir à la fois la puissance de l'intuition géométrique et la capacité du modèle général à réinterpréter les situations déjà connues où les boules et les sphères sont des outils efficaces.

Enfin nous avons vu dans la thèse de Banach que les champs de fonctions étaient réinterprétés en termes d'espaces vectoriels normés, et leurs opérations algébriques sont donc au cœur du processus de généralisation du mathématicien polonais.

Les opérations, que nous qualifions d'algébriques, sont données dès les premiers énoncés qui définissent des espaces vectoriels. Nous disons « qualifier » car Banach lui-même ne fait pas référence à une notion préétablie d'opération algébrique et il redéfinit complètement les propriétés des différentes opérations —addition et multiplication par un scalaire— par une collection d'axiomes.

Nous avons montré comment les axiomes d'espace vectoriel, et en particulier ceux qui définissent les opérations algébriques, avaient été sélectionnés en vue d'établir une théorie des fonctions généralisées que Banach nomme des « opérations ». Les opérations algébriques sont en effet essentielles pour déterminer les différentes propriétés des opérations telles que Banach les envisage. Elles permettent en particulier de construire divers éléments sous formes de séries dont nous avons montré l'importance dans l'élaboration de la théorie des opérations.

La notion de norme introduite par Banach fait aussi intervenir les nombres réels et leurs opérations de façon similaire à celle que nous avons précisée plus haut au sujet de l'« écart » chez Fréchet. Nous avons analysé en détails chez

¹[Fré06, p.18]

²Voir par exemple [Fré53, p.152-153].

Banach le rôle de la norme et sa capacité à *mesurer* grâce à des nombres réels où les opérations algébriques et l'ordre jouent un rôle important.

Enfin chez Banach les champs de fonctions apparaîtront comme des cas particuliers d'espaces vectoriels et les opérations usuelles sur les fonctions seront donc analysées comme des opérations d'espace vectoriel une fois vérifiée leur adéquation aux axiomes.

Ces différents exemples montrent l'importance des opérations algébriques dans les travaux des divers mathématiciens qui œuvrent à concevoir l'analyse fonctionnelle comme champ général. Ils mettent en évidence des conceptions et des utilisations diversifiées des opérations algébriques en liens étroits avec différentes façons de concevoir et d'étudier les opérations fonctionnelles.

LA FONCTION COMME RELATION. Nous avons vu comment Volterra puis Fréchet se saisissaient de la notion de fonction telle que l'avait proposée Dirichlet³ pour définir à leur tour des fonctions généralisées.

C'est à partir de cette notion de fonction que Volterra introduit à son tour des fonctions qui dépendent d'autres fonctions ou des fonctions de lignes. Dans ce cas la variable est elle-même une fonction ou une ligne qui est envisagée sous une forme plus ou moins géométrique. La conception abstraite de fonction comme *relation* entre divers éléments permet aussi à Volterra de voir une fonction de ligne comme fonction de l'infinité de points qui composent la ligne. Cette idée amène à penser qu'une fonction possède une infinité de dérivées et illustre parfaitement la *méthode du passage du fini à l'infini* à laquelle Volterra fait souvent référence.

Fréchet convoque lui-aussi la définition de Dirichlet pour introduire des fonctions sur un ensemble abstrait. Les opérations, qui avaient permis de développer un certain calcul fonctionnel en s'appuyant sur des idées bâties sur le modèle des fonctions analytiques⁴, quittent alors la scène pour laisser place à l'étude des *opérations fonctionnelles*. Celles-ci sont des *fonctions générales*, c'est-à-dire des *correspondances* générales entre les éléments d'un ensemble abstrait et, dans un premier temps, des nombres réels, puis plus tard, des éléments eux-aussi pris dans un ensemble abstrait. Pour Fréchet c'est le choix de cette définition de fonction qui permet de développer une théorie générale dont il envisage dès le départ l'organisation en deux temps : « Il est à présumer que les propriétés de chaque opération dépendront, d'une part des propriétés de l'ensemble dans lequel elle est définie, d'autre part de la nature de la correspondance qui définit l'opération »⁵.

Chez Banach la définition d'opération sur un espace vectoriel est établie à partir de l'idée de *relation* entre des éléments de cet espace et ceux d'un second espace vectoriel. La notation $x\mathfrak{A}y$ qu'il convoque de façon éphémère au début de sa thèse détermine d'emblée une approche abstraite et générale où x et y sont des éléments quelconques d'un espace vectoriel général (chacun dans un espace distinct éventuellement). La relation ainsi définie s'accompagne d'une distribution des rôles joués par les éléments x et y qui reste indépendante de

³On retrouve la définition de Dirichlet dans un texte de 1837. (cf. [Dir37])

⁴Par exemple dans le livre de Pincherle de 1901. (cf. [AP01])

⁵[Fré06, p.4]

leur nature, l'un étant un élément du *domaine* alors que l'autre est dans le *contre-domaine*. La notation adoptée ensuite, $U(x) = y$, conforme à la notation usuelle des fonctions ordinaires, mettra en évidence la différence des rôles joués par les éléments U , x , et y tout en indiquant toujours une conception générale fondée sur le concept de *relation*.

6.1.2 TRANSPOSITION DE LA THÉORIE DES FONCTIONS ORDINAIRES

Une partie des efforts des mathématiciens en analyse fonctionnelle a consisté à transposer certaines idées de la théorie des fonctions ordinaires à l'étude des fonctions généralisées qu'ils mettaient en œuvre. Là encore nous avons examiné plusieurs cas où le parallèle entre la théorie préexistante et la théorie générale est très visible, voire assumé par le mathématicien. Le processus de transposition que nous avons cherché à mettre en évidence n'est pas uniforme et nous l'avons essentiellement illustré grâce à deux cas importants dans lesquels nous avons exploré les approches générales. Nous reprenons ici les grandes lignes des motifs que nous avons établis et dont nous avons étudié la mise en place progressive à travers plusieurs textes.

DÉRIVÉE ET DIFFÉRENTIELLE. Volterra s'efforce dès les premières années de sa carrière de construire un véritable calcul différentiel d'abord pour les substitutions —en termes d'observateurs des matrices dont les coefficients sont des fonctions dérivables— puis pour les fonctions qui dépendent d'autres fonctions et les fonctions de lignes. Il s'appuie sur un schéma classique déjà établi pour les fonctions ordinaires et la notion de quotient différentiel pour définir la dérivée.

Cependant établir la source de la transposition n'en détermine pas les mécanismes ni ne détermine la dynamique de la progression générale du mathématicien. Nous avons examiné en détails plusieurs textes de Volterra qui permettent de préciser diverses façons de sélectionner et de transposer des idées bien identifiées de la théorie des fonctions ordinaires.

Volterra conçoit diverses formes d'accroissement qui mettent à chaque fois en jeu un certain type de voisinage et une approche locale particulière. L'idée d'accroissement est déclinée dans diverses situations qui lui permettent de définir tour à tour un accroissement au voisinage d'un point, d'une courbe ou d'une fonction. Là encore ces éléments sont conçus de façons non uniformes en faisant appel par exemple à une représentation paramétrique ou encore à une vision géométrique plus descriptive.

De plus nous avons montré que la notion de *quotient différentiel* pouvait être interprétée comme l'étude de la variation d'une fonction relativement à la variation d'une autre fonction. Cette idée, commentée par Fréchet et déjà présente chez Jellett, met en avant l'importance du choix de la fonction de référence quand on analyse ainsi le rapport différentiel.

Chez Volterra, dès 1887 et le premier article sur les fonctions qui dépendent d'autres fonctions, le quotient différentiel est de la forme $\frac{\delta y}{\sigma}$ où $\sigma = \int_m^n \theta(x) dx$ représente l'aire entre la courbe de φ , variable de la fonction (généralisée) y , et la courbe de $\varphi + \theta$ où la perturbation est localisée sur un intervalle (mn) .

Volterra choisit de *mesurer* ainsi la perturbation par l'aire entre la courbe et la courbe déformée afin obtenir par un calcul subtil et un passage à la limite⁶ :

$$y(\varphi + \varepsilon\psi) - y(\varphi) = \varepsilon \int_A^B \psi(t) \cdot y'(\varphi, t) dt.$$

Nous avons montré que l'on pouvait, comme le suggère Fréchet, analyser le quotient différentiel en termes de « fonctions qui dépendent d'autres fonctions » ou de « fonctions de lignes » pour reprendre le vocabulaire de Volterra. Le terme $\int_m^n \theta(x) dx$ peut être interprété comme la variation de la fonction (généralisée) $\varphi \mapsto \int_A^B \varphi(x) dx$ qui est additive et fait donc partie des fonctions que Volterra qualifie de « fonctions simples ». Le parallèle avec la notion similaire pour les fonctions ordinaires est établi lorsque l'on pense que leur dérivée est donnée grâce à un quotient différentiel relatif à la variation de la fonction $x \mapsto x$, idée que Jellett avait explicitement mentionnée dès 1850.

L'APPROCHE DE WEIERSTRASS COMME MODÈLE. Fréchet dans sa thèse cherche à transposer les idées qu'il lit chez Weierstraß et en particulier un résultat qui lui paraît tout à fait central qui affirme qu'une fonction continue sur un intervalle fermé borné de \mathbb{R} est bornée et atteint ses bornes. Fréchet mentionne plusieurs fois l'importance de ce résultat dans sa recherche d'une théorie générale pour les opérations fonctionnelles.

Nous avons étudié comment la recherche d'un cadre général qui permettrait de conserver ce résultat avait conduit Fréchet à introduire des éléments que nous qualifions désormais de topologiques pour définir une classe d'espaces abstraits adaptés la plus générale possible.

