



HAL
open science

Formes et fonctions, en théories et en pratiques.

Thomas Lefèvre

► **To cite this version:**

Thomas Lefèvre. Formes et fonctions, en théories et en pratiques. : L'intérêt d'une approche morpho-dynamique fondée sur l'information et la géométrie.. Théorie de l'information et codage [math.IT]. Université de Bretagne Occidentale (UBO), 2014. Français. NNT: . tel-01078992

HAL Id: tel-01078992

<https://theses.hal.science/tel-01078992>

Submitted on 30 Oct 2014

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Copyright



**THÈSE DE DOCTORAT DE
L'UNIVERSITÉ DE BRETAGNE OCCIDENTALE**

sous le sceau de l'Université Européenne de Bretagne

École doctorale SICMA

(Santé, Information, Communication, Mathématiques, Matière – Brest)

Présentée par

Thomas LEFÈVRE

Pour obtenir le grade de

DOCTEUR de l'UNIVERSITÉ DE BRETAGNE OCCIDENTALE

Sujet de la thèse :

**Formes et fonctions, en théories et en pratiques : l'intérêt d'une
approche morphodynamique fondée sur l'information et la
géométrie.**

soutenue le 22 mai 2014

devant le jury composé de :

Prof. Christian ROUX	Directeur-adjoint EMSE / <i>Directeur de thèse</i>
Prof. Jacques DEMONGEOT	PU-PH, Chercheur de l'AGIM, CNRS / <i>Rapporteur</i>
Prof. Bruno FALISSARD	PU-PH, Directeur Inserm U669 / <i>Rapporteur</i>
Prof. Séverine ANSART	PU-PH, Chercheuse Inserm 1101, UBO / <i>Examineur</i>
Prof. Paul BOURGINE	Chercheur du CREA, CNRS-Polytech. / <i>Examineur</i>
Prof. Éric STINDEL	PU-PH, Directeur Inserm 1101, UBO / <i>Examineur</i>



Formes et fonctions, en théories et en pratiques :
l'intérêt d'une approche morphodynamique en
sciences.

Vers une théorie morphodynamique basée sur
l'information et la géométrie.

Thomas Lefèvre

Avril 2014

Table des matières

1	Une introduction	2
1.1	Des problèmes transverses : trois exemples inattendus, une question centrale	6
1.1.1	Médecine, biomédecine et anatomie : <i>Naissance de la clinique</i> , et géométrie des infections	7
1.1.2	Les formes de la société et du fait social : fondations de la sociologie, la <i>distinction</i> et le <i>suicide</i>	12
1.1.3	Entre les deux : psychologie, psychiatrie et santé mentale ou la crise du modèle et de la classification	16
1.1.4	Formulation du problème général	20
1.1.5	Plan général des travaux	23
2	Une brève histoire des théories et modèles en morphogenèse	27
2.1	Les grands noms des théories morphogénétiques	31
2.1.1	Wentworth d'Arcy Thompson : <i>On growth and form</i>	31
2.1.2	Alan M. Turing : morphogenèse et chimie	37
2.1.3	Turing, Gödel et von Neumann : machine universelle, approche algorithmique et automates	46
2.1.4	René Thom et ses prédécesseurs : systèmes dynamiques et théorie des catastrophes	54
2.1.5	Place de la fractalité : Benoît Mandelbrot	57
2.1.6	Une dualité sans cesse rencontrée : discret et continu, objets <i>a priori</i> ou variations topologiques?	61
2.2	Éléments de théorie des systèmes dynamiques – théorie des catastrophes	63
2.2.1	Principes généraux et hypothèses	63
2.2.2	Formulations et définitions	65
2.2.3	Points fixes et notion de stabilité	66
2.2.4	Bifurcations	69
2.2.5	Théorie de Thom et ensemble catastrophe pour un champ gradient	71
2.2.6	Un exemple : modèle de Lorenz	73

2.2.7	Extension : théorie KAM, et formulation en espace Riemannien des ODE	74
2.3	Place des deux grandes théories physiques actuelles : théories quantiques et relativistes	80
2.3.1	Principes généraux en théories quantiques : états, opérateurs, espaces propres et opérateurs d'évolution . . .	80
2.3.2	Principes généraux en théories relativistes : élément d'espace-temps, métrique et géométrie Riemannienne .	92
2.3.3	Les invariants, les groupes de transformation et la géométrie	107
2.3.4	Place du temps dans les 2 théories	111
2.3.5	Evolution à structure fixe <i>vs</i> évolution à structure mouvante	112
2.4	Quelques mots à propos des approches sciences cognitives et des neurosciences	114
2.4.1	Représentations mentales et cerveau	115
2.4.2	Chimie moléculaire, circuits neuronaux et cognition . .	116
2.4.3	Formes, vision et cerveau : neurogéométrie de la vision	118
2.4.4	Objectivité et subjectivité, toujours	119
2.5	Ce que l'on en retient	119
3	La modélisation des systèmes complexes en sciences et en biomédecine	121
3.1	Analyse et synthèse	124
3.1.1	Le réductionnisme, la méthode expérimentale et la norme	125
3.1.2	L'exemple de la recherche biomédicale et les approches statistiques classiques	130
3.1.3	L'exemple de la psychiatrie et des DSM – les modèles biomédical et biopsychosocial	134
3.1.4	Un pas en avant ? L' <i>Epigenetic landscape</i> de Waddington	138
3.1.5	Rechercher la forme et les invariants – méthodes de <i>clustering</i>	143
3.1.6	Géométrie et dynamique	153
3.2	Quelles alternatives pour prendre en compte la complexité ? .	154
3.2.1	De l'absence de définition consensuelle de la complexité	155
3.2.2	Conserver la complexité	157
3.2.3	Restaurer la complexité	159
3.2.4	Le choix du modèle	160
3.3	Causalités, temps et forme	161
3.3.1	Causalité linéaire, proportionnalité directe et boucle ouverte	161
3.3.2	Place du temps – opposition à la forme ?	163
3.3.3	Traitement de l'information, inférence de la causalité, approche bayésienne	165

3.3.4	Modèles allométriques, temps physiologiques et hypothèse de West.	171
3.4	Trois grands modèles possibles	179
3.4.1	Entre onde et corpuscule, numérique et analogique – une histoire qui perdure	179
3.4.2	L'espace plat et le lien probabiliste : la forme ponctuelle	181
3.4.3	L'espace courbe et le lien local : la forme globale . . .	182
3.4.4	L'espace des phases des systèmes dynamiques : la forme comme attracteur	183
3.4.5	Comparaison des 3 modèles	186
3.5	Les techniques associées aux 3 grands modèles	189
3.5.1	Techniques projectives : ACP/PCA et <i>PCA-likes</i> . . .	189
3.5.2	Techniques d'apprentissage des variétés – NLDR . . .	194
3.5.3	Techniques de reconstruction des attracteurs	199
3.5.4	Problème de la dimensionnalité <i>vs</i> complexité	204
3.5.5	Quantification de <i>l'epigenetic landscape</i>	207
4	Capturer, intégrer et traiter l'information : exemples et frontières actuelles des approches par la forme	208
4.1	Prédire ou expliquer	210
4.1.1	Modèles explicatifs	211
4.1.2	Modèles prédictifs	212
4.1.3	Limites des approches synthétiques en statistique : revues et méta-analyses	214
4.2	La géométrie des infections comme démarche prototype . . .	216
4.2.1	Une brève histoire des maladies infectieuses comme pathologie archétypique	217
4.2.2	La place particulière du VIH-SIDA	220
4.2.3	Panorama actuel et concepts en maladies infectieuses .	233
4.2.4	Vers une géométrie des maladies infectieuses : topologie de la causalité	238
4.2.5	Au-delà de la pathologie infectieuse : l'exemple du cancer et de son diagnostic	240
4.3	Modéliser un phénomène infectieux	242
4.3.1	Le modèle compartimental et les épidémies	243
4.3.2	Approche des infections ostéo-articulaires : un cas typique	245
4.3.3	Modélisation du système immunitaire	246
4.3.4	Le modèle de réponse immunitaire non spécifique en 3 vagues de l'infection pulmonaire	247
4.3.5	Modélisation de l'homéostasie osseuse	251
4.3.6	Modèle d'IOA : hypothèses de couplage des systèmes .	255
4.3.7	Intégration de la stratégie thérapeutique au modèle d'infection ostéoarticulaire	257

4.3.8	Complexité du modèle et analyse de sa stabilité structurelle : limites de l'approche	260
4.4	Typologies et dynamiques sociales	261
4.4.1	Existe-t-il des formes sociales ? Premières tentatives en sociologie	262
4.4.2	Exemple en épidémiologie sociale : une typologie de recours aux soins en France	265
4.4.3	Opérationnalité des classes et recherche interventionnelle	269
4.4.4	Entre sociologie des interactions et architecture urbaine : de la taille des villes	272
4.5	Au-delà de la pathologie : les facteurs de risque, les syndromes et autres objets de transition	273
4.5.1	L'exemple des maladies chroniques	278
4.5.2	La pathologie ne saurait plus se baser uniquement sur une donnée uniquement physiologique	280
4.5.3	Polypathologies chroniques : quelles dimensions ?	283
4.6	Santé mentale	286
4.6.1	Psychanalyse et physique, une histoire commune : Mach, Freud et Einstein	287
4.6.2	Bref historique de l'objet et de ses classifications en psychiatrie	288
4.6.3	La tentation de la génétique et des neurosciences : insuffisances et impasses	290
4.6.4	Un cas d'école de <i>big data</i> : approche géométrique, intégrative et dynamique basée sur l'observation	293
4.6.5	Principe de modélisation d'un objet psychiatrique : le modèle de Lorenz et la bipolarité	301
4.7	Intégration et traitement des données : les <i>big data</i> en pratique	311
4.7.1	De l'autre côté du spectre des RCT, la fouille de données (le <i>data mining</i>)	311
4.7.2	Capture de routine et entrepôts de données	314
4.7.3	Concept de CDC	315
4.7.4	SPOT-RISC	318
4.8	Retour sur les IOA	329
4.8.1	Modélisation ABM pour la spatialisation	329
4.8.2	Intégrations multi-échelle, multimodale, patient-spécifique	332
4.9	Ce que l'on en retient	333
5	Cadre théorique pour une morphodynamique basée sur la géométrie et l'information	334
5.1	Quels principes transversaux ?	335
5.1.1	Plan général de la démarche	335
5.1.2	Principes de Fermat et de Maupertuis	339
5.1.3	Principe de moindre action, formulation lagrangienne	341

5.1.4	Principes dynamiques	343
5.1.5	Principe de relativité	344
5.2	Causalité et temps – forme fonctionnelle	345
5.2.1	Structure causale d’un ensemble d’observations échantillonnées	347
5.2.2	Temps interne et mesure des événements causaux	348
5.2.3	Utilisation d’automates informationnels – analyse bayésienne	349
5.3	Géométrie informationnelle	350
5.3.1	Formulation du problème général et enjeux	350
5.3.2	Approche relativiste	351
5.3.3	Forme générale du problème	352
5.3.4	Définition du système observé	352
5.3.5	Définition d’un GAD, type réseau Bayésien	352
5.3.6	Définition de l’espace informationnel associé au GAD	357
5.3.7	Définition de la métrique et de l’intervalle infinitésimal	361
5.3.8	Courbure de l’espace informationnel et origine de la dynamique géométrique	364
5.3.9	Hypothèse du lien entre courbure de l’espace et contenu de l’espace	367
5.3.10	Une autre vision : de l’information locale vers l’Hamiltonien du système	369
5.3.11	Hamiltonien multi-événements ou multi-temps	373
5.4	Aspects opérationnels et problèmes ouverts	374
5.4.1	Reconstruire l’espace informationnel : une approche algorithmique	374
5.4.2	Formulations alternatives basées sur l’action	379
5.4.3	Intégration de l’échelle	380
5.4.4	Quelle interprétation du concept relativiste dans le cadre de la géométrie informationnelle?	381
5.4.5	Le temps physiologique de West	382
6	Conclusions et perspectives	387
6.1	Limites des travaux	388
6.1.1	Autres approches : réseaux booléens, théorie de la viabilité, grammaires génératives	388
6.2	Développement des algorithmes de géométrie informationnelle	390
6.2.1	Une question en suspens : réalisme local <i>vs</i> mécanique quantique	392
6.2.2	Formulation hamiltonienne multi-événements	393
6.2.3	Géométrie non commutative	394
6.3	Tests et interprétations en pratique courante	396
6.3.1	Observations et nouveautés	396

6.3.2	Simulations, données réelles, et développement des algorithmes.	396
6.4	Publications et communications	397
6.4.1	Publiés ou acceptés	398
6.4.2	Soumis	398
6.4.3	En cours	398
6.4.4	Communications orales	398
6.4.5	Autres publications de l'auteur	399
A	Éléments de calcul vectoriel	401
A.1	Outils de calcul vectoriel	401
A.1.1	Espace vectoriel et espace dual - forme linéaire	401
A.1.2	Notation d'Einstein	403
A.2	Outils de géométrie Riemannienne	404
A.2.1	Définition du tenseur métrique	405
A.2.2	Coordonnées curvilignes et symboles de Christoffel . .	406
A.2.3	Espace tangent	406
A.2.4	Dérivation tensorielle et écriture des opérateurs classiques	408
B	Éléments de géométrie différentielle	410
B.1	Variétés et fibrés	410
B.2	Notion de jets	414
C	Géomérodynamique et formalisme Hamiltonien multi-temps	417
C.1	Éléments de géomérodynamique	417
C.1.1	Equations d'Euler-Lagrange à temps unique.	418
C.1.2	Hamiltonien à temps unique.	418
C.1.3	Equations d'Euler-Lagrange multi-temps	419
C.1.4	Equations d'Hamilton multi-temps	419
C.2	Le fibré de jets $J^1(T, M)$ et le fibré dual, support de l'Hamiltonien multi-temps	421

Table des figures

1.1	Géométries euclidienne et non euclidienne	11
1.2	Une réalisation de la bouteille de Klein	18
1.3	Une réalisation du ruban de Möbius	19
2.1	Exemples de transformations géométriques liant différentes espèces, d'après d'Arcy Thompson	32
2.2	Exemples de transformations géométriques liant des espèces de poisson, d'après d'Arcy Thompson	33
2.3	Exemple schématique d'une transformation géométrique liant deux espèces de poisson, d'après d'Arcy Thompson	33
2.4	Exemple schématique d'une transformation géométrique liant deux espèces de poisson, d'après d'Arcy Thompson	34
2.5	Application de transformations géométriques simples à un crâne humain, type d'Arcy Thompson	35
2.6	État proche initial d'un système de diffusion-réaction type Turing	41
2.7	État intermédiaire d'un système de diffusion-réaction type Turing (1)	42
2.8	État intermédiaire d'un système de diffusion-réaction type Turing (2)	43
2.9	État final d'un système de diffusion-réaction type Turing (2D)	44
2.10	État final d'un système de diffusion-réaction type Turing (3D)	45
2.11	Schéma du mécanisme du canard de Vaucanson	48
2.12	Photographie du dispositif du canard de Vaucanson	49
2.13	Longueurs des côtes selon l'échelle d'observation, (Wikipedia, licence CC)	59
2.14	Dimensions d'une courbe selon son degré de tortuosité, d'après <i>Genre-Grandpierre</i>	60
2.15	Une représentation de l'attracteur étrange de Lorenz (Wikipedia, licence CC)	75
2.16	Causalité, signe de l'élément d'espace-temps et cône de lumière (Wikipédia, licence CC)	101

3.1	Représentation schématique de l' <i>epigenetic landscape</i> de Waddington (1)	139
3.2	Représentation schématique de l' <i>epigenetic landscape</i> de Waddington, cellules souches et attracteurs, <i>Science</i>	141
3.3	Représentation schématique de l' <i>epigenetic landscape</i> de Waddington, cellules T, <i>Nature Reviews</i>	142
3.4	Représentation schématique de l' <i>epigenetic landscape</i> de Waddington (affections rhumatismales, <i>Nature</i>)	143
3.5	Exemple simple de groupes à identifier par différents algorithmes de clustering (taille-poids)	146
3.6	Résultats de différents algorithmes de clustering appliqués à un exemple simple, en fonction du nombre de groupes cherchés	147
3.7	Résultats de différents algorithmes de clustering appliqués à un exemple simple, en fonction du nombre de groupes cherchés (suite).	148
3.8	Valeurs de la silhouette pour différents algorithmes de clustering	149
3.9	Valeurs des robustesses de classe pour différents algorithmes de clustering	150
3.10	Complexité topologique : selon la structure des données, comment savoir lequel de a ou c est plus proche de b?	159
3.11	Un exemple de réseau bayésien	168
3.12	Temps absolu et temps physiologique (d'après <i>West</i>)	175
3.13	Les deux axes d'inertie selon l'analyse en composantes principales (Wikipedia, licence CC)	191
3.14	Trois ensembles simples de données, et leur déploiement en deux dimensions	200
3.15	8 algorithmes de déploiement de variétés appliqués à l'ensemble rouleau suisse	201
3.16	8 algorithmes de déploiement de variétés appliqués à l'ensemble rouleau suisse percé	202
3.17	8 algorithmes de déploiement de variétés appliqués à l'ensemble hélice toroïdale	203
4.1	Rétrovirus VIH au microscope électronique, comme communiqué en 1983, dans <i>Science</i>	226
4.2	Rétrovirus VIH au microscope électronique. CDC PHIL	227
4.3	Une représentation de l'espace social selon Bourdieu (d'après <i>La Distinction</i>)	264
4.4	Robustesse des groupes en fonction du nombre de groupe, pour la typologie de recours aux soins	269
4.5	Taille et organisation des villes, d'après <i>Bettencourt</i>	273
4.6	Schéma de principe pour la construction d'un paysage psychologique	297

4.7	Concepts de CDC et de SPOT RISC en pratique clinique de routine	317
4.8	Architecture générale de SPOT RISC	322
4.9	Architecture et départements en charge de SPOT RISC . . .	323
4.10	Principe d'alimentation en données de SPOT RISC	324
6.1	Exemples d'herbes sauvages générées à partir d'un L-système.	389
6.2	Un exemple d'espace des phases d'un réseau booléen (Wikipedia, licence CC)	390

Liste des tableaux

3.1	Distribution de la probabilité jointe de S_{ex}	170
3.2	Tableau comparatif des trois grands modèles théoriques pouvant servir de base pour une théorie de la forme.	187
4.1	Critères de Bradford-Hill pour la causalité.	219
4.2	Critères cliniques (stades) du SIDA, selon l’OMS.	223
4.3	Typologie du suicide, d’après Durkheim (<i>Le suicide</i>).	262
4.4	Les variables de recours aux soins utilisées pour la constitution d’une typologie.	267
4.5	Prévalences de la polypathologie chronique en fonction des définitions.	285
4.6	Méthode étape par étape de construction d’un paysage psychologique.	294
4.7	Tableau comparatif des deux descriptions : actuelle selon le DSM, et selon une hypothèse type paysage psychologique . . .	299
4.8	Critères diagnostiques pour les troubles bipolaires, selon le DSM 5 : 1 des 2 critères principaux (manie ou dépression) et 3 symptômes dits "mixtes".	308
5.1	Distribution de la probabilité jointe de S_{ex}	356

Remerciements

Une page sans doute ne suffirait pas à remercier ou à rendre compte et des personnes ayant compté de loin comme de près dans les travaux ici présentés, et à signaler l'ensemble des influences significatives qui ont pu peser sur la réflexion d'ensemble que nous livrons dans les pages qui suivent. Que soient remerciées toutes les personnes en droit de se sentir concernées : nul besoin ici de lister et de hiérarchiser, inévitablement.

Des remerciements particuliers à :

Christian Roux – l'aventure a commencé il y a maintenant 10 ans, elle ne s'est sans doute pas déroulée exactement comme chacun l'aurait entendu ; cependant, il existe des résultats. Merci à sa patience, son aide, son accueil, jusqu'au bout ;

Éric Stindel – pour sa confiance renouvelée, le goût du projet et de la ténacité, son aide et son soutien ;

Séverine Ansart – en cours de route, pour ce nouvel épisode, fatigant, mais prometteur, de la vie du LaTIM ;

Micheline Rouillard – pour son aide logistique longue portée, depuis plusieurs années.

Merci également à *Pierre Chauvin* et *Isabelle Parizot*, pour l'apport non négligeable à l'exposé de ces travaux que constituent les études que nous menons ensemble en épidémiologie sociale. De même, merci à *Rémi Gagnayre*, pour nos travaux portant sur la polypathologie chronique. Merci enfin à *Patrick Chariot* et *Aude Lepresle* pour nos travaux, qui se poursuivent, en santé mentale et médecine légale.

Et, bien évidemment, mes remerciements aux membres du jury.

À tous ceux, toutes celles qui nous empêchent, consciemment ou inconsciemment, de chercher à y voir clair.

À Jeanne.

À Yves Bizais.

Résumé

La méthode expérimentale et le réductionnisme ont été autant de révolutions dans la façon de penser la science, la production de la preuve et la formalisation de nombreuses disciplines. Au-delà de la sphère strictement scientifique, ils ont également laissé une empreinte profonde dans notre vie quotidienne, notamment avec une façon de penser le travail et l'économie qui en ont dérivé. Néanmoins, nombreux aussi sont ceux qui ont pu critiquer tant l'approche conceptuelle et que son application à des domaines connexes de la science, et qui ont pu notamment prôner une science de la complexité – à tout le moins, qui soit plus respectueuse des réalités. Plus particulièrement, le domaine biomédical, par ses implications larges et ses difficultés d'étude qui lui sont propres, a vu petit à petit se réduire à une poignée de méthodes statistiques fondamentales, et substituer à une pluralité de raisonnements possibles un raisonnement essentiellement corrélational : on y établit des rapprochements, alors que l'on y poursuit une causalité simple et linéaire en général. Nombreux encore, sont les exemples pour lesquels il est possible de montrer qu'une approche réductionniste, à dominante génétique et déterministe, nous mène dans des impasses tant conceptuelles que pratiques. Au travers de ces travaux, nous abordons notamment 3 grands exemples, que sont la conception du normal et du pathologique, en particulier en médecine infectieuse, la sociologie et la santé mentale. Nous expliquons comment une alternative est possible, via une approche que nous nommons approche par la forme (en réalité : tendre vers une forme fonctionnelle). Une telle approche peut être qualifiée d'approche morphodynamique, dans le sens où l'on s'intéresse autant aux formes qu'aux principes régissant leur évolution. Si l'on cherche une alternative aux analyses réductionnistes et linéairement causales, quels seraient alors les outils venant les étayer ? Après avoir recensé les 3 grands systèmes théoriques candidats à la fondation d'une morphodynamique – les théories quantiques, relativistes et des systèmes dynamiques, nous expliquerons en quoi celles-ci nous éclairent et peuvent se rejoindre pour définir la notion de forme et l'expliquer. Par ailleurs, puisque la théorie ne saurait seule faire face à l'arsenal technique du réductionnisme, nous présenterons les techniques associées à chacune de ces théories. Plus encore, nous

exposerons en pratique, les principes d'un système de collecte de données et de traitement de l'information qui est en cours de développement. Très rapidement, nous verrons qu'il manque un élément essentiel aux éléments fournis par les 3 cadres théoriques, de même qu'ils fournissent un même frein : où trouver la causalité, et que faire du temps. Nous terminerons en proposant un cadre théorique pour la morphodynamique, bâti sur la géométrie et l'information, ainsi qu'un algorithme de traitement de données correspondant. Le parti pris sera celui d'une géométrie informationnelle, relativiste, dont les principes d'évolution seront basés d'une part sur les réseaux de causalité, d'autre part sur la notion d'Hamiltonien multi-temps.

Abstract

Science and more especially biomedicine, is still impaired by approaches and methods that have previously and in certain settings proved useful, but also showed since then clear limitations. Reductionism and the classical experimental method are good tools within their proper domain of validity, as any other tools. This domain of validity is the domain of linearity, and proportionality, particularly when it comes to the concept of causality : A causes B, and most of the time, in a relationship of proportionality. In this work, we show that reductionism and classical experiment methods, as inherited from the philosophical and practical views of Descartes and Bernard, led in many cases to dead-ends. 3 examples are chosen and studied : the difference between normal and pathological in infectious diseases, the existence of groups in sociology, and the identification of nosologic entities in mental health. We explain that another way of thinking is possible, thanks an analysis in terms of forms and functions – which we name a morphodynamical approach. In such an approach, we treat in equal respects both the form and the principles that determine this form. Invoking a morphodynamical approach for solving complex problems calls for theoretical and technical developments, so that it does not remain a pure wish. We therefore suggest to review three main theoretical framework that could provide formal elements to build our theory of morphodynamics – namely, the quantum-like frameworks, the relativistic-like and the system dynamics-like ones. We also present the practical techniques that support each type of framework. We end this work by presenting the detailed blue prints of a unified theoretical framework for morphodynamics, based on geometry and information. This framework relies on the use of an informational automaton such as a bayesian network, and the co-construction of what we term its associated informational space. This space is structured by a metric, built from the form of the automaton. From a relativistic point of view, we postulate that the informational space is shaped by its local informational content, as well as the local information flow. The adaptation of an existing algorithm is presented, that can determine the geodesic paths of the informational space.

Avant-propos

Ou : pourquoi il ne sera pas – ou anecdotiquement – question de génétique dans ces travaux sur la morphodynamique.

Le lecteur pourrait s'attendre à ce que l'on aborde plus directement ou massivement, dans les pages qui suivent, le point de vue génétique, notamment en tant que source privilégiée et fondamentale d'information structurante des organismes. Sur ce point, le lecteur sera déçu : il ne sera pas traité de génétique dans les pages qui suivront, pour diverses raisons, non exhaustives. En voici quelques unes : 1) dans un monde où l'on cherche à réduire la complexité superflue ou à tout le moins, à isoler celle irréductible à une compréhension féconde des phénomènes, dénombrer l'ensemble des combinaisons génétiques en aveugle et leur supposer une importance de premier plan dans la structuration et le fonctionnement des organismes paraît aller à l'encontre d'une entreprise simplificatrice, puisque les liens directs de type un gène – une expression phénotypique unique sont clairement l'exception plutôt que la règle ; 2) une grande partie des théories morphogénétiques s'intéressent déjà aux profils d'expression génétique, et il ne nous a pas paru nécessairement pertinent d'en augmenter le nombre par notre contribution (voir par exemple la quantité d'algorithmes de clustering appliqués à la génétique) ; 3) la génétique (et les sciences en « omics ») est un domaine fortement doté dans la recherche actuelle, tant en moyens humains qu'en moyens financiers, toutes proportions gardées – là aussi, nous avons préféré nous engager dans une

voie sensiblement différente ; 4) la génétique est la recherche d'aujourd'hui, fournira peut-être des résultats déterminants demain : à l'heure actuelle, la génétique faillit quasi systématiquement à expliquer de manière convaincante les observations (ce qui est particulièrement vrai en psychiatrie, ou en neurosciences, mais également comme échec à expliquer la majorité des pathologies, notamment les plus banales et courantes) ; 5) si la génétique est un domaine pertinent – c'est-à-dire, si le gène est effectivement l'objet princeps à étudier en science du vivant – alors elle fournira des informations comme d'autres domaines en fournissent, et ces informations pourront être intégrées dans les méthodes que nous aborderons dans ce travail : il est a priori inutile de se focaliser sur l'échelle génétique afin de mener à bien notre raisonnement, ce d'autant que l'attention générale y est déjà fortement attirée.

Si cet avant-propos peut paraître lapidaire, le lecteur pourra noter qu'au travers de ce travail, les arguments cités ci-dessus seront étayés et référencés au cours de l'exposé – il ne s'agit évidemment pas d'une critique gratuite, mais bien plutôt d'un constat sans cesse répété.

Enfin, nous aimerions que le lecteur garde à l'esprit que les travaux exposés à la suite forment un compromis entre le désir et la nécessité de proposer une approche théorique, formelle qui soit bien assise et valide, et le besoin d'exposer les tenants et aboutissants de notre problématique tant d'un point de vue épistémologique, que d'un point de vue éminemment plus pratique – notamment en sciences biomédicales. Notre intention a été également de chercher à montrer combien le même type de question pouvait se retrouver au centre d'approches *a priori* bien différentes. Nous espérons que chacun pourra y trouver son compte.

Chapitre 1

Une introduction

Sommaire

1.1	Des problèmes transverses : trois exemples in-	
	attendus, une question centrale	6
1.1.1	Médecine, biomédecine et anatomie : <i>Naissance</i>	
	<i>de la clinique</i> , et géométrie des infections	7
1.1.2	Les formes de la société et du fait social : fonda-	
	tions de la sociologie, la <i>distinction</i> et le <i>suicide</i>	12
1.1.3	Entre les deux : psychologie, psychiatrie et santé	
	mentale ou la crise du modèle et de la classification	16
1.1.4	Formulation du problème général	20
1.1.5	Plan général des travaux	23

La morphodynamique en théories et en pratiques, au pluriel, sans erreur de notre part. Pourquoi, qu'est-ce que la morphodynamique, en quoi consiste une morphodynamique ? Sans doute est-on plus familier du terme morphogénèse – littéralement genèse ou naissance d'une forme. La morphodynamique serait alors en toute simplicité une dynamique de la forme. La dynamique recouvre les notions de force, de mouvement et d'énergie. Elle ne se contente pas donc de qualifier le seul mouvement, auquel cas nous aurions pu parler davantage de morphocinétique.

On parle souvent de morphogenèse dans le sens de l'apparition et de l'évolution graduelle d'une forme vers une forme destinée, finale, comme il en est question par exemple en embryogenèse, ou en organogenèse. Il existe également un certain nombre de théories dites morphogénétiques, qui nous ont semblé à la fois très attachées à cette notion de destination, mais également à la notion de motifs spatialisés, en deux ou trois dimensions, le temps faisant office de flèche selon laquelle on déroule les clichés à différents instants de la forme d'intérêt. Aussi, nous avons préféré nous détacher partiellement des implicites de la morphogenèse, et lui préférer le terme de morphodynamique, en toute généralité : un principe morphodynamique est un principe régissant l'évolution d'une forme, quelle qu'elle soit, et sans considération exclusive du temps comme facteur d'évolution. Nous y reviendrons.

Il n'est absolument pas neuf de s'intéresser aux problèmes de formes, et l'histoire est déjà riche d'études les concernant, au premier lieu desquelles figure la philosophie. [1] La forme rejoint les débats sur l'esthétisme, la beauté, la nécessité mais aussi la causalité. Les approches par la forme ont également une tendance à lorgner du côté des approches holistiques, globalisantes – le risque principal de théories de la forme devient alors d'être une théorie tellement totalisante qu'elle en devient tautologique, et perd toute portée scientifique d'explication du monde.

A l'opposé, l'histoire, par le biais des sciences, d'un type de philosophie là encore, puis de l'industrie et des technologies, semble avoir d'abord donné raison et privilégié les approches atomistes et réductionnistes. Il a paru plus simple, mais également plus efficace – ce qu'on ne peut aisément dénier dans nombre de domaines – de diviser les problèmes comme les tâches en sous-unités élémentaires, toutes étudiables séparément et individuellement. Le paradigme sous-jacent étant de parier sur une synthèse linéaire du problème : on recolle en juxtaposant les cubes de problèmes les uns aux autres, à l'instar

des pixels d'un écran, afin de déterminer la solution, l'image totale, du problème initial et complexe. C'est là tout le principe attribué à Descartes,[2] et la Méthode analytique, méthode à visée universelle, à l'opposé des idées d'un Pascal, pour qui il n'existe aucune approche universelle, seulement autant de solutions adaptées à des problèmes particuliers. Cette méthode initialement métaphysique puis scientifique, fut traduite dans l'organisation du travail, dont le Fordisme est le représentant le plus pur. Jusque les sciences les plus fondamentales et expérimentales, en particulier la physique des particules ou des hautes énergies, où en marge de la théorie de la relativité générale, l'on considère que le monde matériel est composé de divers grains qui se combinent.

Oui mais. Même jusque dans ce sanctuaire hautement atomiste, il convient d'y regarder de plus près, dès la formalisation au début du XXe de la mécanique quantique. Certes, les problèmes sont décomposés sur des grains de réalité, mais l'assemblage de ces grains n'est plus aussi simple qu'il y paraissait au départ. La contrepartie des théories réductionnistes, sa traduction mathématique, est le postulat de linéarité. Par linéarité, il faut entendre additivité des effets, ou encore principe de superposition : si mon problème se décompose en deux parties A et B, et que je connais les effets de A et de B (disons respectivement C et D), alors je connais les effets de A + B, ce sera C + D. Ce postulat ne tient plus pour nombre de phénomènes, même et surtout physiques ou biologiques, où l'on a pris soin pendant le plus longtemps possible, d'étudier les systèmes dans leur plage de linéarité. Or, il est devenu incontournable d'essayer de comprendre tout le reste de l'iceberg, à savoir, comprendre les problèmes non linéaires.

Ce mouvement semble relativement acquis, et gagne du terrain en sciences plus ou moins dures. Il existe cependant un autre pan du problème, que l'on peut apparenter au problème des approches dites quantitatives *versus* quali-

tatives. Un des exemples les plus criants est donné par le domaine biomédical, au sens large. Prenons le cas d'école que nous étudierons plus loin : les essais randomisés contrôlés, en recherche pharmaceutique. On y juge de l'efficacité d'une molécule sur un critère ponctuel et principal (par exemple : guérison de la personne traitée – peut-on cependant définir de façon non ambiguë la guérison ?) en comparant deux groupes constitués de telle façon qu'ils soient semblables en tout, sauf selon qu'ils aient reçu ou non la molécule. Ces essais ont longtemps été, et demeurent la référence dans le domaine, et y sont la traduction de la méthode expérimentale. Pourtant, ils accusent des limites indépassables, à commencer par leur caractère peu réaliste. Des techniques, qui tiennent presque de l'astuce plus que de la vraie stratégie – mais tous les coups sont permis – sont développées afin d'atténuer les limites de ces approches quantitatives (on définit la guérison sur des critères mesurés, chiffrables) et réductionnistes, mais elles n'apportent aucune réponse vraie et viable.

Ce ne sont que des exemples rapides et somme toute assez génériques que nous délivrons ici : nous en aborderons davantage et de façon plus précise dans cette thèse, dans le but avoué de montrer les limites des approches classiques et reposant sur un même fond théorique, puis d'y proposer une approche de substitution, une approche qui définira de proche en proche ce que nous appellerons une approche par la forme, une approche morphodynamique. Dans ce chapitre introductif, nous aborderons trois exemples distincts et historiques, concernant des domaines relativement séparés : la biomédecine, la sociologie et la psychologie (au sens large : aussi bien psychologie que psychiatrie ou psychanalyse : bref, le versant complémentaire du somatique). Puisque nous ne retracerons pas toute l'histoire de la morphogenèse ou les origines de la morphodynamique, nous nous contenterons d'ancrer notre propos sur quelques événements historiquement pertinents et

étonnamment proches. Enfin, nous proposerons une première formulation du problème général auquel nous nous attaquons par ce travail, dont l'une d'entre elles pourraient se présenter sous la forme de la question suivante : existe-t-il, ou peut-on construire, un cadre général liant l'observateur, l'information, et la morphodynamique ?

In fine, ce travail abordera donc l'intérêt d'une approche différente des approches classiques dans plusieurs domaines scientifiques, par l'illustration des limites de ces approches, puis présentera plus techniquement tant les alternatives théoriques que pratiques sur lesquelles baser une approche différente, et enfin, aboutira à la proposition d'une théorie morphodynamique mère, dont il sera possible de dériver différentes versions à éprouver.

1.1 Des problèmes transverses : trois exemples inattendus, une question centrale

Nous présentons à la suite trois exemples qui seront abordés plus en détails dans un chapitre ultérieur. Nous tenons cependant à les aborder dès ce chapitre introductif, afin de familiariser le lecteur avec leurs spécificités mais également leurs points communs. Car c'est bien cela que nous désirons souligner : l'universalité d'un problème qui se rencontre dans des domaines *a priori* distincts et d'importance équivalente.

Nous aborderons donc à la suite 1) l'archétype du modèle biomédical : la pathologie infectieuse, 2) le problème de la définition de la sociologie en tant que science à travers son objet, le fait social ou le groupe social, 3) une science qui cherche encore son nom et ses frontières, la psychologie.

1.1.1 Médecine, biomédecine et anatomie : *Naissance de la clinique, et géométrie des infections*

La médecine a connu une de ses révolutions lorsque l'on a cherché à mettre en correspondance causes et effets, et traiter la maladie en traitant ses causes. Cette révolution a pu s'accomplir en s'appuyant sur au moins deux éléments : les dissections et autopsies, et la méthode expérimentale – les autres avancées majeures ayant été permises par le développement d'outils plus performants. L'observation y joue donc un rôle majeur, quel que soit le médium permettant l'observation. Le corps et ses pathologies ont alors été de nouveaux territoires à cartographier, à décrire et classer. Tout géographe connaît sa géométrie, mais le clinicien connaît-il la sienne, ou méconnaît-il finalement l'origine de sa science ?

Tradition anatomo-clinique et anatomo-pathologique – les atlas, et les moyens d'observations (le microscope)

Il existe ainsi ce que l'on nomme parfois la tradition anatomo-clinique, dont dérivent finalement les spécialités que sont celles de l'anatomiste ou de l'anatomo-pathologiste. Par observation systématique et comparée des lieux du corps, ils repèrent les variations, les signes qu'il mettra en lien avec une condition désignée comme morbide. La normalité, un bon état de santé sera à cet égard finalement plutôt défini par défaut comme l'absence de signes morbides, plutôt que comme observation *stricto sensu*.

Ainsi, des planches entières d'anatomie sont créées, les corps sont ouverts, disséqués, les structures anatomiques sont décrites en tant qu'objets, mais également situées selon certains rapports, dits rapports anatomiques : l'on définit la place d'un organe relativement aux autres, ce qui, en une logique circulaire et saturante, détermine le corps dans son intégralité.

Il existe ainsi bel et bien des atlas des contrées des corps, de mieux en mieux précisées, confirmées, reproduites, comme il a pu exister une histoire des cartes, notamment maritimes. L'anatomiste est avant tout un géomètre, un cartographe. Il se transforme géologue lorsqu'il tente d'élucider ou de traduire la chronologie des états menant du corps sans altérité significative vers le corps pathologique. Il s'est avant tout servi de ses yeux, mais il a bientôt pu accéder à cet outil emblématique qu'est le microscope : toute une gamme d'échelles lui est alors ouverte, inaccessible jusque lors. Puis, petit à petit, les moyens d'exploration et de visualisation se sont multipliés, notamment grâce à l'avènement plus récent de l'imagerie, type radiographie classique, tomodynamométrie, IRM, ou encore imagerie nucléaire. Autant de moyens de cartographier que de principes d'imagerie. Peu à peu, l'on glisse insensiblement, de la forme vers la fonction, sans qu'il soit toujours simple d'élucider leurs rapports : la physiologie se rêverait comme la science de l'articulation entre la forme et la fonction.

Un premier anniversaire : *Naissance de la clinique*, et la question de la géométrie

Pour autant, en dehors des ingénieurs et techniciens présidant à la conception et à la fabrication de ces procédés de visualisation, il apparaît que l'anatomiste, *a fortiori* l'anatomopathologiste, ne se pose comme problème fondamental de sa science que la question de la géométrie. L'ingénieur se doit de se la poser, mais en des termes décalés des préoccupations du clinicien : chaque outil possède sa plage de fonctionnement, et corollairement ses aberrations et distorsions – ce fut vrai et le demeure dès la mise au point des lentilles et des microscopes optiques. Mais cette question de la géométrie se pose au concepteur avant tout comme la quête de la plus grande ressemblance à ce que verraient l'œil et le cerveau du clinicien, et non ce qui structure *per se* l'entité anatomique ou pathologique. D'une certaine façon, l'imagerie dite

fonctionnelle remet partiellement en question cette conception : la visualisation est plus directement le fruit d'une activité de l'organisme imagé que lorsque l'on utilise l'imagerie morphologique. La notion d'organe, qui réconcilie de manière imprécise forme et fonction, sert alors d'intermédiaire : s'il existe une activité basée sur le métabolisme du glucose, par exemple, en un lieu du corps, alors cette activité a toutes les chances de refléter une réalité anatomique, voire organique.

Il est remarquable que ce n'est pas un clinicien qui s'est posé en des termes directs et explicites la question de la géométrie en médecine. Ce fut un sociologue ou philosophe, voire historien, au choix – ce fut Michel Foucault. Sans détour, Foucault aborde dans son ouvrage « Naissance de la Clinique » (1964) [3] cette question de la géométrie :

« Pour nos yeux déjà usés, le corps humain constitue, par droit de nature, l'espace d'origine et de répartition de la maladie : espace dont les lignes, les volumes, les surfaces et les chemins sont fixés, selon une géographie maintenant familière, par l'atlas anatomique. Cet ordre du corps solide et visible n'est cependant qu'une des manières pour la médecine de spatialiser la maladie. Ni la première sans doute, ni la plus fondamentale. Il y a eu et il y aura des distributions du mal qui sont autres.

Quand pourra-t-on définir les structures que suivent, dans le volume secret du corps, les réactions allergiques ? A-t-on même jamais fait la géométrie spécifique d'une diffusion de virus dans la mince lame d'un segment tissulaire ? Est-ce dans une anatomie euclidienne que ces phénomènes peuvent trouver la loi de leur spatialisation ? Et il suffirait de rappeler, après tout, que la vieille théorie des sympathies parlait d'un vocabulaire de correspondances, de voisinages, d'homologies : termes pour lesquels l'espace perçu de l'anatomie n'offre guère de lexique cohérent. Chaque grande pensée dans le domaine de

la pathologie prescrit à la maladie une configuration dont les requisits spatiaux ne sont pas forcément ceux de la géométrie classique. » (Naissance de la clinique, Chapitre Premier – Espaces et classes).

Par ces quelques phrases, en des termes étonnement bien choisis et très précis, Foucault délimite le problème central qui occupera cette thèse : quelle est donc la géométrie des infections, de la pathologie ? Par extension, quelle est la forme et la dynamique des corps vivants sains et malades ?

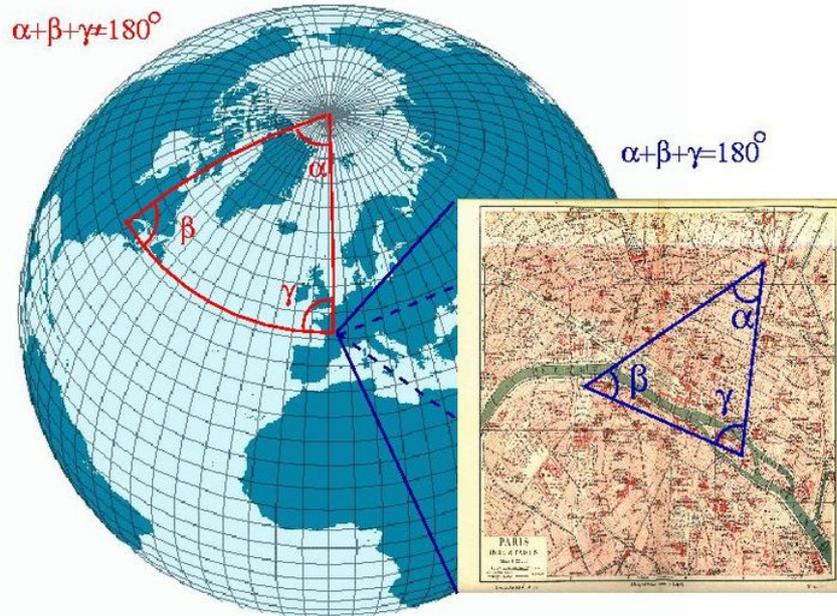
Les premiers éléments d'un champ sémantique très ciblé et technique sont jetés : anatomie euclidienne, voisinages et homologies, géométrie, espace et classes. Ces mots nous accompagneront tout au long du travail, et nous tenterons d'apporter le maximum d'éléments de réponse à la question de Foucault, tout en en démontrant la validité hors de la pathologie infectieuse. Dans quel espace penser la pathologie, et selon quelle structure ?