La progression générale de Fréchet est en partie motivée par la conception d'espaces abstraits qui pourraient jouer un rôle similaire pour les opérations fonctionnelles à celui que jouent les parties de \mathbb{R} dans la théorie des fonctions ordinaires. Il ne s'agit donc pas de définir des espaces pour leurs propriétés intrinsèques seulement, mais de définir des *domaines* pour les opérations fonctionnelles dans une démarche parallèle à celle qui a fait ses preuves pour les fonctions ordinaires. Le processus de généralisation passe alors par la détermination des notions minimales qui sont nécessaires pour obtenir certains résultats que Fréchet souhaite pouvoir démontrer, parmi lesquels figure un théorème similaire à celui de Weierstraß. Nous verrons un peu plus loin les conséquences de cette démarche de laquelle naissent chez Fréchet une certaine conception de l'espace et diverses notions topologiques.

6.1.3 ÉQUATIONS DIFFÉRENTIELLES ET PROBLÈMES INTÉGRAUX

Nous avons souligné depuis l'introduction que plusieurs travaux d'historiens s'étaient attachés à montrer l'importance de l'étude des équations différentielles ou intégrales dans diverses démarches générales en analyse fonctionnelle.

⁶Nous avons ici exprimé le résultat avec des notations modernes. On rappelle que $y'(\varphi, t)$ est la limite du quotient différentiel évoqué plus haut en un point t de la courbe φ .

Notre propre étude a aussi montré qu'elles constituaient une vaste réserve de problèmes qui motivaient certaines approches abstraites.

À travers les travaux de Gregory nous avons illustré et analysé les techniques de séparation des symboles d'opération qui puisent une partie de leur inspiration dans l'analyse des équations différentielles linéaires. Nous avons étudié le lien qu'il existait entre une procédure de séparation des symboles différentiels et l'élaboration d'une vision abstraite qui permet une démarche générale.

Il nous est apparu, comme nous l'avons rappelé plus haut, que l'attribution de différents rôles aux éléments sélectionnés par le mathématicien était essentielle dans cette démarche. La séparation des symboles d'opération s'appuie, dans un domaine où les symboles semblent indifférenciés, sur la distinction entre *symbole-agissant* et *symbole-objet* qui s'élabore parallèlement à la distinction entre symbole différentiel et fonction à dériver.

Nous avons aussi montré que les équations différentielles étaient au cœur de certains travaux de Volterra. En nous appuyant sur des commentaires du mathématicien à ce sujet nous avons rappelé le lien qu'il existe entre l'étude des substitutions et les équations différentielles linéaires. Nous avons de plus montré qu'une grande partie des efforts pour élaborer un calcul différentiel des substitutions était liée chez Volterra au désir de constituer un « fondement » pour une théorie générale des équations différentielles linéaires, comme il l'avait annoncé dès son premier mémoire sur le sujet en 1887. Les domaines qu'il choisit de développer plus spécialement dans l'étude des substitutions sont ainsi façonnés dans la double optique de fournir un tel fondement tout en apportant des outils généraux efficaces.

En ce qui concerne les problèmes intégraux ils sont essentiellement liés à des équations intégrales ou à des recherches d'extrémaux. L'exploration de ces problèmes est omniprésente dans les textes que nous avons étudiés. Ils constituent à la fois une source d'inspiration dans une progression générale, et apparaissent aussi comme des éléments qui valident l'efficacité d'une théorie générale. Nous avons pu voir à plusieurs reprises que les résultats développés par divers mathématiciens qui œuvraient en analyse fonctionnelle étaient formatés pour répondre de façon immédiate à des problèmes intégraux *ordinaires* bien identifiés. Leur analyse détermine en partie le type d'approche générale que développent ces mathématiciens.

Les travaux de Volterra l'amènent à repenser les intégrales du type $\int_a^b f(x)\varphi(x)dx$ comme des fonctions (linéaire et généralisée) de la fonction φ lorsque f est donnée. Son approche consiste donc à analyser un certain type de problèmes en termes de « fonctions qui dépendent d'autres fonctions », et à leur appliquer les résultats obtenus de façon générale. Les problèmes intégraux qui ont été à la source de la démarche générale sont donc relus comme des cas particuliers dans la nouvelle « théorie des fonctionnelles »⁷ de Volterra.

⁷Plusieurs cours de Volterra porteront essentiellement sur le rôle possible de la théorie des fonctions généralisées dans l'étude de nombreux problèmes intégraux. Le cours publié sous le titre *Theory of functionals and of integral and integro-differential equations* en 1930 ou encore celui établi en collaboration avec Pérès en 1936 et intitulé *Théorie générale des fonctionnelles* en sont de parfaites illustrations. (cf. [Vol30] et [Pér36])

De son côté Fréchet a développé certaines de ses idées en suivant une voie initiée par Hadamard, qui consiste à examiner des fonctionnelles données sous formes d'intégrales. Celle-ci sont interprétées comme des opérations linéaires agissant sur des fonctions. Une question vient alors de savoir si toute opération linéaire peut être représentée sous forme d'une intégrale.

Le point de départ est un résultat d'Hadamard de 1903 qui montre que toute opération linéaire U continue, qui agit sur des fonctions continues sur un intervalle (ab) , peut s'écrire $U(f) = \lim_{n \rightarrow +\infty} \int_a^b f(x)K_n(x)dx$ où les fonctions K_n sont continues. L'analyse de ce problème a donné lieu chez Fréchet à une approche d'abord fondée sur la possibilité de développer les fonctions à intégrer en séries avec une convergence uniforme, puis petit-à-petit à une conception générale d'opération agissant sur un ensemble linéaire quelconque. Les différents champs fonctionnels classiques dans l'étude des intégrales sont alors réinterprétés comme des espaces linéaires abstraits.

Enfin Banach indique dès le titre de sa thèse en 1922 qu'il entend donner des applications de son travail général sur les opérations dans les ensembles abstraits aux équations intégrales. Nous avons vu qu'un certain nombre de théorèmes développés dans le cadre général des opérations définies sur un espace vectoriel normé prenaient en effet une forme qui permettait une application immédiate aux équations intégrales classiques de son époque. Plus tard un chapitre de son livre de 1932 sera dédié aux applications de certains résultats obtenus pour les opérations —étude spectrale en particulier— aux équations de Fredholm, de Volterra ou aux équations intégrales symétriques (intégrales à noyaux symétriques).

Ces différents exemples, reconsidérés dans l'optique que nous avons adoptée, montrent que les problèmes liés aux équations différentielles et les problèmes mettant en jeu des intégrales ont été des éléments importants dans la progression des mathématiciens et dans la détermination de leur approche générale. Ils interviennent de diverses façons dont nous avons illustré quelques aspects. Cependant bien que les concepts de dérivation ou d'intégration soient fortement impliqués dans le travail des mathématiciens qui élaborent l'analyse fonctionnelle, ils n'en déterminent pas seuls la dynamique générale. Notre approche a tenté de montrer précisément leur rôle dans diverses démarches générales sans occulter des éléments souvent sous-jacents mais qui déterminent à nos yeux tout autant le type de généralité que l'on peut lire dans les textes que nous avons sélectionnés.

6.2 ESPACES ET OPÉRATIONS

La notion d'*espace* est centrale dans notre conception actuelle de l'analyse fonctionnelle et notre étude a illustré plusieurs façons de mettre ce concept en œuvre. L'espace est toujours le fruit d'une vision abstraite qui engendre un regard réflexif sur un assemblage d'éléments que le mathématicien convoque. Parfois il s'est agi de considérer un champ de fonctions dont on mettait en valeur certains éléments ou certaines caractéristiques. Ailleurs le mathématicien construisait un espace où les éléments étaient de nature quelconque en leur adjoignant une ou plusieurs *opérations* qui déterminaient certaines propriétés

des éléments abstraits : espaces muni d'un mode de convergence, d'un voisinage, d'un écart, d'opérations algébriques, d'une norme, etc. Nous avons rencontré plusieurs situations et de multiples autres existent désormais. De nombreux mathématiciens œuvrant en analyse fonctionnelle ont en effet participé à l'élaboration de diverses notions d'espace dans des démarches générales. On peut trouver dans notre bibliographie plusieurs études qui éclairent des conceptions d'espaces que nous avons laissées de côté. Dans leur prolongement il nous semble que certains travaux, de Hilbert⁸ ou de F. Riesz en particulier, mériteraient une attention nouvelle en se concentrant plus avant sur l'élaboration intriquée des espaces et des opérations.

On peut s'attacher à comparer *a posteriori* différents types d'espaces et déterminer quel mathématicien les a exposés, sous une forme que nous reconnaissons, pour la première fois.

Notre optique est d'un autre ordre et nous avons tenté de mettre en évidence des différences de démarches qui ont mené à la conception des espaces dans plusieurs circonstances. Il s'est agi à chaque fois comprendre la dynamique à l'œuvre dans les processus de sélection et d'organisation qui permettent d'établir une notion abstraite d'espace. En particulier nous avons montré que les espaces de l'analyse fonctionnelle pouvaient être étudiés dans le lien intime qu'ils entretiennent avec diverses façons de concevoir les opérations.

Espaces de fonctions chez divers mathématiciens, classes (L) , (V) ou (E) chez Fréchet, espaces vectoriels, espaces vectoriels normés chez Banach, ainsi que de multiples variantes en fonction des propriétés supplémentaires exigées ou mises de côté, font apparaître, dans une lecture rétrospective, à la fois des éléments algébriques et des éléments topologiques. Un de nos objectifs a été de mettre en évidence quels éléments étaient convoqués, sous quelle forme et comment ils étaient mis en œuvre dans une conception générale. Nous avons déjà longuement commenté et analysé ces éléments et nous résumons donc ici rapidement quelques points essentiels qui lient les opérations aux espaces.

6.2.1 ESPACES LINÉAIRES

La notion d'espace linéaire est apparue fréquemment dans notre analyse. Nous en avons évoqué de nombreux aspects et nous avons comparé diverses façons de les concevoir.

Les différents recours à une structure linéaire que nous avons étudiés sont liés aux opérations essentiellement de deux façons.

D'une part la structure linéaire permet de concevoir d'une manière unifiée diverses collections de fonctions sur lesquelles les opérations fonctionnelles agissent. L'espace linéaire sert alors à sélectionner les éléments essentiels qui permettent l'étude des opérations fonctionnelles. Nous avons souligné, en

⁸Nous soulignons ici la nécessité d'étudier plus avant les travaux de Hilbert en particulier dans leur lien avec l'approche développée par Banach. Ce dernier mentionne en effet les travaux du mathématicien allemand dès sa thèse en indiquant leur généralité : « Ce ne furent que les travaux de Hilbert qui, bien qu'ils traitaient les formes quadratiques à une infinité de variables et non pas les fonctions de ligne, ont apporté des résultats susceptibles à être transférés facilement sur les théorèmes concernant les opérations dont le domaine et le contre-domaine se composent des fonctions de carré intégrables (L) . » (cf. [Ban22, p.133])

commençant par des textes de Peano, la proximité que les mathématiciens organisaient souvent entre une conception abstraite d'espace et l'introduction d'opérations —les « systèmes linéaires » et les « transformations » pour Peano.

D'autre part nous avons mis en évidence l'importance de la *distributivité* dans l'étude des opérations. Celle-ci est apparue sous différentes formes. Dans les exemples tirés de l'algèbre symbolique elle permettait la séparation des symboles d'opérations. Chez Jellett ou Volterra la distributivité est invoquée pour préciser la forme de la variation d'une fonction ordinaire ou généralisée. Elle permet en outre à Volterra de déterminer des « fonctions simples de lignes » dont nous avons montré l'importance à plusieurs titres. Chez Hadamard, Fréchet ou Banach, toujours selon des modalités que nous avons différenciées, la distributivité apparaît d'abord comme propriété des opérations intégrales puis elle est dégagée comme propriété fertile dans l'étude des opérations générales.

Ces exemples montrent que ce qui est retenu dans la conception algébrique de l'espace —sa structure algébrique— est lié, suivant des modalités diverses, à une certaine façon de penser les opérations fonctionnelles ainsi qu'à la possibilité d'effectuer des calculs selon des lois de l'algèbre.

6.2.2 ÉLÉMENTS TOPOLOGIQUES

Chez Volterra les notions de perturbations ou de voisinages sont apparues dans l'étude des fonctions qui dépendent d'autres fonctions et des fonctions de lignes. Nous avons pu montrer que les éléments que nous qualifions désormais de topologiques étaient liés à différentes façons de concevoir les fonctions dans un sens étendu.

Dans un premier temps Volterra choisit une idée de perturbation d'une fonction φ qui peut être donnée sous la forme $\varepsilon\psi$ et qui peut en outre être localisée en un point en prenant ψ nulle en dehors d'un intervalle (contenant l'abscisse du point). La conception d'une perturbation *locale* s'appuie alors sur deux grandeurs : son amplitude et son étendue (mots d'observateur).

Puis dans un second temps nous avons vu que la version géométrique des fonctions généralisées —les « fonctions de lignes »— avait donné lieu à une conception de voisinage tubulaire d'une ligne extrêmement général. Dans ce cas Volterra s'appuie une *estimation* de la surface transversale σ du voisinage tubulaire pour définir des voisinages de plus en plus fin. Nous disons « estimation » car Volterra ne précise pas de façon de *mesurer* cette surface, ni de définir à travers une *mesure* ce que serait un voisinage plus fin qu'un autre. La notion de voisinage est donnée géométriquement et de sa conception découle un type de généralité que l'on peut apprécier à la fois dans la définition du voisinage mais aussi dans son emploi pour définir par exemple la continuité des fonctions de lignes.

Ainsi dès 1887 Volterra définit la continuité grâce à ce type de voisinage tubulaire :

La funzione φ , *funzione delle linee* L , sarà continua se, preso un numero δ piccolo ad arbitrio, potrà trovarsi un intorno S di L tale che i valori di φ corrispondenti a tutte le linee longitudinali di S differiscano dal valore di φ in L meno di δ .

Cette définition montre une façon de tisser un lien entre la conception topologique de l'espace et la continuité des opérations fonctionnelles. D'un côté la définition de continuité dépend du type de voisinage choisi, mais d'un autre côté la théorie des fonctions de lignes se conçoit chez Volterra avec des fonctions de lignes continues et nécessite donc une notion de voisinage adaptée aux lignes.

Dans le cas de Fréchet, et en particulier dans sa thèse⁹, nous avons pu mettre en évidence un processus de sélection qui permet de concevoir un espace abstrait auquel on adjoint une notion de convergence. Là encore les opérations sont centrales dans son approche. L'organisation des idées proposée par Fréchet est clairement énoncée : il s'agit d'étudier des opérations fonctionnelles (continues la plupart du temps) et cette étude nécessite l'établissement préliminaire d'une théorie des ensembles sur lesquels sont définies les opérations. L'étude des ensembles abstraits correspond donc pour Fréchet à l'étude des parties de \mathbb{R} dans l'étude des fonctions ordinaires.

Nous avons vu que la progression de Fréchet débutait par la conception d'un ensemble abstrait E quelconque auquel il adjoint un « procédé » qui permet de dire si une suite converge, c'est-à-dire si elle possède une limite (unique). Les seules exigences que Fréchet retient imposent qu'une suite constante converge (vers le point qui la définit) et que si une suite converge, alors toute sous-suite converge vers le même point de l'ensemble abstrait. Fréchet précise que cette notion de convergence des suites est centrale dans toute étude générale des opérations tout autant qu'elle a été essentielle dans le développement de la théorie des fonctions ordinaires.

La notion de convergence permet à Fréchet de définir dans un premier temps un « élément limite » d'un sous-ensemble $F \subset E$: c'est un élément de E qui est limite d'une suite infinie de points distincts de F . Puis Fréchet définit les ensembles fermés ainsi que plusieurs notions topologiques (définies séquentiellement) sur lesquelles nous ne revenons pas. Retenons seulement ici les notions importantes d'ensembles « compacts » qui correspondent chez Fréchet aux ensembles finis ou à ceux dont toute partie infinie d'éléments contient une suite convergente¹⁰, et les ensembles « extrémaux » qui sont les compacts fermés.

Grâce à cette notion d'ensemble abstrait dans lequel on a un procédé pour dire si une suite converge, Fréchet peut définir d'une façon générale la continuité (séquentielle) des opérations sous la forme : $V(A) = \lim_{n \rightarrow +\infty} V(A_n)$ pour tout élément A limite d'une suite $\{A_n\}$.

Enfin nous avons vu que Fréchet introduisait de nouveaux types d'espaces pour obtenir davantage de propriétés des opérations fonctionnelles. Dès le second chapitre de sa thèse il définit une notion abstraite de « voisinage » donnée par la valeur d'une fonction (qui à deux points de l'espace associe un nombre réel) qui a les propriétés de ce que nous nommons désormais une distance, avec

⁹Il est à noter que les définitions données par Fréchet dans sa thèse ou dans les textes que nous avons étudiés évolueront au cours de sa carrière. En particulier la notion de compact à laquelle nous faisons référence ici est celle de sa thèse.

¹⁰La définition donnée dans la thèse est : « Nous dirons qu'un ensemble est compact lorsqu'il ne comprend qu'un nombre fini d'éléments ou lorsque toute infinité de ses éléments donne lieu à au moins un élément limite. » (cf. [Fré06, p.6]) Ce ne sont que les emplois de la notion de compact qui suivent dans son texte qui permettent de comprendre l'expression « donne lieu ».

cependant, comme nous l'avons rappelé, une condition un peu plus générale que l'inégalité triangulaire classique.

La notion d'« écart » introduite un peu plus tard correspond plus précisément à celle actuelle de distance et permet à Fréchet de donner une définition de la continuité qui est proche de celle que nous utilisons encore aujourd'hui dans les espaces métriques¹¹ :

La condition nécessaire et suffisante pour que U soit une opération continue dans E est que, si A est un élément quelconque commun à E et à E' , on puisse faire correspondre à tout nombre $\epsilon > 0$ un nombre η_A , tel que l'inégalité $(A, B) < \eta_A$ entraîne $|U(A) - U(B)| < \epsilon$ pour tout élément B de E .

Les considérations *topologiques* sont donc pour Fréchet liées à la recherche du cadre le mieux adapté pour développer une théorie des fonctions généralisées, c'est-à-dire de ce qu'il appelle des « opérations fonctionnelles ».

Enfin nous avons vu que chez Banach la conception des espaces vectoriels normés tels qu'ils sont introduits dans sa thèse était adaptée au développement des propriétés des opérations linéaires. D'une part les notions de convergence et de continuité (séquentielle) sont définies à partir de la norme. D'autre part nous avons étudié le recours à une conception géométrique qui s'appuie sur des boules (des sphères dans le vocabulaire de Banach) qui permettent d'établir des propriétés importantes de la norme et permettent de lier ces propriétés avec des constructions d'éléments qui mettent en jeu les opérations algébriques de l'espace.

Nous avons établi que dans les premiers travaux de Banach, les boules et les suites de boules ne servaient pas à déterminer une géométrie de l'espace pour son propre intérêt, mais qu'elles étaient utilisées essentiellement pour construire des points limites de l'espace qui permettent ensuite de démontrer des propriétés des opérations linéaires.

Nous ne cherchons pas ici à reprendre les divers points que nous avons étudiés en détails au cours de notre thèse, mais à préciser qu'il y a un lien fort entre différents choix opérés par les mathématiciens pour concevoir des notions topologiques et leur façon de concevoir les opérations sur des ensembles de fonctions ou des ensembles abstraits. Plus généralement dans les textes que nous avons étudiés une conception d'*espace* ou une *structure*¹² est toujours liée à une conception d'opération fonctionnelle et, conformément à l'idée de Fréchet, l'une ne va pas sans l'autre dans l'élaboration d'une « analyse générale ».