Vers un deuxième anniversaire : la relativité générale et la consécration de la géométrie non euclidienne

En 1915, Einstein publie sa théorie de la relativité générale, 10 ans après sa théorie de la relativité restreinte. [4, 5] Elle sera l'illustration physique de la pertinence non exclusivement mathématique d'une géométrie qui ne soit pas euclidienne, justement. La géométrie d'Euclide [6] reposait initialement sur cinq axiomes, c'est-à-dire cinq présupposés, structure intrinsèque et irréductible dont on peut dériver tous les résultats – les théorèmes – de la géométrie.

Le cinquième postulat a intrigué nombre de mathématiciens, dont Gauss au XIXe. Ce postulat dit que dans un plan, étant donnés une droite et un point distinct de cette droite, il existe une et une seule droite parallèle à cette droite. Il est de fait possible de s'affranchir de ce cinquième postulat, et tout

FIGURE 1.1 – Géométries euclidienne et non euclidienne



de même continuer de définir d'autres géométries qui soient cohérentes. On peut ainsi définir une géométrie dite hyperbolique, où il peut exister une infinité de droites distinctes passant par un point et qui soient parallèles à une autre droite du plan, ou encore une géométrie elliptique, autrement appelée géométrie Riemannienne – il n'y existe aucune droite parallèle. [7]

Cinquante ans après Einstein, Foucault pose la double question de la géométrie définissant la pathologie, et du type de géométrie, qu'il soupçonne non euclidienne. Cinquante ans plus tard encore, soit un siècle, a-t-on, voire peut-on répondre à cette question ? Et surtout, en quoi la fondation d'une autre géométrie a-t-elle pu être nécessaire, que peut-elle apporter que la géométrie conventionnelle ne permet pas ? Il est remarquable que le début du XXe a vu la physique se géométriser, à défaut de pouvoir totalement et directement

s'axiomatiser ainsi que le désirait Hilbert. Les contributions de Minkowski, Lorentz, Poincaré et Einstein furent déterminantes, et ouvrirent la voie à différentes formulations et notions : caractérisation des géométries par les groupes de transformation – donc, par les opérations définissant l'évolution des formes – des invariants de ces géométries, et naissance de la géométrie dynamique, de la théorie du chaos, puis de la stabilité structurelle et la théorie des catastrophes.

1.1.2 Les formes de la société et du fait social : fondations de la sociologie, la *distinction* et le *suicide*

La sociologie est une science jeune. Elle a d'ailleurs débuté son entrée sur la scène scientifique par une crise d'autodétermination : la sociologie, pour s'autoriser à se définir comme science, proposa de se définir par sa méthode, et au passage, détermina son objet – le fait social.[8, 9] Par son domaine d'étude et par ses précautions préliminaires, la sociologie offre de se pencher sur une question qui pourrait sembler lui appartenir, mais qui en réalité s'avère transdisciplinaire : existe-t-il des objets, et si oui, quels sont leurs rapports à la notion d'échelle (d'observation ou de définition)? Plus récemment, les sociologues des sciences nous apprennent que les sciences les plus « dures » demeurent une activité humaine, donc sociale, ce qui n'est pas sans conséquence sur leurs propres objets d'étude – la boucle est bouclée.

La science et l'objet : fondation d'une nouvelle science à partir de sa méthodologie et de son objet

Il est bien entendu possible là aussi, de remonter plus loin le cours de l'histoire et de trouver des penseurs qui seraient précurseurs de ce que l'on appelle désormais la sociologie. Assez classiquement, cependant, l'on a coutume de situer sa naissance en tant que science autonome, aux travaux fin XIXe de Durkheim et Weber, de Marx. [10] Il est à ce stade amusant de constater que

peu auparavant, Comte tentait de définir ce qui sera la sociologie, mais lui prêtait un autre nom : la physique sociale. [11] Le premier geste de Durkheim est de proposer « les règles de la méthode sociologique » ; avec Descartes et Claude Bernard, [12] pas de science sans méthode.

Une particularité tient dans le fait que la sociologie a voulu se démarquer de la psychologie, en faisant s'opposer deux échelles de définition des objets de chacune des science : la sociologie s'occupera d'objets intrinsèquement construits sur un groupe, la psychologie se jouera au niveau de la personne. Les sociologues vont plus loin – du moins Durkheim, en contradiction avec Gabriel Tarde, [13] son rival de l'époque – et décideront que la sociologie ne doit s'occuper que de quantitatif, et d'objectif. Toute mesure qualitative et renvoyant à de la subjectivité se voit exclue par définition. Il faut donc aux sociologues leur lunette comme aux astrologues, leur microscope comme aux microbiologistes. C'est à ce prix et selon ces moyens que le sociologue déterminera son objet : le fait social, sorte d'invariant et dénué de toute subjectivité.

Objets d'échelle et lois d'échelle. Problème de la dynamique et des temps caractéristiques

Si sociologie et psychologie semblent s'antagoniser mutuellement pour se fonder individuellement, il existe au sein même de la sociologie des conceptions opposées quant à la substance fondamentale de l'objet d'étude.

Durkheim considère que la société est un tout supérieur à la somme de ses parties, autrement dit, de la somme des personnes faisant société ; c'est un point de vue plutôt holistique, ou peut-être plus précisément, un point de vue se réclamant des phénomènes dits émergents : il existe des objets à une échelle d'observation supérieure, dont les propriétés ne peuvent s'expliquer

simplement, ou ne sont pas programmés explicitement à l'échelle inférieure, c'est-à-dire au niveau de ses constituants, de ses atomes.

Weber, *a contrario*, propose un point de vue atomiste.[10] Ce sont les interactions atomiques qui produisent de la société, l'activité des personnes qui est société. Dans la continuité de Durkheim, Maurice Halbwachs parlera de morphologie sociale. [14]

Quoi qu'il en soit, la sociologie, *via* sa méthode et la recherche de ses objets, a produit assez rapidement des typologies – des classes, donc – ou plus récemment a tenté de caractériser quantitativement des espaces de définition. Nous y reviendrons : notons ici simplement les travaux de Durkheim, encore, sur le suicide, [10] et de Bourdieu, [15] sur l'espace social. Si dans chaque cas, des groupes peuvent être identifiés, des espaces définis, il se pose toujours la double question suivante : comment situe-t-on une personne dans un groupe sociologique, et le groupe détermine-t-il la personne ou inversement, et ces groupes, ces espaces sont-ils des structures fixes, des invariants de la société, ou bien sont-ils eux aussi des formes dynamiques, mouvantes ? Autrement dit, quels sont les rapports d'échelle entre classes et éléments, et le support est-il un support absolu, fixe, ou bien en mouvement ?

L'objet scientifique, le fait social et le *peer-reviewing* : *Nous n'avons jamais été modernes*

S'il semble que les sciences non "dures" aient un important complexe d'infériorité, ou des interrogations fortes quant à leur droit à s'ériger en science, il est remarquable de constater à quel point les dites sciences dures – mathématiques, physiques essentiellement – ne se sont elles que peu interrogées sur leur légitimité. Or, la science est une activité humaine, et même éminemment sociale, voire politique. En outre, la science permettant précisément, au-delà de la religion ou de la croyance, de fournir des moyens d'action efficace sur

le monde, elle peut rapidement prendre une dimension éthique, puisqu'elle a vocation à transformer notre milieu, notre société et notre être.

Les sciences dures se sont rapidement imposé le devoir d'objectivité, jusqu'à oublier qu'en dehors de l'expérimentation, il n'y a aucun juge absolu de la validité d'une connaissance, d'un savoir.

Nous verrons que les mathématiciens et les logiciens se sont interrogés eux-mêmes et à leur façon, sur la validité externe et interne du savoir qu'ils produisent – ce sera la contribution de Gödel, par exemple.

Néanmoins, l'histoire retient des exemples célèbres qui renseignent sur la production de la norme scientifique, et sur la validation d'un fait objectif. L'un des plus célèbres et fréquemment cité remonte à Mesmer et la théorie de la magnétisation, lorsque la commission académique fut dépêchée pour se prononcer sur l'efficacité de la thérapie inspirée par la théorie de Mesmer, et évacuer la part d'effet induit ou placebo de l'expérience. [16] Cette commission a également, de par le pouvoir et politique et scientifique – par son autorité scientifique et l'institution qu'elle représentait – statué sur ce qui était scientifique, et ce qui ne l'était pas. L'autre exemple célèbre souvent cité est celui de la pompe à vide. Ici, pas d'autorité désignée par les pouvoirs, mais la définition à nouveau du fait scientifique, et la base, finalement, de ce qui est actuellement encore en place pour juger d'un travail scientifique : le jugement des pairs (ou, dans le cas de la production d'articles scientifiques, le *peer-reviewing*, la revue par les pairs). A la suite de Shapin et Schaffer, [17] Bruno Latour a montré [18] comment nous n'avons jamais su séparer, même au plus fort du dogme objectiviste de la science dite moderne, les objets du contexte dans lequel ils sont élaborés, du milieu social où ils naissent. Le fait scientifique est également production de groupes, production sociale. En outre, nous manipulons sans cesse des objets hybrides, et non jamais des

objets « purs », idéaux et totalement découpés du monde qui les entoure. La production du fait scientifique serait ainsi un travail sans relâche d'hybridation et de purification d'une trame sans cesse remaniée, dont une des dimensions est éminemment sociale.

1.1.3 Entre les deux : psychologie, psychiatrie et santé mentale ou la crise du modèle et de la classification

Le fait, ou l'objet, scientifique est donc en partie immergé dans la trame sociale où il naît. C'est déjà en soi une forme complexe, dont la dynamique ne paraît pas simple. Avant même peut-être que la dimension sociale des objets aient été bien mise en lumière, il a souvent été posé la question de l'observateur, notamment dans les schémas expérimentaux, le dogme premier voulant que l'expérimentateur s'efface au maximum de l'expérience. Néanmoins, il n'y a pas d'expérience sans mesure – sans quoi l'expérience n'aurait aucun sens, puisque non évaluée – et donc pas d'expérience sans observation, ni sans observateur, même si celui-ci peut intervenir en aval de l'observation, temporellement parlant.

La place et l'influence de l'observateur ont été assez longuement questionnées au début du XXe, en physique (mécanique quantique), mais également en psychologie, plus spécifiquement en psychanalyse. Faut-il l'éliminer, le gommer, ou au contraire l'intégrer ? Que faire d'une subjectivité supplémentaire alors que l'on tente d'en caractériser une en particulier ?

Modélisation en psychanalyse : deuxième topique Freudienne, la tentative de Lacan (topologie et algèbre)

Autre grande aventure de la fin XIXe, début XXe, simultanée des aventures de la sociologie, du renouveau de la physique, l'aventure psychanalytique. A l'instar de la physique sociale de Comte, la psychologie se figurait pour ses

précurseurs comme une psychophysique (de définition sensiblement différente de l'acceptation actuelle, proche de la neuropsychiatrie), c'est-à-dire, comme une physique de la psyché. Cependant, Freud a œuvré fermement pour que sa science en devenir s'autonomise au maximum des autres sciences, dont la physique, et s'est engagé durablement dans un développement théorique de formulation beaucoup plus littérale que calculatoire. [19]

Il existe cependant des modèles en psychanalyse, qui devient finalement avant tout la science de l'individu en lui-même, et non de la relation intersubjective, dont l'un des plus célèbres est le modèle de la personnalité structuré en ça-moi-surmoi (ce que l'on appelle la deuxième topique). À strictement parler, il s'agit davantage d'un modèle hiérarchique que relationnel. En outre, notons que le mode d'intervention reposant sur les postulats psychanalytiques, met face à lui-même la personne, et l'observateur tend à vouloir d'effacer au maximum (pas de prise de parole, en dehors du champ de vision). De ce point de vue, cette première pratique respecte et partage les critères des autres prétendants au statut de science, en découpant au maximum l'objet d'étude, en l'isolant, en l'autonomisant, et en préconisant un observateur aussi peu perturbateur qu'il est possible.

Plus tard, Lacan s'est essayé à une formulation plus ancrée dans les sciences du calcul, *via* deux grandes voies : l'algèbre, et la topologie. [20] Remarquablement, la topologie qu'il propose repose sur des cas archétypiques de la topologie mathématique : on y trouve par exemple la bouteille de Klein, le ruban de Möbius ou le tore (voir figures 1.2 et 1.3). De fait, le choix de ces objets paraît plus symbolique que réellement opérationnel. Ces sont des objets géométriques remarquables : la bouteille de Klein est un exemple de surface fermée, sans bord et non orientable – il n'est pas possible de lui trouver un intérieur et un extérieur – et le ruban de Möbius est une surface ne possédant qu'une seule face, elle aussi non orientable. Ils ont l'intérêt de

FIGURE 1.2 – Une réalisation de la bouteille de Klein



supposer possible une structuration à la fois géométrique et dynamique, par le mode de construction de ces objets particuliers (découpages, collages des topologistes, et trajectoires induites par la géométrie locale).

L'approche syndromique en psychiatrie, la crise de l'objet et de la classification

La psychiatrie, et l'alternance que l'on connaît avec la neurologie d'abord, puis avec les neurosciences plus largement, s'est distinguée de la psychanalyse comme l'approche et la définition des symptômes, des signes pathologiques chez les personnes, afin d'en définir des groupes diagnostiques. Ces symptômes sont essentiellement des signes subjectifs : sentiments de tristesse, de colère, d'euphorie – ce qui a et continue de constituer une critique majeure du domaine par une partie des médecins et des chercheurs.

L'approche de la santé mentale – ou du moins, de la pathologie mentale – par la psychiatrie est donc initialement et toujours fortement une approche syndromique : un ensemble de signes qu'il semble que l'on trouve plus fréquemment associés ensemble, et définissant le support d'une entité

FIGURE 1.3 – Une réalisation du ruban de Möbius



nosologique séparée des autres (la dépression, la schizophrénie).

Ces classifications souffrent de limitations majeures, et dans leur construction, et dans leur validation, et enfin dans leur intérêt opérationnel – ici, le traitement. Pour exemple, si l'on se posait précédemment la question en sociologie du rapport éléments/classe, du caractère définitoire de l'élément par rapport à sa classe, une question similaire se pose en psychiatrie : quelle est la signification de l'appartenance d'une personne à une classe diagnostique plutôt qu'une autre ? L'interrogation peut paraître étrange, pourtant, elle s'illustre et se justifie de la manière suivante : on oppose en général troubles de l'humeur (troubles bipolaires, par exemple) et troubles psychotiques (schizophrénie par exemple), comme des entités fondamentalement et qualitativement séparées ; en pratique, l'on sait désormais que près de la moitié des personnes ayant reçu un diagnostic de trouble bipolaire, sera reclassée au cours de sa vie dans la catégorie troubles schizophréniques. Si l'on prend le point de vue opérationnel, le point de vue du traitement, que l'on conçoit comme l'intervention sur la ou les causes de la pathologie, l'on

sait également que les molécules alors réservées au traitement des troubles de l'humeur peuvent servir pour traiter les troubles psychotiques et inversement. Ainsi, l'objet, en psychiatrie, et les classifications visant ces objets, souffrent d'un véritable problème de définition.

Les tentatives intégratives, dynamiques et systémiques en psychologie

En dehors de la psychanalyse, de la psychiatrie, il existe nombre d'approches psychologiques, en particulier de ce qui est considéré comme pathologique. Il apparaît que la psychologie a été la plus ouverte conceptuellement, et a tenté d'intégrer et d'adapter un certain nombre de théories plus ou moins éloignées de son domaine initial. La psychologie a par exemple suggéré d'élargir le point de vue de la personne à son environnement immédiat, sa famille ou son entourage. D'étudier les schémas relationnels, le fonctionnement du groupe et de ses éléments en interaction. Il y a eu notamment les rapprochements avec la théorie des systèmes, la théorie de l'information et la cybernétique (théorie du *feed-back*, notamment) – l'école dite de Palo Alto y a fortement contribué. [21, 22, 23, 24]

1.1.4 Formulation du problème général

Que visons-nous au travers de ces travaux, quelle forme générale pour le problème qui traverse les différents domaines et exemples que nous nous apprêtons à traiter? Si nous partons du concept générique et mal défini de forme, nous affinerons notre conception, notamment par le choix d'outils conceptuels et techniques, principalement la géométrie et les approches apparentées aux théories des systèmes et de l'information. Plus largement, nous chercherons un cadre, s'il existe, qui puisse lier la notion d'observateur, d'information et de morphodynamique.

Le concept de forme

Ainsi que nous l'avons avancé, nous souhaitons baser notre réflexion sur le concept de forme. Ou plutôt, nous espérons pouvoir fournir, *in fine*, un cadre et une définition qui puisse rendre compte d'une partie de la multiplicité des acceptions du terme forme. Par notre but et par notre cheminement, nous chercherons ainsi à constituer une approche qui dépasse une approche purement réductionniste sans qu'elle soit totalement et radicalement holistique, ce qui la rendrait peu opérationnelle, et qui puisse aussi bien rendre compte des notions usuelles de formes, que d'aspects plus techniques et pratiques. Notamment, nous chercherons à établir une définition du concept de forme comme problème dual du concept de fonction.

La géométrie comme compromis entre synthèse informationnelle et principe dynamique : morphodynamique et géométrodynamique

À notre connaissance, un des cadres théoriques et techniques les plus à même de rendre compte d'une forme-fonction - nous dirons désormais forme fonctionnelle, et par abus de langage : forme - demeure la géométrie. En effet, la géométrie a continué de bénéficier d'une part de développements et de ponts importants avec les autres domaines des mathématiques (on parle de géométrie différentielle, de géométrie algébrique), a également imprégné de plus en plus les théories physiques (donc de la nature : c'est particulièrement le cas des théories relativistes) jusque se substituer au projet de leur axiomatisation, et d'autre part, la géométrie reste le cadre formel le plus naturel pour expliciter les formes "usuelles", à savoir les formes spatiales (un cube, un cercle...) plus ou moins complexes.

Par ailleurs, les principes géométriques ont su démontrer leur pertinence et leur portée en dehors de toute description purement "statique" : la géométrie peut parfaitement servir de cadre à une théorie dynamique, et ce de manière

autonome. C'est notamment le cas de ce que l'on peut appeler parfois la géométrodynamique. Un exemple célèbre de géométrodynamique est donné par la relativité générale, où la géométrie de l'espace-temps décrit elle-même le mouvement, sans autre donnée nécessaire. Plus important, cette géométrie est déterminée par un champ dit tensoriel. Ce champ représente en l'occurrence la distribution de matière : la présence ou l'absence de matière façonne la géométrie locale. Ainsi, la donnée d'une information - la matière - modifie la géométrie, qui modifie la dynamique possible.

Notre problème général peut donc prendre cette forme : étant donnée un ensemble d'informations (ou d'observations), peut-on définir une forme, et cette forme étant une forme fonctionnelle, peut-on en déduire l'impact, la dynamique sur l'évolution de cette forme ?

Prédire et expliquer : approche hybride ? Permettre la performance et l'intelligibilité

Nous aborderons à quelques reprises le problème dual de la définition et de la prédiction : nous venons de dire que nous chercherons dans le même mouvement à définir une forme, et d'une certaine façon, à en prédire l'évolution. Or, il est vraisemblablement illusoire de vouloir poursuivre simultanément ces deux buts, du moins, si l'on désire le même degré de performance pour les deux. Notre parti pris sera ainsi le suivant : nous visons une théorie basée sur des principes qui, pris indépendamment les uns des autres, sont relativement simples, et leur mise en relation, compréhensible. Ces principes seront formulés et plongés dans un cadre technique riche et relativement complexe, et dont les outils doivent nous permettre précisément d'atteindre des résultats eux aussi intelligibles, mais sans que le cheminement intermédiaire soit particulièrement accessible.

Autrement dit : il s'agit de s'appuyer sur des postulats compréhensibles

et questionnables, traduisibles dans un cadre mathématique, lequel doit permettre d'obtenir des calculs logiquement valides, et non triviaux.

Existe-t-il un cadre général liant l'observateur, l'information et morphodynamique ?

Ultimement, et comme proposition d'extension des approches usuelles, nous tendrons vers la définition d'un cadre qui puisse définir, produire des formes, c'est-à-dire des objets au sens scientifique du terme, lesquels doivent pouvoir être obtenus à l'identique par une même classe d'observateurs. Ces observateurs peuvent mesurer, enregistrer certaines informations, qui doivent servir de support à la définition des formes.

Le but concret est donc le suivant : étant donnés deux observateurs bénéficiant des mêmes capacités d'enregistrement, sont-ils capables de déterminer un objet complexe que l'on appelle une forme, ce de manière reproductible, communicable, et par extension, d'en prédire l'évolution ou les transformations sous l'influence de la variations de certaines informations mesurables ?

1.1.5 Plan général des travaux

Outre ce chapitre introductif, notre exposé se déroulera selon les temps suivants :

Chapitre 2 : Une brève histoire de la morphogenèse

Où il sera abordé essentiellement l'histoire récente des théories morphogénétiques, sur lesquelles nous nous appuieront pour les développements ultérieurs, tant pour les idées que pour le formalisme mathématique. Nous y croiserons les grands noms de la morphogenèse, puis présenterons les trois grands cadres formels bénéficiant d'assises conceptuelles fortes et pouvant servir à expliciter une théorie morphodynamique plus large : les systèmes dynamiques, les théories quantiques et les théories relativistes (géométriques).

Nous donnerons quelques détails sur les approches basées sur la cognition et les neurosciences.

Chapitre 3 : La modélisation des systèmes complexes en sciences et en biomédecine

Où comment, de manière générale, l'on aborde la complexité en sciences et plus spécifiquement en biomédecine : que ce soit par la modélisation, ou par l'analyse de données. On y retracera l'évolution depuis les approches expérimentales statistiques jusque les approches géométriques et dynamiques. Nous y discuterons la notion de complexité, de causalité et de temps en relation avec le concept de forme. Dans la continuité de la présentation des trois grands cadres conceptuels introduits dans le chapitre précédent, nous expliquerons plus précisément en quoi ces cadres fournissent des modèles conceptuels pour la morphodynamique, et les comparerons. Enfin, nous aborderons le premier volet pratique associé à ces trois cadres : de quels outils disposons-nous pour approcher les problèmes de forme en lien avec chacun des cadres ?

Chapitre 4 : Capturer, intégrer et traiter l'information : exemples et frontières actuelles des approches par la forme

Forts de ces premiers éléments conceptuels et pratiques, nous interrogerons comme annoncé en introduction, la place de la géométrie et de la forme dans différents domaines : nous tenterons un premier panorama de la pathologie infectieuse en tant que domaine géométrisable, puis aborderons un cas concret de modélisation d'infection *via* les infections ostéo-articulaires à l'aide des systèmes dynamiques ; nous aborderons ensuite plus concrètement la question des formes sociales, en prenant l'exemple du recours aux soins et de ses facteurs associés ; nous montrerons les limites de la définition actuelle de la pathologie, tant en approche technique qu'en dimension de son

espace de définition en prenant l'exemple de la pathologie chronique ; dernier domaine abordé, nous présenterons comment les insuffisances des modes de classifications en psychiatrie peuvent être traitées par une approche morphodynamique ; puisque le problème de la disponibilité des données et observations est un problème limitant et récurrent, nous expliquerons enfin comment ce que l'on appelle les *big data* fournissent une alternative viable au schéma expérimental classique et fournirons les éléments techniques d'un système de capture et traitement de l'information innovant.

Chapitre 5 : Cadre théorique pour une morphodynamique basée sur la géométrie et l'information

A ce stade, nous aurons tant présenté les cadres initiaux candidats à une théorie générale de la morphodynamique, exposé des cas variés et concrets démontrant les limites des approches classiques et suggérant l'intérêt d'une alternative morphodynamique, que proposé des éléments techniques et pratiques concrets jusque la spécification d'un système de traitement de l'information et sa mise en œuvre. Dans ce chapitre, nous synthétisons les éléments primordiaux que nous pensons nécessaires à la constitution d'un cadre théorique de la morphodynamique, qui soit basée sur l'information, son traitement, et la géométrodynamique – c'est-à-dire, une géométrie définissant la dynamique des éléments la constituant. Pour ce faire, nous introduirons la notion de géométrie informationnelle et de métrique associée à un automate. Nous proposerons dès après un algorithme de détermination de la géométrie informationnelle d'un ensemble d'observations.

Chapitre 6 : Conclusions et perspectives

Ce travail aura été l'occasion d'explorer historiquement, conceptuellement et en pratique la pertinence d'une approche morphodynamique de problèmes généraux . Un premier cadre théorique général aura été proposé, ainsi que

des implémentations techniques s'y rapportant. Pour autant, nous aurons laissé de côté certaines théories ou pratiques, qui auraient pu trouver leur place dans cet exposé ; il restera également des développements théoriques et des validations pratiques à envisager, pour lesquels nous fournissons des éléments d'orientation.

Chapitre 2

Une brève histoire des théories et modèles en morphogenèse

Sommaire

2.1	Les grands noms des théories morphogénétiques	31
2.1.1	Wentworth d'Arcy Thompson : <i>On growth and form</i>	31
2.1.2	Alan M. Turing : morphogénèse et chimie	37
2.1.3	Turing, Gödel et von Neumann : machine universelle, approche algorithmique et automates	46
2.1.4	René Thom et ses prédécesseurs : systèmes dynamiques et théorie des catastrophes	54
2.1.5	Place de la fractalité : Benoît Mandelbrot	57
2.1.6	Une dualité sans cesse rencontrée : discret et continu, objets <i>a priori</i> ou variations topologiques?	61
2.2	Éléments de théorie des systèmes dynamiques – théorie des catastrophes	63
2.2.1	Principes généraux et hypothèses	63
2.2.2	Formulations et définitions	65
2.2.3	Points fixes et notion de stabilité	66
2.2.4	Bifurcations	69
2.2.5	Théorie de Thom et ensemble catastrophe pour un champ gradient	71
2.2.6	Un exemple : modèle de Lorenz	73
2.2.7	Extension : théorie KAM, et formulation en espace Riemannien des ODE	74
2.3	Place des deux grandes théories physiques actuelles : théories quantiques et relativistes	80

2.3.1	Principes généraux en théories quantiques : états, opérateurs, espaces propres et opérateurs d'évolution	80
2.3.2	Principes généraux en théories relativistes : élément d'espace-temps, métrique et géométrie Riemannienne	92
2.3.3	Les invariants, les groupes de transformation et la géométrie	107
2.3.4	Place du temps dans les 2 théories	111
2.3.5	Evolution à structure fixe <i>vs</i> évolution à structure mouvante	112
2.4	Quelques mots à propos des approches sciences cognitives et des neurosciences	114
2.4.1	Représentations mentales et cerveau	115
2.4.2	Chimie moléculaire, circuits neuronaux et cognition	116
2.4.3	Formes, vision et cerveau : neurogéométrie de la vision	118
2.4.4	Objectivité et subjectivité, toujours	119
2.5	Ce que l'on en retient	119

Il serait possible de faire remonter l'histoire de la morphogenèse au moins jusque Aristote, [1] du moins dans des énoncés dont on n'aurait besoin de rien extrapoler ou interpréter pour leur prêter l'intention de traiter de la forme, et de sa dynamique. L'histoire de la morphogenèse présente ce paradoxe qu'elle paraît à la fois ancienne, mainte fois abordée et discutée, et pourtant sans grands noms ni dates qui viennent à l'esprit spontanément. Une raison que l'on pourrait avancer pourrait tenir au caractère complexe et holistique des discours portant sur la forme, et de l'absence de cadre théorique prédominant permettant soit d'expliquer ou de prédire des éléments de réalité non triviaux ou idéaux, idéels.

Par contraste, les succès pratiques et théoriques, philosophiques même, des approches réductionnistes ont pu éclipser ou freiner les développements techniques qui auraient pu sous-tendre une ou des théories de la forme. La méthode expérimentale cristallise en elle la rencontre entre théorie et mise en pratique des idées du réductionnisme, et a servi de point de repère et d'ancrage à la fondation de nombreuses sciences. En effet, les sciences, quelle

que soit la complexité apparente de leur objet d'étude, ont eu pour repli dernier de se constituer science à partir de leur méthode, laquelle devait leur permettre d'isoler des objets de manière reproductible, ainsi que des liens directs et simples entre ces objets. La méthode reine étant la méthode expérimentale, elle a fait fonction d'appui paradigmatique pour la plupart, même si les amputations qu'elle engendrait a pu diviser. Mais les faits sont là : toute science se doit de posséder un laboratoire plus ou moins évident, plus ou moins simple, lequel permet d'isoler et de contrôler des expériences élémentaires – seul le degré de virtualité, c'est-à-dire, le déplacement de l'outil matériel vers l'outil conceptuel, le cadre calculatoire, semble varier dans la définition de ces laboratoires : le biologiste possède sa paillasse, ses petits animaux et ses instruments ; le microbiologiste possède son microscope ; l'épidémiologiste possède ses cohortes et ses modèles de représentation statistiques ; le physicien des hautes énergies possède ses accélérateurs de particule longs de plusieurs kilomètres. . .etc.

Aussi, le but de ce travail n'étant pas tant de retracer une histoire exhaustive des théories morphogénétiques, avec par ailleurs, cette mise en garde introductive quant à ce que nous qualifions ici davantage de morphodynamique, que de se donner des éléments de définir un problème général, un manque récurrent, et de se baser sur des éléments de formalisation existant pour dégager un cadre théorique général, nous n'aborderons dans ce chapitre que peu de noms. Néanmoins, ces noms correspondent à autant d'avancées théoriques déterminantes, souvent fruits d'une approche transdisciplinaire. Ils couvrent en outre une période historique finalement courte – un peu plus d'un siècle désormais – même si bien évidemment, toute avancée se fait sur la base de travaux et d'idées les précédant.

De façon étonnante peut-être, ces différents noms sont associés plus ou moins explicitement au domaine de la morphogenèse, et paraissent pourtant,

tels qu'ils sont souvent présentés, sans lien évident ou direct. Ils partagent néanmoins au moins une de ces deux caractéristiques : ils se sont intéressés au domaine du vivant, tâchant par là de le comprendre, le définir ou le mimer, ou à la complexité.

Nous aborderons ainsi la théorie des transformations géométriques de D'Arcy Thompson, [25, 26] qui a minutieusement mesuré, comparé et rapproché des spécimens du vivant, montrant en quoi des espèces différentes peuvent finalement se déduire morphologiquement les unes des autres par des transformations géométriques ; nous présenterons également Turing, [27, 28, 29] qui servira de trait d'union entre une théorie physico-chimique simple des motifs spatiaux du vivant et une théorie du calcul universel, de l'approche algorithmique de la réalité ; l'approche algorithmique, *via* la représentation du vivant par des automates nous permettra de présenter les travaux de Von Neumann, [30] ainsi que de Gödel ; [31] cette approche intégrative sera contrebalancée par une approche de la complexité par la simplicité, ou comment des considérations simples ou des modèles simples peuvent engendrer des comportements qualitativement complexes : nous parlerons de la stabilité structurelle de Thom, [32, 33] et par extension des systèmes dynamiques. [34] Toutes ces approches posent, d'une façon ou d'une autre, la question de la complexité, et une façon de caractériser la complexité tient dans la dimensionnalité intrinsèque du problème traité : Mandelbrot nous donnera les clés d'une dimension non entière, fractale. [35, 36] Nous finirons sur un invariant qui traverse les sciences et nos propos ici, à savoir la dualité entre discret et continu. Le lecteur pourra consulter avec profit, pour des présentations plus exhaustives des différentes théories de la morphogenèse, l'ouvrage coordonnée par Bourguine et Lesne. [37]

Puisque notre ambition est de repérer des éléments de formalisation sur lesquels il serait possible de s'appuyer pour bâtir une théorie morphody-

namique, nous traiterons plus en détails dans deux sections trois cadres théoriques qui nous serviront ultérieurement : le formalisme des systèmes dynamiques, ainsi que le formalisme des théories quantiques et relativistes.

Un des vœux cher aux pères des théories physiques et psychanalytiques actuelles était de pouvoir rendre compte de la réalité et de l'action des subjectivités, de l'action de la pensée sur le monde, du pouvoir des représentations mentales sur la manière d'appréhender la réalité. Nous parlerons donc dans une dernière section des approches de type sciences cognitives ou neurosciences, qui, paradoxalement, sont actuellement plus en quête d'un nouveau matérialisme ou positivisme que de donner les clés d'un rôle de la psyché et de l'observateur sur le monde physique.

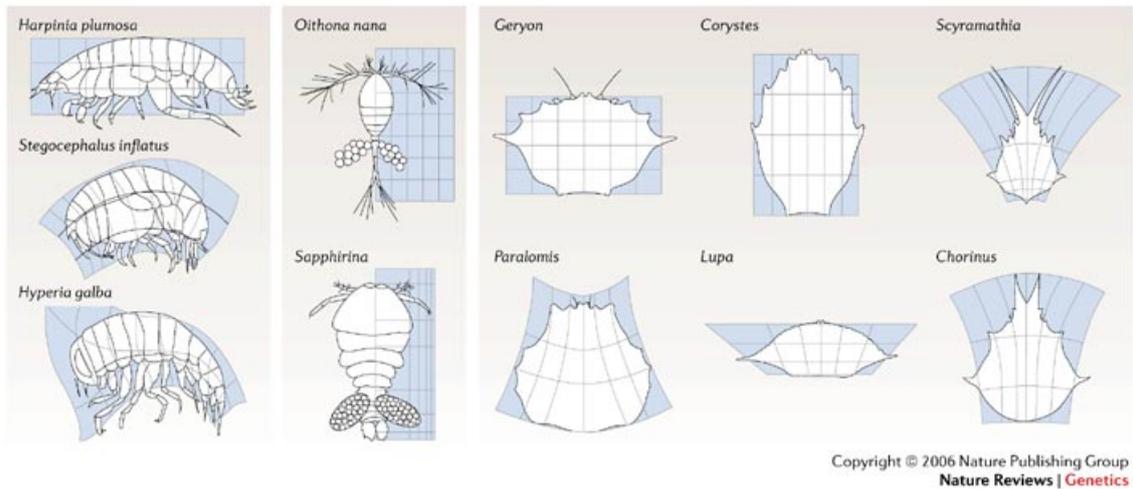
2.1 Les grands noms des théories morphogénétiques

Il semble exister une insatisfaction latente dans le traitement du monde par une approche réductionniste, si l'on en croit les célébrations chaleureuses qu'accompagne chaque nouvelle théorie qui ne relève pas d'un réductionnisme pur. La diversité des approches proposées dans ce qui suit rend compte de la non unicité, jusque ici, des cadres ou outils formels pour expliciter la forme : approche géométrique pure, approche analytique et dynamique, approche calculatoire et algorithmique ou encore approche structurelle et qualitative. Toutes, à des degrés divers, ont cependant opéré un glissement conceptuel propre, intégré un aspect jusque là ignoré ou délaissé, qui a été déterminant.

2.1.1 Wentworth d'Arcy Thompson : *On growth and form*

Quel lien – simple – peut-il exister entre deux espèces de poissons que le biologiste ou le zoologue distinguera l'une de l'autre dans l'arbre phylogénétique ? C'est en quelque sorte la question à laquelle répond d'Arcy Thompson

FIGURE 2.1 – Exemples de transformations géométriques liant différentes espèces, d'après d'Arcy Thompson



dans son ouvrage érudit et complet, *On growth and form* (traduit par *Forme et croissance*). [26]

La réponse apportée déroutera, mais interpellera, les partisans farouches de la causalité et de l'évolution : de simples transformations géométriques permettent d'obtenir une espèce à partir d'une autre, comme il est possible de le voir sur les figures 2.1 et 2.2. D'autres exemples plus schématiques sont donnés en figures 2.3, 2.4 et 2.5.

Actuellement, la théorie des transformations de d'Arcy Thompson pourrait trouver une formulation élémentaire très simple, connue et utilisée par ailleurs en théories physiques. Elle s'écrirait à l'aide de groupes de transformations continues, dans le plan ou l'espace ; autrement dit, elle exploiterait des groupes de Lie.

Les groupes de Lie sont abondamment utilisés en physique, avec des fortunes diverses, mais sont désormais considérés comme un des piliers théo-

FIGURE 2.2 – Exemples de transformations géométriques liant des espèces de poisson, d'après d'Arcy Thompson

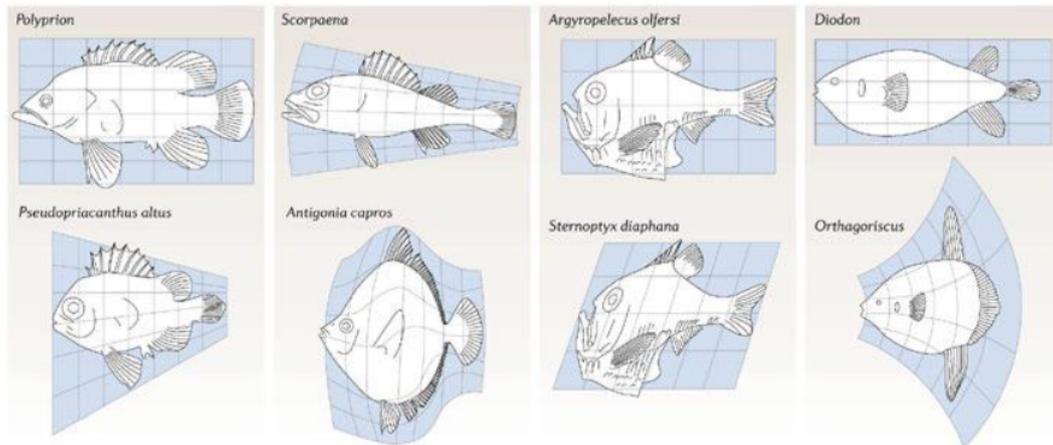


FIGURE 2.3 – Exemple schématique d'une transformation géométrique liant deux espèces de poisson, d'après d'Arcy Thompson

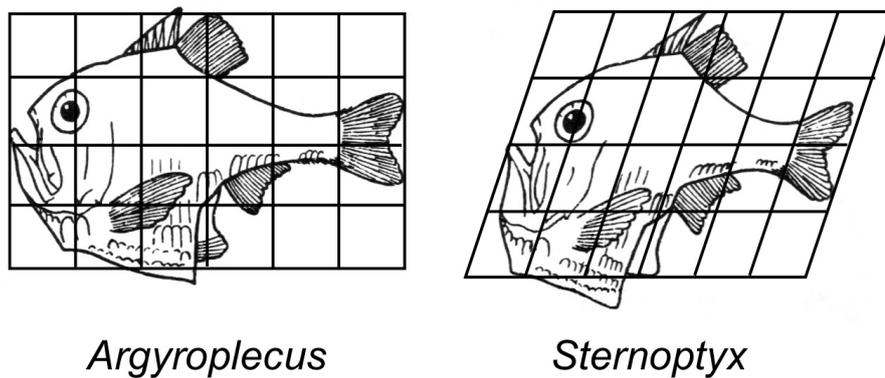


FIGURE 2.4 – Exemple schématique d'une transformation géométrique liant deux espèces de poisson, d'après d'Arcy Thompson

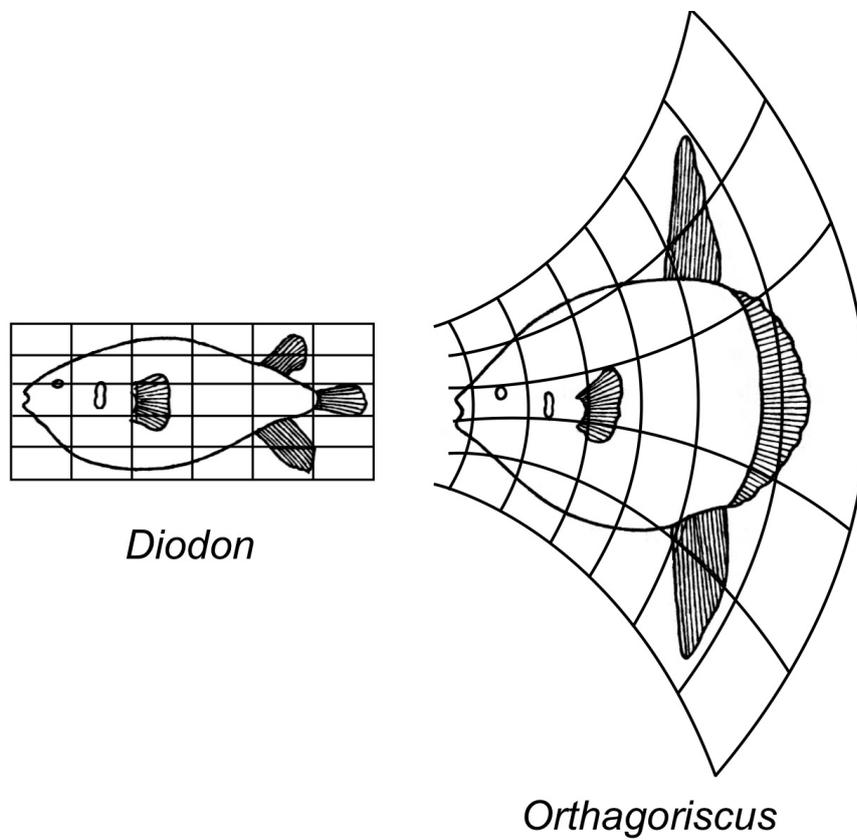
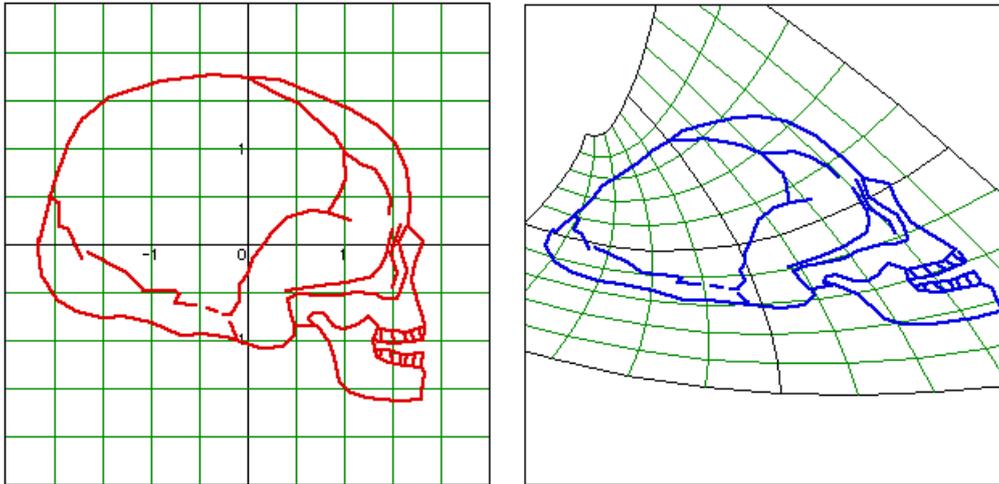


FIGURE 2.5 – Application de transformations géométriques simples à un crâne humain, type d’Arcy Thompson



riques, qui permettent de décrire de manière concise un ensemble de propriétés. [38] Notamment, les théorèmes de Noether, portant sur la conservation de l’énergie, entre autres, s’expriment comme l’obtention de grandeurs invariantes sous l’action de groupes de transformations, par exemple, par translation. [39]

Plus précisément, Noether stipule que pour chaque transformation infinitésimale laissant invariante l’intégrale dite d’action, il existe une grandeur caractéristique du système conservée. Ceci trouve donc des formulations analytiques dans le cadre de théories physiques reposant sur un Lagrangien ou un Hamiltonien – nous y reviendrons.

Pour exemple, l’invariance par translation dans le temps entraîne la conservation de l’énergie du système ; l’invariance par translation selon une direction donnée de l’espace implique la conservation de la quantité de mouvement selon cette même direction.

Des travaux menés au laboratoire (LaTIM, travaux de G. Guillard) précédemment sur le lien forme-fonction, sur la morphologie fonctionnelle, se sont également emparés des groupes de Lie pour caractériser ce lien – en l’occurrence, il s’agissait des quaternions. L’idée était de se doter d’un espace de représentation conjointe pour la forme et la fonction, pour l’étude des articulations anatomiques. [40, 41, 42]

Ainsi, si l’on décrit, initialement comme tout cartésien le ferait, tout corps soumis à développement organique, par un espace muni d’un repère, de coordonnées, alors il est possible d’obtenir un autre corps existant par action de groupe non pas tant sur le corps lui-même, mais bien sur le repère, c’est-à-dire, le système de coordonnées lui-même. En effet, toute la puissance de la théorie de d’Arcy Thompson tient dans la transformation du cadre, de l’espace de définition du corps, et non du corps lui-même. Si la grille initiale était cartésienne, elle dévie très rapidement de cette propriété pour revêtir les caractéristiques d’une variété différentielle.

Ce point de vue est extrêmement intéressant, car il peut laisser supposer que ce qui distingue ici deux espèces tient dans la structure au travers de laquelle s’est manifestée la morphologie. Il existerait de manière sous-jacente, un invariant, une règle vraie pour tout un ensemble de corps, mais dont les observations et réalisations diffèreraient selon le point de vue.

La critique principale, et considérée comme rédhibitoire, faite à la théorie des transformations, tient dans son caractère exclusivement descriptif, sans aucun cadre explicatif ou causal. D’Arcy Thompson le reconnaissait lui-même volontiers, et considérait par ailleurs son ouvrage comme une vaste collection préliminaire d’observations, qui aurait servi d’introduction à une théorie générale. Cependant, ce travail minutieux se doit d’être considéré, et doit pouvoir servir de point de départ solide pour une théorie morphodyna-

mique, en laissant entrevoir d'une part des principes simples et universels, et d'autre part, la possibilité d'une formulation géométrique, relativiste. Reformulée autrement, la critique à lever concernant la théorie des transformations serait la suivante : si deux espèces – deux formes – sont déductibles l'une de l'autre par l'action de groupes de transformation de type groupes de Lie, existe-t-il un autre espace, un espace des paramètres structurant ce ou ces groupes de Lie ?

Il est à noter que cette théorie ne prend pas en compte explicitement le temps, quelle que soit son échelle – échelle de l'évolution, ou échelle de la gestation, de l'embryogenèse. En se focalisant sur une approche purement spatiale et géométrique, la théorie de d'Arcy Thompson a tout pour être *a priori* une théorie sans flèche thermodynamique, et laisse possible toute réversibilité aux processus morphogénétiques sous-jacents. Techniquement, il suffit de voir que les transformations décrites par les groupes sont des transformations inversibles, et deux espèces liées par un ensemble de transformations, finalement isomorphes l'une de l'autre.

Une validation, ou une infirmation, plus complète de cette théorie devra ainsi passer par la recherche d'un paramétrage des groupes de transformations ; la structure liant ces paramètres devra certainement rendre compte d'une certaine forme temporelle, et de conditions de réversibilité. Nous verrons dans la suite que cette approche est relativement unique en son genre, les autres théories favorisant des approches opposées, c'est-à-dire, partir d'un modèle sous-jacent, et en examiner les développements. C'est notamment ce qu'a proposé Turing.

2.1.2 Alan M. Turing : morphogénèse et chimie

Turing a eu une vie courte et brillante, et s'est attaché, comme certains de ces contemporains, à expliquer, ou du moins proposer une représentation

crédible et fonctionnelle minimale, universelle, du vivant.

Plutôt qu'à l'embryogenèse elle-même, ou à la phylogenèse, Turing s'est intéressé aux mécanismes susceptibles d'expliquer la formation de motifs qui paraissent à la fois déterministes – tous les zèbres ont des rayures – et variables – aucun zèbre n'est identique à un autre. Pour ce faire, il a supposé que la survenue de motifs récurrents pouvait être simplement le résultat mécanique quoique d'expression finale et détaillée variable, due aux conditions spécifiques de sa réalisation, d'un processus chimique.

Le constat initial est qu'un organisme est constitué d'un certain nombre de composés chimiques, en quantités réglées ou régulées, et que les réactions impliquant certains de ces composés ne sont d'une part pas toutes possibles, et d'autre part, possiblement concurrentes.

La thèse de Turing concernant la morphogenèse est résumée et publiée dans son article de 1952, *the chemical basis of morphogenesis*. [28] Il renvoie également à ce que l'on appelle en physique à la notion de brisure spontanée de symétrie : il apparaît, à terme, un résultat spatialement non homogène, non uniforme – les zébrures de l'animal – alors que les conditions initiales de répartition des espèces réactives sont elles spatialement uniformes.

Les réactions chimiques mentionnées par Turing suivent ce que l'on a nommé ensuite des règles de réaction-diffusion. Un système de réaction-diffusion est un système qui rend compte de l'évolution spatiale et temporelle des concentrations des composés chimiques considérés. L'évolution est supposée gouvernée par deux mécanismes particuliers : un mécanisme de diffusion, impliquant l'opérateur différentiel Laplacien, appliqué aux concentrations, et un mécanisme de réactions, ici de réactions chimiques, qui gouverne l'évolution des concentrations selon les processus de créations, de dégradations,

d'amplification ou d'inhibition des composés chimiques. Ainsi, la différentielle temporelle des concentrations entre deux instants est le résultat d'une part d'une diffusion spatiale des composés, menant d'autre part à des réactions impliquant ces composés.

Mathématiquement, ces réactions-diffusions se formalisent de la manière suivante :

$$\partial_t q = D \nabla^2 q + R(q)$$

où chaque composante du vecteur $q(x, t)$ désigne la concentration d'un des composés. D est une matrice diagonale de coefficients de diffusion, et R désigne l'ensemble des réactions chimiques locales possibles entre composés. Les solutions de ces équations peuvent présenter des comportements qualitatifs variés, dont la formation de motifs, comme évoqué. Une solution classique de ce type d'équation peut également prendre la forme d'ondes stationnaires ou se propageant.

Dans le cas extrême où l'on ne considère qu'un seul composé, avec un espace unidimensionnel, l'équation se réduit à

$$\partial_t u = D \partial_x^2 u + R(u)$$

connue aussi comme l'équation KPP (Kolmogorov-Petrovsky-Piskounov).

Dans ce cas, il est possible de reformuler cette équation selon une approche variationnelle, de la manière suivante :

$$\partial_t u = - \frac{\delta L}{\delta u}$$

avec L vue comme une énergie libre, donnée par :

$$L = \int_{-\infty}^{\infty} \left[\frac{D}{2} (\partial_x u)^2 - V(u) \right] dx$$

$V(u)$ étant un potentiel tel que $R(u) = \frac{dV(u)}{du}$.

Comme annoncé plus haut, il est possible de chercher des solutions classiques à ce type d'équation, sous la forme d'onde de propagation, de vitesse constante et sans déformation. Ces solutions sont du type $u(x, t) = \hat{u}(\xi)$, avec $\xi = x - ct$, avec c la vitesse de propagation de l'onde. Néanmoins, ces solutions sont des solutions stables, et n'amènent pas à la formation de motifs. En revanche, ce type d'équation peut également accepter des solutions instables : ce sont ces solutions qui engendrent la possibilité de motifs.

Nous montrons, pour quelques pas de l'expérience, comme ce type d'équation peut engendrer des motifs, ici simulée sous Matlab. Les résultats sont donnés par les figures 2.6, 2.7, 2.8, 2.9 et 2.10, ainsi qu'un aperçu 3D du résultat final.

Il est possible ici de construire un parallèle avec ce que propose d'Arcy Thompson, bien que, comme nous l'avons précisé auparavant, la théorie des transformations ne présente pas de mention explicite du temps. Le modèle de Turing prend le contre-pied de celui de d'Arcy Thompson, et part des mécanismes sous-jacents à la formation de motifs. Le référentiel de description est fixe, et ce sont les espèces chimiques qui migrent dans l'espace dans ce point de vue. Néanmoins, si l'on considère que le terme en Laplacien décrit un équivalent de transformation géométrique, l'on peut s'intéresser au terme en $R(u)$, lequel fournit quant à lui un ensemble de règles causales quant à la création et dégradation des composés. Ainsi, il n'est pas question ici de groupe de transformations plus ou moins compliqués, mais d'un seul mécanisme spatial, celui de la diffusion. Ce qui créera la spécificité tiendra dans le modèle de réaction locale entre espèces.

FIGURE 2.6 – État proche initial d'un système de diffusion-réaction type Turing

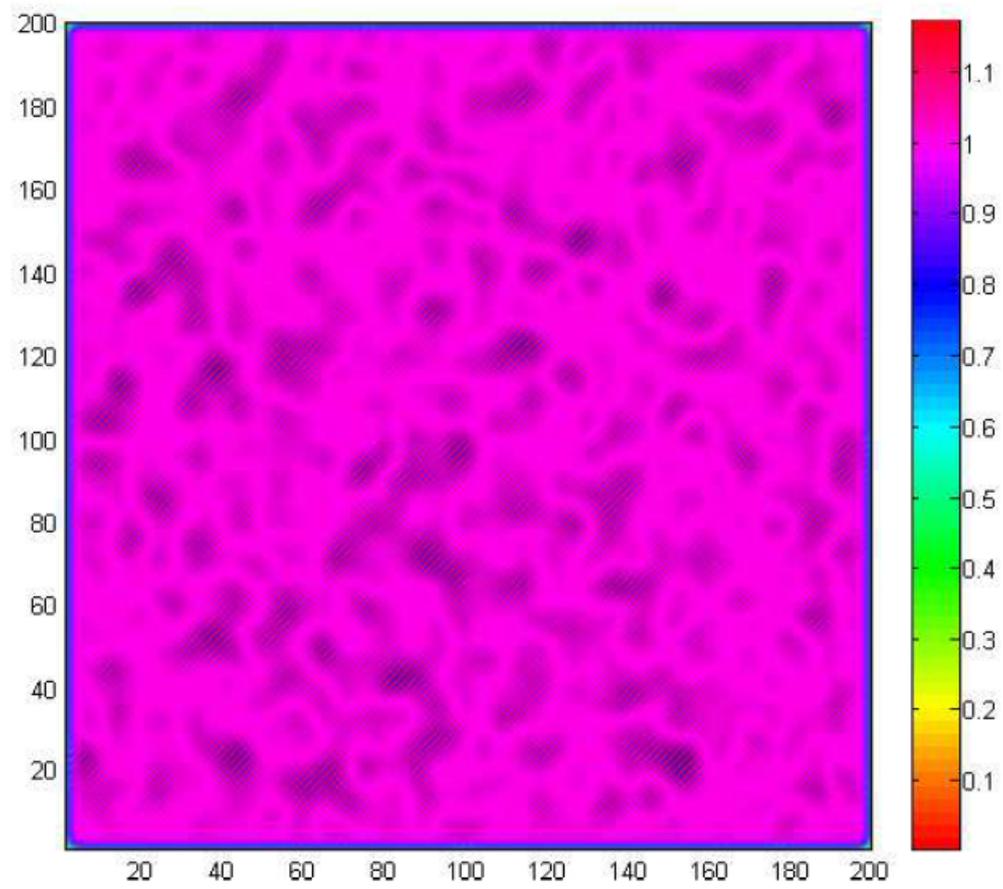


FIGURE 2.7 – État intermédiaire d'un système de diffusion-réaction type Turing (1)

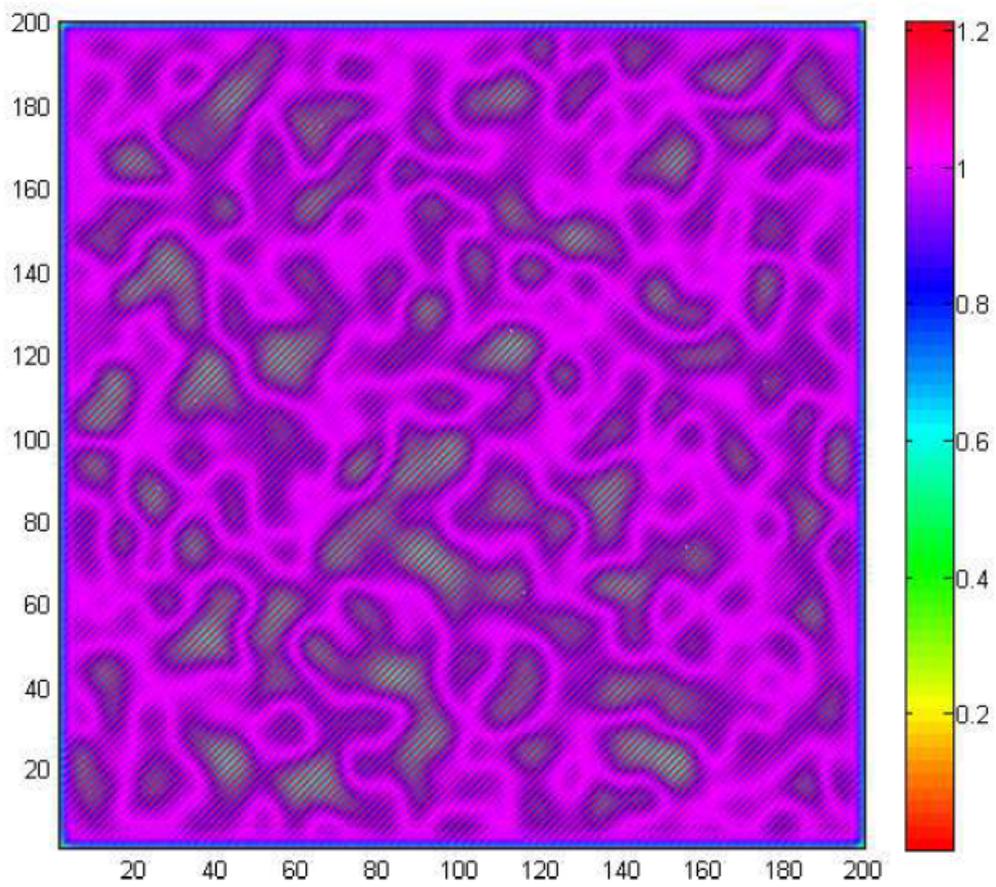


FIGURE 2.8 – État intermédiaire d'un système de diffusion-réaction type Turing (2)

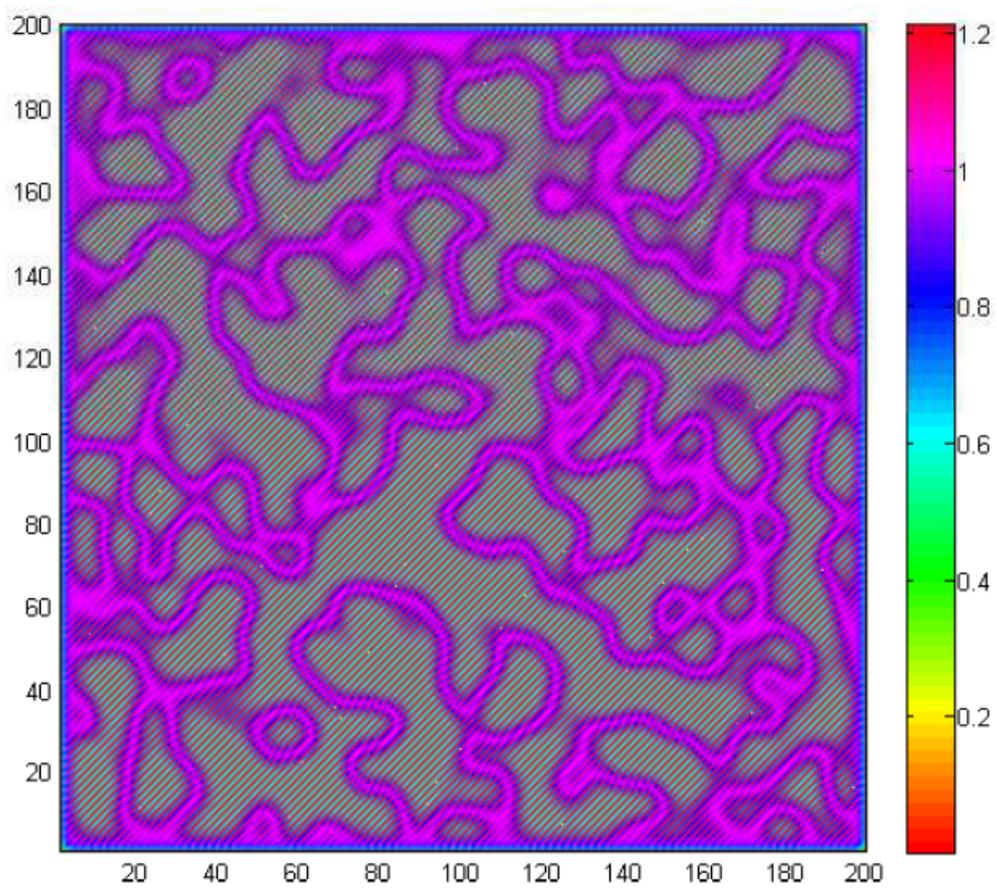


FIGURE 2.9 – État final d'un système de diffusion-réaction type Turing (2D)

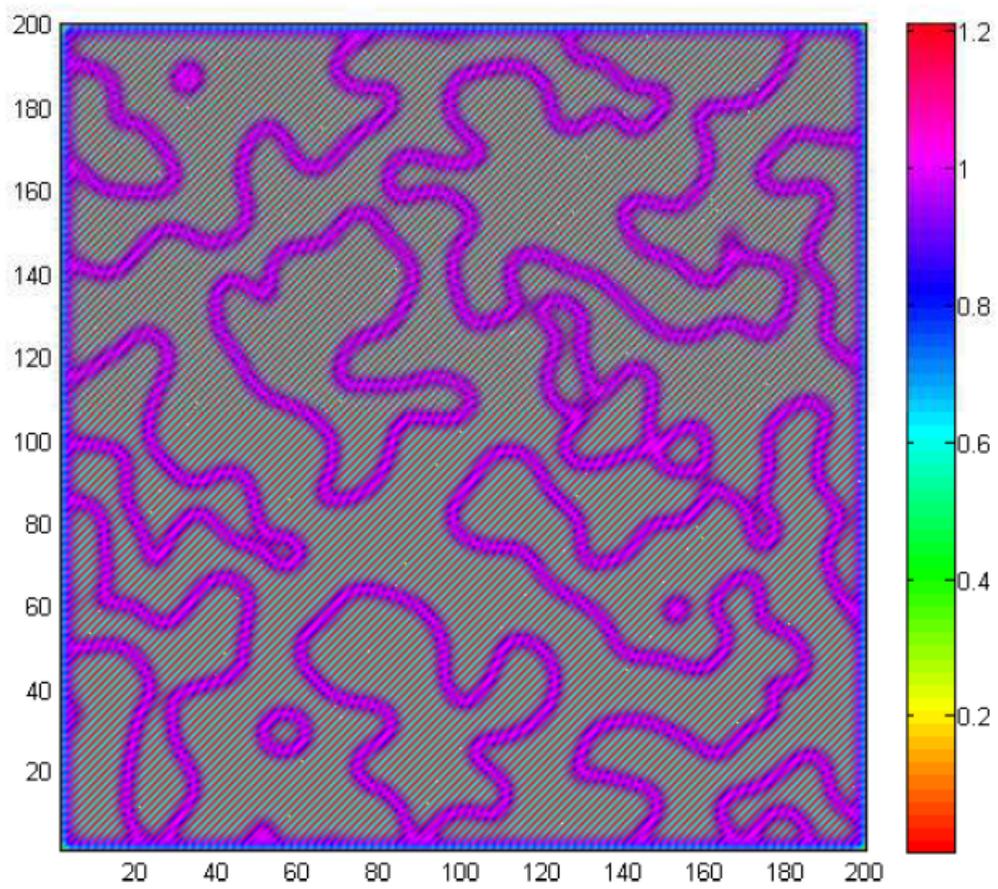
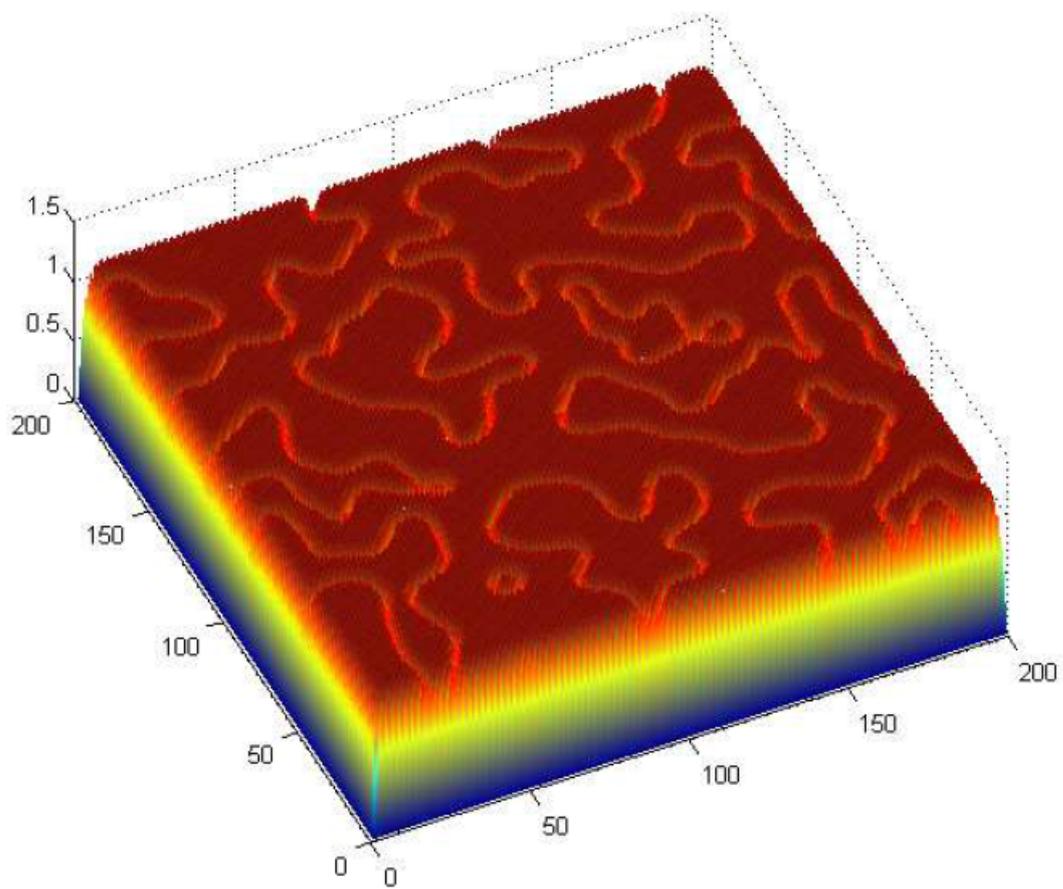


FIGURE 2.10 – État final d'un système de diffusion-réaction type Turing (3D)



Notons que ce modèle proposé par Turing a pour vocation de s'étendre, ainsi que son auteur l'a présenté la première fois, à toute situation où l'apparition d'un motif, qu'il soit coloré, organes ou autre, serait liée à des mécanismes chimiques locaux. Le modèle vaudrait ainsi également pour la gastrulation, par exemple. Les espèces chimiques considérées sont appelées par l'auteur des morphogènes : ils dirigent l'apparition de motifs.

Ce type d'approche a servi de modèle fondateur à la biologie théorique ; néanmoins, les expériences confirmatoires furent relativement tardives, même si elles semblent ne pas infirmer les hypothèses de Turing. [43, 44, 45] L'apparition des simulations numériques, et les améliorations des techniques de biologie moléculaire, la meilleure compréhension des cycles cellulaires ont permis d'investiguer plus avant ce type d'explication.

2.1.3 Turing, Gödel et von Neumann : machine universelle, approche algorithmique et automates

Turing a proposé un modèle de la morphogenèse basé sur la chimie, et a mis en évidence des phénomènes de brisure de symétrie à partir d'un système finalement simple de réaction-diffusion. Le terme en $R(u)$ de l'équation de réaction-diffusion est une forme particulière de formulation des règles régissant localement la création ou la dégradation d'espèces chimiques. En ce sens, ce terme est une expression particulière, entre l'extension ou l'équivalence des forces dans la formulation Newtonienne de la mécanique, et l'expression algorithmique, informatique de règles d'interaction et de production.

Ce terme est constitué de toutes les possibilités de réactions impliquant les espèces considérées, et est muni d'un sens de réaction. Il possède donc une orientation logique préférentielle.

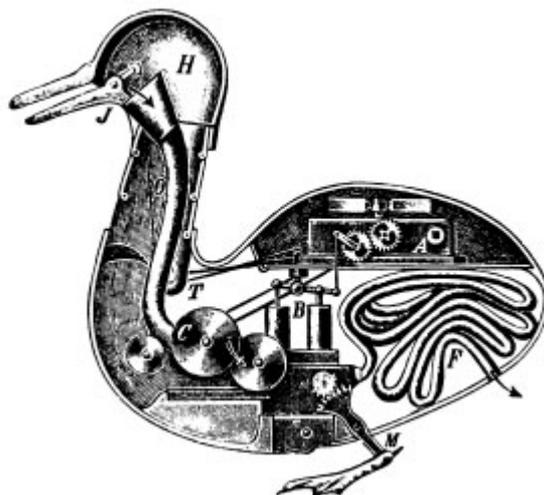
Il n'est certainement pas étonnant de voir que Turing a beaucoup participé à la formalisation de l'informatique, de l'intelligence artificielle et de l'algorithmique. A ce point, nous arrivons davantage à la partie causale, sous-jacente à la morphogenèse, et nous invoquons ici deux autres noms qui ont massivement contribué à cette approche, à savoir Von Neumann et Gödel.

Aux origines de l'ordinateur moderne, il y a eu la volonté d'automatiser, d'autonomiser et de régler le calcul, complexe ou non, avec cette conviction qu'il était possible de sérialiser un certain nombre d'opérations élémentaires afin de déterminer de manière rigoureuse et mécanique, infaillible, la solution d'un problème. L'approche dérive tant de l'axiomatique – déduire de manière logiquement valide des théorèmes à partir de règles supposées vraies *a priori* – que de l'horlogerie. Simultanément, il y a le désir, également de mimer, de comprendre le vivant, dans une ou plusieurs de ses composantes.

Aussi, l'automate, le robot et l'ordinateur sont des avatars plus ou moins complexes venant de temps relativement anciens. Il est par exemple fait mention des automates de Vaucanson, dont le canard digérateur est resté célèbre (voir le schéma du mécanisme en figure 2.11 et la photo du dispositif en figure 2.12). Celui-ci reproduisait en effet la digestion, depuis l'ingestion d'aliments à leur excrétion. Il aurait également été capable de cancaner et de reproduire les mouvements natatoires. Ce canard a été exposé au Palais Royal en 1744.

Dans le registre de l'automatisation du calcul, Charles Babbage et Ada Lovelace (dont un langage informatique a hérité le prénom – ADA) demeurent deux précurseurs incontournables. Babbage s'était lancé à l'époque (à partir de 1812) dans la conception d'une machine à différences, mécanique permettant d'exécuter automatiquement, séquentiellement des calculs par ailleurs fastidieux et sujets aux erreurs d'inattention ou de fatigue de la part d'opé-

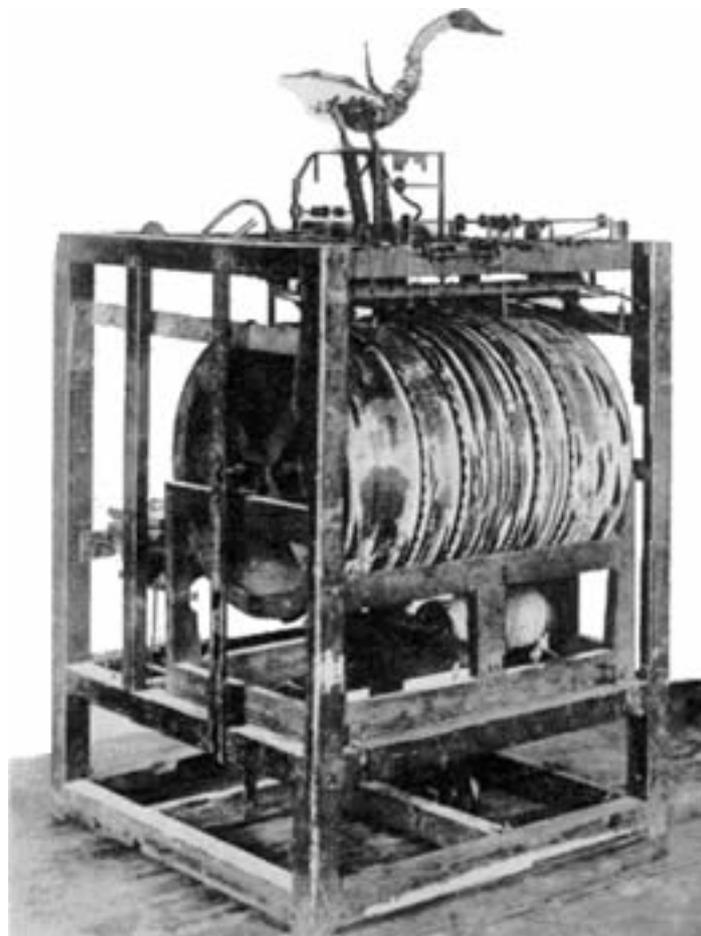
FIGURE 2.11 – Schéma du mécanisme du canard de Vaucanson



rateurs humaines ; la vitesse d'exécution n'était alors, semble-t-il pas une question de premier plan. Ses travaux et réflexions lui auraient été inspirés par ceux tout aussi avant-gardistes de Pascal (la pascaline, machine à calculer) et de Leibniz (la multiplicatrice). Ada Lovelace aurait quant à elle largement épaulé Babbage en tant que mathématicienne talentueuse, et davantage œuvré du côté programmation et algorithmique.

Il est dit que les préoccupations pratiques qui auraient amené ces différents savants à théoriser puis avancer vers la réalisation concrète de tels instruments étaient essentiellement des préoccupations liées à la fiabilité des tables nautiques et astronomiques. Des préoccupations orientées vers la navigation et la prédiction d'observations correctes donc. Des préoccupations similaires ont également motivé le développement des premiers grands calculateurs, ce qui devaient être les premiers ordinateurs, dont l'ENIAC américain (1946), le Z3 allemand (1941) ou le Colossus anglais. Ces ordinateurs ont été conçus pour effectuer des calculs automatiques, sur la base de programmes passés en entrée, essentiellement dévolus à la balistique.

FIGURE 2.12 – Photographie du dispositif du canard de Vaucanson



Parmi les continuateurs de Vaucanson, de Babbage et de Lovelace, nous retrouvons ainsi Turing, et Von Neumann. Turing a théorisé au maximum le concept d'une machine à calculer, au sens large, minimaliste et universelle. Au sens large, car elle devait être capable d'exécuter n'importe quel programme, n'importe quel algorithme, et ce sans aucune faute (*On Computable Numbers, with an Application to the Entscheidungsproblem*). [27]

Classiquement, une machine de Turing est décrite comme constituée d'un ruban, d'une tête de lecture-écriture, d'un registre d'état et d'une table d'actions. Le ruban peut être vu comme un fil d'instructions discrètes, une à la fois, ordonnées séquentiellement, unidimensionnellement, partitionné en cases, lesquelles contiennent chacune une instruction élémentaire. La tête de lecture-écriture se positionne sur une case à la fois, peut lire l'instruction sur le ruban, ou en écrire une. En général, il est considéré que c'est la tête de lecture-écriture qui se déplace le long du ruban, et non le ruban qui bouge. Le registre d'état retient l'état dit courant de la machine ; c'est cet état qui donne le contexte actuel de la machine au moment de la lecture. En ce sens, la mémoire de la machine étant limitée à un état unique, cette description est compatible avec la propriété de Markov (l'état futur ne dépend que de l'état immédiatement précédent et non au-delà). La table d'actions, enfin, fournit les fonctions de base de la machine, qui se traduisent par un déplacement donné (droite/gauche, nombre de cases), et la lecture ou l'écriture, ainsi que le changement d'état. En quelque sorte, cette table est le langage de la machine, et les programmes (les rubans) devront être écrits en prenant en compte ce langage. La table d'actions peut également être représentée par un automate à états finis, décrivant sous quelles conditions l'on passe d'un état à un autre, et quelle action un état peut déclencher.

Une machine de Turing est un septuplet $(Q, \Gamma, B, \Sigma, q_0, \delta, F)$, où :

- Q est un ensemble fini d'états ;
- Γ est l'alphabet de travail des symboles de la bande ;
- $B \in \Gamma$ est un symbole particulier (dit *blanc*) ;
- Σ est l'alphabet des symboles en entrée ($\Sigma \subseteq \Gamma \setminus \{B\}$) ;
- $q_0 \in Q$ est l'état initial ;
- $\delta : Q \times \Gamma \rightarrow Q \times \Gamma \times \{\leftarrow, \rightarrow\}$ est la fonction de transition ;
- $F \subseteq Q$ est l'ensemble des états acceptants (ou finaux, terminaux).

Les flèches dans la définition de δ représentent les deux déplacements possibles de la tête de lecture, à savoir le déplacement à gauche et le déplacement à droite. La signification de cette fonction de transition peut être expliquée sur l'exemple suivant : $\delta(q_1, x) = (q_2, y \leftarrow)$ signifie que si la machine de Turing est dans l'état q_1 et qu'elle lit le symbole x , elle écrit y à la place de x , va dans l'état q_2 , et déplace sa tête de lecture vers la gauche.

On parle de machine de Turing universelle, car l'auteur a montré qu'en réalité, il était possible d'encoder sur le ruban la table d'actions (donc le langage) que l'on désirait en même temps que le programme à exécuter. La machine de Turing est supposée (selon la thèse dite de Church-Turing) pouvoir résoudre n'importe quel problème pouvant se résoudre par l'application d'un algorithme ou tout moyen concret, réel et réaliste de calcul.

De fait, disposant d'un moyen théorique et pratique de caractériser un problème et sa résolution, il a été tenté de proposer une catégorisation de ces problèmes selon les caractéristiques de la machine nécessaire à leur résolution. C'est ainsi que l'on a pu proposer la notion de complexité algorithmique et de calculabilité. Il existe différentes approches et typologies de classes de complexité ; retenons simplement ici que la complexité d'un problème peut s'exprimer selon deux dimensions particulières : le temps et la mémoire nécessaires à la résolution. La possibilité d'exprimer tout problème *via* un

seul formalisme permet ainsi les comparaisons ; par exemple : quelle est la longueur du ruban nécessaire pour le problème A et le problème B, quel temps pour résoudre chacun, quelle taille de la table d'actions ?

Si Turing a été l'un des principaux théoriciens de la calculabilité et de la machine à calculer, il est un autre nom qui est resté célèbre pour ses travaux visant à répliquer le vivant, ou du moins son mode de fonctionnement, sous la forme d'automates. Von Neumann est par exemple l'auteur de la théorie générale et logique des automates. [30] Il est également le co-créateur de l'architecture classique des ordinateurs modernes, où sont séparées la mémoire, l'unité de traitement et une unité de contrôle. Tant le programme à exécuter que les données sur lesquelles appliquer le programme sont stockés en mémoire, les mettant sur un pied d'égalité. On retrouve là le principe de la machine universelle. L'unité de contrôle joue le rôle d'ordonnanceur : elle crée les séquences d'opérations, lesquelles sont disponibles dans l'unité de traitement (appelée unité arithmétique et logique).

Au travers de tous ces différents modèles, l'on retrouve des invariants, que sont la mémoire, la lecture et l'écriture d'informations, et la représentation du traitement de l'information sous la forme d'un automate.

Parallèlement et à la suite, un autre type de théorie s'est développé, mettant en avant un élément supplémentaire, qui n'apparaît pas, du moins explicitement, dans les théories de Turing ou des automates en général. Il s'agit de la notion de rétroaction, notamment de rétroaction négative, comme moyen de régulation. Cette notion traversera la théorie générale des systèmes de Bertalanffy (où l'on parle avant tout d'interactions entre systèmes et sous systèmes), [24] et celle de la cybernétique portée par Wiener. [22] Sa traduction en biologie, particulièrement en biologie de l'évolution et de la physiologie se trouvant sous les plumes d'Henri Laborit [46] et Henri Atlan. [47] C'est

également la période qui voit la formalisation de la théorie mathématique de la communication, de Shannon. [21]

Enfin, à un niveau peut-être plus fondamental, nous soulignons ici les travaux de K. Gödel, lequel nous a légué notamment ses théorèmes d'incomplétude. Ces travaux s'inscrivent dans la quête d'axiomatisation et d'automatisation des théories ne relevant pas explicitement de la logique, en partie à l'appel de D. Hilbert. Ils rejoignent également les travaux de Turing sur la machine universelle et la complexité algorithmique. Un résultat fondamental des théorèmes d'incomplétude est saisissant ; il stipule, dans le cadre de l'arithmétique, que toute théorie pouvant faire "fonctionner" une arithmétique, est mécaniquement *incomplète*, ce qui signifie qu'il est possible, au sein de cette théorie, avec son langage propre, d'énoncer des propriétés, dont il ne sera pas possible de dire si elles sont vraies – ou si leurs contraires sont fausses – avec les outils de cette même théorie. Autrement formulé : il est possible d'exprimer des énoncés syntaxiquement corrects, ayant *a priori* un sens, mais dont il est intrinsèquement impossible de dire s'ils sont exacts ou non. [48, 31] Reformulé encore autrement, en des termes et un contexte qui nous intéressent ici, si l'on se donne un système axiomatique, notamment décrit par un automate, il est possible d'exhiber des formes, des variétés, dont le système sera incapable de nous attester le caractère "véridique". Ce type de constat nous amène à plusieurs réflexions. Premièrement, il met un certain coup d'arrêt à l'ambition d'une théorie parfaitement huilée et automatisée, qui fonctionnerait comme un oracle parfaitement binaire vrai/faux et sur lequel il serait possible de se reposer. Gödel nous apprend que nos systèmes de représentation sont capables de représenter – peut-être même de représenter toutes les possibilités – mais qu'ils présentent tous un point aveugle ; en d'autres termes, un système seul est incapable de se contempler lui-même et de se juger dans son intégralité. Deuxièmement, il en découle donc

une place intéressante et pour la subjectivité (un sujet n'a sans doute pas la capacité de se juger entièrement par ses propres moyens), et partant, pour l'intersubjectivité, dans le sens où un sujet complémentaire, une deuxième voix peut possiblement lever, soit "objectivement", soit par décision ou arbitraire, le caractère indécidable d'un énoncé pour un sujet isolé. Selon nous, les résultats de Gödel, plutôt que de nous rendre pessimiste sur nos capacités, devraient nous inciter à explorer au mieux les systèmes collectifs ou hybrides.

2.1.4 René Thom et ses prédécesseurs : systèmes dynamiques et théorie des catastrophes

Si l'on se résume ici, nous avons vu deux approches diamétralement opposées de la morphogenèse : l'approche de d'Arcy Thompson, purement géométrique, atemporelle et acausale dans sa formulation actuelle, et les approches calculatoires, d'explications locales et atomistes, basées sur une expression différentielle et algorithmique, temporelle, dues à Turing dans ses versions les plus explicites.

Les approches type réaction-diffusion sont des approches caractérisant le problème de morphogenèse au travers d'un observateur ou d'un référentiel fixe, absolu et universel, où se déroulent les diffusions et les réactions. On y suit les évolutions des concentrations des différents éléments au court du temps. Les coefficients régissant la forme des équations, et finalement du programme (les réactions sont autant d'actions qui pourraient constituer une table de machine de Turing), y sont fixés.

Quelles seraient les implications sur les résultats classiquement obtenus de ces équations de réaction-diffusion si l'on faisait varier la structure du programme, donc les paramètres des équations ? Si l'approche de Turing concerne par exemple, 3 éléments en réaction, le problème est globalement

de dimension 3. Néanmoins, si le système d'équations qui régit le phénomène est réglé par disons, 9 paramètres, la dimension du problème associé devient toute autre. Ces paramètres caractérisent la structure du problème, et définissent la géométrie, les qualités topologiques du modèle.

Ainsi que nous le verrons plus loin formellement explicité, il est possible d'écrire un tel système, c'est-à-dire, similaire à un système de Turing, de manière plus générale et ne se limitant pas aux systèmes de réaction-diffusion. Nous abordons ici le formalisme des systèmes dynamiques dans leur plus grande généralité.

Une fonction f y joue alors le rôle de structure du problème, ou de programme. Il est alors certes intéressant de connaître, pour un ensemble de valeurs données et fixées de f , les solutions du problème, mais en termes mathématiques, celles-ci ne représentent en réalité qu'un ensemble négligeable des possibilités. En outre, comment s'assurer que ces solutions "négligeables" ne soient pas des solutions très peu crédibles ou que rarement représentées, voire, qu'elles aient des conditions d'existence suffisamment stables pour pouvoir être observées ?

C'est ici la question de la stabilité structurelle d'un système. L'espace étudié n'est plus tant celui des espèces chimiques et de leurs concentrations au cours du temps, mais celui des paramètres structurant le problème, l'espace des programmes ; il s'agit donc de l'espace de définition du problème.

La théorie des systèmes dynamiques introduit la notion d'attracteur, dont nous verrons plus loin la définition formelle. Un attracteur est une portion de l'espace de description qui attire les différentes trajectoires du système étudié. Par exemple, dans un exemple de Turing, on peut concevoir que fatalement, la concentration en composé B évolue vers un état stable, mais os-

cillant entre deux valeurs possibles, sans jamais se fixer sur un intermédiaire. La trajectoire de B évoluerait donc depuis son état initial jusque une portion de l'espace délimitée par ces deux valeurs, et pourrait orbiter autour d'un point fixe, intermédiaire, de cet espace des concentrations. Naturellement, l'existence d'états stables dynamiques candidate pour une explication, un modèle possible de la morphogenèse : de différents états initiaux, il apparaîtrait intrinsèquement une forme particulière et caractéristique. Néanmoins, là aussi, l'apparition *ex nihilo* d'une telle entité ne renseigne pas directement sur la causalité de cette forme : on constate son existence, et la relie à un type de problème donné.

La question de la stabilité structurelle du système précédemment évoquée peut alors se poser quant aux attracteurs eux-mêmes : ils définissent un état attractant du système (les concentrations), mais ont-ils des conditions d'existence précaires, ou sont-ils eux aussi extrêmement stables, si l'on fait varier les valeurs des paramètres structurant ?

L'étude de la stabilité structurelle a trouvé une expression particulière dans les travaux de René Thom, lequel a explicitement relié les notions de stabilité structurelle, de morphogenèse et de topologie différentielle. [32, 33] Sa théorie des catastrophes (entendre la notion de catastrophe comme une singularité, un lieu de changement qualitatif très brutal) rend compte des différentes formes minimales donnant lieu à un changement de comportement qualitatif du système étudié.

Cette conception permet de qualifier le programme régissant le système ; en ce sens, elle rapproche les modèles de morphogenèse de Turing des modèles de d'Arcy Thompson. Les idées de Thom et la théorie des catastrophes ont beaucoup séduit un auteur particulier, qui s'est attaché à étudier et compilé les exemples de modèles, notamment biologiques, pour lesquels la

théorie pourrait être éclairante (E.C. Zeeman – *Catastroph theory, selected papers*, [49]). Cependant, le réalisme, ou la praticabilité de la théorie des catastrophes en particulier, et non des systèmes dynamiques, ont été sévèrement critiqués, alors peut-être à raison, au titre qu’il n’existerait aucune confirmation expérimentale correspondant au formalisme. La théorie initiale a également été attaquée pour sa restriction quant aux cas décrits, à savoir les modèles à base de gradient. La critique est d’autant plus étonnante que la théorie des catastrophes est classiquement décrite et intégrée dans toute présentation de la théorie des systèmes dynamiques, laquelle dernière est elle en général toujours célébrée ou avancée comme un cadre intéressant ou prometteur pour la biologie théorique, ou la biologie systémique.