6.3 UNE HISTOIRE HÉRÉDITAIRE DE L'ANALYSE FONCTIONNELLE

DES MATHÉMATIQUES ET DES MOTS. Il y a bientôt dix ans Anne Robadey écrivait en conclusion de sa thèse sur la démarche générale dans certains travaux de Poincaré¹³ :

¹¹[Fré06, p.7]

¹²Deux de nos études prolongent certains points que nous avons évoqués dans cette thèse : [Jaë10] et [Jaë15].

¹³[Rob06]

La formation d'un concept mathématique inclut l'élaboration d'une terminologie et d'une façon d'employer cette terminologie : c'est seulement par là que le mathématicien peut faire travailler le concept qu'il a forgé. La rédaction n'est donc pas une activité annexe, littéraire, du mathématicien. C'est une part essentielle de son travail mathématique.

Nous pensons que c'est là un point essentiel et qu'un des objectifs de l'historien des mathématiques est de concevoir une dynamique qui permette de rendre compte, à partir des textes qui lui sont donnés à lire, de l'œuvre d'une terminologie et de son emploi.

Il y a plusieurs niveaux où les textes mathématiques peuvent être étudiés et en particulier l'étude locale d'un fragment, d'une preuve, d'un article ou d'une collection d'articles sont à nos yeux essentiels. En outre la lecture d'un même fragment à divers niveaux ne dit pas forcément la même chose.

C'est avant tout à ce niveau que se tisse le lien logique et la progression des idées. Mais c'est aussi au même niveau, où l'on peut emboîter quelque temps le pas d'un mathématicien, que se détermine le pacte nécessaire que celui-ci tente d'établir avec son lecteur, ce sur quoi il faut s'entendre pour suivre la pensée de l'autre.

Il y a des difficultés inhérentes au type d'analyse qui opère au plus près des textes mathématiques. Il ne s'agit pas en effet de simplement rapporter le *contenu* d'un texte écrit par tel ou tel mathématicien, serait-ce de façon pédagogique ou élégante, ni de définir quel chemin il avait choisi dans une cartographie établie *a posteriori*. Construire une dynamique qui rende compte de la démarche abstraite que les auteurs mettent en place dans leurs textes nécessite de se saisir d'un certain nombre de concepts et de jouer un instant le jeu de langage que propose le mathématicien. Or les mathématiques écrites par les auteurs de notre corpus ne sont souvent pas élémentaires même pour un lecteur de notre époque. Nous avons cherché à établir une présentation équilibrée où les textes sont présents pour permettre à notre lecteur de suivre et de ressentir dans un passage les éléments que nous cherchions à déterminer, de voir à l'œuvre la dynamique du général que nous tentions d'établir.

Cette étude locale que nous avons répétée à plusieurs moments de l'élaboration de l'analyse fonctionnelle a en outre permis de montrer que ce que l'on peut lire comme une progression des mathématiques vers une conception générale pour l'analyse fonctionnelle est en fait l'œuvre de processus et de démarches générales extrêmement divers. Nous avons voulu montrer que cette diversité pouvait elle-aussi être interprétée dans une certaine évolution qui se conçoit à une autre échelle et selon une modalité différente de celle étudiée au niveau local.

L'ANALYSE FONCTIONNELLE. L'expression « analyse fonctionnelle » nous a servi de fil conducteur et nous l'avons souvent utilisée à des fins taxinomiques pour désigner un groupement de travaux qui nous semblaient présenter une certaine cohérence. Cependant notre étude montre que le terme *fonctionnel* désigne un rapport à la notion de fonction qui dépend à la fois du concept convoqué par le mathématicien mais aussi du rôle qu'il lui fait jouer.

Par exemple en 1897 Pincherle définissait le *calcul fonctionnel* en ces termes¹⁴ :

[...] Calcul fonctionnel : on réunirait sous ce titre les chapitres de l'analyse où l'élément variable n'est plus le nombre, mais la fonction considérée en elle-même.

Ici la fonction est donc la variable. Par contraste nous avons établi une évolution de l'utilisation du mot « fonctionnel » en montrant par exemple chez Fréchet que la variable d'une opération fonctionnelle n'était plus une fonction mais un élément abstrait, et que parallèlement une notion généralisée de fonction — pensée à partir de l'idée proposée par Dirichlet — permettait de définir l'opération comme correspondance entre les éléments abstraits. Enfin chez Banach la notion d'*opération* est construite à partir de celle de *relation* et le terme « fonction » disparaît complètement de la scène où il forge sa théorie des opérations. Il réapparaîtra comme nous l'avons vu un peu plus tard pour désigner des opérations particulières à valeurs dans \mathbb{R} sous la forme de « fonctionnelle » (substantif). Enfin de nos jours la notion d'*opérateur* est proche de celle d'« opération » telle que Banach la conçoit vers 1920 et nous avons gardé le substantif « fonctionnelle » pour désigner des applications linéaires à valeurs dans le corps de base de l'espace vectoriel où elles sont définies.

Notre approche devait permettre de rendre compte de ce type d'évolution qui opère à une échelle plus large et au sein de processus complexes qui ne suivent pas un ordre chronologique strict.

La notion d'*opération* telle que nous l'avons précisée en introduction est celle qui nous a servi à désigner tout au long de notre travail diverses conceptions et qui a permis une étude dynamique des idées qui ont trait à l'analyse fonctionnelle. L'idée de Volterra de penser l'hérédité de la physique en termes de fonctions de lignes nous a séduit. Elle permet de penser une dynamique à plusieurs échelles sans embarquer, d'une façon qui la rendrait insaisissable, toute l'histoire de l'espace. Il nous semble que la *ligne*, cette fois-ci émancipée de la physique, permet à nouveau de concevoir précisément une histoire des idées mathématiques qui ne se laisse pas appréhender suivant une temporalité unique ni à travers une dynamique globale.

Loin de vouloir constituer une histoire de l'analyse fonctionnelle notre titre annonçait que notre thèse étudierait « quelques aspects » de cette histoire. La *ligne* que nous avons choisie est celle des *opérations* et elle s'appuie sur les textes que nous avons sélectionnés. De nombreuses autres lignes la croisent en tout point pour constituer un champ d'étude encore vaste à explorer.

PRÉHISTOIRE DE LA THÉORIE DES OPÉRATEURS. Nous finissons cette thèse en mentionnant son commencement qui fût le désir d'étudier l'œuvre de Stefan Banach. Bien que les théorèmes qu'il a établis figurent dans tous les traités d'histoire des mathématiques qui couvrent l'analyse du 20^e siècle, il y a en fait très peu de travaux d'ampleur sur ses recherches.

Banach a publié un peu plus d'une cinquantaine d'articles de son vivant qui constituent un ensemble de textes foisonnant d'idées, de méthodes et de résultats.

¹⁴[Pin97, p.325]

Une des difficultés provient de notre connaissance sous une forme proche de celle proposée par Banach de nombreux théorèmes qu'il a établis pour la première fois. Le livre de 1932, à l'orée duquel nous nous sommes arrêtés dans notre étude, donne un accès quasi immédiat au lecteur moderne à la grande majorité des résultats que nous retenons de l'œuvre de Banach. Pourtant ce livre offre une présentation qui, considérée seule, ne rend pas compte de la progression de Banach qui s'opère sur plusieurs années et se nourrit de l'étude de problèmes très divers.

D'autre part l'organisation générale des chapitres du livre, dont on lit *a posteriori* une cohérence rassurante, ne se laisse pas appréhender en étudiant seulement les articles de Banach. La première chose à comprendre était la notion d'« opération linéaire » ce qui constitua un but principal de cette thèse.

Les éléments que nous avons étudiés s'inscrivent donc dans une préhistoire de ce qui sera la théorie des opérateurs au 20^e siècle. Non seulement nous y avons établi diverses démarches générales dans la constitution de l'analyse fonctionnelle, mais nous avons aussi montré un changement de paradigme qui sourd dans différents textes des premières décennies du 20^e siècle et qui apparaît de façon majeure dans l'œuvre de Banach. C'est ce changement de paradigme, intimement lié à un repensé de la notion de fonction, que nous avons voulu faire ressortir dans une collection de textes de Banach établis avant son livre de 1932. Le point de vue réflexif que nous y avons dégagé, qui s'installe progressivement, permet en effet de concevoir un changement suffisant dans le regard porté sur les espaces vectoriels normés et les opérations linéaires —deux éléments devenus essentiels dans la théorie des opérateurs— pour établir à notre avis une inflexion claire dans l'évolution de l'analyse fonctionnelle et définir avec Banach l'émergence de la théorie des opérateurs du 20^e siècle.

Bibliographie

- [AB02] P. R. ALLAIRE et R. E. BRADLEY, "Symbolical algebra as a foundation for calculus : D. F. Gregory's contribution", *Historia Mathematica*, t. 29, n° 4, p. 395–426, 2002.
- [AP01] U. AMALDI et S. PINCHERLE, *Le operazioni distributive e le loro applicazioni all'analisi*. Bologna : Zanichelli, 1901.
- [Arb00] L. F. ARBOGAST, *Du Calcul des Dérivations*. Strasbourg, 1800.
- [AT05] T. ARCHIBALD et R. TAZZIOLI, "The reception of Fredholm's results on integral equations: preliminary report", *Real Anal. Exchange*, 2005.
- [AT14] —, "Integral equations between theory and practice: the cases of Italy and France to 1920", *Archive for History of Exact Sciences*, vol. 68, no. 5, pp. 547–597, 2014.
- [Arz89] C. ARZELÀ, "Funzioni di linee.", *Rom. Acc. L. Rend.*, (1889).
- [Arz99] —, "Sulle serie di funzioni (parte 1), Sessione del 28 maggio 1899", *Memorie della R. Acc. delle Scienze dell'istituto di Bologna 1899-1900*, V, t. VIII, p. 131–186, 1899.
- [Asc83] G. ASCOLI, "Le curve limite di una varietà data di curve", *Mem. Accad. Lincei*, t. 18, n° 3, p. 521–586, 1883.
- [Bai05] R. BAIRE, *Leçons sur les fonctions discontinues professées au Collège de France*. Gauthier-Villars, 1905.
- [Ban22] S. BANACH, "Sur les opérations", *Fundamenta Math.*, t. 3, 1922.
- [Ban23] —, "Sur le problème de la mesure", *Fundamenta Math.*, t. 4, p. 7–33, 1923.
- [Ban25] —, "Sur le prolongement de certaines fonctionnelles", *Bulletin des sciences mathématiques*, t. 2, p. 301–307, 1925.
- [Ban29a] —, "Sur les fonctionnelles linéaires", *Studia Mathematica*, t. 1, n° 1, p. 211–216, 1929.
- [Ban29b] —, "Sur les fonctionnelles linéaires II", *Studia Mathematica*, t. 1, n° 1, p. 223–239, 1929.
- [Ban36] —, "Die theorie der Operationen und ihre Bedeutung für die Analysis", in *Comptes Rendus du Congrès International des Mathématiciens Oslo, 1936, Tome I : Procès-Verbaux et Conférences Générales*, 1936.

- [Ban87] —, *Theory of linear operations*, E. science publisher B. V., éd., trad. par F. JELLETT. 1987.
- [Ban32] S. BANACH, *Théorie des opérations linéaires*. Instytut Matematyczny Polskiej Akademi Nauk, 1932.
- [Boo47] G. BOOLE, *The Mathematical Analysis of Logic, Being an Essay towards a Calculus of Deductive Reasoning*. Cambridge Eng. : Macmillan, 1847.
- [Boo54] —, *An Investigation of the Laws of Thought on which are Founded the Mathematical Theories of Logic and Probabilities*. Walton et Maberly (Macmillan), 1854.
- [Bor05] É. BOREL, *Leçons sur les fonctions de variables réelles et les développements en séries de polynômes*, R. par M. FRÉCHET, éd. Gauthier-Villars, 1905, Avec des notes de P. Painlevé et H. Lebesgue.
- [Bot86] U. BOTTAZZINI, *The higher calculus : a history of real and complex analysis from Euler to Weierstrass*. Springer-Verlag, New York, 1986, p. vi+332, Translated from the Italian by Warren Van Egmond.
- [BG13] U. BOTTAZZINI et J. J. GRAY, *Hidden harmony—geometric fantasies*, sér. Sources and Studies in the History of Mathematics and Physical Sciences. Springer, New York, 2013, p. xviii+848, The rise of complex function theory.
- [Cal95] B. CALÒ, "Sulle operazioni funzionali distributive.", *Rom. Acc. L. Rend. (5)*, t. 4, n° 2, p. 52–59, 1895.
- [Con28] CONGRESSO, *Atti del Congresso Internazionale dei Matematici. Bologna 3-10 Settembre 1928*, 1 t. Nicola Zanichelli, 1928, p. 267–274.
- [Cri04] T. CRILLY, "The Cambridge Mathematical Journal and its Descendants: the Linchpin of a Research Community in the Early and Mid-Victorian Age", *Historia Mathematica*, t. 31, n° 4, p. 455–497, 2004.
- [Des11] S. E. DESPEAUX, "'Very Full of Symbols': Duncan F. Gregory, the Calculus of Operations, and the Cambridge Mathematical Journal", in *Episodes in the History of Modern Algebra (1800-1950)*, J. J. GRAY et K. H. PARSHALL, éd. American Mathematical Society, 2011, p. 49–69.
- [Die81] J. DIEUDONNÉ, *History of functional analysis*, sér. North-Holland Mathematics Studies. North-Holland Publishing Co., Amsterdam-New York, 1981, t. 49, p. vi+312, Notas de Matemática [Mathematical Notes], 77.
- [Dir29] L. J. G. DIRICHLET, "Sur la convergence des séries trigonométriques qui servent à représenter une fonction arbitraire entre des limites données.", *Journal für die reine und angewandte Mathematik*, t. 4, p. 157–169, 1829.
- [Dir37] —, "Über die Darstellung ganz willkürlicher Functionen durch Sinus- und Cosinusreihen", *Repertorium der Physik*, t. Bd. I, p. 152–174, 1837.
- [DF11] J. D. DOLLARD et C. N. FRIEDMAN, *Product Integration with Application to Differential Equations*. Cambridge : Cambridge University Press, 2011, 286 p.

- [Dor95] J.-L. DORIER, "A general outline of the genesis of vector space theory", *Historia Mathematica*, t. 22, p. 227–261, 3 1995.
- [Dud79] R. DUDA, "The origins of the concept of dimension", *Colloq. Math.*, t. 42, p. 95–110, 1979.
- [Dud96] —, *Fundamenta mathematicae and the Warsaw school of mathematics*. Ed. Maison des Sci. de l'Homme, Paris, 1996, p. 479–498.
- [Dud04] —, "On the origins of functional analysis and the Lvov school", *Comment. Math. Prace Mat.*, n° Tomus specialis in Honorem Juliani Musielak, p. 5–45, 2004.
- [Dud05] —, "The discovery of Banach spaces", in *European mathematics in the last centuries*, Univ. Wrocław, Wrocław, 2005, p. 37–46.
- [Dud09] —, "Facts and myths about Stefan Banach", *Wiad. Mat.*, t. 45, n° 2, p. 281–294, 2009, ISSN : 2080-5519.
- [Dud13] —, "Leaders of Polish mathematics between the two world wars", *Comment. Math.*, t. 53, n° 2, p. 105–112, 2013.
- [Dur90] M.-J. DURAND-RICHARD, "Genèse de l'algèbre symbolique en Angleterre : une influence possible de J. Locke", *Revue d'Histoire des Sciences*, t. 43, n° 2-3, p. 129–180, 1990.
- [Dur96] —, "L'école algébrique anglaise : les conditions conceptuelles et institutionnelles d'un calcul symbolique comme fondement de la connaissance", in *L'Europe mathématique/Mathematical Europe*, C. GOLDSTEIN, J. GRAY et J. RITTER, éd. Maison des Sci. de l'Homme (Paris), 1996, p. 445–477.
- [Dur97] —, "L'impact des travaux de l'école algébrique anglaise dans les journaux scientifiques autour de 1830", *Rivista di Storia della Scienza. Serie 2. Università degli Studi di Roma "La Sapienza". Dipartimento di Matematica*, t. 3, n° 2, p. 119–156, 1997.
- [Dur07] —, "Opération, fonction et signification de Boole à Frege", *Cahiers Critiques de Philosophie*, t. 3, p. 99–128, 2007.
- [Dur11] —, "Peacock's *Arithmetic* : an Attempt to Reconcile Empiricism to Universality", *Indian Journal of History of Science*, t. 46, n° 2, p. 251–311, 2011.
- [Dyc82] W. DYCK, "Gruppentheoretische Studien", *Mathematische Annalen*, t. 20, 1882.
- [Fej03] L. FEJÉR, "Untersuchungen über Fouriersche Reihen", *Mathematische Annalen*, t. 58, n° 1-2, p. 51–69, 1903.
- [Fic63] G. FICHERA, "L'analyse fonctionnelle en Italie", *Cahiers d'Histoire Mondiale*, t. 7, p. 407–416, 1963.
- [Fou11] M. FOUCAULT, *Michel Foucault, entretien avec Claude Bonnefoy*. EHESS, 2011.
- [Fou22] J. FOURIER, *Théorie analytique de la Chaleur*. Firmin Didot, 1822, p. 639.
- [Fré03] M. FRÉCHET, *Lettre à Volterra*, 1903.

- [Fré04a] —, "Généralisation d'un théorème de Weierstrass", *Comptes rendus hebdomadaires des séances de l'Académie des Sciences*, t. 139, p. 848–850, Séance du 21 novembre 1904.
- [Fré04b] —, "Sur les fonctions de lignes fermées.", *Ann. Sci. Éc. Norm. Supér.* (3), t. 21, p. 557–582, 1904.
- [Fré04c] —, "Sur les opérations linéaires", (Note 1), *Transactions of the American Mathematical Society*, t. 5, p. 493–499, 1904.
- [Fré05] —, "Sur les opérations linéaires", (Note 2), *Transactions of the American Mathematical Society*, t. 6, p. 134–140, 1905.
- [Fré06] —, "Sur quelques points du calcul fonctionnel", *Rendiconti del Circolo Matematico di Palermo*, t. XXII, p. 1–72, 1906.
- [Fré07] —, "Sur les opérations linéaires", (Note 3), *Transactions of the American Mathematical Society*, t. 8, p. 433–446, 1907.
- [Fré13] —, *Lettre à Volterra*, 1913.
- [Fré28a] —, "L'analyse générale et les espaces abstraits", in *Atti del Congresso Internazionale dei Matematici : Bologna del 3 al 10 de settembre di 1928*, (Bologna), c. RENDICONTO DEL CONGRESSO, éd., t. 1, 1928, p. 267–274.
- [Fré28b] —, *Les espaces abstraits et leur théorie considérée comme introduction à l'analyse générale*. Paris : Gauthier-Villars (Collection de monographies sur la théorie des fonctions)., 1928.
- [Fré53] —, *Pages choisies d'analyse générale*, sous la dir. de P. GAUTHIER-VILLARS, sous la dir. de L. NAUWELAERTS, III t. 1953.
- [Gra10] I. GRATTAN-GUINNESS, "D-company : the rise and fall of differential operator theory in Britain, 1810s–1880s", *Arch. Int. Hist. Sci.*, t. 60, p. 477–528, 2010.
- [Gra84] J. D. Gray, "The shaping of the Riesz representation theorem: a chapter in the history of analysis", *Arch. Hist. Exact Sci.*, vol. 31, no. 2, pp. 127–187, 1984.
- [Gra01] J. J. GRAY, "Mathematics and natural science in the nineteenth century : the classical approaches of Poincaré, Volterra and Hadamard", in *Changing images in mathematics*, sér. Stud. Hist. Sci. Technol. Med. U. BOTTAZZINI et A. D. DALMEDICO, éd., t. 13, From the French Revolution to the new millenium, Routledge, London, 2001, p. 113–135.
- [GP11] J. J. GRAY et K. H. PARSHALL, *Episodes in the History of Modern Algebra (1800-1950)*. American Mathematical Society, 2011.
- [Gre37] D. F. GREGORY, "On the solution of linear differential equations with constant coefficients", *Cambridge Mathematical Journal*, t. 1, p. 22–35, oct. 1837.
- [Gre38a] —, "Notes on Fourier's Heat", *Cambridge Mathematical Journal*, t. 1, p. 104–108, mai 1838.