Les systèmes dynamiques explorent ainsi un problème sous deux angles complémentaires, l’un plus quantitatif et l’autre plus qualitatif. La dimension du problème est alors différente selon l’angle adopté. Il est notable de constater que de la même façon qu’il n’est pas nécessaire d’avoir un système initial très complexe pour observer des phénomènes complexes, il n’est pas nécessaire que la structure du système soit de haute dimensionnalité pour qu’apparaissent des catastrophes, c’est-à-dire des variations topologiques brusques. Dans les deux cas, se pose la question de la dimensionnalité d’un problème ; si une haute dimensionnalité n’est *a priori* pas directement synonyme de complexité, la quantification plus fine de cette dimensionnalité peut révéler un aspect de cette complexité.

2.1.5 Place de la fractalité : Benoît Mandelbrot

Il existe une autre façon d’estimer la dimensionnalité, au sens de sa complexité, d’un problème, ou plus particulièrement de la forme de ses solutions ou de ses comportements. Intuitivement, nous savons qu’entre une ligne parfaitement droite, tirée entre deux points au plus court, et une ligne aux

détours multiples et aléatoires, reliant ces deux mêmes points, il existe une différence qualitative. Dans les deux cas, une seule coordonnée sera certes nécessaire et suffisante pour se situer sur la ligne, étant donné un point d'origine de référence. Néanmoins, comment qualifier la qualité qui les distingue ? S'il était nécessaire de décrire à un interlocuteur ne disposant pas d'une visualisation des deux courbes, il est certain que nous ne proposerions pas la même explication pour la construction de ces deux courbes.

D'une façon un peu différente, la question de l'observateur, du référentiel et de la façon de mesurer un phénomène, quantitativement et qualitativement se pose ici aussi face à ces deux courbes. C'est la question qu'a également posé Benoît Mandelbrot, lorsqu'il nous a interrogés sur la longueur « réelle » des côtes bretonnes ou anglaises. Il s'avère que la longueur des côtes n'est pas un invariant, une donnée fixe, mais un objet d'échelle. Autrement dit, la longueur des côtes, ou la longueur de nos deux courbes précédentes, dépend de l'outil de mesure utilisé pour mesurer cette longueur, essentiellement parce qu'elles sont presque partout non différentiables (les côtes sont très escarpées).

Mandelbrot reprend une approche chère à Einstein quant il s'était agi d'expliquer les phénomènes de dilatation ou de contraction des longueurs en relativité, de manière plus générale, d'expliquer que les longueurs mesurées dépendaient de l'observateur et de l'instrument utilisé pour les mesurer (il a été question d'allumettes pour mesurer des unités de longueur). [5] Mandelbrot propose alors que, selon la taille de la règle ou de l'allumette servant à mesurer les côtes, on captera davantage de détails que la règle sera petite, et la longueur totale, mesurée à une échelle inférieure, s'en trouvera d'autant augmentée (voir figure 2.13). [36, 50]

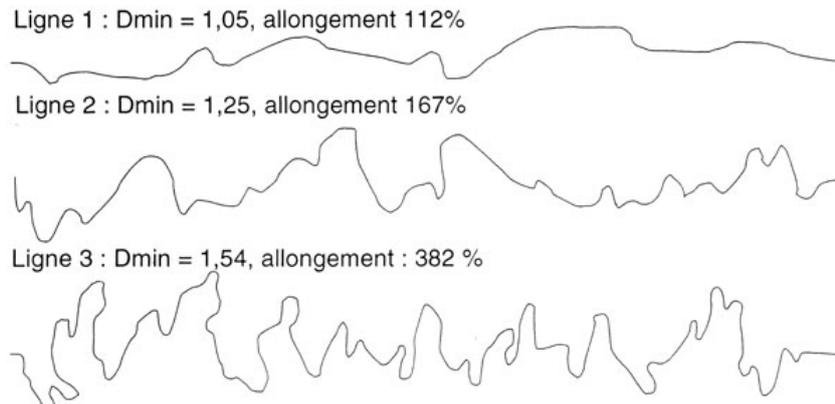
FIGURE 2.13 – Longueurs des côtes selon l'échelle d'observation, (Wikipedia, licence CC)



Si l'on reprend le cas de nos deux courbes, l'une rectiligne, l'autre accidentée en de nombreux endroits, et que l'on désire ne serait-ce qu'en communiquer les longueurs respectives à notre interlocuteur, que se passe-t-il ? La courbe rectiligne, quel que soit l'instrument de mesure, affichera la même longueur totale. En revanche, pour la courbe accidentée, selon la taille (la résolution) de l'instrument de mesure, la longueur sera vraisemblablement différente.

En suivant son raisonnement jusqu'au bout, Mandelbrot proposera une extension de la notion de dimension, qu'il baptisera dimension fractale. Comme son nom peut l'indiquer, cette dimension n'est pas, en toute généralité, un entier, mais un réel. De fait, la dimension fractale de la courbe rectiligne serait de 1, quand la dimension fractale de la courbe accidentée serait vraisemblablement différente de 1, et inférieure à 2 (par exemple : 1.3). En d'autres

FIGURE 2.14 – Dimensions d'une courbe selon son degré de tortuosité, d'après *Genre-Grandpierre*



termes, le caractère fractal de la dimension renseigne sur la « capacité » de la courbe considérée à remplir le plan : une dimension de 1.3 rendrait compte d'une courbe moins « compliquée », remplissant moins le plan qu'une courbe de dimension fractale 1.7 (voir figure 2.14).

Parmi les exemples de systèmes dynamiques que nous aborderons, nous verrons qu'il est possible de caractériser certains attracteurs, et donc, certaines formes, par leur dimension fractale ; c'est le cas des attracteurs dits étranges. Différentes approches coexistent quant à la façon de définir des dimensions non entières. Nous verrons dans la section dédiée aux techniques et algorithmes disponibles pour l'étude des formes à partir d'ensemble de données différents estimateurs de la dimension fractale. Quoiqu'il en soit, la notion de dimension fractale est intimement liée à une notion géométrique et topologique, et permet de qualifier plus précisément certains phénomènes complexes.

2.1.6 Une dualité sans cesse rencontrée : discret et continu, objets *a priori* ou variations topologiques ?

Il existe une notion duale qui traverse la majorité des théories, explicitement ou non, et ce de façon assumée et simultanée depuis l'avènement de la mécanique quantique. En effet, avant la formulation quantique, et son élaboration quasi concomitante de celle de la relativité générale, il semble que cette dualité dont nous voulons parler ait existé, mais de manière successive et périodique, un aspect prenant le pas sur l'autre pendant un temps, puis éventuellement l'inverse.

Il s'agit de ce que l'on appelle la dualité onde-corpuscule, mais qui peut également se matérialiser comme une dualité discret-continu, ou encore analogique-numérique. *Stricto sensu*, ces dualités ne sont pas directement équivalentes, mais se font néanmoins écho les unes aux autres. La dualité onde-corpuscule désigne le fait que la matière, notamment à ses plus petites échelles, peut se décrire aussi bien sous forme d'éléments individualisables – on pensera donc aux particules, aux atomes, aux grains de réalité que l'on peut faire remonter jusque les auteurs latins de l'antiquité – que sous forme ondulatoire, c'est-à-dire, comme un processus continu et étendu dans l'espace. Le cas le plus connu étant le cas de la nature de la lumière, dont il a pu être dit alternativement qu'elle était corpusculaire – des grains de lumière – ou bien une onde se propageant dans l'espace. Einstein, Planck puis De Broglie ont finalement amené à reconnaître que la lumière pouvait être les deux, selon le point de vue descriptif : les interférences et les figures de diffraction s'expliquent par des mécanismes ondulatoires, mais l'effet photo-électrique est également une bonne représentation de la réalité, et implique l'émission de photons comme éléments de base de la lumière – la particule lumineuse.

Quand on en arrive à parler de formes, ce type de dualité apparaît rapidement : instinctivement, une forme est vue comme un objet, en tant qu'abstraction (le poisson, l'idée du poisson, la classe des poissons : un pois-

son est un animal satisfaisant un certain nombre de critères qui le séparent du mammifère, par exemple) ou en tant qu'instance, itération unique dans sa réalisation (un poisson spécifique, né à une date et un lieu donnés, avec une histoire particulière). Nous sommes naturellement enclin à traiter une forme comme une catégorie, et en soi, comme un phénomène séparable et séparé d'autres, identifié par disjonction voire par essence. Pour autant, il est possible de voir de la continuité, voire du diffusif dans le concept de forme. De la continuité dans l'évolution, par exemple, même s'il l'on peut penser qu'il existe des sauts ou des ruptures historiques et évolutionnaires (mais sont-ce de réels sauts, ou bien des bifurcations, des variations topologiques brusques, mais continues ?), ou de la continuité dans l'organogenèse, le développement d'un organisme. De la continuité également entre deux instances, deux formes relevant du même idéal, mais étant des variations l'une de l'autres que l'on peut déduire par des processus continus.

Que dire en outre, de la conception que l'on a hérité des théories de la communication, en premier lieu, de celle de Shannon-Weaver ? L'un des théorèmes fondamentaux concerne précisément les conditions sous lesquelles il est possible de reconstruire sans perte un signal dit continu à partir d'un ensemble de ses points - d'échantillons ; on parle du théorème fondamental de l'échantillonnage. Si un signal possède un spectre de fréquences fini, alors il est possible d'échantillonner ce signal à une fréquence qui doit être au moins 2 fois aussi grande que la fréquence maximale de ce spectre. De fait, là encore, il devient équivalent, dans ce cadre théorique, de se donner une description continue du signal ou une description discrète adaptée.

Encore formulé autrement, les formes sont-elles des objets donnés *a priori*, et pour lesquels il peut exister des lois d'évolution les reliant, expliquant sous quelles conditions il est possible de sauter de l'une à l'autre, ou bien sont elles la manifestation de variations topologiques d'un tissu continu ? Ce type de considérations renvoie aussi inexorablement aux problèmes de classifica-

tion et de catégorisation : y a-t-il un sens à chercher à identifier des groupes dits homogènes dans des données ? Si oui, lequel ? Nous verrons plus loin que ce type de problème est particulièrement prégnant en santé mentale, où le modèle prédominant actuel est celui de la classification dure et étanche des pathologies, les seuls liens permettant de passer de l'une à l'autre étant donnés par la notion de facteur de risque. Pourtant, les approches continues ne manquent pas, voire holistiques, comme la Gestalt en psychologie. A l'interface, existe par ailleurs la question de la séparabilité effective des formes qui peuvent paraître distinctes, mais pour lesquelles on échoue régulièrement à trouver des règles simples ou compliquées qui permettent de dire à coup sûr : ce domaine de l'espace appartient à la forme A, celui-ci à la forme B. Quiconque se sera essayé dans sa vie à faire du traitement d'image, par exemple, saura à quel point la notion de forme pure et séparée ou séparable relève de l'utopie. Quiconque a également tenté de séparer deux effets moyens d'un médicament et d'un placebo sur une affection donnée aura été confronté à des problèmes de résolution, de séparabilité des gaussiennes : les formes, au mieux si elles sont idéalement séparables, sont le plus souvent enchevêtrées les unes dans les autres, soit par nature, soit par le point de vue adopté pour les observer, par exemple.

2.2 Éléments de théorie des systèmes dynamiques

– théorie des catastrophes

2.2.1 Principes généraux et hypothèses

Nous abordons ici la question des systèmes dynamiques, dont nous avons déjà commencé à évoquer le sujet dans les sections précédentes. Nous traitons de cette théorie avant toute autre, pour différentes raisons. Premièrement, elle donne un cadre général, ou disons même qu'elle rétablit une généralité qui avait échappé à l'études de systèmes physiques amputés de leur complexité

de base, très proche initialement de ce que l'on connaît en modélisation mathématique usuelle, à savoir une modélisation différentielle, variationnelle. Nous avons discuté précédemment des modèles de Turing, qui sont des modèles à équations différentielles : les systèmes dynamiques leur offrent un cadre parfaitement adapté.

Les systèmes dynamiques, dont nous donnons dès après la formulation mathématique de base, ont pour ambition de décrire tout système qui varie dans le temps, et dont les variables clés sont reliées entre elles par une fonction particulière, linéaire ou non. La théorie élargit ainsi le cadre relativement fermé des systèmes physiques qui répondaient des lois fondamentales de la mécanique, ou lois de Newton. Dans ces modèles, l'accélération du système est postulée directement proportionnelle aux forces qui s'y appliquent. Les forces en question étant de manière générale, les forces classiques, type gravité ou force électromagnétique – souvent, des forces dérivant d'un potentiel.

Evidemment, un des buts de la théorie est de permettre de décrire totalement les trajectoires temporelles du système étudié. Néanmoins, et de la sorte de ce que nous venons de discuter juste avant, il existe dans cette théorie une approche duale qui en fait sa richesse. La théorie des systèmes dynamiques va s'attacher à chercher les points fixes, s'ils existent, du système, et à en qualifier la stabilité. Ces points fixes, ou les attracteurs du système, peuvent être vus comme des formes en ce sens où elles sont des propriétés invariantes du système, privilégiées, et observables : ce ne sont pas des états quelconques. Outre la recherche de ces points fixes, la théorie, *via* ce que l'on appelle la théorie de Thom ou des catastrophes, va s'attacher à examiner aussi la stabilité dite structurelle du système : on ne s'intéressera pas aux variables observées, mais aux évolutions possiblement drastiques de la forme des solutions selon les valeurs des paramètres structurant le système (par exemple : un coefficient de proportionnalité d'une variable).

La présentation des éléments théoriques qui suivent est essentiellement

empruntée à l'ouvrage plutôt complet sur le sujet des bifurcations et des systèmes dynamiques. [34]

2.2.2 Formulations et définitions

Nous commençons par définir formellement ce qu'est un système dynamique. Un système dynamique est la donnée de :

$$\frac{dx}{dt} = \dot{x} = f(x, t, \nu) \quad (2.1)$$

où l'on a

$$x \in U \subseteq \mathcal{R}^n, t \in \mathcal{R}, \nu \in V \subseteq \mathcal{R}^p, f \in C^r(U)$$

avec U l'espace dit des phases, V l'espace des paramètres.

Orbite d'un système dynamique

Une solution de l'équation (2.1) est appelée une orbite. C'est une courbe dans \mathcal{R}^n . On la notera :

$$x_0 = x(t_0)$$

$$x(x_0, t)$$

Il est possible que f ne dépende pas explicitement du temps t ; dans ce cas, le système dynamique est dit autonome. Par suite, l'orbite $x(x_0, t)$ ne dépend pas de t_0 .

Flot d'une orbite

Pour une orbite $x(x_0, t)$, il est possible de définir ce qu'on appelle le flot de x . On note ϕ_t le flot de x , et est défini tel que :

$$\begin{aligned}\phi_t : \mathcal{D} &\longrightarrow \mathcal{R}^n \\ x_0 &\longmapsto x(x_0, t)\end{aligned}\tag{2.2}$$

où l'on a $x_0 = x(0) \in \mathcal{D}$, et $\phi_t(x_0) /$

$$\begin{aligned}\phi_t(x_0) &\in \mathcal{C}^r \\ \phi_0(x_0) &= x_0 \\ \phi_{t+s}(x_0) &= \phi_t[\phi_s(x_0)]\end{aligned}\tag{2.3}$$

2.2.3 Points fixes et notion de stabilité

Point fixe d'un système dynamique

Étant donné un système dynamique, on s'intéresse certes à ses trajectoires, ses solutions donc – ses orbites – mais également à des aspects plus structurels. En effet, quelle que soit la forme de la solution, elle est, sauf chaos parfait, généralement en grande partie très régulière. En particulier, une telle solution peut avoir des points dits points fixes, ou stationnaires, ou encore critiques. Ces 3 termes sont équivalents. Les points fixes d'un système dynamique sont les points pour lesquels le système n'a plus d'accroissement, ne présente plus de variations autour d'un point, dans le temps. On définit donc un point fixe \bar{x} de f par :

$$f(\bar{x}) = 0\tag{2.4}$$

Dans le cas où un système admet un ou plusieurs points fixes \bar{x}_i , la question qui se pose alors est de déterminer leur éventuelle stabilité. En effet, un système peut très bien passer à un moment donné par un point où son accroissement entre t et $t + dt$ s'annule ; ceci ne signifie cependant pas qu'il soit destiné à y rester indéfiniment, si on lui communique par exemple une force extérieure, aussi petite soit-elle.

Stabilité d'un point fixe

Nous définissons ici deux types de stabilités pour les points fixes : la stabilité (simple), et la notion de stabilité asymptotique – on dit d'un tel point qu'il est asymptotiquement stable. Un point fixe \bar{x} est stable si et seulement si :

$$\forall \varepsilon > 0, \exists \delta > 0 / |x(0) - \bar{x}| < \delta \Rightarrow |x(t) - \bar{x}| < \varepsilon \quad (2.5)$$

Est asymptotiquement stable un point fixe \bar{x} vérifiant quant à lui :

$$\exists \delta_0 / 0 < \delta_0 < \delta / |x(0) - \bar{x}| < \delta_0 \Rightarrow \lim_{t \rightarrow \infty} x(t) = \bar{x} \quad (2.6)$$

Système linéarisé autour d'un point fixe et Jacobien en un point

La définition d'un système dynamique est très générale, comme on a pu le voir, nonobstant une exigence de régularité sur f dans le cas continu (nous ne nous intéresserons pas ici aux systèmes dynamiques discrets). Cependant, en dehors d'applications *a priori* très pathologiques, par construction, comme une courbe fractale (qui se définit aussi comme n'admettant pas de tangente, quel que soit le point de sa courbe), ce sont des conditions souvent admissibles.

En revanche, cette grande généralité fait la richesse et la difficultés des systèmes dynamiques : un grand nombre d'entre eux, considérés intéressants, sont des systèmes non linéaires en les variables x . Leur étude se rapporte alors en général à des études locales, où l'hypothèse d'une linéarisation du système peut être acceptable. Dans le cas de l'étude de la stabilité des points fixes de f , on a recours à ce genre de technique : on étudie le système linéarisé autour du point fixe.

Concrètement, étant données les conditions de régularité exigées de f , on écrit le développement de Taylor généralisé de f au voisinage du point fixe \bar{x} . Pour simplifier cette première écriture, on se rapportera au cas d'un point fixe, éventuellement à un changement d'origine prêt, tel que $\bar{x} = 0$. Le développement de Taylor généralisé de f s'écrit alors :

$$f(x) = Df(0)x + \frac{1}{2!}D^2f(0)(x, x) + \frac{1}{3!}D^3f(0)(x, x, x) + \dots \quad (2.7)$$

avec

$$Df(x)x = \sum_i \frac{\partial f(x)}{\partial x_i} x_i \quad (2.8)$$

$$D^2f(x)(x, x) = \sum_{i,j} \frac{\partial^2 f(x)}{\partial x_i \partial x_j} x_i x_j \quad (2.9)$$

et ainsi de suite, dans les limites de différentiabilité de f . Le premier terme $\frac{\partial f(x)}{\partial x_i}$, que l'on peut récrire $\left(\frac{\partial f_i}{\partial x_j}\right)_i$, est la matrice Jacobienne de f . Elle intervient notamment lorsque l'on veut passer d'un système de coordonnées à un autre, pour une fonction suffisamment régulière.

Si l'on considère que x est petit, c'est-à-dire, proche de 0, ou dans le cas général, que $x - \bar{x}$ est petit, on approche $f(x)$ autour de x par le premier terme du développement de Taylor généralisé :

$$f(x) \approx Df(0)x \quad (2.10)$$

et f est donc linéarisé en la variable x autour du point fixe. En outre, si l'on considère que $Df(0)$ possède n valeurs propres dans \mathcal{C} , distinctes 2 à 2, on peut écrire la solution x comme :

$$x = \sum_i c_i a^{(i)} e^{\lambda_i t} \quad (2.11)$$

où e est l'exponentielle, $a^{(i)}$ le i^{me} vecteur propre de valeur propre associée c_i . On a alors les cas suivants :

- si $\forall \lambda_i, \Re(\lambda_i) < 0$, et alors \bar{x} est asymptotiquement stable ;
- si $\exists \lambda_i / \Re(\lambda_i) = 0$, et les autres λ_j sont telles $\Re(\lambda_j) < 0$, alors \bar{x} est un centre ou un point elliptique stable ;
- si $\exists \lambda_i / \Re(\lambda_i) > 0$, alors \bar{x} est instable ;
- si $Df(0)$ n'a pas de valeur propre nulle ou imaginaire pure, \bar{x} est un point hyperbolique ;
- si $\exists i, j / \Re(\lambda_i) < 0, \Re(\lambda_j) > 0$, alors \bar{x} est un point selle ;
- si $\forall \lambda_i, \Im(\lambda_i) = 0$ et toutes ont même signe, alors \bar{x} est un noeud. S'il est stable, il s'agit d'un puits, s'il est instable, il s'agit d'une source.

L'analyse des valeurs propres permet ainsi de classer et catégoriser le genre du ou des points fixes.

2.2.4 Bifurcations

Nous abordons ici rapidement une notion fondamentale en systèmes dynamiques, à savoir les bifurcations que peuvent présenter de tels systèmes.

Définition d'une bifurcation

On parle de bifurcation lorsqu'il apparaît un changement de type topologique de la trajectoire d'un système dynamique. Ce changement est sous l'influence direct de la variation de paramètres dont dépend la trajectoire. Autrement dit, la variation d'un ou de plusieurs paramètres apparaissant dans la formulation de la trajectoire d'un système dynamique peut entraîner une variation topologique de cette trajectoire.

Plus formellement, on parle de bifurcation lorsque les portraits de phases cessent d'être homéomorphes, pour des valeurs des paramètres différentes.

Supposons le système dynamique paramétré suivant :

$$\frac{dx}{dt} = f(x, \nu)$$

où $x \in X \subset \mathbb{R}^n$ et $\nu \in D \subset \mathbb{R}^p$. Un point $\nu \in D$ est dit structurellement stable s'il existe un voisinage W de ν , tel que pour tout $\nu' \in W$, il existe un homéomorphisme (une application continue et inversible) qui applique les trajectoires du système :

$$\frac{dx}{dt} = f(x, \nu')$$

sur celles du précédent, et préservant le sens du temps. Ces deux systèmes sont topologiquement équivalents. Dans le cas contraire, ν est un point de bifurcation. Il est possible de rechercher l'ensemble des points de bifurcation, caractérisé par la codimension de la bifurcation.

Codimension associée à une bifurcation

Supposons que l'ensemble des points de bifurcation soit défini par k conditions de la forme :

$$b_1(\nu) = b_2(\nu) = \dots = b_k(\nu) = 0$$

avec k tel que $1 \leq k \leq P$. On dit alors que la bifurcation est de codimension k . L'ensemble des points de bifurcation est une variété V_{p-k} de dimension $p-k$, appelée ensemble de bifurcation. Il est possible de définir ce que l'on nomme le déploiement universel du système, présentant la forme :

$$\frac{dx}{dt} = f(x, \mu, a)$$

où l'on a $\mu \in R^k$, $a \in R^{p-k}$ et $\mu = 0$

Notion de catastrophe et d'ensemble catastrophe

Prenons alors un point fixe du même système, $x_0(\nu_0)$, autrement dit, vérifiant :

$$f(x, \nu) = 0$$

Si l'on a la condition suivante vérifiée :

$$\det \left| \frac{\partial f_i}{\partial x_j} \right| (x_0, \nu_0) \neq 0$$

alors il existe une solution unique $x = x(\nu)$ dans le voisinage de x_0 , de classe C^r et vérifiant $x = x(\nu_0) = x_0$. Le point $\nu = \nu_0$ est structurellement stable.

A contrario, si l'on a la condition :

$$\det \left| \frac{\partial f_i}{\partial x_j} \right| (x_0, \nu_0) = 0$$

alors le point ν_0 est appelé point catastrophe. L'ensemble $K \subset R^p$ des points catastrophes ν_0 est appelé ensemble catastrophe. Un ensemble catastrophe est ainsi un ensemble de bifurcation.

2.2.5 Théorie de Thom et ensemble catastrophe pour un champ gradient

Nous finissons cette partie théorique des systèmes dynamiques par les éléments clés de la théorie de R. Thom, renommée théories des catastrophes.

Thom a étudié les ensembles catastrophes d'un certain type de systèmes dynamiques, à savoir les systèmes à gradient. Le système

$$\frac{dx}{dt} = f(x, \nu)$$

se reformule, dans le cas d'un système à gradient, de la façon suivante :

$$f(x, \mu) = -\nabla V(x, \nu)$$

Autrement dit, on suppose l'existence d'un potentiel $V(x, \nu)$ dont dériverait la fonction $f(x, \nu)$. De la sorte, l'ensemble catastrophe devient caractérisé par les conditions suivantes :

$$\frac{dx}{dt} = -\nabla V(x, \nu)$$

$$\nabla V(x_0, \nu_0) = 0$$

$$\det \left| \frac{\partial^2 V(x_0, \mu_0)}{\partial x_i \partial x_j} \right| = 0$$

En théorie des catastrophes, si le déterminant $\det \left| \frac{\partial^2 V(x_0, \nu_0)}{\partial x_i \partial x_j} \right|$ est de rang $n - k$, on parle de catastrophe de corang k . De même, on parle plutôt de singularité, en lieu et place de point fixe.

Un résultat important de la théorie des catastrophes est un théorème dû à Thom, affirmant que pour un système gradient, de codimension inférieure ou égale à 4, le corang d'une catastrophe est alors au plus de 2. Thom a ainsi classé et donné la forme des 4 catastrophes de corang 1 et des 3 catastrophes de corang 2 pour les systèmes gradient.

Ce résultat est un, sinon le plus connu des résultats de la théorie de Thom. Il a également été largement critiqué, et continue de l'être pour ses restrictions, en particulier, le fait que le théorème ne s'applique qu'aux systèmes gradient. L'hypothèse d'un potentiel dont dériverait le système dyna-

mique est certes contraignante, mais toutes proportions gardées et à l'aune des autres théories physiques, cette hypothèse est très fréquente et a permis d'établir de nombreux résultats utiles. En outre, il est également possible qu'il n'existe pas de résultats analogues ou directement établissables dans le cas général, i.e. sans gradient. Les limites doivent donc être établies de la validité du théorème, et il incombe davantage à ceux qui l'invoquent d'en respecter les conditions d'applications, que de fustiger son auteur.

2.2.6 Un exemple : modèle de Lorenz

Nous profitons de cette section pour introduire un modèle archétypique de système dynamique, dû à Lorenz et introduit en 1963 – il s'agit du modèle météorologique de Lorenz. [51]

Système dynamique différentiel de Lorenz

Le système différentiel de Lorenz s'écrit :

$$\begin{cases} \frac{dx}{dt} = \sigma(y(t) - x(t)) \\ \frac{dy}{dt} = \rho x(t) - y(t) - x(t)z(t) \\ \frac{dz}{dt} = x(t)y(t) - \beta z(t) \end{cases}$$

Dans ces équations, σ , ρ et β sont trois paramètres réels strictement positifs. σ est le nombre dit de Prandtl et ρ est le nombre de Rayleigh.

Les variables x , y et z représentent l'état du système à un instant t . Le système de Lorenz est issu d'une volonté de simplification drastique des équations de mécanique des fluides initiales, celles-ci étant trop complexes à traiter pour les calculateurs de l'époque. x est proportionnel au mouvement de convection, y est proportionnel à la différence de température entre les courants ascendants et descendants, et z est proportionnel à l'écart du profil de température vertical par rapport à un profil linéaire.

On pose souvent $\sigma = 10$, $\beta = 8/3$; ρ restant variable. On sait depuis Lorenz que le système présente un comportement chaotique pour $\rho = 28$.

Points d'équilibre Les points d'équilibre, ou points fixes, du système sont les solutions (x, y, z) constantes du système différentiel, c'est-à-dire : $(\frac{dx}{dt}, \frac{dy}{dt}, \frac{dz}{dt}) = (0, 0, 0)$. Il en existe au plus trois :

- le point fixe trivial $(0, 0, 0)$, vrai pour toutes valeurs de σ, ρ et β .
- les deux points fixes symétriques : $(-\sqrt{\beta(\rho - 1)}, -\sqrt{\beta(\rho - 1)}, \rho - 1)$
et : $(\sqrt{\beta(\rho - 1)}, \sqrt{\beta(\rho - 1)}, \rho - 1)$, qui n'existent que lorsque $\rho > 1$.

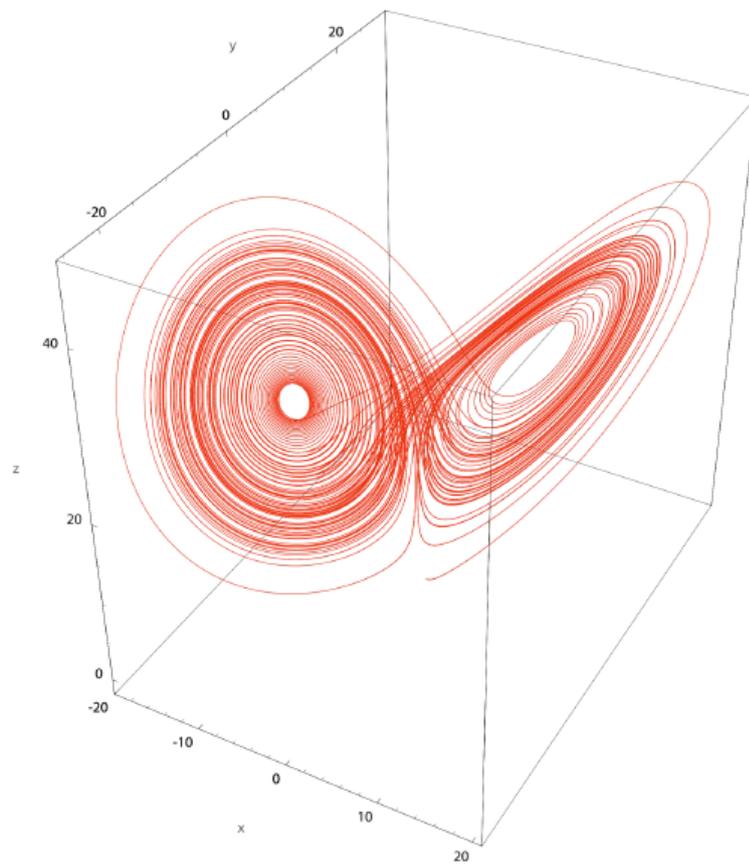
Comme annoncé, le modèle de Lorenz présente pour des valeurs particulières de ses paramètres, un attracteur particulier, nommé attracteur étrange. C'est le fameux attracteur en forme "d'ailes de papillon", que nous reproduisons en figure 2.15.

2.2.7 Extension : théorie KAM, et formulation en espace Riemannien des ODE

Il existe des résultats qui étendent la théorie des systèmes dynamiques, ou qui à tout le moins la complètent. Nous ne faisons qu'en citer l'existence, comme le théorème KAM, pour Kolmogorov-Arnold-Moser. [34, 52, 53] Plus proche de ce qui nous intéressera dans nos travaux, le prolongement du formalisme général des équations différentielles ordinaires (et leur pendant, les équations différentielles partielles) par un espace géométrique plus large, de type Riemannien.

Nous montrons ici, *via* un exemple simple de système différentiel, comment il est possible d'étendre le formalisme usuel, et d'interpréter les résultats comme des objets géométriques.

FIGURE 2.15 – Une représentation de l'attracteur étrange de Lorenz (Wikipedia, licence CC)



Géométrie de Riemann-Lagrange des espaces de jets et application en cancérogenèse

Il a été proposé récemment une construction géométrique appelée géométrie de Riemann-Lagrange des espaces de jets du premier ordre, comme extension aux espaces de jets de premier ordre de la géométrie lagrangienne sur les fibrés tangents. Ce type de construction permet d'interpréter les solutions des équations différentielles ordinaires – ODE en anglais – comme des géodésiques de l'espace construit. [54, 55] Nous rapportons ici, pour un exemple simple de modèle de prolifération de cellules cancéreuses, comment un système différentiel ordinaire peut s'écrire dans un tel espace. Ces développements et exemple sont repris de [54].

Géométrie de Riemann Lagrange d'un modèle populationnel de cellules cancéreuses

Le modèle rapporté ici est celui développé par Garner en 2006, [56] et est constitué d'un système différentiel ordinaire à 4 paramètres. Une telle population repose sur une combinaison de cellules proliférantes, quiescentes ou apoptotiques ; leurs proportions relatives déterminent la croissance tumorale. Il a notamment été proposé que plusieurs formes de cancers dérivent initialement d'une seule cellule anormale ou d'une sous population très restreinte de telles cellules. Ces cellules sont appelées des cellules souches cancéreuses.

Les hypothèses sur lesquelles reposent le modèle sont les suivantes :

1. les cellules cancéreuses sont soit proliférantes soit quiescentes ;
2. les cellules peuvent perdre leur capacité de mitose, et passent de proliférantes à quiescentes ;
3. les cellules quiescentes peuvent soit devenir proliférantes soit entrer en apoptose.

Le système présente deux états, à savoir P , le nombre de cellules proliférantes et Q le nombre de cellules quiescentes. Leur évolution est donnée par le système suivant :

$$\begin{cases} \frac{dP}{dt} = P - P(P + Q) + F(P, Q), \\ \frac{dQ}{dt} = -rQ + aP(P + Q) - F(P, Q), \end{cases} \quad (2.12)$$

$$F(P, Q) = \frac{hPQ}{1 + kP^2}, \quad r = \frac{d}{b}, \quad h = \frac{A}{ac}, \quad k = \frac{Bb^2}{c^2}, \quad (2.13)$$

où

- a est une constante sans dimension, mesurant la prise nutritive des cellules quiescentes ou proliférantes ;
- b est le taux de mitose des cellules proliférantes ;
- c dépend de la consommation des cellules proliférantes, et est le taux de cellules passant de proliférantes à quiescentes, en cellules par jour ;
- d est le taux d'apoptose parmi les cellules quiescentes, par jour ;
- A taux d'accroissement initial du nombre de cellules passant de l'état quiescent à l'état proliférant lorsque P est petit ;
- A/B taux de décroissance du nombre de cellules passant du nombre de cellules passant de l'état quiescent à l'état proliférant lorsque P devient grand.

Le flot de la population cellulaire cancéreuse, selon une géométrie de Riemann-Lagrange, sur l'espace de 1-jets $J^1(T, R^2)$ est alors donné par les éléments suivants :

(i) La connexion canonique non linéaire sur l'espace $J^1(T, R^2)$ induite par le flot présente les composantes locales :

$$\hat{\Gamma} = \left(0, \hat{N}_{(1)j}^{(i)}\right), \quad i, j = \overline{1, 2}$$

où, si

$$F_P = \frac{hQ(1-kP^2)}{(1+kP^2)^2} \quad \text{et} \quad F_Q = \frac{hP}{1+kP^2}$$

sont les dérivées partielles d'ordre 1 de la fonction F , alors on a

$$\begin{aligned}
\hat{N}_{(1)1}^{(1)} &= \hat{N}_{(1)2}^{(2)} = 0, \\
\hat{N}_{(1)2}^{(1)} &= -\hat{N}_{(1)1}^{(2)} = \frac{1}{2} [(2a+1)P + aQ - (F_P + F_Q)] = \\
&= \frac{1}{2} \left[(2a+1)P + aQ - \frac{hQ(1-kP^2)}{(1+kP^2)^2} - \frac{hP}{1+kP^2} \right].
\end{aligned} \tag{2.14}$$

(ii) toutes les composantes de la connexion canonique de Cartan généralisée $C\hat{\Gamma}$ induite par le flot s'annulent ;

(iii) Toutes les composantes du d-tenseur de torsion $\hat{\mathbf{T}}$ de la connexion canonique de Cartan généralisée $C\hat{\Gamma}$ induite par le flot sont nulles, sauf

$$\begin{aligned}
\hat{R}_{(1)21}^{(1)} &= -\hat{R}_{(1)11}^{(2)} = a + \frac{1}{2} (1 - F_{PP} - F_{PQ}), \\
\hat{R}_{(1)22}^{(1)} &= -\hat{R}_{(1)12}^{(2)} = \frac{1}{2} (a - F_{PQ} - F_{QQ}) = \frac{1}{2} (a - F_{PQ}),
\end{aligned} \tag{2.15}$$

où

$$F_{PP} = -\frac{2hkPQ(3-kP^2)}{(1+kP^2)^3}, F_{PQ} = \frac{h(1-kP^2)}{(1+kP^2)^2} \text{ and } F_{QQ} = 0$$

sont les dérivées partielles de second ordre de la fonction F .

(iv) Toutes les composantes du d-tenseur de courbure $\hat{\mathbf{R}}$ de la connexion canonique de Cartan généralisée $C\hat{\Gamma}$ induite par le flot s'annulent ;

(v) La 2-forme distinguée induite par le flot a l'expression suivante :

$$\hat{\mathbf{F}} = \hat{F}_{(i)j}^{(1)} \delta x_1^i \wedge dx^j, \tag{2.16}$$

où

$$\delta x_1^i = dx_1^i + \hat{N}_{(1)k}^{(i)} dx^k, \forall i = \overline{1,2},$$

et ses composantes $\hat{F}_{(i)j}^{(1)}$, $i, j = \overline{1,2}$, sont données par

$$\begin{aligned}
\hat{F}_{(1)1}^{(1)} &= \hat{F}_{(2)2}^{(1)} = 0, \\
\hat{F}_{(2)1}^{(1)} &= -\hat{F}_{(1)2}^{(1)} = \frac{1}{2} [(2a + 1)P + aQ - (F_P + F_Q)] = \\
&= \frac{1}{2} \left[(2a + 1)P + aQ - \frac{hQ(1-kP^2)}{(1+kP^2)^2} - \frac{hP}{1+kP^2} \right].
\end{aligned} \tag{2.17}$$

(vi) L'énergie géométrique de type Yang-Mills induite par le flot est donnée par

$$EYM(P, Q) = \frac{1}{4} \left[(2a + 1)P + aQ - \frac{hQ(1-kP^2)}{(1+kP^2)^2} - \frac{hP}{1+kP^2} \right]^2 \tag{2.18}$$

Bien entendu, nous ne précisons pas ici un certain nombre d'objets mathématiques, comme la connexion, la torsion ou l'énergie de Yang-Mills – qui se rapportent par analogie à la théorie physique éponyme. Ce sont ces objets qui permettent d'introduire de la géométrie dans la description du système différentiel initial.

2.3 Place des deux grandes théories physiques actuelles : théories quantiques et relativistes

2.3.1 Principes généraux en théories quantiques : états, opérateurs, espaces propres et opérateurs d'évolution

Les éléments théoriques que nous rapportons, et discutons, dans les sections qui suivent peuvent être retrouvés dans divers ouvrages, dont en particulier. [57, 58, 59, 60, 61]

Quantification et fréquences

Le coeur de la mécanique quantique, s'il n'en existe qu'un, tient dans l'hypothèse historique de Planck quant à la quantification de l'énergie des atomes. Selon lui, cette énergie ne peut qu'être un multiple d'une grandeur directement proportionnelle à la fréquence de leur rayonnement. La constante de proportionnalité est ainsi devenue une des constantes fondamentales de la physique - il s'agit de la constante de Planck h . L'énergie devient alors :

$$E = h\nu$$

avec ν la fréquence de rayonnement. De son côté, De Broglie proposera une version plus générale de ce type de relation, en supposant que la quantité de mouvement P est elle-même directement proportionnelle au vecteur d'onde \vec{k} , la constante de proportionnalité étant là encore la constante de Planck :

$$\vec{P} = h\vec{k}$$

C'est alors le renouveau du principe de dualité onde-corpuscule ; en effet, l'inspiration de De Broglie lui serait venue après que Einstein aurait postulé que l'énergie minimale de Planck $h\nu$ est l'énergie du photon. Une

description ondulatoire, à l'inverse de ce mouvement de quantification, sera alors recherchée, donnant entre autres naissance à l'équation de Schrödinger (initialement, construite pour l'atome d'hydrogène).

La mécanique quantique s'appuie sur un certain nombre de principes, que nous ne passerons pas ici en revue de manière exhaustive. Parmi ceux qui nous intéressent pour une possible description de la réalité en tant que formes, mentionnons le principe d'état quantiques et le principe de superposition.

États et superpositions

Un système quantique est décrit par des états isolables et distincts, caractéristiques de ce système. Ces états sont observables, et reproductibles pour différents observateurs. Pour un système simple, par exemple mono particule comme l'électron, il lui a été découvert une dimension spécifique, à savoir le spin (expérience de Stern et Gerlach, 1922). Le spin de l'électron ne présente que deux états possibles (+ ou $-\frac{1}{2}$). Ces états correspondent donc aux formes fondamentales de ce système particulier qu'est l'électron.

De manière plus générale, un système physique quelconque peut être par exemple caractérisé par un état fondamental $|\phi\rangle$ ou un état $|\psi\rangle$. Ces deux états sont des états séparés, distincts et exclusifs, décorrélés. Néanmoins, est-ce à dire que le système étudié est nécessairement dans l'état $|\phi\rangle$ ou l'état $|\psi\rangle$ à tout instant, et ce de manière certaine ? Autrement dit, sans que l'on fasse intervenir d'observation permettant de lever l'incertitude sur l'état actuel du système, quel est l'état le plus général de celui-ci ? Le principe de superposition énonce que cet état quelconque est un état linéairement composé, à savoir :

$$\alpha|\phi\rangle + \beta|\psi\rangle$$

Où α et β sont deux nombres complexes quelconques. En termes mathématiques, cela revient à dire que l'ensemble des états possibles d'un système physique quelconque est décrit par un espace vectoriel, et que n'importe quel état du système est un vecteur de cet espace. On peut notamment considérer que les deux états fondamentaux $|\phi\rangle$ et $|\psi\rangle$ en sont une base privilégiée.

L'interprétation la plus répandue de cet état quelconque ne se résume pas à dire que l'observateur, tant qu'il n'a pas effectué de mesure de ce système, est ignorant de l'état actuel du système. Il est davantage admis que la superposition caractérise une indétermination intrinsèque du système à un instant quelconque ; c'est également, pour le cas des états "positions" (dans l'espace usuel), dire que la particule ou le système n'est pas localisé, mais bel et bien étendu, du moins "en probabilité".

Ce concept de non localisation "en probabilité" tient par ailleurs de ce qui est appelé la règle de Born. Plus précisément, la règle de Born porte sur l'interprétation des coefficients α et β - les coefficients de la combinaison linéaire d'états fondamentaux - ces coefficients étant directement liés à la probabilité pour le système de se trouver effectivement dans l'état $|\phi\rangle$ ou $|\psi\rangle$.

Notamment, la probabilité pour que la mesure effectuée sur notre système nous donne le résultat "le système est dans l'état $|\psi\rangle$ " est proportionnel à β^2 .

Plus généralement, pour un système dont le vecteur d'état est une combinaison linéaire d'états fondamentaux ou distincts $(|i\rangle)_{i \in \mathbf{N}}$, la probabilité pour que le résultat de la mesure soit celui correspondant au fait que le système soit dans l'état $|i\rangle$ est donnée par :

$$P_i = \frac{|\alpha_i|^2}{\sum_i |\alpha_i|^2},$$

où les α_i sont les coefficients linéaires du vecteur d'état. Cependant, si l'on ne prend pas garde à certaines subtilités, cette écriture peut nous donner

des probabilités différentes pour un même système.

Projections et probabilités

Il est possible de donner une définition plus rigoureuse et mieux bâtie de la règle de Born. En particulier, nous n'avons pas scrupuleusement imposé aux états $|\phi\rangle$ et $|\psi\rangle$ d'être indépendants l'un de l'autre au sens d'un produit scalaire dont on munirait l'espace vectoriel décrivant les états du système. Il convient ainsi de se donner un tel produit scalaire, et de se donner une base d'états fondamentaux qui soient orthogonaux les uns aux autres (i.e, dont les produits scalaires sont deux à deux nuls, si les états sont différents).