- [Gre38b] —, "On the real nature of symbolical algebra", *Transactions of the Royal Society of Edinburgh*, t. 1, p. 122–144, oct. 1838.
- [Gre38c] —, "On the solution of linear equations of finite and mixed differences", *Cambridge Mathematical Journal*, t. 1, p. 54–63, fév. 1838.
- [Gre38d] —, "On the solution of partial differential equations", *Cambridge Mathematical Journal*, t. 1, p. 123–133, mai 1838.
- [Gre39] —, "On the impossible logarithms of quantities", *Cambridge Mathematical Journal*, t. 1, p. 226–238, mai 1839.
- [Gre41] —, *Examples of the Processes of the Differential and Integral Calculus*. Cambridge Eng. : J. & J.J. Deighton, 1841.
- [Gre42] —, "On a difficulty in the theory of algebra", *Cambridge Mathematical Journal*, t. 3, p. 153–160, nov. 1842.
- [Gre43] —, "On the solution of certain functional equations", *Cambridge Mathematical Journal*, t. 3, p. 239–248, mai 1843.
- [Gre65] —, *The Mathematical Writings of Duncan Farquharson Gregory*. Cambridge Eng. : Deighton, Bell and Co., 1865, With a Biographical Memoir by Robert Leslie Ellis.
- [GJM15] A. GUERRAGGIO, F. JAËCK et L. MAZLIAK, "Lines on the horizon, Hadamard and Fréchet, readers of Volterra", in *Le regard des mathématiciens français sur l'Italie — 1815-1930*, sér. Trends in the history of science. Springer, to appear, 2015.
- [GP13] A. GUERRAGGIO et G. PAOLONI, *Vito Volterra*. Springer, 2013.
- [Had02] J. HADAMARD, "Sur les dérivées des fonctions de lignes", *Bulletin de la S.M.F.*, t. 30, p. 40–43, 1902.
- [Had03] —, "Sur les opérations fonctionnelles", *Comptes rendus hebdomadaires des séances de l'Académie des Sciences*, t. 136, p. 351–354, Séance du 9 février 1903 1903.
- [Hah27] H. HAHN, "Über lineare Gleichungssysteme in linearen Räumen.", *Journal für die reine und angewandte Mathematik*, t. 157, p. 214–229, 1927.
- [Hei70] E. HEINE, "Ueber trigonometrische Reihen", *Journal für die reine und angewandte Mathematik*, t. 71, K. W. BORCHARDT, éd., p. 353–365, 1870.
- [Hel12] E. HELLY, "Über lineare Funktionaloperationen.", German, *Wien. Ber.*, t. 121, p. 265–297, 1912.
- [Hil00] D. HILBERT, "Sur le principe de Dirichlet", *Nouvelles annales de mathématiques*, 3^e sér., t. 9, 1900.
- [Hil04] —, "Grundzüge einer allgemeinen Theorie der linearen Integralgleichungen", *Nachrichten von der Gesellschaft der Wissenschaften zu Göttingen, Mathematisch-Physikalische Klasse*, 1904.
- [Jac81] C. JACOBI, C. G. J. *Jacobi's gesammelte Werke, Vorlesungen über Dynamik*, First edition 1866, C. W. BORCHARDT, K. WEIERSTRASS, E. LOTTNER et A. CLEBSCH, éd. G. Reimer (Berlin), 1881.

- [Jaë10] F. JAËCK, "Éléments structurels en analyse fonctionnelle : trois notes de Fréchet sur les opérations linéaires", *Archive for History of Exact Sciences*, t. 64, n° 4, p. 461–483, 2010.
- [Jaë15] ———, "Generality and structures in functional analysis : the influence of Stefan Banach", in *The Oxford handbook on generality in mathematics and sciences*. Sous la dir. de K. CHEMLA, sous la dir. de R. CHORLAY, sous la dir. de D. RABOUIN. 2015, to appear.
- [Jel50] J. H. JELLETT, *An Elementary Treatise on the Calculus of Variations*. J. McGlashan, 1850, 410 p.
- [Jel60] ———, *Die Grundlehren der Variationsrechnung*. Eduard Leibrod, 1860.
- [Kop71] E. KOPPELMAN, "The Calculus of Operations and the Rise of Abstract Algebra", *Arch. Hist. Exact Sci.*, t. 8, p. 155–242, 1971.
- [Lag88] J.-L. LAGRANGE, *Mécanique analytique*. Desaint (Paris), 1788.
- [Leb04] H. LEBESGUE, *Leçons sur l'intégration et la recherche des fonctions primitives professées au Collège de France*. Gauthier-Villars, 1904.
- [Lin94] E. LINDELÖF, "Sur l'application de la méthode des approximations successives aux équations différentielles ordinaires du premier ordre", *Comptes rendus de l'Académie de Sciences de Paris*, t. 114, p. 454–457, 1894.
- [Mar48] E. MARCZEWSKI, "Sur l'oeuvre scientifique de Stefan Banach. II. Théorie des fonctions réelles et théorie de la mesure", *Colloquium Mathematicae*, t. 1, n° 2, p. 93–102, 1948.
- [MT09] L. Mazliak and R. Tazzioli, *Mathematicians at war*, ser. Archimedes: New Studies in the History and Philosophy of Science and Technology. Springer, New York, 2009, vol. 22, pp. x+194, Volterra and his French colleagues in World War I, With a preface by Ivor Grattan-Guinness.
- [Mon73] A. F. MONNA, *Functional analysis in historical perspective*. John Wiley & Sons, New York-Toronto, Ont., 1973, p. viii+167.
- [Mon75] ———, "Banach's théorie des opérations linéaires", *Nieuw Arch. Wisk. (3)*, t. 23, n° 1, p. 67–71, 1975.
- [Mon78] ———, "La théorie des fonctions, réflexions historiques", *Nieuw Arch. Wisk. (3)*, t. 26, n° 1, p. 41–53, 1978.
- [Mon82] ———, "Volterra et les fonctions de lignes : un centenaire", *Nieuw Arch. Wisk. (3)*, t. 30, n° 3, p. 247–257, 1982.
- [Moo10] E. H. MOORE, *Introduction to a form of general analysis, The New Haven Mathematical Colloquium. Lectures held September 5th to 8th, 1908*. Yale University. Yale University Press, 1910.
- [Nov68] L. NOVÝ, "L'école Algébrique Anglaise", *Revue de Synthèse*, t. III°S. n° 49-52, p. 211–222, 1968.
- [Orl48] W. ORLICZ, "Sur l'oeuvre scientifique de Stefan Banach. I. Théorie des opérations et théorie des séries orthogonales", *Colloquium Mathematicae*, t. 1, n° 2, p. 81–92, 1948.

- [Pan91] M. PANTEKI, "Relationships between algebra, differential equations and logic in England 1800-1860.", Ph.D. Middlesex University, England, déc. 1991.
- [Par11] K. H. PARSHALL, "Victorian algebra : the freedom to create new mathematical entities", in *Mathematics in Victorian Britain*, R. FLOOD, A. C. RICE et R. WILSON, éd. Oxford University Press, 2011, p. 339–356, 446–450.
- [Pea20] G. PEACOCK, *A Collection of Examples of the Applications of the Differential and Integral Calculus*. Cambridge Eng. : Printed by J. Smith, sold by J. Deighton et sons, 1820.
- [Pea30] —, *A Treatise on Algebra*. Cambridge Eng. : J. et J.J. Deighton, 1830.
- [Pea42] —, *A Treatise on Algebra. Vol. I: Arithmetical Algebra*. J. & J.J. Deighton, Cambridge, G. F. & J. Rivington, et Whittaker & Co. London, 1842.
- [Pea45] —, *A Treatise on Algebra. Vol. II: Symbolical Algebra and its Applications to the Geometry of Positions*. J. & J.J. Deighton, Cambridge, 1845.
- [Pea88] G. PEANO, *Calcolo Geometrico, Secondo l'Ausdehnungslehre di H. Grassmann, preceduo dalle Operazioni della logica deduttiva*. Fratelli Bocca Editori, 1888.
- [Pér36] J. PÉRÈS, *Théorie générale des fonctionnelles*, sér. Collection de monographies sur la théorie des fonctions. Paris : Gauthier-Villars, 1936.
- [Pie80] J.-P. PIER, "Historique de la notion de compacité", *Historia Mathematica*, t. 7, p. 425–443, 4 1980.
- [Pie07] A. PIETSCH, *History of Banach spaces and linear operators*. Springer Science & Business Media, 2007.
- [Pin97] S. PINCHERLE, "Mémoire sur le Calcul Fonctionnel Distributif", *Mathematische Annalen*, t. 49, p. 325–382, 1897.
- [Poi87] H. POINCARÉ, "Sur les résidus des intégrales doubles", *Acta mathematica*, t. 9, p. 321–380, 1887.
- [Poi11] —, "Sur la théorie des quanta", *Comptes rendus hebdomadaires de l'Académie des sciences de Paris*, t. 153, p. 1103–1108, 1911.
- [Rie76] B. RIEMANN, *Bernhard Riemann's gesammelte mathematische Werke und wissenschaftlicher Nachlass*. Teubner, 1876, 132 p.
- [Rie10] F. RIESZ, "Untersuchungen über Systeme integrierbarer Funktionen", *Mathematische Annalen*, p. 449–497, 1910.
- [Rie16] F. RIESZ, "Über lineare Funktionalgleichungen", *Acta Math.*, t. 41, n° 1, p. 71–98, 1916.
- [Rob06] A. ROBADEY, "Différentes modalités de travail sur le général dans les recherches de Poincaré sur les systèmes dynamiques", Thèse de doctorat, Université Paris 7 - Diderot, 2006.