Ainsi, pour toute famille de vecteurs orthogonaux $(|i\rangle)_{i \in \mathbf{N}}$, il existe au moins une famille de vecteurs $(|i'\rangle)_{i \in \mathbf{N}}$ et de scalaires $(\lambda_i)_{i \in \mathbf{N}}$ tels que $|i\rangle = \lambda_i |i'\rangle$ et tels que pour toute combinaison linéaire des $(|i'\rangle)_{i' \in \mathbf{N}}$, la probabilité pour que le résultat de mesure soit le même que si le système avait été dans l'état $|i'\rangle$ est :

$$\frac{|\alpha_i|^2}{\sum_i |\alpha_i|^2},$$

où les α_i sont les coefficients linéaires du vecteur d'état dans la base $(|i'\rangle)_{i \in \mathbf{N}}$. Il est bien entendu possible de normaliser les vecteurs d'états de telle sorte que le dénominateur soit égal à un. La règle de Born devient alors :

$$P_i = |\alpha_i|^2,$$

ou encore :

$$P_i \propto |\alpha_i|^2,$$

où le coefficient de proportionnalité se calcule en tenant compte de la relation de normalisation portant sur les P_i : $\sum_i P_i = 1$, puisque l'on désire une probabilité totale de 1. On voit que le calcul de la probabilité associée

à un état fait intervenir la notion de projection du système sur cet état, à supposer que la base d'états est bien orthogonale : le produit scalaire de l'état global du système par l'état fondamental dont on veut mesurer la probabilité nous donnera bel et bien α_i^2 .

À ce stade, il devient important de comprendre ce qu'il peut se passer lorsque l'on désire réaliser plusieurs mesures sur un même système, notamment des mesures portant sur des états ou des propriétés de types différents. Il est possible de mener une première expérience, laquelle nous donnera une mesure concernant l'état du système ; cette mesure nous donne α , et l'on peut considérer que le système est dans l'état $|\alpha\rangle$. Il est possible que cet état soit stable, n'évolue pas dans le temps. Les mesures réitérées avec le même dispositif donneront alors toujours le même résultat. Cependant, si l'on désire effectuer aussi une mesure concernant une autre grandeur de ce système, il n'est pas certain que l'on puisse déterminer avec certitude l'état du système selon cette autre grandeur. L'exemple classique concerne les mesures de positions et d'impulsions (disons, pour simplifier ici, de vitesse) d'un système quantique. Nous savons, classiquement, que la vitesse dérive de la position, selon le temps. Ainsi, la connaissance de la position, avec certitude, devrait, pourrait-on penser, entraîner la connaissance aussi certaine de la vitesse.

Or, il n'en est rien : plus la mesure portant sur la position est certaine, plus grande sera l'incertitude portant sur la vitesse, et inversement. Techniquement, cela vient précisément du fait que ces deux grandeurs ne sont pas indépendantes, mais conjuguées (l'une est liée à l'autre par une transformation, en l'occurrence une transformation de Fourier), et que les amplitudes de probabilités (les coefficients de linéarité) soient des ondes. Le résultat en est que l'observation conjointe de ces deux grandeurs conjuguées ne peuvent se faire avec une certitude parfaite pour les deux grandeurs. L'incertitude irréductible liant ces grandeurs est donnée par l'inégalité d'Heisenberg – ou relation d'incertitude. Cette relation est la même, formellement, que pour

toutes quantités liées par une transformation type Fourier ; c'est ainsi le cas, en théorie du signal, pour le temps et la fréquence.

Opérateurs et observables

De manière plus générale et formelle, prenons deux processus de mesure, soient A, et B. Pour A, on se donne une base d'états fondamentaux $(|\alpha_i\rangle)_{i \in \mathbf{N}}$. Le principe de superposition nous donne alors la forme des états quelconques de ce système :

$$|\phi\rangle = \sum_i \phi_i |\alpha_i\rangle$$

Si l'on prend maintenant le processus B, rien ne nous permet de dire, à ce niveau, qu'il n'existe pas dans la base des états fondamentaux de A ainsi que dans l'ensemble de leurs combinaisons linéaires, des états qui soient fondamentaux pour B. Dans un tel cas, comment peut-on déterminer le résultat de mesure dû au processus A correspondant à un état fondamental de B ?

L'approche probabiliste de Born suggère que le résultat de la mesure doit être donné par l'espérance (la moyenne) mathématique des α_i , autrement dit, on aurait le résultat de mesure α tel que :

$$\alpha = \sum_i P_i \alpha_i = \sum_i |\phi_i|^2 \alpha_i$$

Si l'on se place dans le sous-espace vectoriel engendré par les états $|\alpha_i\rangle$, on peut comme dit précédemment se munir d'un produit scalaire pour lequel la base des $(|\alpha_i\rangle)_{i \in \mathbf{N}}$ est orthonormée. En théorie quantique, le produit scalaire est manipulé un peu différemment des cas classiques rencontrés en mathématiques et en algèbre. On le décompose en deux parties, dites bra et ket (pour *bra-ket*, crochets, en anglais, les crochets $\langle | \rangle$ étant une manière fréquente de noter le produit scalaire). Les bras, notés $\langle \beta|$, sont les éléments de l'espace dual de l'espace des kets, ici les $|\alpha_i\rangle$. La relation de dualité donne :

$$\langle \alpha_i | \alpha_j \rangle = \delta_{ij},$$

avec δ_{ij} le symbole de Kronecker, valant 1 si et seulement si $i = j$ et 0 sinon. Ceci permet de récrire l'espérance mathématique :

$$\begin{aligned} \alpha &= \sum_i \phi_i^* \phi_i \alpha_i = \sum_{i,j} \phi_i^* \phi_j \alpha_j \delta_{ij} = \sum_{i,j} \phi_i^* \phi_j \alpha_j \langle \alpha_i | \alpha_j \rangle = \sum_{i,j} \langle \alpha_i | \phi_i^* \phi_j \alpha_j | \alpha_j \rangle \\ &= \sum_j \langle \phi | \phi_j \alpha_j | \alpha_j \rangle = \langle \phi | \sum_j \phi_j \alpha_j | \alpha_j \rangle \end{aligned}$$

On voit apparaître un opérateur linéaire : celui dont les vecteurs propres sont les $(|\alpha_i\rangle)_{i \in \mathbf{N}}$ et dont les valeurs propres associées sont les α_i . Cet opérateur \mathbf{A} est appelé observable associé au processus de mesure A . L'observable permet de récrire plus simplement la relation donnant l'espérance :

$$\alpha = \langle \phi | \mathbf{A} | \phi \rangle$$

De la sorte, nous arrivons finalement à une expression qui d'une part, ne fait plus apparaître une dépendance explicite en une base $(|\alpha_i\rangle)_{i \in \mathbf{N}}$ spécifique, et d'autre part, fait intervenir la notion d'observable, notion centrale de la mécanique quantique. En effet, on considérera que à toute mesure est associée un tel opérateur, lequel, appliqué à un état quelconque du système, donne l'espérance du résultat de mesure. Par ailleurs, les vecteurs propres de cet opérateur donnent une base orthogonale d'états fondamentaux du système par rapport à cette mesure.

De la même façon, un telle observable \mathbf{A} peut s'écrire en fonction de ses vecteurs propres et valeurs propres ; c'est en outre un opérateur auto-adjoint :

$$\mathbf{A} = \sum_i \alpha_i |\alpha_i\rangle \langle \alpha_i|$$

Une telle formulation nous ramène aux considérations précédemment énoncées, à savoir qu'une opération de mesure sur un système, étant donné un type de mesure, s'apparente à une projection sur une base particulière. Le résultat de cette projection donne le résultat de mesure, et a une traduction

probabiliste.

Enfin, comme nous l'avons sous-entendu, il est possible de se donner plusieurs observables pour un même système, si l'on désire capturer de ce système différents types d'informations. Dans le cas où l'on dispose de suffisamment d'observables pour décrire tout résultat de mesure, on parle d'ensemble complet d'observables qui commutent, et c'est dans l'espace hermitien généré par les vecteurs propres de ces observables que l'on travaille.

Équation de Schrödinger et Hamiltonien

Nous venons de voir les principes structurant l'espace dans lequel l'on travaille en théorie quantique. Cet espace étant donné, nous n'avons pas à ce stade de principe dynamique permettant de rendre compte de l'évolution d'un système donné. L'équation réglant l'évolution d'un système quantique est donnée par l'équation de Schrödinger :

$$H|\psi(t)\rangle = i\hbar \frac{\partial |\psi(t)\rangle}{\partial t} \quad (2.19)$$

donnée ici sous sa forme dépendante du temps. H est l'Hamiltonien du système considéré. La forme indépendante du temps lie l'Hamiltonien aux niveaux d'énergie du système. Elle s'écrit :

$$H|\varphi_n\rangle = E_n|\varphi_n\rangle \quad (2.20)$$

L'Hamiltonien est constitué de l'énergie cinétique et de l'énergie potentielle du système – autant dire, de l'énergie liée au mouvement relatif du système et l'énergie liée à ce que le système, dans sa "position" ou son état actuel, peut accéder selon son évolution. Il existe ainsi une notion de transfert d'un type d'énergie en un autre, dont la totalité se conserve (dans le cas d'un système qui ne subit pas de "fuite" d'énergie de quelque façon que ce soit). Rappelons-nous que le potentiel constitue la "source" d'une force si

celui-ci n'est pas uniforme en espace.

Par ailleurs, rappelons que la notion de *moindre action*, à la Maupertuis (voir chapitre 5), qui consiste à concevoir tout mouvement comme le choix, parmi l'infinité de possibilités se présentant *a priori* à la nature, de l'action consommant le "moins d'énergie", ou plutôt, le "plus économe". Ceci révèle la nature d'optimalité sous-tendant nombre de raisonnements en physique. Formellement, ce principe de moindre action s'applique à une quantité appelée Lagrangien, mais qui est en réalité équivalente au Hamiltonien (sous certaines conditions) : le principe appliqué au Lagrangien fait apparaître les équations du mouvement connues par ailleurs.

L'Hamiltonien s'écrit donc de la sorte :

$$H = K + V \quad (2.21)$$

avec

$$K = \frac{P^2}{2m} \quad (2.22)$$

où K est l'énergie cinétique, V l'énergie potentielle, P est l'*impulsion* et m la masse. Sous cette forme, 2 particularités qui nous semblent fondamentales, apparaissent :

- H , décrit en termes de ressources les possibilités du système, avec un transfert entre mouvement et potentialité,
- \hbar , déterminant à la fois le quantum, la division élémentaire d'énergie, et l'incertitude fondamentale.

Par ailleurs, le \hbar définit la longueur dite longueur de *Planck* l_p , vue comme la longueur d'espace la plus petite, la maille la plus fine possible de l'espace-temps. Elle est également interprétable comme fréquence limite ou de cou-

pure selon des considérations de bande passante :

$$l_p = \sqrt{\frac{\hbar G}{c^3}} \quad (2.23)$$

où G est la constante universelle de gravitation, c la célérité de la lumière dans le vide.

De fait, si l'on étend le principe de la mécanique quantique à des systèmes autres que proches de l'infiniment petit (ou de très haute énergie, ce qui est équivalent), nous disposons d'une équation valable *a priori* à des *résolutions* variées : il suffit de ne pas se limiter à \hbar , et de relâcher ce paramètre pour en faire une variable d'échelle ou de résolution.

Enfin, une particularité, un principe central des théories quantiques est "caché", n'est pas directement visible à la vue de l'équation de Schrödinger, hors la présence du \hbar . Il s'agit des principes d'incertitude, ou communément appelés principes d'Heisenberg. On en tire 2 considérations importantes :

- Certaines observables ne peuvent être résolues simultanément, comme la mesure de la position et de la vitesse d'un mobile ; cette propriété se résume par la *non-commutativité* des observables,
- Il existe bel et bien une limite de résolution en-deçà de laquelle il n'est pas possible de descendre, pour un paramètre de résolution fixé.

Pour résumer, l'équation de Schrödinger et ses dérivations peuvent se synthétiser en 3 propriétés :

- L'Hamiltonien, ou transfert d'énergie sous 2 formes duales,
- Un paramètre de résolution,
- Une relation de non-commutativité.

Par ailleurs, nous avons affirmé qu'il était, entre autres, permis d'imaginer des équations de type Schrödinger qui seraient relatives à une échelle, et non limitée par l'échelle quantifiée de Planck. Il s'avère que Nottale, dans le cadre de la relativité d'échelle, [62] a proposé une telle équation, que nous retranscrivons ici :

$$\mathcal{D}^2 \Delta \psi + i\mathcal{D} \frac{\partial \psi}{\partial t} - \frac{\varphi}{2m} \psi = 0 \quad (2.24)$$

où \mathcal{D} est un paramètre d'échelle.

Les propriétés de résolution et de non-commutativité proviennent de la façon de décrire et de décomposer les signaux sur l'espace général. Cela tient au type de transformation qui lie les quantités ne commutant pas entre elles, dans le cas de la mécanique quantique, la transformée de Fourier. Il existe cependant d'autres façons décomposer ou de transformer les signaux d'un espace vers un autre (par exemple, la transformation en ondelettes).

Opérateur d'évolution

L'équation de Schrödinger est centrale aux théories quantiques. Elle fournit le principe d'évolution des systèmes quantiques. En dehors de cette équation, il est possible de définir un opérateur dit d'évolution, dont l'application à un état quelconque du système permet de calculer l'état à un instant t ultérieur.

Considérons un Hamiltonien \hat{H} composé de deux termes :

$$\hat{H} = \hat{H}_0 + \hat{V}(t)$$

où la dépendance temporelle est contenue dans $\hat{V}(t)$.

Quand $\hat{V}(t) = 0$, le système est complètement connu par ses kets $|n\rangle$ propres et ses valeurs propres E_n : $\hat{H}_0|n\rangle = E_n|n\rangle$

L'opérateur d'évolution est noté $U(t, t_0)$. Nous avons alors la relation suivante, liant l'état à t et l'état à t_0 :

$$|\psi(t)\rangle = U(t, t_0) |\psi(t_0)\rangle$$

où

- $|\psi(t)\rangle$ représente le ket au temps t
- $|\psi(t_0)\rangle$ représente le ket au temps t_0

Pour le bra, on a alors la relation suivante :

$$\langle\psi(t)| = \langle\psi(t_0)|U^\dagger(t, t_0)$$

U^\dagger représente l'opérateur adjoint de U . L'opérateur U a les propriétés suivantes :

- c'est un opérateur linéaire
- $U(t_0, t_0) = 1$
- il est transitif en la variable temporelle : $U(t_2, t_1)U(t_1, t_0) = U(t_2, t_0)$
- c'est un opérateur unitaire ($U^\dagger U = U U^\dagger = 1$).

Les trois premières propriétés sont des conséquences évidentes de l'équation d'évolution du premier ordre. La dernière propriété vient de ce que la probabilité totale doit être conservée par l'équation d'évolution.

Comme le système est donné par l'équation de Schrödinger, on a :

$$i\hbar \frac{\partial}{\partial t} |\psi(t)\rangle = \hat{H} |\psi(t)\rangle,$$

soit :

$$i\hbar \frac{\partial U(t, t_0)}{\partial t} = \hat{H} U(t, t_0)$$

Dans le cas d'un système quantique dont l'opérateur Hamiltonien \hat{H} est indépendant du temps, l'opérateur d'évolution s'écrit alors :

$$U(t, t_0) = e^{-\frac{i}{\hbar}\hat{H}(t-t_0)}$$

Ainsi, pour un système de type quantique, au sens où les états de ce système évoluent selon une équation de type Schrödinger, il est possible de construire un opérateur, basé sur l'Hamiltonien du système, qui permette de calculer un état à un instant t arbitraire étant donné un état à un temps t_0 . Il est également possible de construire un opérateur d'évolution dans le cas où l'Hamiltonien dépend du temps; nous n'en présenterons pas ici la construction.

Ce dernier résultat concernant la mécanique quantique est clé : en effet, il signifie que notre système est entièrement connu si l'on se donne d'une part ses états propres, et d'autre part son Hamiltonien. La forme et le principe d'évolution permettant de passer d'un état à un autre est par ailleurs accessible.

2.3.2 Principes généraux en théories relativistes : élément d'espace-temps, métrique et géométrie Riemannienne

Maintenant que nous avons vu les principaux aspects élémentaires - et qui nous intéressent ici - des théories quantiques, ainsi que des systèmes dynamiques, nous nous proposons de traiter les éléments clés des théories relativistes - ou plus généralement, géométrodynamiques. Les ouvrages traitant "simplement" des théories relativistes ne sont pas légion, qu'on ait des notions plus ou moins avancées de physique ou de mathématique. En outre, le formalisme d'Einstein est certes concis et autorise des manipulations abstraites relativement aisées, mais il est également difficile de trouver des références faisant un lien didactique entre formalisme mathématique usuel, notamment en géométrie différentielle, et formalisme physique. Enfin, la majorité des livres semblent préférer une présentation formelle, lourde, qu'une

approche plus sensible. Pour ce qui suit, nous avons pris le temps de recourir à différents ouvrages et références, et pris le parti de ne pas expliciter tous les intermédiaires techniques, lesquels obscurcissent plus le propos et les hypothèses de travail qu'ils ne rendent réellement service. Pour la relativité restreinte, consulter : [63, 4, 5, 64] ; pour la relativité générale, se référer à [4, 65, 66, 5, 64].

En théories relativistes, l'observateur a un autre statut qu'en mécanique quantique, où il s'agit finalement plus de l'observabilité qui y est traitée. L'observateur est représenté par un système de coordonnées, ou autrement dit, une règle pour mesurer les longueurs et une horloge pour mesurer le temps. Tous les observateurs ont les mêmes instruments, et ce qui distingue deux observateurs est leur déplacement relatif l'un par rapport à l'autre. Cet ensemble de caractéristiques définissent un référentiel R de l'observateur. Tout l'enjeu des théories relativistes tient dans son postulat d'invariance : les lois physiques doivent être les mêmes pour tout le monde, quel que soit l'observateur. Ce postulat induit différents résultats, dont les relations de passage entre deux référentiels qui seraient en mouvement relatif non nul. La relativité des mesures est une conséquence logique de l'invariance des lois physiques.

Les principaux éléments nous intéressant sont alors les relations de passage entre référentiels, que l'on appelle transformation(s) de Lorentz, mais également l'existence d'un élément d'espace-temps qui soit invariant par cette transformation ; nous verrons les implications de l'invariance sur la mesure des distances et du temps, qui peuvent être contractés ou dilatés ; la mesure du temps étant devenue relative, nous aborderons la notion de temps propre.

Ces premiers éléments concernent le ciment initial de la théorie relativiste. Nous aborderons en prolongement la notion d'intervalle d'espace-temps, et

en relation, la notion de causalité *via* le concept de cône de lumière. À partir de là, nous pourrions aborder la relativité générale, qui étend ces concepts à un cadre moins limité. Nous y lierons la notion d'intervalle d'espace-temps infinitésimal à celle de moindre action et de géodésique.

Transformation de Lorentz

Afin de comprendre ou de sentir l'idée d'un premier principe de relativité, nous commencerons par envisager le cas de deux observateurs, en mouvement relatif l'un par rapport à l'autre, à la vitesse \vec{v} . Plus précisément et formellement, nous considérons donc deux référentiels, R' et R , dont nous supposons que la vitesse de R' par rapport à R est donc \vec{v} . Il serait possible d'envisager une vitesse quelconque, mais nous choisirons pour les besoins d'une démonstration didactique un cas simplifié, qui n'ôte rien aux résultats : nos deux référentiels, caractérisés respectivement par les axes (O', x', y', z') et (O, x, y, z) , sont tels que les axes sont parallèles et en particulier, les axes $O'x'$ et Ox sont communs et parallèles à la vitesse \vec{v} .

En outre, on supposera qu'au temps initial, $O = O'$, c'est-à-dire lorsque les horloges fixées à chaque référentiel indiquaient toutes deux $t = 0$ et $t' = 0$.

La question, dans ce cas où chaque référentiel a donc la capacité de positionner un élément en espace et en temps avec son propre matériel, lequel est similaire pour les deux, et initialisé conjointement, est de savoir quelles sont les relations de passage entre référentiels qui permettent non pas de préserver les résultats de mesure, mais de préserver les lois de la physique. Autrement dit, afin qu'étant données les mesures effectuées dans deux référentiels concernant le même phénomène, les équations régissant le comportement de la nature donnent le même résultat, conservent la même forme. Si ce principe n'est pas respecté, cela signifierait que l'action de la gravité, par exemple, ne s'exprimerait pas de la même façon dans un référentiel ou dans un autre, favorisant certains observateurs plutôt que d'autres.

La préservation de la forme des lois de la physique mènent à la transformation dite de Lorentz, lesquelles sont données par : [67, 68]

$$\begin{cases} ct = \gamma(ct' + \beta x') \\ x = \gamma(x' + \beta ct') \\ y = y' \\ z = z' \end{cases}$$

où c est la vitesse de la lumière dans le vide, β et γ sont des coefficients sans dimension définis par

$$\beta = v/c, \quad \gamma = \frac{1}{\sqrt{1 - \beta^2}}$$

On constate ainsi que d'une part, temps et espace sont liés si l'on considère des observateurs en mouvement relatif, et d'autre part, que la transformation permettant de conserver la forme des lois physiques fait intervenir la vitesse de ce mouvement relatif, en rapport avec une constante (la vitesse de la lumière), qui joue en quelque sorte un rôle symétrique de la constante de Planck, dans le sens où elle fournit une sorte de grandeur limite, de coupure. Enfin, ces relations sont non linéaires en la vitesse.

Il est possible de modifier la formulation de ces transformations, afin de mettre en évidence un type de rotation particulière, faisant non pas intervenir les fonctions trigonométriques usuelles de type cosinus et sinus, mais leurs équivalents hyperboliques. Les transformations de Lorentz s'expriment alors par l'utilisation d'un paramètre θ , qui est un angle de rotation dans l'espace dit de Minkowski. Ce paramètre est défini par :

$$\tanh \theta = v/c \equiv \beta$$

soit

$$\theta = \operatorname{atanh}(v/c) \equiv \operatorname{atanh} \beta$$

Ce qui mène à :

$$\gamma = (1 - \beta^2)^{-1/2} = (1 - \tanh^2 \theta)^{-1/2} = \cosh \theta$$

et

$$\begin{cases} ct = ct' \cosh \theta + x' \sinh \theta \\ x = ct' \sinh \theta + x' \cosh \theta \end{cases}$$

On note que l'obtention de la transformation inverse s'effectue en changeant β en $-\beta$, et donc θ en $-\theta$. Ce type de relation et de manipulation sont caractéristiques d'opérations de groupes, sur lesquels nous reviendrons lorsque nous parlerons des groupes de Lie.

Dilatation des durées et contraction des longueurs

La transformation de Lorentz, qui garantit à ce stade l'invariance des lois de la nature lorsqu'elles s'expriment dans des référentiels en mouvement de translation l'un par rapport à l'autre, engendre deux phénomènes intéressants concernant les deux types de mesure effectuée dans ces référentiels : les mesures portant sur les durées – ou intervalles de temps – et les mesures portant sur les longueurs – ou intervalles d'espace.

Considérons dans un premier temps les intervalles de temps mesurés dans deux référentiels différents, lesquels sont donc en mouvement l'un par rapport à l'autre, comme supposé dans la section portant sur la transformation de Lorentz. Annonçons dès ici ce que l'on constate et théoriquement, et expérimentalement : l'horloge attachée à un référentiel en mouvement relatif par rapport à une horloge attachée à un référentiel "fixe" semblera fonctionner plus lentement que l'horloge du référentiel fixe, ces deux horloges étant

pourtant synchronisées et identiques.

Dans tous ce qui suivra concernant les théories relativistes, nous écrirons les intervalles d'une grandeur donnée G comme ΔG – c'est-à-dire, comme étant $G_2 - G_1$ – tandis que les éléments différentiels ou intervalles infinitésimaux seront notés dG .

Considérons un référentiel R , dans lequel deux événements se produisent au même endroit, spatialement parlant, mais non au même instant. Il existe donc un intervalle de temps Δt tel que $\Delta t \neq 0$. En outre, pour ces deux événements, nous avons également : $\Delta x = \Delta y = \Delta z = 0$. L'horloge attachée à R affiche donc deux instants différents.

Prenons désormais le référentiel R' , en mouvement de translation uniforme, c'est-à-dire, à vitesse constante par rapport au référentiel R . Les deux événements étudiés ne se dérouleront pas au même endroit, cette fois-ci, selon ce nouveau référentiel, puisqu'il est mobile par rapport à R . Mais que peut-on dire de la durée $\Delta t'$ mesurée dans R' , par rapport à la durée Δt de R ? La transformation de Lorentz nous permet d'écrire :

$$\Delta t' = \gamma(\Delta t - v\Delta x/c^2),$$

et si $\Delta x = 0$, on a même :

$$\Delta t' = \gamma\Delta t$$

Comme $\gamma > 1$, on a $\Delta t' > \Delta t$: ainsi, depuis l'horloge attachée au référentiel R' en mouvement par rapport à R , le temps écoulé entre les deux événements sera mesuré plus grand que s'il est mesuré par l'horloge de R . C'est pourquoi on parle de dilatation des durées ou de ralentissement de l'horloge en mouvement.

Il est intéressant de constater alors que, le mouvement des référentiels étant un mouvement relatif de l'un par rapport à l'autre et vice-versa, l'hor-

loge de R sera vue comme ralentie depuis R' , mais l'horloge de R' sera également vue comme ralentie depuis R . On peut reformuler cette propriété de la façon suivante : chaque référentiel perçoit ce qui provient du passé de l'autre référentiel, s'il s'en réfère à sa propre horloge.

Expérimentalement, les premières vérifications de cette propriété étonnante ont pu être menées en 1960 par des physiciens, en accélérant des atomes d'un cristal radioactif, en en augmentant la chaleur. La fréquence des rayons gamma émis par le cristal a ainsi diminué, ce qui correspond à un ralentissement du temps propre des atomes accélérés par rapport aux atomes non accélérés. Les mesures effectuées ont montré un accord à la théorie correct, puisque présentant 10% de marge d'erreur avec les prédictions.

Considérons maintenant une règle de longueur L immobile dans notre référentiel R' , et orientée dans la direction de la vitesse relative entre les référentiels R' et R . De manière symétrique à notre problème d'horloge, nous supposons que le référentiel R possède une règle strictement identique, attachée à R . Supposons que la règle de R' soit mesurée à l'aide de celle de R , alors qu'elle "passe" et dépasse R . La mesure effectuée sera aussi surprenante que celle effectuée pour les durées : la règle apparaîtra plus courte que si elle était mesurée dans R' .

En reprenant la transformation de Lorentz, et en prenant $L = \Delta x'$, nous avons ici :

$$\begin{cases} c\Delta t' = \gamma(c\Delta t - \beta\Delta x) \\ \Delta x' = \gamma(\Delta x - \beta c\Delta t) \\ \Delta y = \Delta y' \\ \Delta z = \Delta z' \end{cases}$$

Dans R , $\Delta t = 0$, et il vient alors que $L = \Delta x' = \gamma\Delta x$, et que $L' = \Delta x = \gamma^{-1}L$. Par conséquent, nous avons bien une longueur $L' < L$.

Notons par ailleurs que dans notre cas, nous avons un déplacement dans le sens de la règle ; il en découle que pour des mesures orthogonales au déplacement, les longueurs suivant y et z demeurent inchangées.

Nous voyons ainsi que les mesures de durées ou de longueurs ne sont pas invariantes par changement de référentiels, quand bien même ces référentiels seraient définis initialement strictement de la même manière, possédant les mêmes instruments de mesure organisés similairement. La seule différence entrant en jeu étant un mouvement très simple, celui d'une translation rectiligne uniforme.

Voyons maintenant ce qu'il peut advenir pour la définition d'un type d'intervalle un peu différent, un intervalle liant temps et espace, plutôt que considérer des intervalles séparément de genre temps et de genre espace.

Intervalle d'espace-temps

Nous considérons à nouveau un référentiel donné R . Dans ce référentiel, nous localisons ce que l'on appelle des événements, e_1 et e_2 , formulation pour simplement désigner deux points déterminés par leurs coordonnées spatiales et temporelle : $e_1(x_1, y_1, z_1, t_1)$ et $e_2(x_2, y_2, z_2, t_2)$.

On définit l'intervalle d'espace-temps qui les sépare dans R par le carré de cet intervalle :

$$(\Delta s)^2 = c^2(t_2 - t_1)^2 - (x_2 - x_1)^2 - (y_2 - y_1)^2 - (z_2 - z_1)^2$$

ce qui s'écrit avec les notations utilisées précédemment :

$$\Delta s^2 = c^2\Delta t^2 - \Delta x^2 - \Delta y^2 - \Delta z^2 = c^2\Delta t^2 - \Delta l^2$$

Ce nouvel intervalle, quant à lui, est invariant par la transformation de Lorentz, ce qui signifie que la mesure de cet intervalle d'espace-temps dans un référentiel R sera la même que dans un référentiel en mouvement R' . En

effet, nous avons :

$$c^2\Delta t^2 - \Delta x^2 - \Delta y^2 - \Delta z^2 = c^2\Delta t'^2 - \Delta x'^2 - \Delta y'^2 - \Delta z'^2.$$

Notons à ce stade que l'on parle communément de "carré" de l'intervalle d'espace-temps, alors même que l'on n'est pas assuré de la positivité de cette grandeur Δs^2 . Ce n'est en aucun cas une limite, et cela constitue précisément la différence avec une distance euclidienne classique.

C'est justement le signe de cet intervalle Δs^2 qui permet de classer deux événements l'un par rapport à l'autre, de leur attribuer une nature particulière. Le mode de représentation usuel pour représenter les types d'événements est celui du "cône de lumière". Nous expliquons à la suite cette appellation et l'intérêt de ce cône. Il permet notamment de raisonner en termes de causalité, de lien entre deux événements.

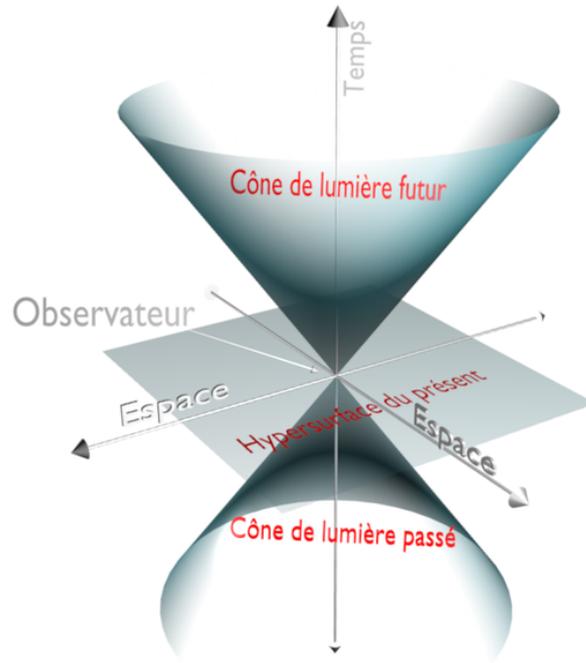
Dans cette approche, temps et espace sont donc liés, et ce de façon pertinente en termes d'invariance par transformation de Lorentz. S'ils ne sont pas parfaitement symétriques, le temps et l'espace n'en sont pas moins connectés, ce qui permet finalement d'exprimer les durées ou les distances aussi bien en unités de longueur qu'en unités de temps. Ceci est notamment possible du fait de l'invariance d'une vitesse limite – considérée comme la vitesse de la lumière dans le vide.

Cône de lumière et causalité

Considérons un événement e_0 au sens précédemment décrit, d'un point de l'espace-temps localisé par ses coordonnées spatiales et temporelle, dans le référentiel R . Cet événement nous sert ici d'événement de référence, auquel comparer d'autres événements.

Trois cas de figure peuvent alors se présenter, selon le signe et la nullité de

FIGURE 2.16 – Causalité, signe de l'élément d'espace-temps et cône de lumière (Wikipédia, licence CC)



l'intervalle d'espace-temps séparant e_0 d'un autre événement. Nous prenons donc un intervalle défini par : $\Delta s^2 = c^2 \Delta t^2 - \Delta l^2$.

Les événements séparés par un intervalle Δs^2 tel que $\Delta s^2 = 0$ sont ceux qui sont à une distance spatiale Δl et une distance temporelle Δt de e_0 telles que $\Delta l / \Delta t = c$. C'est-à-dire que ces événements ne peuvent être joints depuis e_0 que par un message allant à la vitesse de la lumière – autrement dit, les deux événements peuvent être reliés par un rayon de lumière dans le vide, lequel peut porter une information. La nullité de l'intervalle Δs^2 entraîne $l/t = c$, laquelle égalité donne l'équation du bord à trois dimensions d'un cône de révolution dans un espace à quatre dimensions (espace-temps). Cette équation limite, départageant les cas où l'intervalle va être négatif ou positif strictement, définit le cône de lumière. Nous en reproduisons en figure 2.16 le principe.

Voyons maintenant le cas où les événements sont séparés par un intervalle tel que $\Delta s^2 > 0$. Il s'agit des événements espacés d'une distance spatiale Δl et une distance temporelle Δt de e_0 telles que $\Delta l/\Delta t < c$. Dans ce cas, les événements peuvent être reliés physiquement par des messages se propageant à une vitesse inférieure à celle de la lumière dans le vide. C'est la condition pour que, en théories relativistes, il y ait une possible relation causale entre les deux événements. Cette condition définit ici l'inégalité $l/t < c$, laquelle donne l'équation de l'intérieur à quatre dimensions d'un cône dans un espace à quatre dimensions.

Ce cône, ou disons, ce sablier, se décompose ainsi en deux parties, supérieures et inférieures, à partir de l'événement de référence e_0 : la partie supérieure de l'intérieur du cône contient tous les événements futurs que l'on peut joindre à partir de e_0 , tandis que la partie inférieure de l'intérieur du cône contient tous les événements passés à partir desquels on pouvait joindre e_0 .

Enfin, examinons le cas où l'intervalle est tel que $\Delta s^2 < 0$. Cela concerne les événements qui sont à une distance spatiale Δl et une distance temporelle Δt de e_0 telles que $\Delta l/\Delta t > c$. Ces événements ne peuvent être reliés depuis e_0 par un processus physique, si l'on considère qu'il n'existe aucun phénomène se propageant plus rapidement que la vitesse de la lumière dans le vide. En relativité, il ne peut y avoir de relation causale liant ces événements. De manière complémentaire au cas positif, ce type d'intervalle définit l'inégalité $l/t > c$, qui correspond à l'équation de l'extérieur à quatre dimensions d'un cône dans un espace à quatre dimensions.

La représentation par cône de lumière définit géométriquement les relations causales possibles entre deux événements, selon le signe de l'intervalle d'espace-temps les reliant, ce signe étant invariant par transformation de Lorentz, donc ne dépendant pas de l'observateur. Il s'agit donc d'une notion fondamentale, qui relie géométrie et causalité.

Temps propre

Nous prenons le temps ici d'aborder une notion qui nous sera utile, notamment pour mieux comprendre certains liens que nous ferons entre la théorie que nous proposons et des hypothèses émises par d'autres auteurs. La notion de temps propre se rapporte au temps tel que l'horloge d'un référentiel l'affiche. Par exemple, le temps propre d'une particule est le temps propre d'une horloge qui lui serait attachée – autrement dit, il s'agit du temps qui s'écoule dans un référentiel où la particule est immobile. De la sorte, tout observateur de cette particule dans un référentiel en mouvement relatif par rapport à la particule considèrera que le temps propre de la particule est plus lent que son temps propre à lui. Le temps propre est classiquement noté τ .

Dans le référentiel de la particule, la variation de temps propre de cette particule est $\Delta\tau$ et les variations de ses coordonnées spatiales sont nulles $\Delta x_0 = \Delta y_0 = \Delta z_0 = 0$. Dans un autre référentiel, les intervalles respectifs sont cependant Δt et $\Delta x; \Delta y; \Delta z$.

L'invariance du carré de l'intervalle d'espace-temps entraîne que $\Delta s^2 = c^2\Delta\tau^2 = c^2\Delta t^2 - \Delta x^2 - \Delta y^2 - \Delta z^2$. De fait, on a $\Delta\tau = \frac{\Delta s}{c}$: le temps propre et l'intervalle d'espace-temps sont égaux, au coefficient $\frac{1}{c}$ près. Le temps propre est donc invariant par changement de référentiel. Plus précisément, on a la relation $\Delta\tau = \Delta t \cdot \sqrt{1 - \frac{v^2}{c^2}} < \Delta t$. Le temps propre est bien contracté par rapport au temps mesuré par un observateur en mouvement relatif.

On note le cas particulier d'une particule se déplaçant à la vitesse de la lumière : celle-ci n'a pas de temps propre – ou dit différemment, le temps ne s'écoule pas pour elle : si $v = c$ et $\Delta\tau = \Delta t \cdot \sqrt{1 - \frac{v^2}{c^2}}$ alors $\Delta\tau = 0$. Il est néanmoins notable que les objets concernés par ce cas sont précisément des objets de masse nulle. Nous y reviendrons, mais de la même façon que l'on peut caractériser une distance ou un temps en unités de longueur ou de temps, il est également possible de mesurer le temps en unité de masse. Ainsi, l'absence de masse paraît logiquement liée à l'absence de temps propre.

Relativité générale et principe de moindre action : géodésiques

Les notions que nous avons abordées jusque là concernaient essentiellement des résultats issus de la relativité restreinte - mais ces principes sont fondateurs des théories relativistes, dans le sens d'un intervalle invariant sous l'action d'une transformation. Notamment, si nous avons étendu la dimension de l'espace à quatre, en considérant un espace-temps, la géométrie de ce type d'espace demeure "plat".

En relativité, il est postulé un principe d'équivalence – dit d'Einstein – qui pousse un peu plus la correspondance entre géométrie et loi physique. Plus précisément, le principe d'équivalence stipule que l'action de la gravité est localement équivalente au choix d'un référentiel accéléré.

Dans sa version "générale", la relativité essaie de gommer au maximum les différences entre temps et espace, aussi la formulation change un peu pour devenir plus générique et manipulable conceptuellement. De manière générale, un événement devient caractérisé par ses coordonnées $(x'_0; x'_1; x'_2; x'_3)$, et l'on considèrera, plutôt que des intervalles de mesures, une formulation équivalente mais en termes d'élément infinitésimal d'espace-temps $ds^2 = (x_0)^2 - (x_1)^2 - (x_2)^2 - (x_3)^2$. Cette expression peut s'écrire de manière plus condensée encore, en faisant intervenir un objet que l'on appelle métrique de l'espace-temps, notée $g^{ij}(x')$. Cette métrique dépend donc des coordonnées, et représente ni plus ni moins que les coefficients d'une combinaison linéaire en les coordonnées. De façon assez informelle, on obtient : $ds^2 = g^{ij}(x')x'_i x'_j$. Ce type de notation est repris en annexes de ce document.

De fait, l'action de la gravité s'exprime par la géométrie de l'espace-temps, et inversement, la forme de la métrique matérialise l'action de la gravité en tout point de l'espace-temps. Plus rigoureusement, on bascule dans le cadre de la relativité générale lorsque l'on définit concrètement l'élément infinitésimal d'espace-temps, structurant la géométrie de celui-ci. Il y est alors considéré, pour un objet étant soumis à la gravité, de masse m , que

l'action de cette gravité sur l'objet est donnée par la relation :

$$ds = -mc\sqrt{g^{ij}dx_i dx_j}$$

où on suppose que $g^{ij} = g^{ji}$ sans rien enlever à la généralité.

À partir de cette relation, et en conservant l'idée que l'élément ds correspond également au temps propre de l'objet concerné, il est possible de déterminer la forme du mouvement de cet objet dans l'espace-temps, selon un principe de moindre action (développé davantage au chapitre 5 de nos travaux).

En effet, en définissant l'action S par la relation :

$$S = -mc \int_A^B ds$$

on déduit classiquement les équations du mouvement en minimisant cette action, ce qui revient ici à maximiser le temps propre de l'objet entre deux événements A et B. Une telle opération définit ce qu'on appelle les géodésiques, c'est-à-dire les chemins de plus courte distance entre deux points, là où l'on se place dans un espace qui ne comporte pas nécessairement de lignes droites au sens euclidien du terme – on peut penser aux méridiens du globe terrestre, reliant les pôles suivant le chemin le plus direct, mais n'étant pas des segments "droits".

Par ailleurs, l'action étant définie, on peut également se donner un Lagrangien L_0 , tel que :

$$\frac{dS}{dt_0} = L_0 = -mc\sqrt{g^{ij}V_i V_j},$$

les V_i représentant les composantes de la "vitesse" – la quadri-vitesse, en toute rigueur.

Les équations dites d'Euler-Lagrange sont alors mobilisables :

$$\frac{d}{dt_0} \frac{\partial L_0}{\partial V_k} - \frac{\partial L_0}{\partial x_k} = 0$$

Il en découle :

$$\frac{d}{dt_0} \left(\frac{2 \cdot g^{ik} V_i}{2 \cdot \sqrt{g^{ij} V_i V_j}} \right) - \frac{\partial^k g^{ij} \cdot V_i V_j}{2 \cdot \sqrt{g^{ij} V_i V_j}} = 0 ,$$

ainsi que :

$$\dot{V}_m + \Gamma_m^{ij} V_i V_j = 0$$

que l'on peut aussi écrire :

$$\frac{d^2 x_k}{ds^2} + \Gamma_k^{ij} \frac{dx_i}{ds} \frac{dx_j}{ds} = 0$$

Cette dernière équation est la forme fréquemment rencontrée des équations des géodésiques. L'existence d'une métrique ne se résumant pas en tout point de l'espace à une métrique euclidienne assure que l'espace peut se "déformer" localement, puisque de proche en proche, les éléments infinitésimaux de mesure peuvent varier et changer d'orientation, par exemple.

Nous n'explicitons pas davantage les détails des calculs ni le formalisme, qui dépasseraient largement les limites de ce travail. Des éléments plus précis peuvent être trouvés en annexes, ainsi que dans le chapitre 5. En revanche, nous précisons un peu plus la place de la géométrie dans ce qui suit, en introduisant la notion d'espaces tangents.

Principe d'inertie dans l'espace courbe

Plus techniquement, pour que les outils mentionnés plus haut, notamment la notion de vitesse et de métrique, de dérivation, puissent être opérationnels dans un espace qui n'est plus plat, il devient nécessaire d'introduire en tout point de l'espace, la notion d'espace tangent. Autrement dit, en chaque point

de l'espace courbe, on lui associe un espace local plat. C'est dans cet espace local que l'on travaille, et que les opérations que l'on désire effectuer sont bien définies.

Ces espaces tangents sont des espaces vectoriels de dimension 4, sont munis de leur base naturelle $(\vec{e}^0; \vec{e}^1; \vec{e}^2; \vec{e}^3)$. Pour un point M tel que $M(x_0; x_1; x_2; x_3)$, on écrit en général la base correspondante comme $\vec{e}^i = \left(\frac{\partial x_j}{\partial x_i} \right)_{j=0,1,2,3}$; ce que l'on écrit également souvent $\vec{e}^i = \frac{\partial}{\partial x_i}$.

Nous avons donné plus haut les équations des géodésiques : celles-ci portent sur les coordonnées $\frac{dx_i}{dt_0}$ ou $\frac{dx_i}{ds}$ de la quadri-vitesse le long de la géodésique. En revanche, elles ne fournissent aucune information quant aux variations d'un "quadri-vecteur" \vec{e}^i d'un point à un autre de l'espace, ni même pour la dérivation du quadri-vecteur vitesse défini comme $\vec{V} = V_i \vec{e}^i$.

C'est ici que l'on peut invoquer le principe d'inertie, et le reformuler dans le cadre de la relativité générale : le long d'une géodésique, et en l'absence d'intervention extérieure, le (quadri-)vecteur vitesse d'un objet est constant.