- [Ruf05] C. RUFFIEUX, "La naissance du concept de structure algébrique en Grande-Bretagne dans la première moitié du 19^{ème} siècle. Influence des philosophes de l' "école Écossaise du Sens Commun" ", Doctorat (thèse n° 3622, français), Faculté des Sciences de l'Université de Genève, 2005.
- [Ser68] J. A. SERRET, *Cours de calcul différentiel et intégral, Tome premier. Calcul Différentiel*. Gauthier-Villars, 1868.
- [Ser15] F.-J. SERVOIS, "Analyse transcendante. Essai sur un nouveau mode d'exposition des principes du calcul différentiel", *Annales de Mathématiques Pures et Appliquées*. [Annales de Gergonne], t. 5, p. 93–140, 1814/15.
- [Sie82] R. SIEGMUND-SCHULTZE, "Die Anfänge der Funktionalanalysis und ihr Platz im Umwälzungsprozeß der Mathematik um 1900", *Arch. Hist. Exact Sci.*, t. 26, n° 1, p. 13–71, 1982.
- [Sie03] —, "The origins of functional analysis", in *A history of analysis*, sér. Hist. Math. T. 24, Amer. Math. Soc., Providence, RI, 2003, p. 385–407.
- [Tay82] A. E. Taylor, "A study of Maurice Fréchet. I. His early work on point set theory and the theory of functionals", *Arch. Hist. Exact Sci.*, vol. 27, no. 3, pp. 233–295, 1982.
- [Tay85] —, "A study of Maurice Fréchet. II. Mainly about his work on general topology, 1909–1928", *Arch. Hist. Exact Sci.*, vol. 34, no. 4, pp. 279–380, 1985.
- [Tay87] —, "A study of Maurice Fréchet. III. Fréchet as analyst, 1909–1930", *Arch. Hist. Exact Sci.*, vol. 37, no. 1, pp. 25–76, 1987.
- [Taz97] R. TAZZIOLI, "The role of differential parameters in beltrami's work", *Historia Mathematica*, t. 24, n° 1, p. 25–45, 1997.
- [Vol30] V. VOLTERRA, *Theory of functionals and of integral and integro-differential equations*. London, Blackie & Son., 1930.
- [VH38] V. VOLTERRA et B. HOSTINSKY, *Opérations Infinitésimales Linéaires*. Paris : Gauthier-Villars, 1938.
- [Vol13] V. VOLTERRA, *Leçons sur les fonctions de lignes : professées à la Sorbonne en 1912, Recueillies et rédigées par J. Pérès*. Gauthier-Villars, 1913.
- [Vol54] —, *Opere matematiche memorie e note, Volume primo 1881-1892*. Accademia nazionale dei Lincei, 1954, t. 1/5.
- [Vol56] —, *Opere matematiche memorie e note, Volume secondo 1893-1899*. Accademia nazionale dei Lincei, 1956, t. 2/5.
- [Vol57] —, *Opere matematiche memorie e note, Volume terzo 1900-1913*. Accademia nazionale dei Lincei, 1957, t. 3/5.
- [Vol60] —, *Opere matematiche memorie e note, Volume quarto 1914-1925*. Accademia nazionale dei Lincei, 1960, t. 4/5.
- [Vor11] M. VORMS, *Qu'est-ce qu'une théorie scientifique?*, Philosophie des sciences. Vuibert, 2011.

- [Win13] M. WINTER, "Les principes du calcul fonctionnel", *Revue de Métaphysique et de Morale*, t. 21, n° 4, p. 462–510, 1913.

Références des textes de Gregory, premières éditions (Chapitre 1)

Note. Le *Cambridge Mathematical Journal* était publié en trois fois au rythme de l'année universitaire (novembre, février, mai), puis republié sous forme de volumes rassemblant six parties tous les deux ans : vol. 1 (1837–1839) ; vol. 2 (1839–1841) ; vol. 3 (1841–1843) ; vol. 4 (1843–1845).

- [Gre37] D. F. GREGORY, "On the solution of linear differential equations with constant coefficients", *Cambridge Mathematical Journal*, t. 1, p. 22–35, oct. 1837.
- [Gre38a] —, "Notes on Fourier's Heat", *Cambridge Mathematical Journal*, t. 1, p. 104–108, mai 1838.
- [Gre38b] —, "On the real nature of symbolical algebra", *Transactions of the Royal Society of Edinburgh*, t. 1, p. 122–144, oct. 1838.
- [Gre38c] —, "On the solution of linear equations of finite and mixed differences", *Cambridge Mathematical Journal*, t. 1, p. 54–63, fév. 1838.
- [Gre38d] —, "On the solution of partial differential equations", *Cambridge Mathematical Journal*, t. 1, p. 123–133, mai 1838.
- [Gre39] —, "On the impossible logarithms of quantities", *Cambridge Mathematical Journal*, t. 1, p. 226–238, mai 1839.
- [Gre41] —, *Examples of the Processes of the Differential and Integral Calculus*. Cambridge Eng. : J. & J.J. Deighton, 1841.
- [Gre42] —, "On a difficulty in the theory of algebra", *Cambridge Mathematical Journal*, t. 3, p. 153–160, nov. 1842.
- [Gre43] —, "On the solution of certain functional equations", *Cambridge Mathematical Journal*, t. 3, p. 239–248, mai 1843.

Table des figures

Multiplication par $a(\cos x + (-)^{\frac{1}{2}} \sin x)$ chez Gregory (Ch. 1)	44
Intégrale curviligne et intégrale double chez Volterra (Ch. 2)	82
Découpe d'un accroissement Φ pour les fonctions de lignes (Ch. 2)	111
Voisinage par translation (Ch. 2)	125
Concaténation des lignes $L_1 + L_2$ (Ch. 2)	129
Voisinage et addition des lignes (Ch. 2)	129

Première lettre de Fréchet à Volterra (Automne 1903)

Monsieur le professeur

J'ai été engagé par M. Borel à m'adresser à vous de sa part pour obtenir quelques renseignements au sujet d'une question mathématique qui me préoccupe. C'est pourquoi j'ai pris la liberté de vous écrire cette lettre à laquelle j'espère que vous voudrez faire bon accueil. Permettez-moi d'abord de me présenter rapidement afin de vous faire connaître votre correspondant.

Je viens de sortir de L'École Normale Supérieure de Paris après avoir passé l'agrégation de mathématiques. Monsieur Hadamard et Monsieur Borel m'ont fait l'honneur, chacun de leur côté, de me charger de la rédaction de leurs derniers cours intitulés « Calcul des Variations » et « Séries de Polynômes » et qui paraîtront prochainement. J'ai ensuite l'intention de préparer le doctorat de mathématiques. Mais je n'ai pas encore choisi le sujet de ma thèse. Cependant, j'ai été vivement intéressé par la lecture de mémoires que vous avez publiés sur les fonctions de lignes et c'est à ce propos que je vous écris. Tout d'abord, il me semble me rappeler que vous dites dans votre mémoire des Acta Mathematica de 1888 que vous avez l'intention de le continuer. Est-ce que cette suite à jamais paru ?

J'aurais beaucoup de plaisir à la connaître. Cela m'intéresse aussi par cette raison que j'ai obtenu pendant mon séjour à l'École quelques minuscules propositions qui se trouveraient peut être dans ce second mémoire.

Voici les résultats (sans démonstrations) que je voudrais vous soumettre si vous avez la patience de me lire : On peut considérer la dérivée d'une fonction analytique $f(z)$ comme la limite du rapport $\frac{\Delta f(z)}{\Delta z}$ de la variation de $f(z)$ à la variation de la fonction analytique la plus simple : z . Ne pourrait-on pas définir de même la dérivée d'une fonction de ligne fermée U_L par la limite $\frac{\Delta U_L}{\Delta V_L}$, V_L étant une fonction de lignes déterminée une fois pour toutes (par exemple l'aire pour une courbe fermée plane).

On supposerait que cette limite est indépendante de la manière dont toute la ligne variée L' tend vers L (je me borne dans la suite aux lignes fermées). J'ai trouvé que cela ne peut arriver que si U_L est une fonction (au sens ordinaire du mot) de V_L . Par conséquent cette extension ne conduirait pas à un résultat nouveau. Mais la forme positive de cette proposition m'a semblé intéressante et je l'ai généralisée. J'appelle avec vous δU_L la différentielle de U_L lorsque L fait partie d'une famille F de courbes à un paramètre.

Si l'on a quelle que soit la famille F :

$$\delta U_L = A_L^{(1)} \delta K_L^{(1)} + \dots + A_L^{(n)} \delta K_L^{(n)}$$

$A_L^{(i)}, K_L^{(i)}$ étant certaines fonctions de lignes continues, quand même on ne serait sûr de cette relation que si $\delta K_L^{(1)}, \dots, \delta K_L^{(n)}$ ne sont pas tous nuls, la relation ne pourra avoir lieu que si U_L est une fonction au sens ordinaire du mot de $K_L^{(1)}, \dots, K_L^{(n)}$. Le théorème est aussi vrai si L désigne une surface ou une multiplicité quelconque.

Ce qui précède m'a fait mieux comprendre combien vous aviez eu raison de prendre votre définition qui consiste au fond à généraliser l'expression de la variation de l'intégrale : $\int Pdx + Qdy + Rdz$.

Je ne considère donc dans la suite que des fonctions de lignes fermées qui jouissent de cette propriété que l'on a :

$$(1) \quad \delta U_L = \int_L (U'_x \delta x + U'_y \delta y + U'_z \delta z) ds$$

U'_x, \dots désignant des quantités fonctions de la ligne L et du point (d'arc s).