Autrement dit, on a : $d\vec{V} = \vec{0}$, ce qui entraîne : $d\vec{V} = \vec{0} = dV_i \cdot \vec{e}^i + V_i \cdot d\vec{e}^i = -\Gamma_i^{jk} dx_j V_k \cdot \vec{e}^i + V_i \cdot d\vec{e}^i$

Le quadri-vecteur vitesse initial étant quelconque, on obtient :

$$d\vec{e}^i = \Gamma_k^{ij} dx_j \vec{e}^k$$

Cette dernière expression montre bien que les vecteurs de base de l'espace tangent associé à un point donné M peuvent se déformer, et cette déformation est liée directement à la métrique, dont dépendent les Γ_k^{ij} .

2.3.3 Les invariants, les groupes de transformation et la géométrie

Nous venons de le voir, tant en théories quantiques qu'en théories relativistes, les invariants sont des notions centrales pour la caractérisation d'objets identifiables de manière stable et reproductible, en lien avec soit

le concept d'observable, soit d'observateur, ce qui n'est pas nécessairement directement équivalent. De manière plus indépendante du cadre théorique, quantique ou relativiste, il est possible de donner une formulation abstraite et générale de la notion d'invariance.

Pour ce faire, il est nécessaire de disposer de deux éléments dans la théorie concernée : la définition d'une *action*, laquelle d'une façon ou d'une autre, explicite la possibilité d'un mouvement relatif, et la donnée d'un type de transformation.

Il est alors possible de parler de grandeurs invariantes sous l'action d'une transformation – en général, sous l'action d'un objet qu'on appelle groupe. En théories physiques, la mathématicienne E. Noether a énoncé des théorèmes portant son nom, de portée très générale, et reliant l'existence d'une conservation d'une grandeur donnée avec l'invariance de l'action sous une transformation particulière. [39] Plus spécifiquement :

Théorème de Noether : *À toute transformation infinitésimale qui laisse invariante l'intégrale d'action correspond une grandeur qui se conserve.* Autrement dit :

Si on se donne $q^{i(s)}$, un jeu de coordonnées généralisées qui dépendent continûment d'un paramètre s , si le Lagrangien L est indépendant de s , c'est-à-dire $L(q^{i(s)}, \dot{q}^{i(s)}, t) = L(q^i, \dot{q}^i, t)$ avec $q^{i(0)} = q^i$, alors on sait que l'intégrale I telle que :

$$I(q^i, \dot{q}^i) = \left. \frac{\partial L}{\partial \dot{q}^i} \frac{dq^{i(s)}}{ds} \right|_{s=0}$$

est une intégrale première, c'est-à-dire que $\frac{dI}{dt} = \frac{\partial I}{\partial t} + \{I, H\} = 0$, où H désigne l'hamiltonien associé à L et où $\{I, H\}$ est le crochet de Poisson.

Le théorème s'applique ainsi aux théories décrites soit par un Lagrangien, soit par un Hamiltonien, ce qui est le cas de la plupart des théories en physique. Ce type de résultat souligne à nouveau l'intérêt de pouvoir se doter

d'un Hamiltonien, puisque celui-ci permet d'accéder à la caractérisation des invariants du système.

Symétries d'espace-temps

Noether a ainsi montré que les invariances par transformations portant sur l'espace d'étude, plus spécifiquement, des transformation de type translation ou rotation, engendraient la conservation de grandeurs clés en physique :

- l'invariance par translation dans le temps entraîne la conservation de l'énergie ;
- l'invariance par translation dans l'espace selon une direction x^i entraîne la conservation de la quantité de mouvement p_i dans la même direction ;
- l'invariance par rotation dans l'espace entraîne la conservation du moment angulaire.

Là encore, la portée des résultats est générale. Elle fait notamment intervenir les groupes de transformations. Un grand type de groupes de transformations rencontrés sont les groupes de Lie. [38]

Groupes de Lie

Définition : Une structure algébrique G est un groupe de Lie réel ou complexe lorsque :

- G est une variété différentielle réelle ou complexe ;
- G , munie de deux fonctions : la multiplication définie par $\times : G \times G \rightarrow G$, et l'inversion définie par $- : G \rightarrow G$, est un groupe ;
- les applications de multiplication et d'inversion sont analytiques, réelles ou complexes.

Il est également possible de définir un groupe de Lie comme une variété différentielle munie d'opérations de groupe différentiables, ou même seulement continues. La dimension d'un groupe de Lie est définie comme sa

dimension en tant que variété.

Il est également possible de définir des algèbres de Lie, notamment en utilisant la notion d'espace tangent, comme déjà mentionné en relativité. Dans ce cas, on définit un espace tangent en l'élément neutre de G . L'extension des groupes aux algèbres permet précisément d'exprimer les invariances, d'un champ de vecteurs par exemple, sous l'action d'un groupe de Lie.

Exemple simple de groupe de Lie : Un exemple simple est le groupe des matrices de rotation plane, noté $SO(2, R)$:

$$\begin{pmatrix} \cos \lambda & -\sin \lambda \\ \sin \lambda & \cos \lambda \end{pmatrix}.$$

Il est paramétré par un seul angle λ : sa variété est donc unidimensionnelle – un cercle.

Il s'agit bien d'un groupe : l'inverse d'un élément de paramètre λ est donné par l'élément de paramètre $-\lambda$ et le produit des éléments de paramètres λ et μ est donné par l'élément de paramètre $\lambda + \mu$.

Nous n'irons pas plus en détails là non plus, le point essentiel étant qu'il existe des objets mathématiques bien identifiés et connus, qui permettent de caractériser d'une part des transformations à appliquer sur un objet, des formes, et d'autre part, de caractériser des invariants. Rappelons-nous simplement ici ce que nous avons suggéré, concernant la théorie des transformations de d'Arcy Thompson, et l'utilisation des groupes de Lie. Le paramètre λ dont il est question ici pour paramétrer les rotations, pourrait être un des paramètres "causaux", composites ou non, rendant compte des transformations proposées par d'Arcy Thompson.

2.3.4 Place du temps dans les 2 théories

Nous venons d'aborder les 2 grandes théories - outre la théorie des systèmes dynamiques - prédominant, sous une forme ou une autre (théorie quantique, théorie quantique des champs, théorie des cordes, théories des membranes, théorie supersymétrique...etc.) en physique théorique et mathématique. Une des variables fondamentales entrant en jeu est la variable temporelle, et celle-ci n'a absolument pas le même rôle dans les deux cas.

En théorie quantique, l'espace de représentation est un espace plat, sans autre structure que celle de son spectre – en ce sens, les théories quantiques sont des théories projectives et spectrales : on caractérise les systèmes par leurs états propres, qui sont censés être autant de formes pures et élémentaires. Le temps y est une variable indépendante, et possède un caractère absolu : il est le même pour tout observateur, partout et à n'importe quel moment. Plus encore, le temps sert habituellement de moyen commode pour délimiter une causalité, selon le postulat que l'effet ne peut précéder la cause. Ce qui se traduit usuellement par le fait que toute information se propage dans une étendue à vitesse finie, et non instantanément. Cependant, en mécanique quantique, la causalité ne revêt pas exactement les mêmes caractéristiques, puisque d'une part, le phénomène d'intrication autorise la téléportation quantique et la détermination instantanée de l'état d'une particule dès lors que l'on mesure l'état d'une particule intriquée à la première, et que d'autre part, même si elle prête toujours à débat, la contrefactualité paraît pouvoir y exister – la capacité qu'a le futur à influencer le passé, à la manière qu'une particule "possible" interfère avec elle-même lorsqu'elle doit franchir une fente d'Young sur les deux. L'observateur, en théorie quantique, a un statut différent de celui échu en théorie relativiste. Il y est pourtant central, puisque toute mesure est liée à une observable, un opérateur rendant une observation. La notion même d'observation détermine l'état d'un système quantique : la question de l'observateur idéal qui n'interfère pas avec l'expé-

rience en méthode expérimentale est donc mise à mal. Enfin, puisque il existe des particules que rien ne distingue *a priori*, et que celles-ci se comportent également comme des ondes, qu'elles ont une étendue, qu'elles sont délocalisées jusque leur mesure effective, ces particules n'ont pas de trajectoire bien identifiée et unique, si bien qu'il existe des scénarii où le temps ne suffit pas non plus à séparer plusieurs possibilités : quand deux particules indiscernables se croisent, comment savoir si elles se sont entrechoquées, ou si elles se sont simplement croisées ? Ne risque-t-on pas de prendre une particule pour l'autre, si cela a un sens ?

Alternativement, le temps relativiste a une toute autre valeur, et perd à la fois son indépendance des autres variables descriptives, et prend également une importance particulière. Le principe de relativité amène à considérer que la forme élémentaire, car invariante et permettant la formulation de lois identiques pour tous, est une forme liant espace et temps. Cette forme n'est pas quelconque, car elle revêt une relation d'intervalle, d'écart entre temps et espace. Les signaux, en relativité générale, se propage à une vitesse finie, éventuellement égale à celle de la lumière dans le vide, mais ne sauraient être instantanément interactifs. Aussi paradoxal que cela puisse sembler, la formulation d'une invariance des lois mène à la relativité du temps : il devient différent selon l'observateur, selon le mouvement relatif des observateurs. Le temps s'incorpore à la trame descriptive, et ne saurait s'en extraire.

Ce rôle différent dévolu au temps dans ces deux théories amène également à les considérer selon un autre angle, celui dont la façon la dynamique y est perçue.

2.3.5 Evolution à structure fixe vs évolution à structure mouvante

La mécanique quantique tire partie de l'analyse spectrale afin de déterminer ses propres invariants ou ses propres points fixes : un état propre est un

état pour lequel la mesure sera proportionnelle à l'état et stable par répétition de la mesure. Ainsi, le monde quantique, à tout le moins sa structure, est un monde à repères fixes, et la réalité s'y décrit alors tant par la possibilité de superposer ces états propres (des combinaisons linéaires d'états propres) que par la notion d'opérateur d'évolution pour ce qui concerne l'évolution temporelle d'un système, d'un état vers un autre, intriqué ou pur. D'une certaine façon, la complexité des systèmes quantiques tient dans la possibilité de figures d'interférences, dans les termes non linéaires des observations.

La relativité, quant à elle, dans sa formulation générale, construit un espace dont la géométrie même régit les déplacements possibles : le mouvement est pure géométrie. Néanmoins, cette géométrie n'a pas à demeurer fixe, en toute généralité, et est modelée par la distribution de masses dans l'étendue. Les masses pouvant elles aussi se déplacer, nous comprenons rapidement que la géométrie de l'espace est mouvante, fruit d'interactions permanentes. Ainsi, le temps faisant partie intégrante de la géométrie de l'espace, sa perception peut varier sans cesse selon l'observateur.

Si l'on reprend l'exemple des transformations de d'Arcy Thompson, l'équivalent quantique serait de dire qu'il existe différents types de poissons, bien identifiés et "purs". L'évolution de l'un vers l'autre se ferait alors selon des superpositions possibles de poissons "purs", et réglée par un opérateur d'évolution dépendant du temps. La grille de description serait fixe. L'équivalent relativiste, lui, serait celui d'un seul poisson, qu'importe sa forme de référence, les autres pouvant se déduire de celui-ci par des transformations continues de l'espace, donc de la grille de description. Dans le premier cas, il y aurait autant d'états propres que de poissons suffisamment différents. Dans le second, il existe une forme invariante, qui serait la définition d'un poisson en toute généralité, et les différentes observations que l'on pourrait en faire donneraient les différents poissons observés.

Enfin, selon une analyse plus inspirée des systèmes dynamiques, les dif-

férentes formes ou espèces de poissons peuvent s'interpréter comme autant d'états stables, au moins localement d'un même système dynamique. Cette approche, conformément à ce que l'on a expliqué des interprétations duales en termes de points fixes et de stabilité structurelle, se situerait alors à mi-chemin d'un schéma explicatif à structure fixe ou à structure mouvante : si les points fixes ou attracteurs peuvent matérialiser des formes distinctes – des espèces différentes – qu'en serait-il de la stabilité structurelle de ces points - et donc de ces espèces ? Les variations structurelles sont bel et bien des variations de la structure d'analyse, rejoignant en cela la perspective relativiste. Cependant, l'analyse de la stabilité structurelle fournit une méthode qui aboutit à la caractérisation de l'influence des paramètres d'un système sur sa structure, mais n'explique pas pourquoi ou sous quelle impulsion ces paramètres devraient changer. *A contrario*, la dynamique géométrique de la relativité a pour postulat que la structure change directement en lien avec la présence de masse : la structure doit changer.

2.4 Quelques mots à propos des approches sciences cognitives et des neurosciences

Au travers de ce premier chapitre, nous avons abordé tant les aspects historiques que théoriques des grandes théories morphogénétiques, que les formalismes candidats à une théorie unifiée de la morphodynamique. Nous n'avons pas réellement traité du pendant subjectif de ces théories, au sens de la subjectivité humaine, à tout le moins au sens de l'organe ou de la fonctionnalité qui appréhende, traite et forme ? ou bien reconstitue ? les formes dont nous parlons ici. Nous n'avons pas traité du cerveau, et de ses capteurs, de ses satellites sensoriels, qui l'alimentent en informations et le structurent.

Cette section sera courte, et n'a pour d'autre ambition que de mentionner certaines approches ou extensions, lesquelles prennent effectivement le point

de vue du cerveau.

2.4.1 Représentations mentales et cerveau

Notamment, nous n'avons pas traité, du côté des théories psychologiques au sens large, la manière dont est envisagée la forme. Bien sûr, les philosophes nous ont parlé d'objet en soi, de substance, de perceptions, ou encore d'objets idéaux, platoniciens. Parmi les approches plus récentes, nous pouvons cependant citer ce que deux auteurs au moins ont nommé comme la "psychologie de la forme". [69, 70] Historiquement, nous ne pouvons ainsi pas faire l'économie de mentionner la théorie Gestaltiste.

La Gestalt est une théorie intéressante dans le sens où elle opère un certain renversement avec la majorité des théories d'inspiration davantage réductionniste. En Gestalt, la forme n'est jamais forme isolée ou isolable, elle est, en quelque sorte, un découpage de l'esprit, pour des raisons soit de perceptions, soit, finalement, de nécessité d'opérationnalisation de l'espace. Les formes séparées sont le produit d'un processus de catégorisation. En dehors de ce découpage parfois "nécessaire", toute forme individualisée pour une raison X ou à un instant t fait en réalité partie d'une forme plus grande, dont les contours eux-mêmes ne sont pas parfaitement définis. Il n'existe pour ainsi dire pas d'objet en soi, mais des objets en relations, en contexte, lesquels se définissent par leurs interrelations aux autres formes.

Nous voulons profiter de cette section pour mentionner l'essai d'un auteur particulier, B. Falissard, lequel a proposé pour tentative de réconciliation entre neurosciences et psychanalyse, une représentation entre système dynamique et connectivisme – dans le sens où les concepts ou les objets perçus sont identifiables ou représentables par la notion d'attracteurs ; ces attracteurs étant des régions de plus basse énergie, dans un espace de représentation assis sur des automates cellulaires, les cellules étant autant de neurones pouvant être "activés" ou "inactivés". Ceux-ci sont évidemment

interconnectés et communiquant. L'auteur a qualifié cet essai de *Gedankenexperiment*, d'expérience de pensée – ce qui se justifie par le caractère prononcé de modélisation et de simplification ; cet exemple nous fournit par ailleurs une excellente illustration de modèle explicatif, voire même descriptif, et non prédictif. [71] Nous pensons, et proposerons plus loin, qu'il est possible d'asseoir ce type d'idées sur des données réelles, en toute généralité.

Si nous anticipons un peu sur la suite, notamment les techniques de réduction de dimension non linéaires ou de déploiement de variétés (mathématiques, géométriques), certaines approches, entre perception "cérébrale", et appréhension du monde en contexte, c'est-à-dire, faisant office d'intermédiaire entre une théorie qui dément tout objet en soi et une théorie qui ne considère que des catégories, des objets idéaux isolables, ont émergé, principalement pour des questions tournant autour de la représentation de grands ensemble de données, de visualisation. Il a ainsi été proposé que la perception, des attitudes d'un visage, de ses positions dans l'espace et le temps, par exemple, pouvait s'appuyer sur une représentation géométrique, continue, de basse dimension – l'espace géométrique en question n'étant pas trivialement plat comme un espace euclidien. [72]

2.4.2 Chimie moléculaire, circuits neuronaux et cognition

Néanmoins, la psychologie de la forme, si elle possède certainement encore des adeptes, n'est sans nul doute pas la théorie ou l'ensemble de théories dominante actuellement quand il s'agit d'aborder le cerveau et son fonctionnement. Selon les disciplines, le fonctionnement du cerveau – et les comportements censés en découler directement – sera explicité en termes de chimie moléculaire – c'est le cas en particulier en psychiatrie, comme l'atteste par exemple les aspects moléculaires du comportement suicidaire ou de la dépression, notamment *via* la sérotonine et ses récepteurs [73, 74, 75, 76] – ou, de plus en plus, en termes de circuits neuronaux, dont on rêve qu'ils soient

mieux caractérisés et dévoilés, par exemple par l'optogénétique, plutôt que par simplement les techniques IRM [76, 77, 78, 79].

En psychiatrie, ou plus largement dans le champ de la santé mentale, certains auteurs et chercheurs plaident précisément pour des bases moléculaires, fonctionnelles et neurologiques, connectivistes, des pathologies. [80, 81]

Et puis il y a également les approches davantage à la manière de Turing, ou plus généralement, de type traitement du signal et de l'information. Nous pouvons notamment citer deux grands types d'approches, l'une consistant à vouloir voir en un cerveau une architecture modélisable par une machine de Turing [82], l'autre visant à modéliser le comportement de stockage et récupération des informations à partir de réseaux de neurones, notamment en réseaux organisés autour de la notion de cliques de neurones – c'est le pari de C. Berrou et V. Gripon, dans une tradition de traitement du signal. [83, 84]

Aussi logiques ou intelligibles ces approches puissent être, nous pensons qu'elles souffrent chacune d'un certain nombre de limitations. La chasse aux molécules, au-delà de l'intérêt éventuel de trouver des événements corrélés, semblent promise à peu de succès – l'exemple précis de la sérotonine est un exemple criant : il existe des antidépresseurs fonctionnant très exactement à l'inverse l'un de l'autre, concernant la sérotonine, avec des résultats cliniques similaires. Les débats sur l'ocytocine ne sont pas moins édifiants, tant cette molécule (de l'altruisme, de la fidélité, de l'attachement, etc.) est censée expliquer nombre d'états, complexes ou non. Quant aux approches connectivistes, elles ne font que prolonger les approches plus macroscopiques que sont les représentations par aires fonctionnelles du cerveau. Très sans doute, l'information est encodée d'une manière qui dépend en partie de la mise en réseau des ressources support de l'information – de là à supposer que ces connexions relient nécessairement des régions parfaitement identifiées, au rôle totalement fixé et bien circonscrit, le pas semble délicat à franchir. En-

fin, pour ce qui est des approches par machine de Turing, dans le sens d'une identification directe entre machine et cerveau, ce sont des travaux préliminaires potentiellement porteurs, mais dont le développement tant théorique que pratique est si lourd qu'ils risquent de rigidifier et de stériliser une grande partie de ce que l'idée aurait pu accoucher.

2.4.3 Formes, vision et cerveau : neurogéométrie de la vision

On considère volontiers, désormais, que le cerveau, en tant qu'organe qui manipule de l'information, a besoin d'être alimenté en signaux plus ou moins structurés, souvent, régulièrement, sans quoi son développement et son maintien en seraient hypothéqués. L'un des "branchements" les plus directs, et certainement les plus "haut débit", au cerveau est celui le reliant les organes de la vision. De fait, vouloir considérer non pas d'un côté les yeux, de l'autre le cerveau, avec entre deux uniquement une ligne de transmission, mais bel et bien l'ensemble, semble une idée prometteuse.

Un auteur (J. Petitot), connu par ailleurs pour s'intéresser aux modèles de type catastrophe, au sens de Thom, a ainsi proposé d'étudier ce qu'il a nommé la "neurogéométrie" de la vision. Il s'agit d'une conséquente architecture théorique mêlant géométrie sous-riemannienne, invariants par actions de groupes, analyses par ondelettes et études par stabilité structurelle. Selon nous, ce genre d'approche, certes très technique et peu accessible à la majorité, ce qui peut en rendre l'utilisation rédhitoire, paraît plus adaptée et fondée que les approches parcellaires ou rigides dont nous avons discuté juste avant. Les outils mathématiques que nous mobiliserons, en particulier en fin de ces travaux, sont communs pour la plupart à ceux mobilisés par Petitot – nous pensons par exemple au formalisme à base de 1-jets, de groupes de transformation, de géométrie riemannienne. Nous n'avons cependant ni la volonté ni le loisir d'exposer en détails et de manière facilement abordable les développements de cet auteur. [85]

2.4.4 Objectivité et subjectivité, toujours

Quoi qu'il en soit, aucune de ces approches ne prend ou parvient à prendre en compte un point qui demeure certainement central. Il s'agit d'un point dual : donner un cadre permettant de définir l'objet ou la forme, de telle manière à ce qu'elle puisse remplir sa fonction scientifique voulue, et rendre compte de la place non seulement de la subjectivité dans la définition de la forme, mais également de l'intersubjectivité – en quoi la communication, la mise en relation d'organes traitant des informations, de consciences partageant un même espace langagier, social, culturel, structurent ou déterminent les formes possibles. La question des usages en est par ailleurs une question duale : en quoi les usages révèlent-elles les formes que nous manipulons ?

2.5 Ce que l'on en retient

Au cours de ce chapitre historique, descriptif et analytique des noms de la morphogenèse et des principes qui leur sont rattachés, nous avons dressé un panorama, certes non exhaustif des théories morphogénétiques, mais sans doute suffisamment large et représentatif pour ne pas s'être privés de ressources fondamentales. Nous avons ainsi fait remarquer tant les points communs que les oppositions récurrentes entre les différentes approches. Nous serions tenté, caricaturalement, de synthétiser ce panorama par l'existence d'une triade composée d'un côté des théories quantiques, aux objets et temps absolus, de l'autre des théories relativistes aux objets invariants déduits de principes posés comme absolument valides et au temps relatif ; entre les deux, l'approche systèmes dynamiques, empruntant *a priori* un temps absolu à la théorie quantique, et l'approche, dans le cas continu au moins, davantage similaire à la théorie relativiste, consistant à s'appuyer sur l'analyse d'une géométrie non linéaire. L'approche informatique, ou plus largement, reposant sur des concepts d'information et de communication – nous dirons informa-

tionnels – traverse également, en pointillés, ces différents modèles, notamment grâce aux contributions doubles de Turing. Par ailleurs, la géométrie paraît pouvoir fournir soit des extensions naturelles, soit des ponts entre ces différentes théories, comme le suggère par exemple le prolongement des systèmes différentiels classiques des systèmes dynamiques par un formalisme non-euclidien, riemannien, nous rapprochant en cela des théories relativistes. Enfin, la place de l'esprit, de la conscience et plus spécifiquement, dans une optique d'interaction et de communication, de l'intersubjectivité, demeure encore floue, alors qu'elle est pressentie depuis longtemps comme fondamentale.

Chapitre 3

La modélisation des systèmes complexes en sciences et en biomédecine

Sommaire

3.1	Analyse et synthèse	124
3.1.1	Le réductionnisme, la méthode expérimentale et la norme	125
3.1.2	L'exemple de la recherche biomédicale et les approches statistiques classiques	130
3.1.3	L'exemple de la psychiatrie et des DSM – les modèles biomédical et biopsychosocial	134
3.1.4	Un pas en avant ? L' <i>Epigenetic landscape</i> de Waddington	138
3.1.5	Rechercher la forme et les invariants – méthodes de <i>clustering</i>	143
3.1.6	Géométrie et dynamique	153
3.2	Quelles alternatives pour prendre en compte la complexité ?	154
3.2.1	De l'absence de définition consensuelle de la complexité	155
3.2.2	Conserver la complexité	157
3.2.3	Restaurer la complexité	159
3.2.4	Le choix du modèle	160
3.3	Causalités, temps et forme	161

3.3.1	Causalité linéaire, proportionnalité directe et boucle ouverte	161
3.3.2	Place du temps – opposition à la forme?	163
3.3.3	Traitement de l’information, inférence de la causalité, approche bayésienne	165
3.3.4	Modèles allométriques, temps physiologiques et hypothèse de West.	171
3.4	Trois grands modèles possibles	179
3.4.1	Entre onde et corpuscule, numérique et analogique – une histoire qui perdure	179
3.4.2	L’espace plat et le lien probabiliste : la forme ponctuelle	181
3.4.3	L’espace courbe et le lien local : la forme globale	182
3.4.4	L’espace des phases des systèmes dynamiques : la forme comme attracteur	183
3.4.5	Comparaison des 3 modèles	186
3.5	Les techniques associées aux 3 grands modèles	189
3.5.1	Techniques projectives : ACP/PCA et <i>PCA-likes</i>	189
3.5.2	Techniques d’apprentissage des variétés – NLDR	194
3.5.3	Techniques de reconstruction des attracteurs	199
3.5.4	Problème de la dimensionnalité <i>vs</i> complexité	204
3.5.5	Quantification de <i>l’epigenetic landscape</i>	207

Que la médecine soit un art ou une science, que la biologie, en tant que science de la vie, peine à définir son objet central – qu’est-ce que la vie? – des préoccupations récurrentes se dégagent de ces domaines : les moyens valides de production de la preuve, et la conscience d’une complexité immanente. Concernant la médecine au sens large, il n’est pas inutile de garder à l’esprit que l’histoire a réuni sous un même toit deux grands métiers aux approches distinctes : les barbiers-chirurgiens et les médecins, réunis autour de l’éviction de la pathologie. Cette pathologie qui, longtemps, tuait ou ne tuait pas : voilà où résidait finalement, le critère principal de jugement. La question de savoir si l’intervention prodiguée par les uns ou les autres était formellement la cause d’un rétablissement ou au contraire d’une aggravation était davantage affaire de consensus professionnel ou de jugement personnel, de réputation de l’intervention ou de celui qui la délivrait. Nous avons de-

puis hérité de l'*Evidence Based Medicine* – *EBM*, la médecine fondée sur les preuves, actuellement discutée et discutable. [86, 87, 88, 89, 90]

Néanmoins, lorsque des progrès majeurs se sont imposés : découverte et utilisation de la pénicilline, découverte de l'échelle microbiologique et place de l'hygiène minimale, amélioration de l'état nutritionnel des populations, concernant le critère principal de mortalité, les choses étaient suffisamment tranchées pour que l'on puisse attribuer un mérite direct aux interventions, du moins pour un périmètre d'observation (mort, vivant) et temporel restreint.

Les révolutions n'arrivent pas tous les jours. La recherche, l'industrie et l'économie ont besoin de stabilité pour fonctionner, telles qu'on les conçoit encore actuellement. Le politique nécessite, face aux impératifs d'allocations de ressources vécues comme limitées ou rares, de prendre des décisions, si possible motivées. Là où l'on pouvait motiver un choix comme la répercussion d'une commande divine, on tâche désormais d'invoquer la science et ses objets. La méthode expérimentale s'est développée dans ce contexte, et s'est imposée comme le plus haut standard en sciences biomédicales. Certainement, elle a pensé vaincre la complexité en multipliant *ad nauseam* les expériences, en contrôlant au-delà de toute raison le maximum de variabilité indésirables, ne laissant libres que le paramètre à évaluer et le paramètre mesurant l'intervention. Ainsi, nous croulons sous les rapports analytiques, chaque jour, évalués et publiés selon la même méthode, dans des milliers de journaux scientifiques, là où il serait légitime d'attendre quelque raisonnement synthétique – et si l'exhaustivité n'est pas possible, de se doter de principes permettant d'estimer ou de calculer à partir d'un ensemble réduit de preuves fermes et invariantes, les données nécessaires à d'autres contextes.

Nous verrons donc dans la suite quelles sont les connaissances et méthodes principales invoquées par les sciences et la biomédecine en particulier, sur quelles initiatives, peu nombreuses, il est possible de compter pour la suite. Nous expliquerons alors quelle est la place de la complexité, comment la conserver et comment la restaurer le plus fidèlement possible dans les études.

Un des concepts qui nous empêchent de penser correctement la complexité – et la simplicité – est le concept de causalité : dans quelle mesure est-il nécessaire, qu’entend-on ou que sous-entend-on, en général, quand on parle de causalité en biomédecine ? Notamment, nous verrons quelle est la place du temps dans notre conception de la causalité, et comment l’un peut être rapidement et abusivement pris pour l’autre.

Les grands modèles actuels, essentiellement hérités de la physique et des mathématiques et dont nous avons donné les aspects formels plus haut, proposent un abord de la complexité et de la causalité *a priori* divergeant. Ils peuvent cependant servir de modèles à notre réflexion quant à la recherche d’un cadre unique et théorique pour la morphodynamique. Nous expliquerons en quoi ces modèles peuvent être intéressants, en quoi ils divergent et en quoi ils partagent plusieurs points communs. Enfin, nous présenterons les techniques mobilisables et correspondant aux grands modèles retenus : il s’agit là d’une constante de ce travail, identifier d’une part la théorie, la construire, et d’autre part, en proposer la traduction technique, de telle sorte que la théorie soit testable et opérationnelle.

3.1 Analyse et synthèse

Nous voyons dans cette partie l’approche actuellement prédominante en sciences, à savoir celle héritée de la méthode expérimentale et du réductionnisme. La place de la norme ou de la normativité – par exemple, la

normativité biologique – y est examinée. Nous proposons ensuite d’aborder la traduction classique de l’approche expérimentale en biomédecine, *via* l’usage extensif des statistiques. Nous revenons alors sur une des branches de la médecine en crise, notamment dans sa tentative d’adapter les critères objectifs et expérimentaux à son domaine : la psychiatrie. Nous y voyons en particulier quel a été son choix méthodologique de synthèse des résultats expérimentaux.

Nous donnerons un premier aperçu d’une initiative à part dans le domaine de la biologie, et plus précisément, dans le domaine dit de l’épigénétique : le modèle de Waddington, initialement et éminemment géométrique. À partir de là, nous exposerons les alternatives à la recherche et à l’établissement d’objets invariants par les techniques statistiques classiques : la recherche de formes invariantes et stables par classes d’observateurs. Enfin, en liaison avec la notion de forme invariante, nous esquisserons un nouveau rapprochement entre forme, géométrie et dynamique.

3.1.1 Le réductionnisme, la méthode expérimentale et la norme

Si nous ne devons citer que trois noms dans cette partie, ils seraient les suivants : Descartes pour sa méthode [2] – le réductionnisme – Bernard pour sa méthode expérimentale, [12] et Canguilhem pour son travail sur le normal et le pathologie – sur la normativité biologique. [91]

Avant toute critique de l’approche proposée par Descartes, qui a su fructifier comme on le sait désormais, il convient de dire qu’elle a sans doute été avant tout un réflexe intellectuel salvateur, du moins en première intention. En effet, le principe de réductionnisme a permis de formaliser, d’ériger en rempart méthodologique, une approche unique face à des problèmes en apparence trop conséquents ou compliqués pour que l’esprit d’une seule personne puisse les traiter rigoureusement et avec l’assurance de résultats fiables et

reproductibles. L'idée de saisir l'entièreté du problème afin de le découper en multiples tâches élémentaires, simples à traiter et si possibles génériques afin que leur étude initiale en soit rentable par la multiplication systématique de leur analyse, a ouvert des horizons certainement insoupçonnés.

Comme nous reviendrons sur cette notion un peu plus loin, le concept de réductionnisme implique en pratique et en général, un concept connexe, celui de la linéarité. Par ce biais, réductionnisme et complexité s'opposent. La linéarité, au sens large, désigne le caractère simplement additif des effets élémentaires en termes de causes. Autrement dit, si deux causes élémentaires impliquent chacune un effet donné, et ce, en général, selon un lien de proportionnalité, alors la somme de ces deux causes, pondérées ou non, aboutit mécaniquement à la somme des effets, pondérés ou non également.

En formalisme mathématique ou de la théorie des systèmes, on parle d'applications linéaires ou de systèmes linéaires. Cette propriété se traduit alors de la manière suivante :

Soit f une application de E dans F où E et F sont deux espaces vectoriels à gauche sur un corps K .

L'application f est dite linéaire (ou morphisme de K -espaces vectoriels) si :

$$\forall x \in E, \forall y \in E, f(x + y) = f(x) + f(y),$$

et

$$\forall \lambda \in K, \forall x \in E, f(\lambda x) = \lambda f(x).$$

Une application f possédant la première propriété est dite additive, et, pour la seconde, homogène. Elle possède ces deux propriétés à la fois si et seulement si :

$$\forall (x, y) \in E^2, \forall (\lambda, \mu) \in K^2, f(\lambda x + \mu y) = \lambda f(x) + \mu f(y),$$

ou plus simplement, si et seulement si :

$$\forall (x, y) \in E^2, \forall \mu \in K, f(x + \mu y) = f(x) + \mu f(y).$$

Autrement dit, f préserve les combinaisons linéaires, c'est-à-dire :

pour toute famille finie $(v_i)_{i \in I}$ de vecteurs et pour toute famille $(\lambda_i)_{i \in I}$ de scalaires (i.e. d'éléments de K), $f(\sum_{i \in I} \lambda_i v_i) = \sum_{i \in I} \lambda_i f(v_i)$.

La méthode expérimentale est celle, par excellence, du laboratoire, comme pour le physicien qui se rapportera, pour l'étude mécanique de son système, à un référentiel local, si possible isolé et fermé. Il s'agit donc d'une situation extrêmement artificielle, visant à éliminer le maximum de perturbation extérieure et bruyante l'expérience d'informations et de données qui ne seraient pas pertinentes pour qualifier le mécanisme que l'on cherche à élucider.

Le principe est simple : pour un système à étudier, éventuellement sous l'influence de n facteurs dont on ignore ou ne considère pas les contributions mutuelles au fonctionnement du système, on fixe – on dit que l'on contrôle l'expérience – $n - 1$ de ces facteurs, de telle sorte qu'ils ne puissent varier durant l'expérience. En ce cas, c'est que l'on cherche à étudier l'influence du n ème facteur sur le fonctionnement du système. En réalité, l'effet que l'on cherche à observer est en général de deux sortes possibles : soit un effet qualitatif – la présence ou l'absence du facteur n engendre la présence ou l'absence d'une observation O d'intérêt – soit un effet quantitatif – on cherche à savoir, pour une variation du facteur n de Δn , quelle est la variation engendrée ΔO sur la variable O .

Tout ceci est tout-à-fait probant lorsque l'on fait l'hypothèse d'une indépendance des n facteurs les uns vis-à-vis des autres. Lorsqu'ils sont liés, il devient en pratique bien compliqué de neutraliser les variations des $n - 1$ facteurs lorsque l'on fait varier le facteur n . A l'autre bout de l'expérience, l'autre limitation tient dans le fait que l'on qualifie un système par une seule

observation, en l'occurrence O . Quand il s'agit de déterminer si une personne est morte ou non, O est assez facilement observable. Quand on se place dans des situations moins extrêmes, ce peut devenir compliqué de résumer un système à une seule observation, en supposant que la quantité en question soit en effet observable, et quantifiable.

Ce dispositif peut paraître anodin, mais il s'y joue, sans que ce soit explicitement dit, une définition de la causalité. En effet, la conclusion d'une telle expérience sera invariablement la conclusion comme quoi le facteur n est causal de l'observation O – si l'expérience est positive, bien sûr, et inversement. Cette méthode demeure le dispositif favorisé, considéré comme producteur du plus haut grade de preuve en biomédecine et biologie. Ca l'est pour juger de l'efficacité d'une thérapeutique, notamment médicamenteuse, ce sont les fameux essais randomisés contrôlés, ou pour juger de la causalité d'un facteur en biologie ; c'est là par exemple tout l'intérêt de posséder des animaux clonés : on contrôle certains facteurs par leur réplique à l'identique chez plusieurs personnes.

La randomisation, c'est-à-dire, le tirage au sort des personnes participant à une expérience, tient lieu de contrôle de la variabilité inter individuelle, en attribuant au hasard par exemple à deux groupes distincts (celui qui va recevoir le traitement, et celui qui va recevoir un placebo), des personnes relativement comparables initialement, mais qui diffèrent autour de la norme, quelle soit en âge, en poids, en taille, etc. Ainsi, le hasard brassant sans préférence ces caractéristiques, on s'assure que, moyennant un peu de bruit et d'imprécision dans les résultats finaux, il n'y ait pas d'autres caractéristiques que celle testée (le médicament) qui varient de manière déséquilibrée entre les deux groupes.

Le biologiste est peut-être bien en mal de pouvoir donner une définition satisfaisante de ce qu'est la vie, il n'a toutefois pas à devoir se prononcer en pratique sur cette question, en dehors d'un questionnement éthique. Il partage en outre le même questionnement que le médecin, quand il doit définir ce qu'est la santé, ou ce que sont le normal et le pathologique. Lorsque la seule préoccupation, ou du moins la plus fréquente et prégnante, était de savoir si l'on allait mourir ou rester en vie, la question de ce que peut être la santé ne se posait pas tant, et la santé était sans doute davantage perçue comme l'absence de maladie hypothéquant ses propres jours. La normalité, par ailleurs, dans l'hypothèse de fortes prévalences de certaines maladies et d'une espérance de vie significativement plus courte qu'actuellement, n'était peut-être pas synonyme ou n'impliquait pas l'absence de maladie.

La dichotomie prévalant encore actuellement est donc davantage celle de la normalité et de la pathologie, l'anormalité étant éventuellement davantage à ranger dans la pathologie, ne serait-ce qu'avec une psychiatrie constituée comme science de la santé mentale et des comportements. Certainement, l'anormalité a également à faire avec l'échelle sociale, et non pas uniquement individuelle, en tant que définition des comportements socialement normaux et adaptés, lesquels, s'ils ne relèvent pas de la pathologie, relèvent alors des lois, et de la justice.

Canguilhem s'est intéressé à ce que l'on pourrait nommer normal ou pathologique en médecine. Plus spécifiquement, il a relevé l'existence d'une normativité, plus que d'une normalité. En particulier, la normativité biologique, qui est extensible à toute autre mesure destinée à qualifier et quantifier le fonctionnement humain ou d'un organisme, sa physiologie. La normativité biologique définit de fait un fonctionnement normal : la mesure n'est pas trop éloignée de la moyenne observée en population, elle est, disons, dans les limites d'une variabilité susceptible d'être normale, compatible avec un

fonctionnement correct de l'organisme, et n'est attribuable qu'à la variabilité inter individuelle. Si l'écart à la moyenne est trop important –deux, trois écarts types? – la mesure est anormale, le reflet d'un fonctionnement pathologique. Toute la question est alors de savoir si cet état pathologique peut être en soi simplement une déviation morbide, menant vers un dérèglement potentiellement létal de l'organisme, ou s'il est possible qu'il existe des états stables, différents de la norme. Partant de ce principe, la dichotomie normal/pathologique est-elle tout le temps valide, ou bien existe-t-il des états plus nombreux, tout autant stables et régulés que peut l'être l'état dit physiologique?

3.1.2 L'exemple de la recherche biomédicale et les approches statistiques classiques

Ainsi, les trois concepts énoncés ci-dessus permettent de caractériser l'état de la recherche biomédicale : réduction d'un problème à une quantification unique d'intérêt, contrôle des variabilités notamment par l'affectation aléatoire des effectifs et interprétation en termes d'écarts à la norme. La translation des résultats obtenus dans une étude de ce genre est censée être assurée lorsque l'on a effectué l'étude sur une population la plus générique possible. On parle de validité externe : l'échantillon étudié étant similaire par ses caractéristiques à la population générale, alors les résultats seront les mêmes dans la population générale.

En industrie ou en sondage, cette approche se résume à déterminer ce que l'on appelle un plan d'expérience et un plan d'analyses statistiques : pour une efficacité maximale, le plan d'expérience (ici, la façon de choisir l'échantillon et de le randomiser) doit être conçu de telle sorte que le plan d'analyse lui soit adapté et réciproquement. [92]

Les avantages de l'essai randomisé contrôlé (RCT) tiennent dans la qualité de la preuve, si elle a un sens, et dans la possibilité qu'il offre d'appliquer de manière rigoureuse un formalisme mathématique adapté. Ainsi, une expérience bien contrôlée peut permettre de trouver des différences entre deux traitements qui seraient éventuellement douteuses en l'absence de quantification des effets aléatoires et de la dispersion des observations entre individus. Néanmoins, cet avantage est aussi certainement son plus grand aveu de faiblesse : de manière générale, a-t-on besoin de mettre en évidence une différence précise et infime d'efficacité entre deux traitements, à supposer par exemple, que déterminer une différence d'efficacité de 5% ait un sens ?

Finalement, un tel arsenal mathématique serait intéressant s'il existait une nécessité à mettre en lumière des variations précises. Or, la quantification de l'incertitude, de la dispersion des mesures, que l'on cherche par ailleurs à contrebalancer par un grand nombre d'effectifs, n'a d'intérêt poussé que lorsqu'il s'agit de démêler des variations suffisamment petites pour être susceptibles d'être douteuses à l'œil nu.

Allons plus loin, et discutons également le critère selon lequel l'on va considérer qu'il existe une différence entre deux groupes. En RCT, le critère sera comme en détection de signal (historiquement, de signal radar, pour déclencher une alarme), d'être au-delà d'un seuil ou non. On appelle cela le risque de première espèce : la probabilité de conclure qu'il existe une différence due au facteur étudié entre les deux groupes, alors que cet effet n'est qu'un effet dû au hasard. Il va sans dire que ce critère est fixé de manière arbitraire ; en général, il est fixé à 5% : 5 pour cent de chance que ce ne soit pas attribuable à l'effet du facteur. 5 pour cent, cela correspond ni plus ni moins à une chance sur vingt, certainement pas l'idée que se ferait tout un chacun d'un événement rare. Ces 5% sont ce que l'on appelle couramment en recherche le « petit p », ou la *p value*. Une différence entre deux groupes est alors dite

significative statistiquement si le petit p est strictement inférieur à cinq pour cent, par convention.

En outre, cette interprétation des cinq chances sur cent n'est pas nécessairement la plus adaptée, ou vraie, dans tous les cas. Ce serait notamment la plus adaptée si nous ne disposions pas d'un échantillon, mais que nous travaillions en population totale. En outre, ce pourcentage qualifierait davantage une incertitude type "bruit", plus qu'un risque attribuable au hasard.

De manière plus générale, si l'on exclut les interprétations de type bayésien, la discussion de l'inévitable "petit p " doit au minimum s'envisager différemment selon le design de l'étude, ou disons, l'origine des données analysées. Caricaturalement, on pourra considérer que l'approche à la Neyman-Pearson, reposant sur la définition d'une hypothèse (en fait : d'une hypothèse dite nulle, et d'une hypothèse alternative, H_0 et H_1), et détermination d'un seuil de rejet de l'hypothèse nulle, est adaptée aux RCT, où il aura précisément été prévu de tester une hypothèse – par exemple une différence d'efficacité entre deux groupes égaux "toutes choses par ailleurs". La valeur du petit p , dans ce cas, n'a d'importance que dans le sens où elle a été fixée, arbitrairement, comme la valeur au-delà de laquelle on ne veut prendre "le risque" de rejeter l'hypothèse nulle. C'est un seuil, et en tant que tel, il importe seulement de savoir si l'on se situe au-dessus ou en-dessous de la valeur de ce seuil. L'autre possibilité est d'interpréter le petit p selon l'approche dite de Fisher, qui lui donne de l'importance à la valeur du petit p . Dans cette optique, plus le p est petit, plus la différence entre groupes est dite "significative". Cette interprétation n'est pas adaptée, *a priori*, aux études randomisées. Elle convient davantage aux études que nous qualifierons "d'épidémiologiques", davantage au sens d'observationnelles. [93, 94, 95]

Il devient de plus en plus évident que les utilisations de calculs statistiques, en la situation actuelle, mènent à des études dont la qualité ne croît pas. Vraisemblablement, la majorité des études biomédicales avancent des

conclusions fausses étant données les hypothèses initiales, le design mobilisé, les outils utilisés et les interprétations qui en sont faites. [96, 97, 98, 99, 100, 101, 102, 94]

Venons en aux outils utilisés pour quantifier l'incertitude, pour modéliser les mesures effectuées dans le cadre d'une étude. Il existe deux grands cas de figure classiques : la comparaison de moyennes, et la comparaison de proportions. Il existe des tests dits paramétriques et des tests dits non paramétriques. Globalement, les tests paramétriques recouvrent les tests pour lesquels des hypothèses sont faites sur la distribution statistique de la variable mesurée. Dans l'écrasante majorité, il s'agit d'une hypothèse de normalité : la mesure suit une distribution gaussienne ou normale. Plus encore qu'une hypothèse de normalité, il est souvent demandé que la variance des distributions soient égales ou peu différentes. Cette hypothèse est supposée peu restrictive, ce qui, en pratique, ne se vérifie pas si aisément. L'argument fréquemment invoqué, et généralement hors de contexte, est que l'on peut se justifier de l'usage d'une distribution normale en raison du théorème central limite. Il est également à souligner que l'on ne dispose d'aucun test réellement fiable de la normalité d'une distribution, et qu'il est très difficile, à l'œil, de vérifier le caractère effectivement gaussien ou non d'une distribution.