Si l'on part de cette définition on voit que U'_x, U'_y, U'_z sont les projections d'un vecteur U' (normal à L) qui est bien déterminé connaissant U_L . Car si L' coïncide avec L sauf par un arc limitant une petite aire S d'un plan \mathbf{P} le rapport $\frac{U_{L'} - U_L}{\Delta S}$ a pour limite la projection du vecteur U' sur le plan \mathbf{P} . Je suis ainsi une marche un peu différente de la vôtre mais qui en découle directement. La formule (1) peut aussi s'écrire :

$$(2) \quad \delta U_L = \int_L (\text{projection du vecteur } A \text{ sur } n) \times \delta S$$

n la normale à la surface engendrée par L et A un vecteur qu'on peut placer sur une normale quelconque à U' (mais qui varie avec cette normale).

Je me suis demandé si le vecteur U' pouvaient être choisi à l'avance pourvu qu'il soit normal à L , ou bien encore si le vecteur A n'est pas a priori (connaissant L) soumis à certaines conditions. Plusieurs remarques confirment cette opinion :

I. Si dans la formule (2) A est un vecteur qui est déterminé en un point $M(x, y, z)$ quelle que soit la ligne L qui y passe, on a nécessairement l'équation vectorielle $\nabla A = 0$ et cela suffit pour qu'il y ait une fonction U_L satisfaisant à (2). Cette fonction est (ce que j'appellerai une fonction simple) de la forme

$$U_L = \int_L P(x, y, z) dx + Qdy + Rdz + \text{const.}$$

Plus généralement la condition nécessaire et suffisante pour qu'une fonction U_L soit fonction au sens ordinaire de n fonctions simples est que l'une des déterminations du vecteur A soit de la forme :

$$A_{L,M} = \sum_{i=1}^{i=n} K_L^{(i)} A_M^{(i)}$$

$K_L^{(i)}$ étant des nombres fonctions de la ligne L seulement (et non du point M) et $A_M^{(i)}$ étant des vecteurs fonctions du point M seulement (et non de la ligne L) tels que l'on ait $\nabla A_M^{(i)} = 0$.

II. Au contraire supposons que A soit le même tout le long de L . Alors cela n'est possible que si l'on a :

$$A_i = \frac{\partial}{\partial S_i} \varphi(S_1, S_2, S_3) \quad (i = 1, 2, 3)$$

A_1, A_2, A_3 étant les projections de A et S_1, S_2, S_3 étant les aires des projections de L , φ une fonction ordinaire. D'ailleurs cela suffit pour qu'il y ait une fonction U_L (déterminée à une constante près).

III. Il ne suffit donc pas de satisfaire à l'équation $U'_x dx + U'_y dy + U'_z dz = 0$ pour que U'_x, U'_y, U'_z correspondent réellement à une fonction U . Autrement dit, on ne peut pas prendre arbitrairement deux des fonctions U'_x, U'_y, U'_z de L et de M .

Mieux encore il suffit de se donner la seule combinaison :

$$K_{L,M} = xU'_x + yU'_y + zU'_z \quad (x, y, z \text{ coordonnées de } M)$$

pour que U soit complètement déterminé (si le problème est possible) à une constante près (en admettant que U_L tende vers un nombre fixe U_0 quand L se réduit à l'origine).

J'ai aussi généralisé votre proposition concernant l'équation fonctionnelle : $U_{L+L'} = U_L + U_{L'}$ ($L + L'$, somme géométrique des contours orientés L, L') :

Si une fonction U_L (satisfaisant encore à (1)) possède un théorème d'addition, c'est une fonction (au sens ordinaire du mot) d'une fonction simple.

J'entends par théorème d'addition, que l'on a

$$U_{L+L'} = \varphi(U_L, U_{L'})$$

φ étant une fonction ordinaire dérivable.

J'ai démontré le théorème précédent en supposant ce qui est bien naturel que $\frac{U_L - U_0}{S}$ a une limite différente de zéro lorsque L se réduit à un point M au moins d'une certaine manière.

Comme je vous l'ai déjà expliqué ces observations m'ont été suggérées il y a 2 ans par la lecture de vos mémoires dont Monsieur Hadamard avait parlé dans son cours. Je n'ai plus eu le temps de m'en occuper depuis. Mais avant de m'y remettre, je désirerais vivement connaître votre opinion pour savoir s'il vaut la peine que je reprenne (ou plutôt que je commence) mes recherches. Je serais aussi très heureux si vous vouliez bien m'indiquer dans quel sens il serait le plus utile de les porter. Cette question est connexe des travaux de Monsieur Pincherle que je désirerais aussi connaître.

Je vous demande mille pardons de vous avoir importuné aussi longtemps ; je ne me le serais certainement pas permis si Monsieur Borel ne m'avait pas parlé en termes aussi flatteurs de votre bienveillance pour les jeunes étudiants. Je ne parle pas, bien entendu, de la haute estime en laquelle vous tiennent les mathématiciens français.

Veillez recevoir, Monsieur le professeur, l'assurance de mon profond respect.

Maurice Fréchet
Agrégé de mathématiques
7 Rue Bausset, XV^e, Paris
France

Lettre de Fréchet à Volterra (30/07/1913)

Monsieur et cher collègue

Je vous prie de m'excuser de ne pas vous avoir répondu plus tôt. Je suis occupé à faire passer des examens qui ne me laissent pas une minute.

Je n'avais pas pensé en vous envoyant une épreuve que vous auriez des rectifications à faire. J'avais en effet le souvenir d'être resté en accord avec ce que vous dites dans votre livre sur les équations intégro différentielles, des remarques de M. Hadamard. Aussi j'avais envoyé le « bon à tirer » en même temps que je vous envoyais l'épreuve en double.

Il n'est donc pas possible de corriger maintenant mon article.

Cependant, comme les *Transactions of the American Math. Soc.* ont accepté un article que j'ai écrit sur un sujet connexe, je pourrais dans ce second article corriger ce qu'il y avait d'inexact dans le premier.

J'ai certainement lu les articles que vous me citez ; mais je ne m'étais plus souvenu de ce que vous me dites des points singuliers. C'est cela qu'il y aurait lieu de faire remarquer dans mon second article. D'autre part — et je vous serais reconnaissant de me dire votre avis sur ce point — l'observation de M. Hadamard me semble cependant comporter quelque chose de nouveau : l'idée d'imposer à la dérivée fonctionnelle la condition abstraite de vérifier la condition $V_{f_1+f_2} = V_{f_1} + V_{f_2}$ tandis que vous imposiez à cette dérivée une forme déterminée : soit la forme intégrale soit celle qu'on obtient en ajoutant les termes relatifs aux points singuliers. Et cela n'est pas la même chose. J'ai démontré en effet, en utilisant un résultat de F. Riesz que toute opération linéaire portant sur des fonctions continues (non nécessairement dérivables) est décomposable d'une seule manière (ce terme se précisant facilement) en trois parties : les deux premières de la forme que vous considérez :

$$\int_a^b f(x)\alpha(x)dx \quad \alpha(x) \text{ sommable au sens de Lebesgue}$$
$$\sum_1^\infty f(n)A_n \quad \sum A_n \text{ absolument convergente}$$

(et peut-être sont-elles un peu plus générales α n'étant supposé que sommable) et la troisième d'une forme tout nouvelle étant une intégrale de Stieltjes

$$\int_a^b f(x)du(x) \text{ c'est-à-dire } \lim \sum f(\xi_i)[u(x_i) - u(x_{i-1})]$$

où u est non seulement à variation bornée comme dans le théorème de Riesz mais encore étant une fonction continue et à dérivée nulle sauf peut-être dans un ensemble de points de mesure nulle.

Il me semble donc bien que l'observation de M. Hadamard est bien nouvelle.

D'autre part l'idée qui constitue le fond de mon article et que je crois aussi nouvelle consiste à adopter le point de vue de M. Hadamard mais en l'appliquant non à la variation de $f(x)$ mais à son accroissement ; autrement dit de définir non la variation de l'opération c'est-à-dire sa dérivée dans une famille de fonctions $f(x)$ à un paramètre mais sa différentielle, la fonction $f(x)$ étant variable d'une façon absolument quelconque. Je ne ramène pas la question à la théorie des fonctions d'une variable qui serait un paramètre ϵ dont dépendrait la fonction $f(x)$; je reste dans le Calcul Fonctionnel.

Vous voudrez bien me dire si je ne me trompe pas en considérant ainsi que les observations de M. Hadamard et celles que je fais dans mon article constituent un perfectionnement de la Théorie que vous avez eu le grand mérite de fonder.

Je ne connais pas l'article de M. Callo¹ celui-ci dont vous me parlez. Pour moi le résultat essentiel en ce qui concerne les opérations continues et distributives est celui de M. Riesz : toute opération continue et distributive peut s'écrire :

$$U_f = \int_a^b f(x) d\alpha(x)$$

où α est une fonction à variation bornée absolument quelconque et où l'intégrale est prise au sens de Stieltjes.

Ce théorème a été mis sous d'autres formes. Je vous en ai signalé une dans la page précédente. Une autre est due à M. Lebesgue. On peut écrire :

$$U_f = \int_0^1 f(\beta(y)) \times \theta(y) dy$$

$\beta(y)$ et $\theta(y)$ étant des fonctions sommables, $a \leq \beta(y) \leq b$, et l'intégrale étant prise au sens de Lebesgue.

Vous connaissez sans doute ces propositions, je vous les signale au cas où vous n'en auriez pas pris connaissance.

Agréez Monsieur et cher collègue, l'assurance de ma très haute considération.

M. Fréchet

Mon adresse reste Faculté des Sciences Poitiers.

¹Nous pensons qu'il peut s'agir en fait de Benedetto Calò, qui fut assistant de Volterra de 1895 à 1896 alors que celui-ci occupait la chaire de Mécanique Rationnelle de l'université de Turin. Calò, qui a fait une courte carrière universitaire, a publié quelques résultats intéressants dont l'un porte le titre « Sulle operazioni funzionali distributive » en 1895, où il étend des idées de Pincherle. (cf. [Cal95])