Le RCT est donc une expérience extrêmement étroite, artificielle, et d'application théorique et réelle finalement restreinte. On ne s'y occupe que d'un couple facteur-observation, l'ensemble des autres caractéristiques mesurées étant randomisé et contrôlé. Le plan d'expérience fait pour l'essentiel de la validité des résultats. L'outil statistique n'intervient que dans la perspective de prendre en compte et quantifier l'incertitude et la dispersion.

L'étude observationnelle couvre l'autre côté du spectre, ne répond pas à un plan d'analyse très contraint, et fait porter la responsabilité du contrôle de la

variabilité sur les outils statistiques utilisés. Par son caractère observationnel, elle rend davantage compte de la réalité, puisque capture en situation de vie réelle des données diverses. Elle ne peut évidemment pas se s'affranchir de biais liés au mode de recrutement de la population étudiée, ou à l'observateur. Les outils statistiques qui y sont utilisés sont limités, mal utilisés en pratique, et tentent de répondre à des questions non ou mal posées.

Dans les deux cas, la critique globale qui peut être faite est double : une triple approche (réductionniste, expérimentale et normative) discutable et un manque de généralisabilité des résultats.

3.1.3 L'exemple de la psychiatrie et des DSM – les modèles biomédical et biopsychosocial

Les méthodes dont nous venons de parler sont actuellement les méthodes prévalant en recherche biomédicales au sens large. Elles concernent par ailleurs essentiellement ce que l'on appelle la recherche quantitative, et non qualitative (représentée typiquement par les enquêtes biographiques, les entretiens...).

Nous avons déjà esquissé un premier tableau du domaine de la psychologie, psychiatrie, psychanalyse ou santé mentale. Partagée entre neurologie et psychanalyse, entre symptômes dits objectifs et symptômes rapportés, dits subjectifs, la santé mentale fait office de domaine médical qui se cherche encore, voire en retard sur les autres spécialités. Il est possible, sans doute, de voir le problème sous les deux angles opposés : soit la santé mentale a erré plus longuement que disons, la cardiologie, soit elle a eu la "chance" de ne pas s'engouffrer directement corps et âmes dans une approche purement biomédicale.

En effet, dans la veine de ce que Canguilhem a pu décrire en tant que normativité biologique, les disciplines médicales se sont engagées dans une quête

de quantification de tout : l'existence d'une pathologie, ses degrés de sévérité, les corrélations inter observations, les échelles d'évaluation, les critères diagnostiques, les scores prédictifs...etc. La description clinique étant sans cesse chahutée par la contribution plus ou moins pertinente mais toujours grandissante et insistante des examens dits complémentaires : imageries variées, biologie et marqueurs biologiques, ADN et tests génétiques, analyses moléculaires... Si certaines mesures ont une assise théorique et physiologique bien établie, et sont d'intérêt difficilement contestable, comme par exemple, la pression artérielle ou la glycémie, une vaste partie des analyses, ou du moins, leur utilisation, est davantage douteuse, notamment par manque de cadre interprétatif adéquat.

Ainsi donc, la santé mentale, jusque peu, sous ses airs schizophréniques entre vision somatique pure et vision psychanalytique pure, ne partageait pas cet état de fait avec les autres disciplines. Ce qui, notamment dans le cadre d'une recherche se standardisant et s'internationalisant de plus en plus, posait problème : comment inclure, par exemple, des patients dans une étude si tout le monde n'est pas d'accord sur les critères diagnostiques à retenir ? Il était nécessaire de se mettre d'accord sur un ensemble stable et consensuel de critères et de pathologies – c'est là l'ambition du DSM (*Diagnostic and Statistical Manual of Mental Disorders*) et de ses versions successives. En particulier, un des buts affichés de la troisième version était d'atteindre une fiabilité diagnostique équivalente à celle des autres domaines médicaux.

Les responsables successifs du DSM se sont alors donnés comme objectifs de donner des bases objectives aux catégories diagnostiques recensées en santé mentale. [103, 76, 77, 104, 105, 106] De fait, les chercheurs se sont engagés à la poursuite de mesures biologiques, microbiologiques, de neuro-imagerie, de génétique pour étayer leurs catégories. C'est l'époque des études dites *genome-wide* : on étudie le génome entier de personnes chez

qui on a diagnostiqué un trouble mental particulier, et l'on cherche des associations entre gènes et pathologie. On y découvre ainsi des dizaines de gènes candidats à une pathologie, et la presse grand public n'hésite pas à parler de découverte du gène de la schizophrénie, de la dépression, etc. [107, 108, 109, 110, 111, 112, 113] C'est également la période pendant laquelle il a été mis en évidence que les catégories diagnostiques n'avaient rien de stable ni dans le temps, ni entre observateurs, sans présupposer de la cause de cette instabilité. Pour exemple, il a été montré que sur une dizaine d'années, presque la moitié des personnes chez qui il a été diagnostiqué un trouble bipolaire se voyait diagnostiquer une schizophrénie plus tard, et inversement. Théoriquement, ces deux catégories sont certainement les plus indépendantes et opposées de la classification. [103, 114, 108, 115, 116] D'autres études sont venues entamer sérieusement la possibilité 1) soit que les catégories actuelles soient correctement identifiées, 2) soit qu'il existe un substrat biomédical, en particulier génétique soutenant ces catégories, 3) soit même que l'idée de catégorie plutôt que de continuum soit pertinente.

Assez brutalement, nous sommes donc passés d'un modèle centré sur la personne, et même sur son fonctionnement supposé intérieur, à un modèle biomédical pur, où l'individu devient mesurable et quantifiable en toute objectivité – les troubles mentaux ne sont que le reflet direct de variations génétiques pathologiques ou prédisposant, ou encore d'une infection, d'une pathologie somatique autre, ou simplement d'une configuration neuronale défaillante (en quantité, en qualité).

Le problème répété de l'analyse et de la synthèse : l'analyse est conduite sur des objets restreints, pire, sur des catégories *a priori*, non stables, et la synthèse sur des avis d'experts, spécialisés dans une catégorie particulière.

Récemment, deux auteurs ont renouvelé leur croyance en la nécessité de dresser au mieux la carte des comportements humains, dans leurs inter-

actions, désirant également redonner de la place au rôle des émotions. Ce serait en observant et en analysant la variété de ces comportements, que l'on pourrait revenir à une conception plus saine de ce qui peut être supposé comme pathologique – dans le sens de : nécessitant une intervention, une aide à proposer. Notamment, ces auteurs rappellent que c'est en connaissant la fonction de certaines fonctions physiologiques, leurs limites à ne pas excéder et leur domaine de fonctionnement stable, que l'on peut décider d'un traitement et du type de traitement en médecine non psychiatrique ; ils appellent ainsi à mieux (re)connaître, entre autres, la fonction des émotions dans le fonctionnement humain et social, dans une perspective davantage ressemblant au modèle biopsychosocial. [117, 118]

Quelle que soit l'approche retenue, néanmoins, se posent inlassablement les mêmes questions : classification étanche ou non, floue ou dure, complexité linéaire ou non linéaire qui relie les différentes observations et personnes, quelle doit être, *in fine*, la dimension de notre espace de représentation ? Nous préférierions, dans une optique d'opérationnalité de cet espace, qu'il soit de basse dimension, voire même aisément visualisable, ce qui nous emmène définitivement loin des approches modernes, notamment génomiques. Si l'on schématise, actuellement, les partis pris se répartiraient en deux cas : soit de très grands ensembles de données, dont la mécanique intrinsèque serait essentiellement linéaire et corrélationnelle, statistique, soit des modèles de basse dimensionnalité, mais de dynamique volontiers plus souvent non linéaire. Au final, y a-t-il seulement une approche préférable ou meilleure, ou peut-on un instant même imaginer, d'une certaine manière, que ces deux approches, si proprement menées, puissent être soit équivalentes, soit duales, à la manière de la dualité que nous rencontrons sans cesse – onde et corpuscule ?

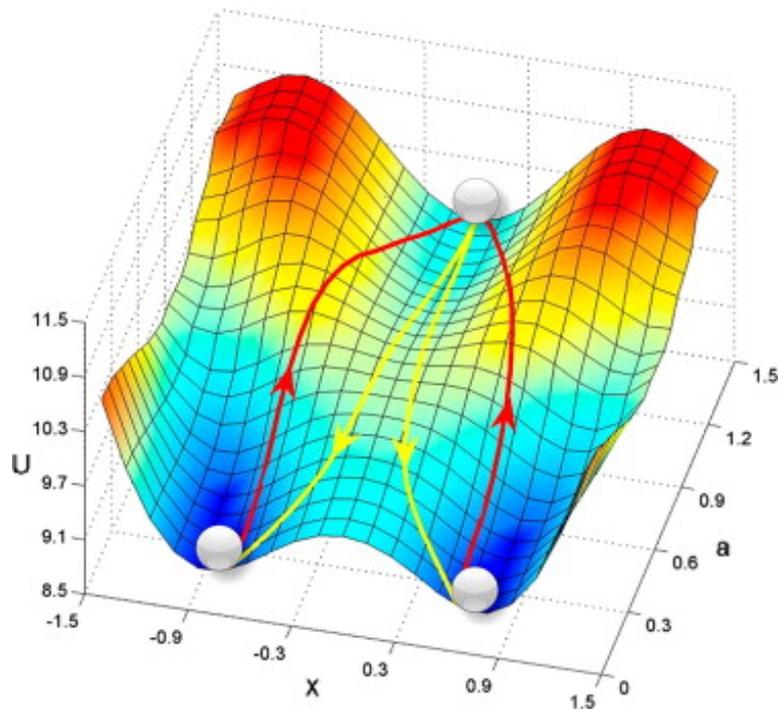
3.1.4 Un pas en avant ? L'*Epigenetic landscape* de Waddington

En 1942, le biologiste Conrad H Waddington proposa de dépasser, ou de compléter, la vision "tout génétique" du vivant en introduisant la notion d'épigénétique. [119, 120, 121, 122, 123] Aujourd'hui, on définit l'épigénétique comme les mécanismes intervenant à l'échelle moléculaire, qui agissent sur le génome et la régulation de l'expression des gènes. Ces mécanismes peuvent être activés par les interactions de l'organisme avec l'environnement, ponctuellement ou tout au long de son histoire individuelle. Ces mécanismes laissent intouché l'ADN lui-même. [123, 124, 125, 126, 127, 128, 129, 130]

Plus spécifiquement, Waddington avait proposé la notion d'*epigenetic landscape* pour rendre compte de l'influence d'événements non génétiques, de nature environnementale au sens large, sur le devenir et l'expression des informations génétiques - sur l'utilisation du code. Initialement, il a décrit ce "paysage" de manière tout-à-fait figurée, en représentant ce qui s'apparente à un terrain avec des reliefs, et des "billes", pouvant rouler, se mouvoir sur ce terrain. L'idée sous-jacente étant de symboliser par un comportement mécanique, des évolutions plus ou moins probables ou réalisables pour une cellule ou un individu donné. Le principe demeure le même que pour toute représentation mobilisant la notion de potentiel, ou de barrière de potentiel favorisant ou faisant obstacle au mouvement d'un mobile – ce mobile bénéficiant d'une inertie particulière (cf figure 3.1).

Si l'approche originelle peut paraître relativement métaphorique, elle pour elle l'avantage de ne pas restreindre l'évolution d'un organisme à son supposé déterminisme génétique, ni ses possibilités à la nature de son code. À ce stade, il serait même difficile de préciser quelle serait la part relative, si cela a un sens, la plus importante dans l'évolution d'un organisme, de

FIGURE 3.1 – Représentation schématique de l'*epigenetic landscape* de Waddington (1)



son code génétique ou de ses interactions avec l'environnement. L'apport de Waddington est théoriquement et potentiellement considérable, dans ce sens où il propose une approche intégrative et systémique.

À l'instar des modèles de Turing, les premières études s'intéressant à tester ou reconstituer un paysage épigénétique pour un organisme donné ont tardé à apparaître. Par ailleurs, si l'on commence à voir quelques essais, lesquels se voient en général ouvrir les colonnes des revues les plus prestigieuses, ceux-ci restent très frileux dans le panel d'événements qu'ils pourraient intégrer. Les études restent très ancrées sur l'échelle moléculaire, plus particulièrement sur les phénomènes de méthylation des histones (cf figures 3.2 et 3.3). Là où Waddington ouvrait la voie à une révolution quant à la manière de penser la génétique et l'évolution du vivant, l'apparition de moyens techniques permettant de tester et reconstruire de tels paysages n'a pas suffi, pour le moment, à faire penser réellement différemment les chercheurs. Nous restons dans des perspectives d'élargissements incrémentaux, pour ne pas dire infinitésimaux, de la recherche dans le domaine, de la même façon qu'après le séquençage du génome, et son impasse actuelle, nous sommes remontés d'un cran en échelle et en modalité, pour s'intéresser au protéome, et de suite, à tous les "*omics*".

De manière remarquable, un auteur s'est très récemment intéressé à l'impact présumé du modèle de Waddington, et a pu mettre en relation ce modèle avec ceux, par exemple de Thom. [131] Autre exception trop rare, des auteurs se sont - superficiellement - intéressés à la définition des pathologies rhumatologiques sur la base d'un paysage épigénétique (cf figure 3.4). [132]

FIGURE 3.2 – Représentation schématique de l'*epigenetic landscape* de Waddington, cellules souches et attracteurs, *Science*

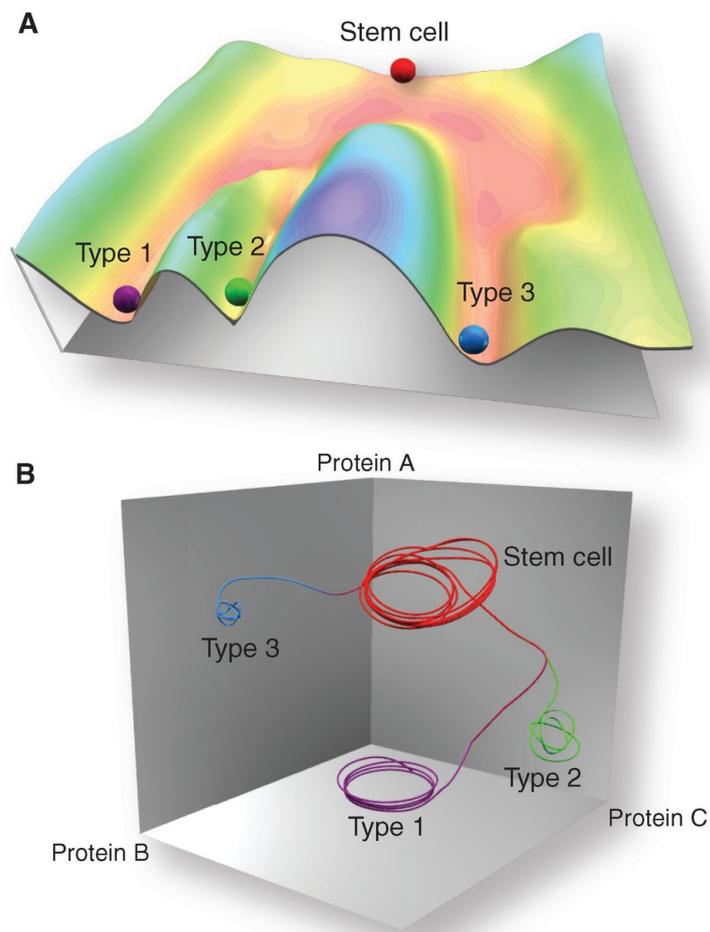
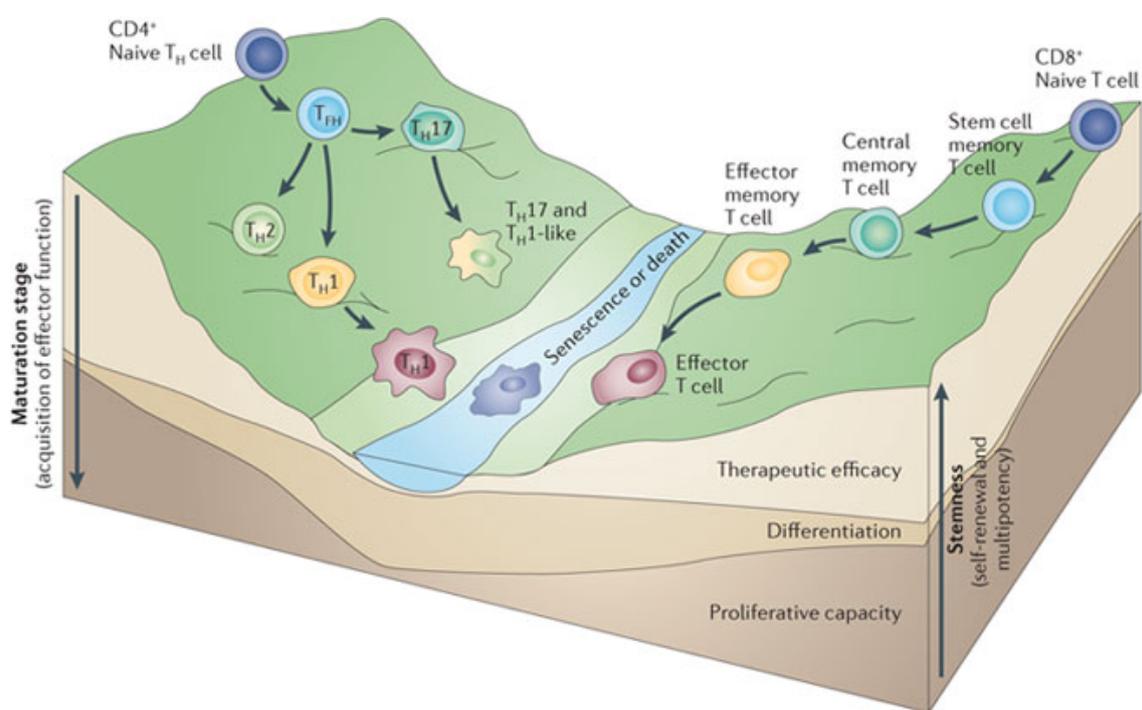
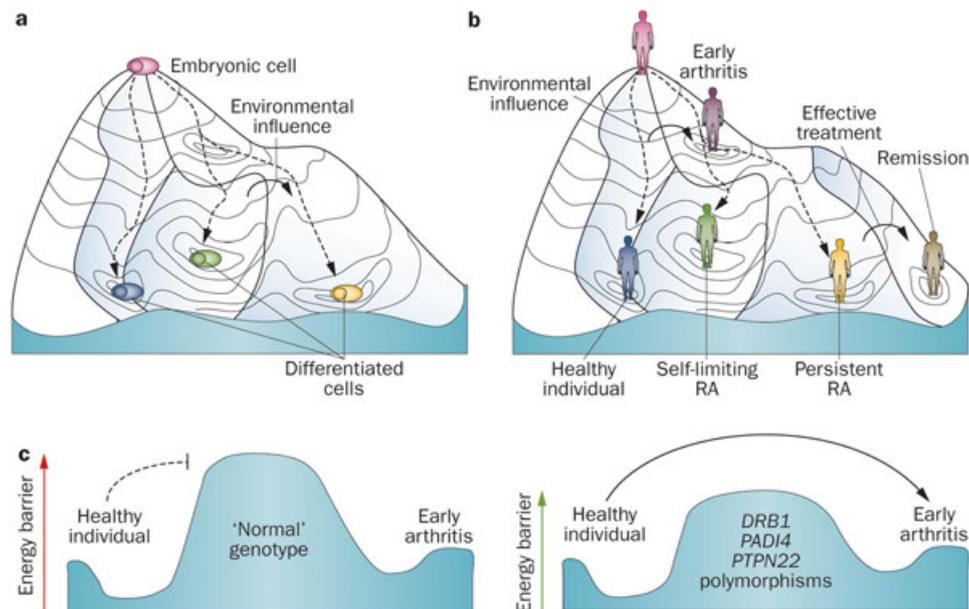


FIGURE 3.3 – Représentation schématique de l'*epigenetic landscape* de Waddington, cellules T, *Nature Reviews*



Nature Reviews | Cancer

FIGURE 3.4 – Représentation schématique de l'*epigenetic landscape* de Waddington (affections rhumatismales, *Nature*)



3.1.5 Rechercher la forme et les invariants – méthodes de *clustering*

Dans le cas de l'*epigenetic landscape*, les objets sont vus comme des états possibles d'un même objet initial, et reliés dans le temps par des trajectoires continues. Ces objets sont identifiables assez aisément, car vus comme des objets morphologiquement, géographiquement et fonctionnellement différents à la même échelle. Ces objets ont une forte présomption d'existence, même si leur existence n'exclut pas l'existence d'autres états non identifiés jusque là. L'autre atout, tient dans l'identification visuelle, par le biologiste – encore le rôle ici du microscope.

En toute généralité, il peut être compliqué d'identifier des objets à partir de mesures variées, sans les qualifier plus ou moins arbitrairement par une observable unique plutôt qu'une autre. C'est particulièrement vrai si l'on

cherche à identifier des objets invariants qui soient caractérisés à partir de mesures séparées, mais reliées de façon complexe en réalité.

Il existe tout un ensemble de techniques, en traitement du signal ou en traitement de données - en *data mining* ou fouille de données - dont le but est précisément l'identification de groupes au sein de ces données. Le problème général est alors le suivant : étant donné un ensemble d'observations, existe-t-il des groupes homogènes, au sens d'un sous ensemble de ces observations, et si oui, lesquels sont-ils ? Les techniques de fouilles de données répondant à ce type de problème sont appelées techniques de *clustering*, de *pattern recognition* ou encore de classification. [133, 134, 135] Les experts divergent d'avis sur le meilleur terme à adopter selon les situations ; pour ce qui nous concerne, nous nous contenterons de prendre pour équivalentes ces dénominations.

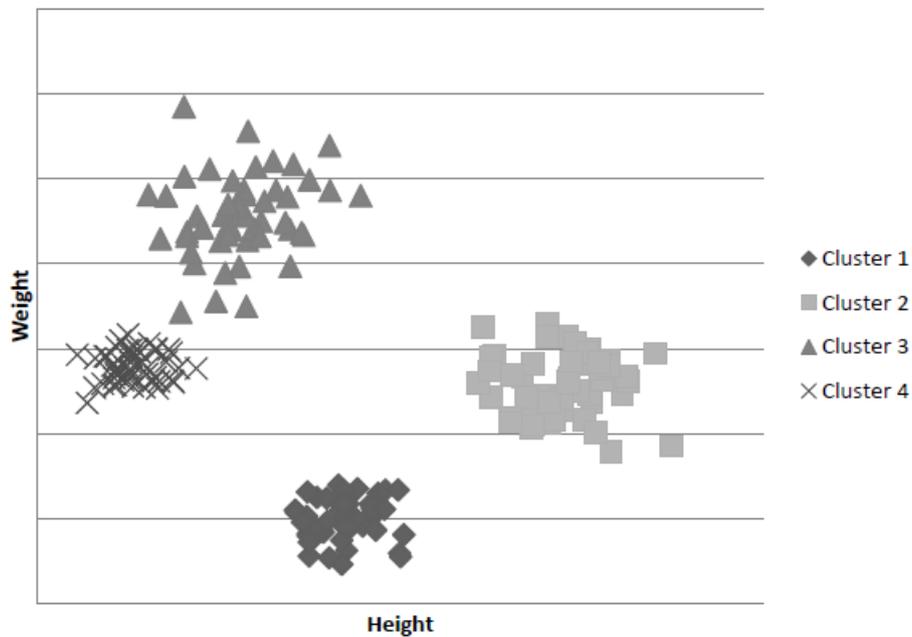
Bien souvent, cependant, la fouille de données se borne régulièrement à ne répondre qu'à la deuxième partie du problème : on cherche dans les données des groupes les plus homogènes possibles. Cette pratique n'est pas sans poser de potentiels problèmes, en particulier dans une démarche de recherche de preuves de l'existence de tels groupes ou formes. Il est en effet virtuellement toujours possible d'appliquer un algorithme de *clustering* à un ensemble de données, et que celui-ci produise un résultat, donc des groupes. En dehors du fait de savoir si ces groupes sont correctement délimités et respectent la "réalité" des formes à retrouver, ainsi que le nombre correct de ces groupes (il n'est pas dit, par exemple, que l'échantillon sur lequel on travaille est représentatif de tous les groupes existants), se pose avant tout la question l'existence : les groupes découverts dans les données peuvent aussi bien être de purs artefacts. Pour pallier ce problème, divers indices ont été proposés, le plus souvent totalement heuristiques, mais s'avérant à l'expérience plutôt pertinents, et qui donnent une idée du degré d'artifice des groupes découverts

dans les données.

Nous ne décrivons pas ici d'algorithmes de *clustering*, ceux-ci étant très nombreux et relativement variés. Néanmoins, nous désirons en expliquer les principes et limites communs à la majorité d'entre eux. La plus part des techniques souffrent de deux grandes faiblesses : il faut spécifier à l'algorithme le nombre de groupes à identifier (mais certains se proposent de les déterminer automatiquement), et il est nécessaire, implicitement ou explicitement, de se donner un principe de similarité ou de dissimilarité - c'est-à-dire qu'il faut préciser la façon dont on caractérise la ressemblance entre deux individus, leur distance relative dans l'espace de description. En général, une distance est utilisée, par exemple la distance euclidienne ou la distance de Manhattan. Cette distance règle la géométrie de l'espace, et peut ainsi produire des résultats différents pour un même algorithme, si la nature des données est telle qu'elle relève davantage d'une distance que d'une autre. Pour illustration, il suffirait de penser à la représentation d'un trafic routier urbain. Si l'on veut mesurer les distances entre deux points d'une ville dont les rues sont perpendiculaires les unes aux autres, et que l'on n'est pas un oiseau, la réalité veut que les distances les plus vraisemblables soient des sommes de segments (les longueurs des rues à parcourir), et non les distances "à vol d'oiseau". Dans ce cas, la distance adaptée est la distance de Manhattan, et non la distance euclidienne.

Le choix d'un algorithme plus qu'un autre, et plus généralement, le choix d'une mesure particulière de similarité, va contraindre plus ou moins les formes possibles des groupes reconstruits. Dans le cas des algorithmes apparentés à l'algorithme dit des k moyennes, on sait par exemple que les groupes auront globalement des formes sphéroïdes. D'autres classes d'algorithmes sont plus opérationnels, reposant sur des principes non linéaires (cartes auto-adaptatives, SVM ou *Support Vector Machine* [135, 136]) lors-

FIGURE 3.5 – Exemple simple de groupes à identifier par différents algorithmes de clustering (taille-poids)

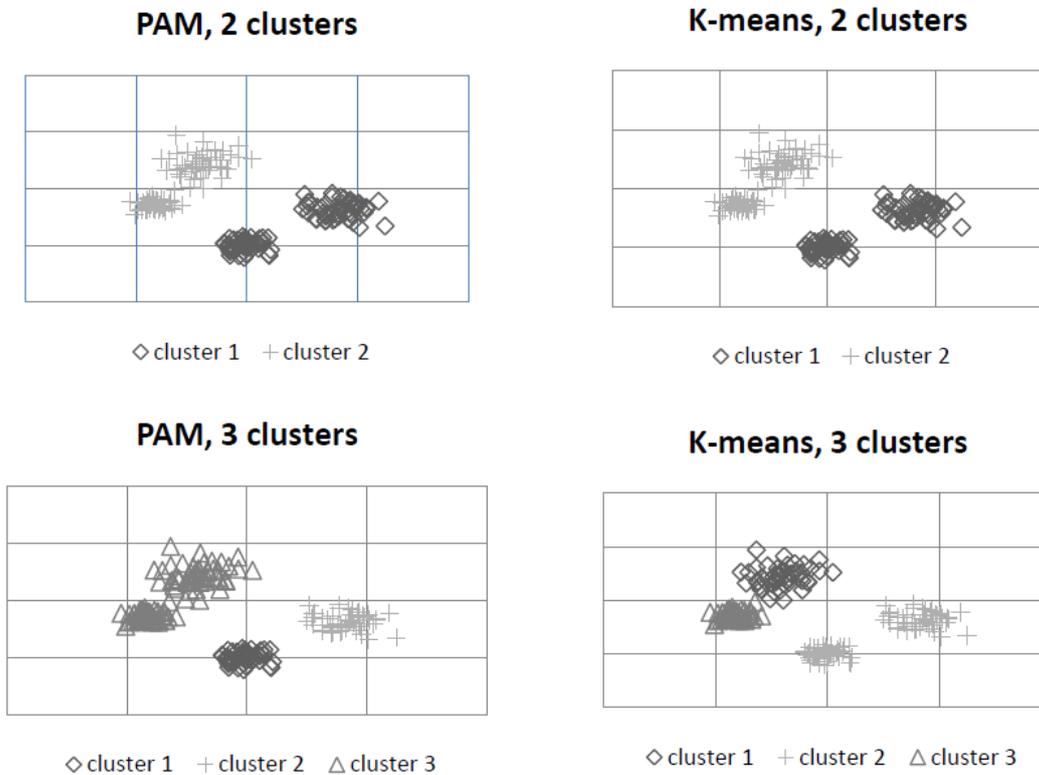


qu'il s'agit de démêler des formes plus complexes ou un peu interpénétrées (mais non partiellement recouvrantes).

Afin d'illustrer l'intérêt d'une approche par les techniques de *clustering* dans la découverte de formes stables au sein de données, nous prenons ici un exemple artificiel et simple. Il s'agit de la simulation des mesures de taille et poids de 4 groupes différents – ce qui se matérialise dans notre cas par le tirage aléatoire de réalisations selon 8 lois normales de moyennes et variances distinctes (une loi pour la taille, une loi pour le poids, pour 4 groupes, voir figure 3.5).

Par construction, il existe bien un "mécanisme" générant 4 formes différentes. Le but est alors ici, par l'application de différentes techniques de *clustering* sur le jeu de données, de voir comment certains indices nous ren-

FIGURE 3.6 – Résultats de différents algorithmes de clustering appliqués à un exemple simple, en fonction du nombre de groupes cherchés

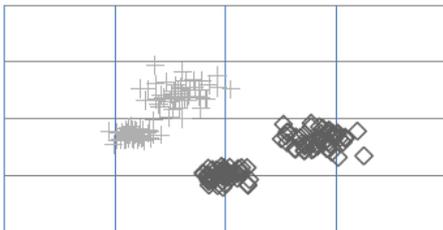


seignent sur la caractère d'existence de ces groupes (indice dit de la silhouette), [137] mais aussi sur le nombre "optimal" de groupes existants dans les données, notamment *via* une approche dite de la robustesse des groupes.[138] Les résultats sont donnés par les figures 3.6 et 3.7, les indices étant donnés dans les figures 3.8 et 3.9.

Il apparaît ainsi que pour un même jeu de données, l'application de différentes techniques et l'utilisation d'indices d'existence et de recherche du nombre de groupes mène à des résultats cohérents et globalement convergents. Deux remarques à ce stade : il faut bien entendu garder à l'esprit ici d'une part le caractère artificiel de l'exemple, d'autre part le fait que les

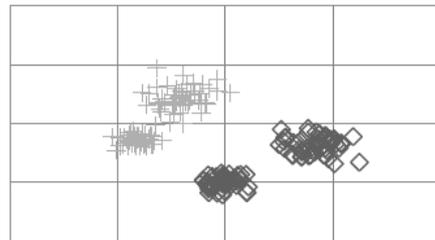
FIGURE 3.7 – Résultats de différents algorithmes de clustering appliqués à un exemple simple, en fonction du nombre de groupes recherchés (suite).

PAM, 2 clusters



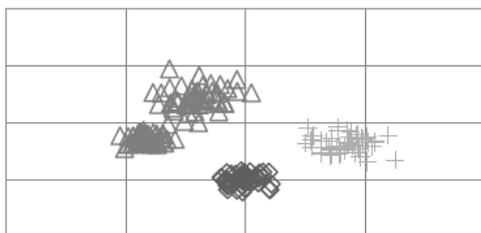
◇ cluster 1 + cluster 2

K-means, 2 clusters



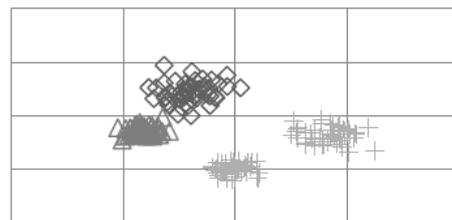
◇ cluster 1 + cluster 2

PAM, 3 clusters



◇ cluster 1 + cluster 2 △ cluster 3

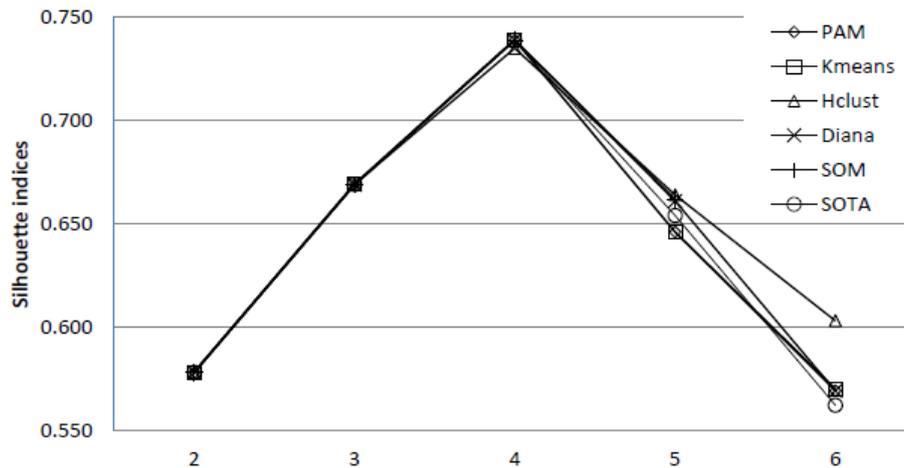
K-means, 3 clusters



◇ cluster 1 + cluster 2 △ cluster 3

FIGURE 3.8 – Valeurs de la silhouette pour différents algorithmes de clustering

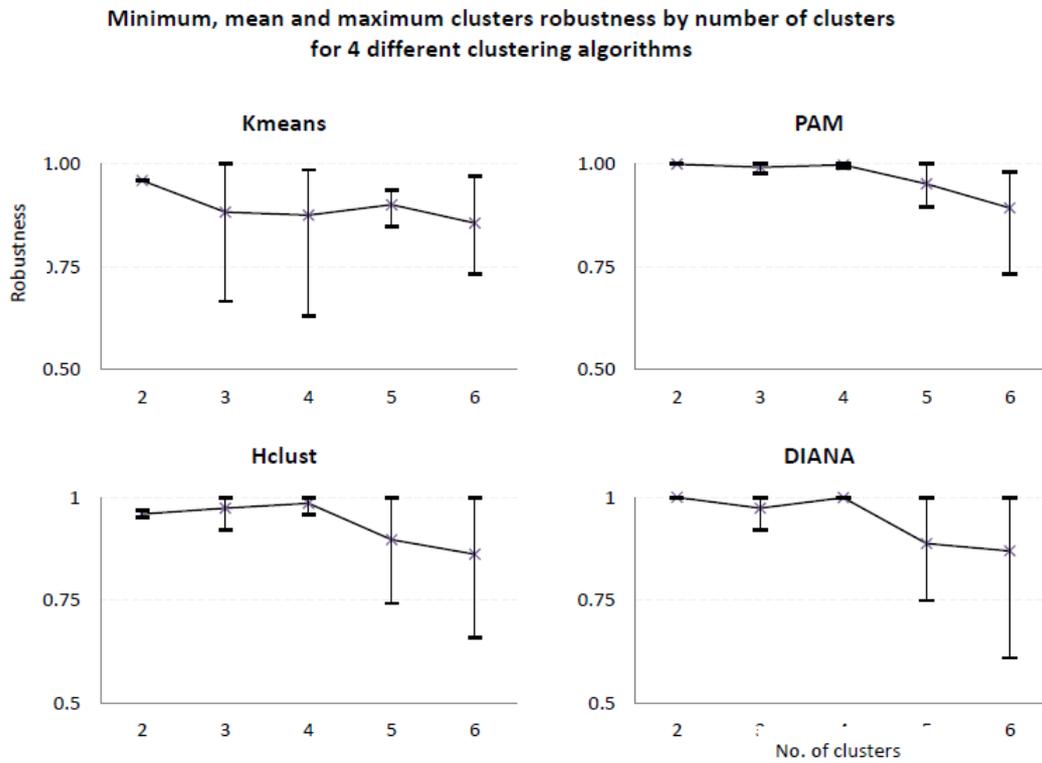
Silhouette indices by number of clusters according to 6 different clustering algorithms



groupes soient relativement bien séparables ; par ailleurs, notons que la notion de forme peut être soumise à subjectivité et à une notion de d'échelle : est-il fondamentalement faux de considérer qu'il puisse n'exister que deux groupes, à une échelle donnée, plutôt que 4 ? Nous savons que 4 groupes ont été générés par 4 couples de lois de même forme mais de paramètres distincts, et par conséquent, nous désirons identifier 4 groupes. Pour autant, si l'on oublie le sens que l'on a artificiellement attribué aux mesures pour les besoins de l'exemple (poids et taille), il est tout à fait concevable que le même jeu de données, dont on ignore tout des mécanismes de production sous-jacents, puisse être interprété comme recelant 2 groupes, eux-mêmes décomposables éventuellement en 2 sous groupes.

Malgré leurs limitations, les techniques de *clustering* demeurent des techniques largement sous exploitées, notamment en domaine biomédical. Pourtant, elles proposent des approches naturellement multivariées, sans les *a*

FIGURE 3.9 – Valeurs des robustesses de classe pour différents algorithmes de clustering



PAM: Partitioning Around Medoids; Hclust: Hierarchical Clustering; DIANA: Divisive Analysis;
SOM: Self Organizing Map; SOTA: Self Organizing Tree Algorithm

priori forts sous-tendant les approches statistiques classiques (et très rarement validées, sinon même vérifiées, ou vérifiables), et sont un premier pas intéressant vers la reconstruction possible de réalités complexes, aux observations interdépendantes, sans hypothèse sur la nature de leurs relations. Les faiblesses de ces techniques pourraient en outre être largement atténuées par le raisonnement qui prévaut partout ailleurs, à savoir qu'une seule expérience ne fait et ne doit pas être prise comme preuve définitive d'un fait scientifique. Aussi, les résultats obtenus par l'application de techniques de *clustering* peuvent tout autant bénéficier et d'une reproductibilité sur d'autres données similaires, et de confirmation ou d'infirmité par d'autres études, mobilisant d'autres techniques. Inversement, l'expérience montre qu'une analyse par *clustering* de données par exemple biomédicales, par le biais de différents algorithmes, est très souvent et largement soutenue par des analyses statistiques classiques ultérieures - par exemple, les tests de comparaisons de moyennes ou de proportions classiques entre individus de deux groupes différents et identifiés par *clustering* sont la majeure partie du temps très significatifs ; or, ces différences n'ont pas été postulées *a priori*, mais découvertes naturellement, et enfin, en relation avec d'autres observations, non de manière isolée.

Il n'est pas exact, en dernier ressort, de dire des techniques de *clustering* ou de reconnaissance de forme, qu'ils opèrent sans *a priori* sur les données. Certes, la majorité s'affranchissent d'hypothèses restreintes sur la distribution de telle ou telle variable, comme l'hypothèse de normalité si chère à de nombreuses approches statistiques, et les algorithmes sont essentiellement catégorisés selon le type de codage de l'observation : selon qu'elle soit catégorielle, ordinale ou continue. On sait également que les algorithmes de *clustering* sont plus généraux que les approches statistiques usuelles - notamment, il est possible de montrer qu'un démêlage de gaussiennes (c'est-

à-dire, chercher à identifier dans un ensemble de données, des profils qui correspondent à des distributions normales, de moyennes et potentiellement de variances différentes) correspond à une implémentation particulière des k moyennes (plus spécifiquement, à l'aide d'une descente de gradient).

Toutefois, une technique doit reposer sur un ou des principes particuliers. D'une façon générale, une grande partie des techniques est fondée sur la notion de similarité, ou de façon complémentaire, de dissimilarité : des personnes sont classées ensemble selon leur ressemblance. La notion de similarité est elle-même basée soit sur des mesures *ad hoc*, créées à cet effet, soit sur des distances, connues par ailleurs (distance euclidienne, de Manhattan, distance infini...etc. Distance est bien pris ici au sens de norme, mathématiquement parlant, ce que ne sont pas nécessairement des mesures de similarités).

Associé à cette notion de similarité, il existe un principe de maximisation ou de manière équivalente de minimisation d'une fonction de coût ou d'utilité, locale ou globale, qui sert de critère d'arrêt à l'algorithme. *In fine*, la géométrie des objets identifiés sont dépendants de ces principes. Par exemple, le classique algorithme des k moyennes implique de manière générale une géométrie sphérique, isotrope des objets retrouvés. Ce type d'algorithmes est incapable, dans l'espace d'opération, d'identifier des groupes potentiellement entremêlés, donc des objets aux frontières complexes dans cet espace de configuration donné. La question de la géométrie "juste" des objets à identifier se pose donc.

Pour finir, nous avons déjà évoqué les deux faiblesses principales qui peuvent être reprochées aux techniques de *clustering*, à savoir leur dépendance à la mesure de similarité et la nécessité de préciser le nombre de groupes à retrouver. Il existe des techniques pour s'affranchir de cette dernière faiblesse, comme l'approche par la robustesse des groupes. [138] Nous devons cepen-

dant mentionner deux autres écueils possibles et fréquents : la sensibilité aux conditions initiales, et la possibilité de minima locaux. Ces derniers sont l'apanage de certaines approches de problèmes d'optimisation, et nous ne nous attarderons pas dessus. Quant aux conditions initiales, le problème est bien entendu contournable par l'exécution répétée de l'algorithme sur les mêmes données, mais par exemple présentées aléatoirement d'une itération à l'autre. Cette dernière remarque nous amène précisément à un aspect important des techniques de *clustering* : ces approches sont d'autant plus précieuses qu'elles sont, au final, des approches tant quantitatives que qualitatives. Elles procurent des formes, des invariants des données recueillies, et autorisent également des calculs.

Ces techniques s'avèrent des techniques de choix pour la caractérisation des données par leurs formes homogènes. En ce sens, et ce d'autant que certaines techniques sont des techniques d'analyse spectrale, elles peuvent fournir l'équivalent des états fondamentaux ou états propres que nécessitent les théories d'inspiration quantique, afin de structurer leur espace.

3.1.6 Géométrie et dynamique

Partant des techniques prévalant en sciences biomédicales, les RCT devant les autres, les études observationnelles les suivant désormais de près, nous en avons montré les limitations quand il s'agit de rendre compte d'une réalité plus complexe, et de conduire une synthèse non triviale des faits. Les techniques de reconnaissance de forme sont nombreuses et d'efficacité variable, et reposent essentiellement sur des hypothèses géométriques.

Les RCT ont pour eux de tendre vers une forte présomption de causalité entre deux observations, sans mention explicite du temps comme médiateur de la causalité. Cette causalité est néanmoins réduite à un horizon très restreint, puisqu'elle ne renseigne pas sur sa validité lors de variations

conjointes d'autres facteurs, et sous entend une indépendance entre facteurs. Disons qu'un RCT peut apporter une preuve de causalité locale, dans une approximation linéaire et point à point (une cause, un effet). A l'opposé, les approches par la forme permettent de rendre compte d'une réalité plus riche, multivariée, certainement plus proche de la réalité et plus respectueuse de toute complexité sous-jacente. Ce que l'on gagne en description d'un objet plus complexe et invariant, on le perd en causalité et en principe d'évolution temporelle. Il manque donc ici certainement une étape intermédiaire, une sorte d'hybride qui permette dans le même mouvement d'identifier des objets invariants et complexes, stables, et leur principe d'évolution, selon un schéma causal.

D'une certaine manière, l'intuition de Waddington, qui rejoint celle de D'Arcy Thompson et d'Einstein quant à un principe géométrique régissant l'existence et l'évolution des formes, semble nous indiquer la voie vers un tel hybride. Cependant, dans l'utilisation qui en est actuellement faite, le concept de paysage épigénétique privilégie l'approche probabiliste et "absolue", dans le sens où elle entend relier des objets identifiés par ailleurs (différents types de cellules, par exemple). La (re)construction du paysage tient alors davantage de la construction des chemins reliant ces objets, que de se donner *ab initio*, des règles d'évolution permettant d'expliquer ou de prédire l'existence et la forme, la fonction d'objets non connus par ailleurs. De fait, l'approche actuelle demeure intéressante, mais pourrait souffrir de limitations liées à son usage plus qu'à son concept initial.

3.2 Quelles alternatives pour prendre en compte la complexité ?

Comme nous venons de le discuter, la majeure partie des études et dispositifs expérimentaux actuels visent à l'analyse, sous contraintes parfois peu

réalistes, et au mépris de toute complexité entre objets élémentaires isolés. Les enjeux au-delà de l'analyse sont doubles : être capable de préserver la complexité lors de l'analyse, et de la restituer lorsque l'on en fait la synthèse. L'étape suivant ce respect de la complexité serait de s'intéresser à la validité externe des résultats issus de l'analyse : comment passe-t-on d'un contexte à un autre tant soit peu différent ? Peut-on même le faire ?

3.2.1 De l'absence de définition consensuelle de la complexité

Ce dont on peut être certain, semble-t-il, est que la complexité n'est pas la complication. La complication, le caractère compliqué d'un problème, tel que vu à l'aune du réductionnisme, tient davantage dans une complexité qui serait purement extensive : il existe beaucoup de parties élémentaires différentes à assembler pour reformer le problème initial. Cela s'apparenterait à un problème de ressources mémoires insuffisantes, en tant que clones d'une même unité de mémoire répliquée n fois, ou d'unités de traitements trop peu nombreuses.

Plusieurs définitions de la complexité ont été proposées. L'une de celles les plus connues étant certainement la complexité algorithmique.

La complexité algorithmique concerne essentiellement les approches visant, à la Turing à répondre de manière mécanique à certaines questions. Ces questions sont en général des questions soit de décisions (de réponse oui/non), soit d'existence ou de recherche de solutions spécifiques (de type, rechercher tous les éléments vérifiant une propriété donnée). Le premier problème tient à savoir si une telle réponse peut être obtenue, en théorie comme en pratique, et si oui, avec quelle efficacité relative. Nous disons : efficacité relative, car il s'agit de déterminer la quantité de ressources en temps et en espace nécessaire à l'obtention de cette réponse – le temps et l'espace nécessaires peuvent ainsi dépendre du support matériel d'exécution, en pratique.

De fait, la complexité d'un problème revient à demander quelle quantité de mémoire sera nécessaire, et quelle quantité de temps. Il existe alors des ensembles différents pour qualifier la complexité d'un programme, selon la relation liant temps d'exécution et taille du problème, par exemple. On distingue ainsi les problèmes de temps d'exécution logarithmique en la taille, polynomial, exponentiel, etc.

On rencontre également la notion de complexité au sens de Kolmogorov, qui elle aussi s'appuie sur la notion d'algorithme et de machine de Turing. Néanmoins, son approche est sensiblement différente. Il s'agit ici de qualifier la complexité du problème selon la taille de la plus petite machine de Turing nécessaire pour formuler ce problème.

Dans un sens moins algorithmique, et plus conceptuel, la notion de complexité peut se trouver, plus ou moins bien explicitée, au coeur de différentes disciplines ou concepts, comme les systèmes complexes, les phénomènes dits d'auto-organisation, d'émergence... La complexité n'est cependant pas identifiable à la quantité brute des éléments du système que l'on considère. Elle met souvent l'accent plutôt sur la quantité et la qualité des inter-relations, que sur les éléments eux-mêmes. Elle va à contre-sens de toute approche isolationniste. On peut notamment se référer aux développements d'E. Morin, qui y a consacré plusieurs volumes de son oeuvre, reprenant à son compte le titre de Descartes, *La méthode*. [139]

En sciences physiques, on pourrait être tentés également d'assimiler du moins partiellement, complexité et prédictibilité d'un système : le comportement d'un système chaotique sera volontiers qualifié de complexe dans le sens où, malgré une formulation initiale simple, les trajectoires de ce système peuvent devenir totalement imprévisibles, et ce très rapidement, sans pour autant qu'il n'y ait aucune régularité, ou bien que l'on puisse qualifier le système de purement aléatoire. La complexité n'est donc pas opposée au déterminisme, mais plus sûrement à la linéarité.

Quelle que soit la définition retenue, il est frappant – et rassurant ? – de constater qu’une telle complexité émerge précisément de la conjonction entre des qualités élémentaires qui peuvent paraître très basiques, et des contraintes que nous qualifierons de pratiques : les contraintes en temps et en espace – vraisemblablement, de manière quasiment équivalentes, en information ou en énergie – et la capacité pour le système de s’écrire et de lire de l’information de manière dynamique.

Quoi qu’il en soit, lorsque l’on s’attaque à un problème dont rien ne laisse supposer qu’il soit simple et linéaire, il paraît nécessaire de respecter cette complexité sous une forme ou une autre.

3.2.2 Conserver la complexité

Dans une optique où l’on ne sait rien des liens qui unissent les différentes observations faites d’un système – lequel système est arbitrairement défini par l’observateur et ses capacités de mesure – ni des principes d’organisation qui peuvent les régir, il doit toutefois demeurer possible de manipuler les données sans en perdre la complexité intrinsèque.

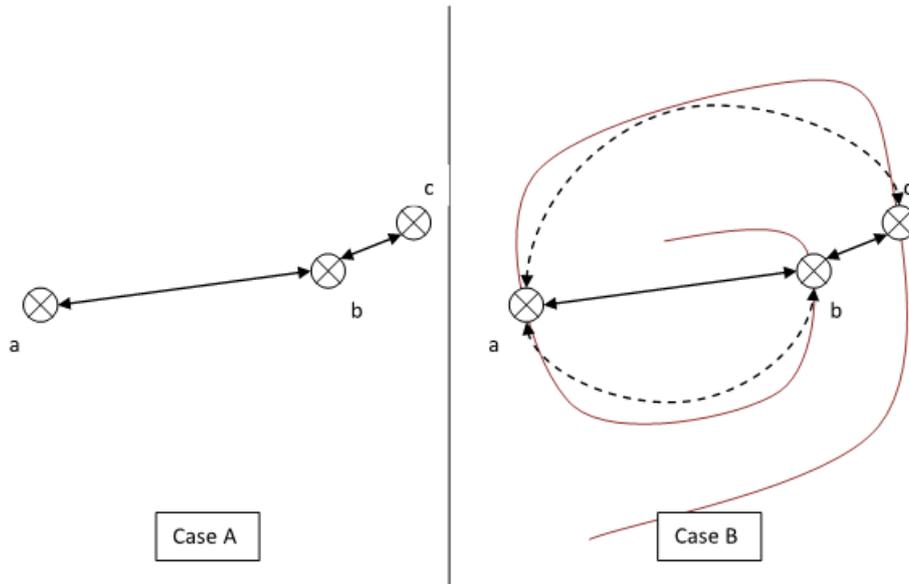
Un exemple simple tient dans notre incapacité actuelle à identifier, automatiquement, des corrélations entre variables qui ne soient pas des corrélations linéaires. Récemment, des approches ont été proposées pour pallier cette limitation, et détecter automatiquement des corrélations qui ne soient pas uniquement linéaires. Cependant, les auteurs restent prudents, et considèrent que ce qui devient possible l’est essentiellement en raison de l’accroissement des capacités de calcul. [140, 141] L’approche proposée est une approche basée sur l’information, et a été nommée MIC, pour *Maximum Information Criterion*. Sans surprise, il repose sur des considérations empruntées à la théorie de l’information, que nous ne détaillerons pas ici. De fait, ce type d’identification d’associations entre variables, qu’il est encore

nécessaire de pleinement qualifier afin de comprendre ce que l'on manipule et ce que l'on peut en attendre, par sa capacité à retrouver les corrélations linéaires que l'on sait déjà traiter, mais également d'autres types de relations, peut être considéré comme une méthode préservant un certain type de complexité. Elle préserve la complexité des données dans le sens où elle ne traite pas les associations de paires de variables comme nécessairement linéaires, notamment sur toute la gamme de leurs valeurs possibles – il n'est en effet pas exclu que, par "morceaux", il existe des associations linéaires entre variables, localement, alors que globalement, leur relation n'a rien de linéaire.

Il peut être intéressant de conserver un autre type de complexité – qui n'exclut pas le type de complexité dont nous venons de parler – que nous appellerons une complexité topologique, une complexité de voisinage. En effet, une question récurrente et de fond, est de savoir, étant donnés plusieurs individus, dans quelle mesure il est possible d'affirmer que tel individu A est plus proche de tel individu B , plutôt que d'un individu C – et ce, seulement pour un certain ensemble de caractéristiques définissant ces individus. Le problème peut également se formuler de manière complémentaire : dans quelle mesure différent-ils ? Sous réserve de se doter du "bon" espace ou sous-espace de représentation pour positionner ces individus, en l'absence d'information sur les relations qu'entretiennent les variables définissant les individus, il est difficile de deviner la structure particulières des données.

Comme l'illustre la figure 3.10, dans le cas où l'espace ambiant est "plat", les distances se font à vol d'oiseau, et il est simple de dire qui de a ou c est plus proche de b (cas A). C'est la technique adoptée couramment, en aveugle. Cependant, dans le cas B , l'espace est structuré comme une spirale, et les relations de voisinage s'en trouvent modifiées. Ce qui était vrai dans le cas A ne l'est plus nécessairement dans le cas B . Nous verrons plus loin comment il peut être possible de respecter cette complexité topologique.

FIGURE 3.10 – Complexité topologique : selon la structure des données, comment savoir lequel de a ou c est plus proche de b ?



3.2.3 Restaurer la complexité

Préserver la complexité, quelle qu'elle soit, constitue une étape souhaitable. Elle n'est pas suffisante dans tous les cas. En effet, si l'on considère que les données observées ont été manipulées, éventuellement transformées et réduites de telle sorte que leur complexité a été correctement préservée, qu'en est-il de la suite des opérations ?

Au moins deux grands problèmes se posent à nous : d'une part, préserver la complexité est rassurant, mais ne nous renseigne pas sur la qualité de cette complexité, sur la forme des relations liant différentes variables, par exemple ; d'autre part, si l'on désire résumer l'information de données, de telle manière qu'elle soit intelligible ou visualisable, ou encore qu'elle puisse se modéliser afin de fournir un modèle prédictif performant, sur quels indices, s'ils existent, allons-nous nous reposer pour nous guider ?

Notamment, pour prendre un exemple relativement concret et direct,

comment déterminer le nombre de variables (d'informations) nécessaires et suffisantes pour qualifier notre système ? Et sous l'hypothèse que ce nombre soit connu, quelles opérations effectuer sur les variables nous occupant afin de récupérer les "vraies" variables pertinentes ? Les statisticiens ont en général une réponse à proposer pour ce type d'interrogation, en particulier *via* les analyses en composantes principales et assimilées. Là aussi, nous y reviendrons plus loin, et montrerons combien cette réponse peut s'avérer insatisfaisante.

3.2.4 Le choix du modèle

Au final, les deux questions que nous venons de nous poser : préserver et restaurer la complexité, mènent à ce troisième problème qu'est le choix du modèle de représentation des données. Même sous des hypothèses les plus faibles possibles, toute représentation de l'information sous-entend l'utilisation d'un modèle, de postulats particuliers. L'un des dilemmes récurrents est celui de savoir si l'on doit observer, échantillonner l'information, sous l'hypothèse d'observations continues, ou discrètes, et dans ce dernier cas, si l'encodage de ces données doit être catégoriel, ou ordonné, par exemple. Il est d'usage de vouloir traiter ces différents types d'observations avec des techniques différentes et appropriées, déterminées par leur domaine d'application et de validité.

Ce type de choix peut paraître trivial dans certains cas – la conception des techniques qui leur sont dédiées, déjà moins. Cependant, d'autres choix de modélisation peuvent rapidement apparaître si l'on ne se cantonne pas à une approche qui nous est familière ou préférée. Par exemple, la modélisation la plus courante actuellement lorsque l'on manipule des données, notamment des données d'enquête ou d'études, est d'appliquer un traitement statistique, ou stochastique si l'on considère une variable d'évolution temporelle. Le postulat, selon le modélisateur, sera de considérer que la bonne façon de ma-

nipuler ces données est d'incorporer la notion de probabilité. Soit que cette probabilité sert à pallier un manque d'information – hasard de modélisation – soit que cette probabilité sert à modéliser un comportement aléatoire intrinsèque. Dans chaque cas, le hasard est présent et traité comme tel : une composante prévisible uniquement en loi.

Pourtant, il est possible de supposer qu'un même système observé, dont le comportement nous paraît aléatoire, ne l'est pas en raison d'un hasard irréductible, mais est "programmé", de manière déterministe, par un système mécanistique qui peut parfois même s'écrire extrêmement simplement, et avec peu de variables. C'est le cas des systèmes exhibant un chaos déterministe, comme le système de Lorenz que nous avons présenté plus haut. Si ce système est observé par des mesures échantillonnées, on pourra penser que son comportement est très aléatoire, même s'il est éventuellement confiné dans une région de l'espace. Pourtant, le support des échantillons sera bel et bien une trajectoire continue, déterministe. Quel est, dans ce cas, le "bon" modèle ?

3.3 Causalités, temps et forme

Nous avons déjà abordé le lien qui unit causalité, temps, et forme. Nous nous proposons de développer ici un peu plus avant ces notions.

3.3.1 Causalité linéaire, proportionnalité directe et boucle ouverte

Un monde simple voudrait qu'à chaque effet - sous la condition, comme pour la question de la forme, que l'on puisse isoler un effet "pur" - il puisse correspondre une cause directe et proportionnelle, linéairement liée à l'effet. Le réductionnisme y trouverait alors toute sa justification, puisque la division des tâches complexes en sous ensembles de tâches élémentaires permettrait

d'étudier ces relations de causalité simples. La reconstitution du problème complexe s'accomplirait ainsi selon un principe de superposition des relations causales élémentaires : à une combinaison linéaire de causes correspondrait une combinaison linéaire d'effets, chacun des effets étant eux-mêmes reliés linéairement à une cause.

On considère assez volontiers que ce type de relation, en boucle ouverte, peut suffire à décrire localement, selon des relations infinitésimales entre variables, le comportement d'un système. À l'opposé, la cybernétique ne s'embarrasse pas de considérations locales ou infinitésimales, mais pose d'emblée le principe de rétroaction – donc de boucle fermée, entre l'entrée et la sortie d'un système – comme fondamental, notamment pour expliciter la persistance des organismes dans leur état vivant, en bref, l'homéostasie. Selon que la rétroaction est négative ou positive, les effets attendus sur le système peuvent varier drastiquement. Il est fréquent d'envisager des systèmes à rétroaction négative, précisément dans un but de contrôle de l'erreur entre la commande d'entrée et le résultat, et donc de correction de l'entrée à l'instant d'après. Cependant, les systèmes à rétroaction positive présente parfois des propriétés étonnantes, notamment d'auto-organisation.

La rétroaction entraîne une causalité en régime permanent qui se suffit à elle-même, puisque l'effet nourrit la cause, et inversement. Il est également possible d'envisager ce type de rétroaction dynamique, adaptable par ailleurs selon le temps et le lieu, et mieux encore, à un niveau infinitésimal. C'est en fait ce qu'énonce, en toute généralité l'équation métrique de la relativité générale : la métrique est déterminée par la masse locale, et la masse se meut selon la forme de la métrique – cette relation étant écrite sous forme différentielle et locale. La relativité nous offre ainsi un excellent exemple de cybernétique généralisée.

Le pendant quantique est plus perturbant, puisque l'interaction sert, en quelque sorte, de rétroaction : une fois des particules corrélées, la mesure de

l'une entraîne la détermination de l'autre, indépendamment de la distance les séparant. Si la relativité nous autorise à écrire une cybernétique dynamique et locale, la mécanique quantique nous offre une cybernétique délocalisée.

3.3.2 Place du temps – opposition à la forme ?

Le temps sert fréquemment et de base d'échantillonnage, et d'axe de représentation privilégié – un axe chronologique, permettant de placer des événements les uns par rapport aux autres. Il permet d'indexer les objets selon une seule dimension, ce qui est pratique. Néanmoins, sa mesure peut poser problème, dans certains cas *a priori* extrêmes comme nous l'avons vu en relativité (il faut que la vitesse de déplacement relative des observateurs soit grande en proportion de c), puisqu'elle est relative, et non invariante. Substituer le temps propre au temps relatif est intéressant, notamment parce qu'il permet dans le même mouvement de rendre compte de la causalité, localement.

Le fait est que la dimension temporelle, prise comme axe indépendant et unidimensionnel, mais également linéaire et homogène dans son écoulement, nous empêche peut-être d'envisager la forme correcte à observer, en ne considérant que des objets spatialisés et leur dynamique ou cinétique dans l'espace, en fonction du temps. Étonnamment, la perspective de manipuler des objets isolés au cours d'un temps absolu relève davantage d'une conception quantique, quand bien même le principe d'intrication quantique qui en fait la spécificité, ne s'applique pas – ou ne perdure pas – au-delà d'une certaine échelle, et pour des mélanges de corps différents.

Aussi, les "vraies" formes, les plus pertinentes pour leur cohérence interne, leur stabilité en termes de définition et de fonction, sont-elles celles que l'on veut voir persister spatialement à l'identique, à une classe de transformations spatiales près (une homothétie, une symétrie spatiale) ? Ou bien ne faut-il pas lier le temps à leur résolution spatiale, selon une relation à déter-

miner ? On pourrait également proposer qu'une forme ne peut être aussi bien résolue en temps et qu'en espace, à la manière d'une relation d'incertitude en théorie quantique, et qu'il est nécessaire soit d'accepter un compromis entre ces deux facettes, soit de privilégier la persistance spatiale, et réduire la fenêtre temporelle d'observation, ou le suivi dans le temps, et d'accepter des transformations non conservatives quant à la résolution spatiale de cette forme.

Par ailleurs, si l'on décide de lier le temps et l'aspect spatial de la forme, de quel temps parlerions-nous ? Il semble relativement acquis que, pour des formes différentes, et à même horloge, les trajectoires et devenir spatiaux de ces formes – disons une plaque de marbre et un steak – divergeront assez drastiquement. Certainement, pour être pertinent, faudrait-il alors choisir la bonne base de temps.

En ce sens, le choix d'un temps absolu et d'une forme spatialisée isolée semble s'opposer à la notion même de forme en tant qu'invariant, car menant systématiquement à la transformation irréversible de cette forme, et selon des évolutions plus ou moins rapidement divergentes selon les objets. En d'autres termes, étudier les formes selon la flèche du temps, laquelle est, à l'échelle macroscopique, souvent motivée par des considérations thermodynamiques, c'est se soumettre peut-être sans trop d'élégance ou de soin aux lois de l'entropie. Nous ne nions bien entendu pas l'existence d'un principe entropique, mais voulons envisager que son application puisse s'envisager de manière plus adaptée, plus locale, que le simple énoncé qu'une forme ne pourra persister en son état "spatial" en raison de l'augmentation inexorable de l'entropie.

3.3.3 Traitement de l'information, inférence de la causalité, approche bayésienne

La place du temps dans chacun des grands modèles est contradictoire, ou pose au minimum problème, puisqu'elle n'est pas similaire selon les points de vue. La tentation serait grande d'éliminer le temps tel qu'on l'utilise jusque lors. De fait, seule la relativité a été une tentative probante de redéfinir la place du temps.

Il n'est pas nouveau d'essayer soit d'éliminer le temps comme variable, soit de suspecter qu'il n'est l'expression que de mécanismes plus élémentaires et fondamentaux. Principalement, l'approche thermodynamique et la définition de l'entropie ont fourni les meilleurs espoirs de substitution. Puisque il semble exister une tendance à l'accroissement naturel de l'entropie, nous tiendrions là un principe directionnel, orientant à la manière d'une inertie, les déplacements et organisations d'un système sans (nouvelles) contraintes. C'est notamment le cas dans un système en boucle ouverte, où précisément, il n'existe pas de rétroaction – ou de force, de principe actif – permettant de contrer la tendance au désordre.

Puisque il existe une tendance, une direction privilégiée, alors on pourrait y calquer ou y corréler une grandeur de type temps, par commodité calculatoire et pertinente dans nombre de cas pratiques. L'entropie au sens thermodynamique ou au sens informationnel (entropie de Shannon), est une notion basée sur le dénombrement des configurations ou états possibles. Quel serait le lien entre un tel dénombrement, et la notion de temps ?

Par ailleurs, l'entropie est une grandeur définie sur un ensemble fermé de possibilités, sur un alphabet plus ou moins complexe. Cette notion d'alphabet ou d'ensemble des possibles ouvre néanmoins la voie vers une définition plus fine d'une potentielle flèche temporelle, sans que nous devions nous encombrier directement d'un temps donné voire absolu. Par ailleurs, physiquement, les manière de compter le temps sont, de toute façon, essentiellement

des techniques de comptage d'événements discrets : il suffit de connaître la fréquence relative de l'événement pour compter en unités de temps les intervalles séparant des événements plus "rares".

Ainsi, l'information et le dénombrement peuvent servir de base à la construction de variables temporelles relatives, et de fait servir de référence à l'évaluation de toute relation de causalité. D'un point de vue informationnel, en outre, il a été proposé une approche du lien de causalité entre deux observations qui ne mentionne pas explicitement la variable temporelle. La causalité repose dans ce cadre sur la circulation de l'information, ou autrement dit, sur la possibilité, étant donnée une observation, d'en déduire une autre, et non l'inverse. Bien entendu, cette causalité peut s'avérer plus faible que la causalité usuellement recherchée, car elle dépend des observations qui ont été prises en compte - et il est possible qu'une observation critique, plus fortement "causale" d'une autre, ait été omise.

Ce type d'approche a été particulièrement développée par Pearl, [142, 143] et s'est plus spécialement matérialisée dans la construction des réseaux de causalité, du type des réseaux dits bayésiens. De fait, dans la suite de cette section, nous allons nous attacher à présenter ce type particulier de réseau, lequel repose notamment sur une représentation graphique des dépendances entre observations – un graphe acyclique dirigé (GAD). Une partie des éléments théoriques ci-dessous peuvent également être retrouvés sous une forme relativement synthétique dans. [144]

Définition d'un GAD, type réseau bayésien

L'idée ici est de représenter le système étudié par la distribution de probabilité jointe de l'ensemble des observations du système. La structure minimale de cette distribution est approchable *via* l'hypothèse d'indépendance conditionnelle, et permet de décrire la structure informationnelle du système

par un graphe acyclique dirigé (GAD). De fait, les différentes observations faites sur le système, et les liens orientés associant ces variables permettent d'interpréter le GAD comme un automate particulier.

Le "moteur" logique qui conduira à la construction de l'automate informationnel tient dans la formulation du théorème de Bayes, que nous rappelons ici :

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)} \quad (3.1)$$

Où A et B correspondent à deux variables aléatoires, et P dénote leur probabilité, $P(.|.)$ leur probabilité conditionnelle.

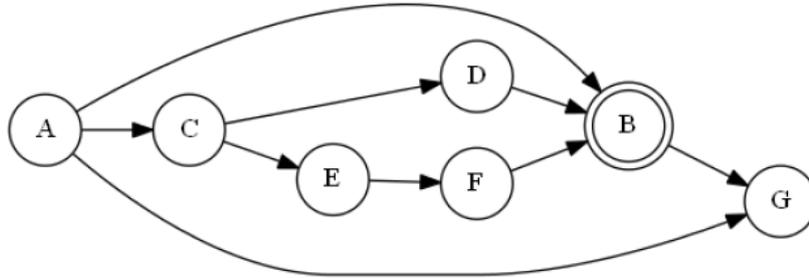
Données d'un graphe et d'une probabilité jointe décrivant le système

Soit l'ensemble O des observations du système S . O est constitué de n observations suivant p dimensions. Il est possible de considérer chaque p dimensions comme une variable aléatoire discrète, avec n tirages. A partir de ces p variables, notées $\{X_i\}_i$, il est possible de représenter l'ensemble de S par la distribution de probabilité jointe $P(X_1, \dots, X_p) = P(S)$. Parmi les techniques permettant d'approcher cette distribution de probabilité, nous pouvons utiliser les techniques Bayésiennes, en particulier l'approche des réseaux Bayésiens. Un réseau bayésien RB est la donnée de 2 éléments, à savoir un graphe G et la distribution de probabilité jointe sur S notée $P(S)$. On a donc par définition :

$$RB = \{G, P(S)\} \quad (3.2)$$

Le graphe G peut être dirigé ou non. Dans notre cas, nous nous munirons d'un graphe dirigé. On impose en outre à celui-ci d'être acyclique, c'est-à-dire de ne présenter aucune "boucle" au sein du graphe. Un tel graphe est

FIGURE 3.11 – Un exemple de réseau bayésien



appelé graphe acyclique dirigé (GAD). En outre, puisque nous construisons un graphe, nous nous dotons de la notation suivante afin de représenter l'arête joignant X et Y , X et Y étant deux variables observées, chacune représentées par un noeud du graphe :

$$X \rightarrow Y$$

Ce qui signifie ici que l'on circule de X vers Y et non l'inverse. Dans un GAD type réseau bayésien, l'orientation de l'arête symbolise le sens de la relation de causalité entre deux variables. Dans l'exemple précédent, il s'agit donc de considérer que la connaissance de X entraîne la connaissance de Y . Tout l'intérêt du réseau bayésien vient alors du fait de "squelettiser" la distribution de probabilité $P(S)$ du système, squelette qui sera donné par l'allure du GAD et ses arêtes. Cette squelettisation est possible grâce à la propriété d'indépendance conditionnelle, laquelle va réduire le nombre de relations point à point entre variables du système. Nous donnons en figure 3.11 un exemple quelconque de GAD.

Notion d'indépendance conditionnelle

La propriété d'indépendance conditionnelle se définit assez simplement. Soient X , Y et Z trois variables du système étudiés. On dit que X est indépendant de Z conditionnellement à Y si la distribution de probabilité de X conditionnellement aux variables Y et Z est égale à la distribution de probabilité de X conditionnellement à Y . Autrement dit :

$$P(X|Y, Z) = P(X|Y) \quad (3.3)$$

On écrit la propriété d'indépendance conditionnelle de X par rapport à Z sachant Y de la façon suivante :

$$X \perp Z|Y \quad (3.4)$$

Par ailleurs, un réseau bayésien est doté de la propriété de Markovianité : chaque variable X du système est conditionnellement indépendante des autres variables qui ne sont pas ses descendants, sachant les parents de X . Voyons comment ces propriétés simples permettent d'écrire la probabilité jointe du système $P(S)$.

Écriture de la probabilité jointe avec l'hypothèse d'indépendance conditionnelle

Suivant les propriétés d'indépendance conditionnelle et de Markovianité, la probabilité jointe $P(S)$ peut s'écrire :

$$P(S) = P(X_1, \dots, X_p) = \prod_{i=1}^p P(X_i|X_{i-1}, \dots, X_1) \quad (3.5)$$

En effet, X ne dépend que de ses prédécesseurs. Allons une étape plus loin : tous les X_{i-1}, \dots, X_1 ne sont pas nécessaires, si l'on considère que l'on peut s'en tenir qu'à la dépendance aux parents directs de X_i . Nous noterons

ici les parents de X_i par $P(C(X_i))$, $C(\cdot)$ pour "cause" de. Nous pouvons alors écrire :

$$P(S) = \prod_{i=1}^p P(X_i|C(X_i)) \quad (3.6)$$

Automatiquement, nous nous apercevons que la structure "utile" de la probabilité $P(S)$ s'allège considérablement, et devient par ailleurs plus lisible et structurée. Cette structuration se fait en termes de circulation de l'information et de causalité locale. Nous allons déterminer un exemple très simple à trois variables qui nous servira tout au long de l'exposé d'exemple compagnon pour illustrer les propriétés dont nous avons besoin.

Présentation d'un exemple compagnon simple : modèle X,Y,Z

Le modèle de réseau bayésien que nous proposons ici est volontairement simple si ce n'est simpliste ; cependant, il est suffisant pour permettre de tester les concepts avancés. Soit un système S_{ex} constitué de trois variables observées binaires 0/1, X , Y et Z . Nous avons donc $S_{ex} = \{X, Y, Z\}$. Le GAD associé à S_{ex} est le suivant :

$$Y \rightarrow X \leftarrow Z$$

Y et Z sont donc les "causes" de X . Un tel modèle est décrit par sa table de probabilité jointe. La table associée à S_{ex} serait ainsi de la forme :

TABLE 3.1 – Distribution de la probabilité jointe de S_{ex}

	$P(Y = 1)$		$P(Y = 0)$	
	$P(Z = 1)$	$P(Z = 0)$	$P(Z = 1)$	$P(Z = 0)$
$P(X = 1)$	X_{111}	X_{110}	X_{101}	X_{100}
$P(X = 0)$	X_{011}	X_{010}	X_{001}	X_{000}

Nous avons introduit les probabilités sous la forme de $X_{X=i,Y=j,Z=k}$, avec

$\{i, j, k\} \in \{0, 1\}$ car la variable à décrire est X , qui dépend des variables Y et Z .

Interprétations en termes d'automate

Le modèle S_{ex} peut donc se décrire par son GAD associé, et $P(S_{ex})$ peut se réduire à :

$$P(S_{ex}) = P(X|Y, Z).P(Y).P(Z) \quad (3.7)$$

Aussi, à chaque événement observé concernant Y ou Z , nous gagnons de l'information sur la valeur de X . Notons également que :

$$Y \perp Z$$

Ainsi, un tel automate informationnel permet de décrire le comportement local d'un système pour lequel nous allons observer l'occurrence ou non d'événements particuliers, ici Y ou Z . Il ne s'agit donc pas à proprement parler d'un automate tel que l'on en rencontre en robotique ou en automatique. Plus loin, nous verrons que nous attribuons des rôles particuliers aux variables telles que Y et Z , si l'on considère que les variables dont il nous intéresse de décrire l'évolution sont des variables telles que X .

3.3.4 Modèles allométriques, temps physiologiques et hypothèse de West.

Un type de modèle que nous n'aborderons pas de manière approfondie, mais dont il peut être nécessaire de signaler l'existence et l'intérêt, est le modèle allométrique. [145, 146, 147] Parmi les relations d'invariance que l'on peut être amené à postuler, il est possible de chercher à vérifier s'il n'existe pas des relations vérifiées à différentes échelles d'observations – ainsi qu'il est possible d'observer pour les fractales autosimilaires, où un motif élémentaire

se trouve répété aux différentes échelles d'observation. On parle également, dans ce cas, de lois d'échelle.

Les relations allométriques ont été assez largement abordées et discutées en sciences de la vie, notamment pour lier morphologie, croissance et métabolisme entre individus d'espèces différentes. L'idée sous-jacente étant là encore d'essayer d'identifier un type de loi qui soit vérifiée pour tout le vivant, du moins dans sa forme, ce qui peut alors fournir un espace de représentation et d'interprétation universel.

Définition d'une relation allométrique. L'allométrie peut ainsi s'écrire sous la forme :

$$Y = b.X \exp(k) \tag{3.8}$$

où, par exemple,

- X peut désigner la taille du corps, la masse corporelle ou sa surface ;
- Y peut désigner la taille de l'organe présentant une croissance différentielle ;
- b est une constante de proportionnalité caractéristique de la loi allométrique ;
- k est une constante allométrique, caractérisant la croissance de Y en fonction de X .

On lit au travers de ce type de relation qu'il existe un rapport constant entre le taux de croissance de la taille globale de l'organisme X et le taux de croissance de la taille Y d'un organe donné, ou d'une activité métabolique partielle ou globale.

Pour des questions de représentations et de simplicité d'interprétations, la relation exponentielle est souvent transcrite sous sa forme linéaire en le logarithme :

$$\log(Y) = \log(b) + k \cdot \log(X) \quad (3.9)$$

On dispose alors d'une relation linéaire, d'ordonnée à l'origine $\log(b)$ et de pente k .

Exemple de la morphométrie des poissons. Une utilisation des lois allométriques fréquemment rencontrée dans la littérature tient dans la recherche d'une relation entre la masse et une taille caractéristique, en l'occurrence, la masse et la longueur de poissons d'espèces différentes. Il s'avère que statistiquement, une majorité de poissons suivent une loi telle que :

$$Masse = a \text{ (longueur du corps) } \exp(b),$$

où

- a est caractéristique de la forme du corps (avec 0,1 pour les poissons de petites tailles, à géométrie plutôt courbe ou sphérique, et 0,01 pour les poissons plus longilignes).
- b est une constante dont les valeurs empiriques avoisinent 3.

Il est possible de se référer à l'article *Morphometrics and Allometry in Fishes*, [146] traitant de la morphométrie et de la morphologie des poissons, et montrant qu'il est possible de bâtir de telles relations allométriques. Un des intérêts de ce type de modélisation, outre le fait de mettre en évidence une forme d'invariant, cette fois, en échelle, est qu'il établit de fait une loi – statistique – qui rende compte du lien entre morphologie et nutrition, ou métabolisme, par exemple.

En outre, nous désirions traiter rapidement des modèles allométriques afin de signaler une hypothèse concernant une conception particulière du temps, en lien avec notre démarche. Il s'agit du temps dit physiologique. L'idée, ici aussi, est de trouver une base de représentation qui permette l'expression de relations invariantes. Le lecteur pourra se référer à l'ouvrage de Laguës et Lesnes, sur les invariances d'échelles, relativement exhaustif.

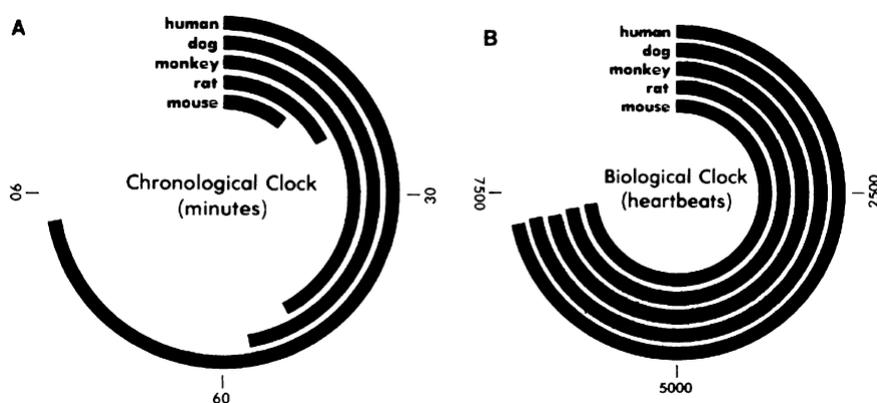
[148]

Temps physiologique, hypothèse de West La notion de temps physiologique n'est pas totalement neuve, et est largement issue des études concernant les lois développementales des animaux, notamment les animaux à sang froid. [149] L'idée d'un temps physiologique, à mettre en relation avec la notion de métabolisme propre, ou de taux métabolique basal, suppose que le temps pertinent pour discuter des fonctions d'un organismes, dont sa croissance, sa durée de vie mais également ses différentes fonctions physiologiques, n'est pas le temps chronologique absolu mesuré par une horloge atomique de précision, mais est davantage à chercher selon un rythme qui serait propre à l'organisme étudié. L'analogie avec l'ordinateur est ici aussi facile : les différents ordinateurs ont leur fréquence de fonctionnement, et il est plus pertinent de qualifier les performances d'un ordinateur selon le nombre d'instructions qu'il est capable de traiter par cycle d'horloge, son horloge, que de donner un temps d'exécution d'un programme donné sans préciser son support d'exécution – à architecture identique, un programme s'exécutera dans le même ordre et avec les mêmes instructions sur deux supports différents. Dans leur référentiel propre, l'exécution du programme prendra le même "temps", mesuré selon leur horloge respective – il s'agit bien d'un espace où il existe un invariant puisque le même programme, la même loi donnera le même résultat, selon le même cheminement. Cependant, un observateur extérieur ne mesurera pas nécessairement la même performance.

Pour en revenir aux organismes vivants, on peut prendre l'exemple de la pharmacodynamique de certains xénobiotiques : on caractérise la vitesse de "traitement" d'une substance étrangère à un organisme par sa demie-vie d'élimination ; c'est-à-dire, le temps nécessaire à l'obtention d'une concentration (sanguine, par exemple) égale à 50% de la concentration initiale.

Pour des espèces différentes, pour un homme et une souris, et pour la

FIGURE 3.12 – Temps absolu et temps physiologique (d'après *West*)



même concentration de produit initiale, cette demie-vie ne sera pas la même si l'on regarde "extérieurement" l'évolution de cette concentration en fonction du temps d'une horloge.

Cependant, il est possible d'examiner cette même demie-vie en fonction de rythmes propres aux organismes, bien que basés sur un fonctionnement partagé ; par exemple, on peut mesurer le temps physiologique en fonction du nombre de battements cardiaques par minute. En ce cas, la demie-vie se "normalise", et devient la même quelque soit l'espèce (cf figure 3.12).

Un auteur (BJ West) s'est intéressé au cours des vingt dernières années à la représentation et à l'interprétation des signaux physiologiques selon des approches un peu différentes des analyses classiques. En particulier, ses études ont essentiellement porté sur le calcul fractionnaire et les lois d'échelles comme reflet d'une complexité physiologique "normale" – la perte de complexité se traduisant alors comme une perte d'interrelations entre échelles, et donc un décrochage de certaines fonctions physiologiques, donc un aspect "pathologique". Il s'est ainsi penché sur la variabilité du rythme cardiaque, du rythme respiratoire, de la marche, entre autres. [50, 150, 151, 152, 153]

Dans un papier récent, il émet une hypothèse liant les modèles allomé-

triques, la production d'entropie et le temps physiologique. [149] Nous en exposons ici les principaux résultats.

Certains auteurs se sont intéressés au taux métabolique basal (BMR - *Basal Metabolic Rate*), pour différentes espèces, et ont cherché à montrer que le rapport entre le volume et le temps chronologique était proportionnel à la masse élevé à un exposant :

$$\frac{Volume}{temps} = M^b$$

Il a également été examiné la notion de temps physiologique, dénoté τ , et envisagé comme une variable aléatoire, lequel serait une fonction de la masse totale de l'organisme (TBM - *Total Body Mass*) m , et respectant une loi d'échelle, comme suit :

$$\tau(\lambda m) = \lambda^v \tau(m)$$

L'hypothèse sous-jacente correspondante à cette relation serait qu'il existerait une génération d'entropie constante à chaque cycle de temps physiologique, traduction d'une efficacité optimale des réseaux physiologiques. Thermodynamiquement, de même qu'informationnellement, tout processus, toute forme engendre par son maintien une entropie. Le postulat ici est donc que dans un espace de représentation intrinsèque à l'organisme, et en régime permanent et "normal", la génération d'entropie est la même à chaque cycle de temps physiologique.

Les relations allométriques peuvent également se formuler de la manière suivante, en toute généralité :

$$Y_{ij} = aM_{ij}^b$$

où j correspond à un individu d'une espèce i donnée.

On cherche ici une relation de ce type, liant le temps physiologique à une autre grandeur, à savoir la masse :

$$\tau_{ij} = aM_{ij}^\beta$$

ce qui se traduit pour les valeurs moyennes $\langle x_i \rangle$, moyennes pour une espèce donnée i :

$$\langle \tau_i \rangle = \alpha \langle M_i \rangle^\beta$$

Une variable aléatoire X vérifiant la relation d'échelle voulue :

$$X(\lambda t) = \lambda^\nu X(t)$$

doit vérifier une densité de probabilité donnée par une fonction de la forme :

$$P(x, t) = \frac{1}{t^\theta} F_x\left(\frac{x}{t^\theta}\right)$$

où F est à déterminer.

Par ailleurs, la probabilité pour que le temps entre deux événements soit compris entre τ et $\tau + d\tau$ pour un organisme de TBM m est donné par $P(\tau, m)d\tau$, avec $P(\tau, m)$ donné par :

$$P(\tau, m) = \frac{1}{m^\nu} F_\tau\left(\frac{\tau}{m^\nu}\right)$$

Si l'on considère l'entropie de Shannon $S_x(y)$ telle que :

$$S_x(y) = -k \int P(x, y) \ln P(x, y) dx$$

et en posant $x = \tau$ et $y = m$,

on obtient la variation d'entropie $\Delta S_\tau(m)$ telle que :

$$\Delta S_\tau(m) = S_\tau(m) - S_0 = vk \ln m$$

avec S_0 le niveau d'entropie de base :

$$S_0 = -k \int F_\tau(z) \ln F_\tau(z) dz$$

on a par ailleurs :

$$\Delta S_m(t) = \theta \ln t$$

La pente de la création d'entropie serait la suivante :

$$\frac{d\Delta S_\tau(m)}{dt} = \frac{vk}{m} \frac{dm}{dt} = \frac{v\theta k}{t}$$

L'hypothèse d'invariance de la génération d'entropie, en temps chronologique et en temps physiologique mène à :

$$\frac{d\Delta S_\tau(m)}{d \langle \tau \rangle} = \frac{d\Delta S_m(t)}{dt},$$

relation vraie si et seulement si :

$$\theta = \frac{1}{\beta}$$

$$v = \alpha M^\beta$$

En l'état, il s'agit bel et bien d'une hypothèse, dont certains aspects sont discutables. Néanmoins, cette hypothèse a le mérite de tenter de jeter un pont entre temps physiologique, entropie d'un système, et invariance d'échelle. Nous reviendrons sur cette hypothèse quand nous aborderons, dans le chapitre dédiée à la construction de notre théorie, la question du temps propre d'un système observé.

3.4 Trois grands modèles possibles

Étant données les contraintes que nous avons posées jusque là, nous aborderons dans cette section les trois grands types de modèles que nous pensons pouvoir exploiter, d'une manière ou d'une autre, afin d'en tirer des éléments, une structure pour une définition de la forme, et un cadre pour la morphodynamique. Nous voyons ainsi assez rapidement comment les paradoxes ou oppositions clés que nous avons déjà identifiés sont traités – ou centraux – de ces modèles. C'est particulièrement le cas des dualités onde-corpuscule ou numérique-analogique, mais également de trois manières de caractériser la forme, comme objet spécifique de ces modèles.

Bien entendu, il existe d'autres approches, ou il serait envisageable de discuter de principes alternatifs, de variantes de ceux qui gouvernent les trois modèles que nous discuterons. Néanmoins, nous avons retenu ces modèles tant pour leur généralité, leurs fondations solides, que pour le subtil alliage de points communs et d'antagonismes qui les relie. C'est par ailleurs en circonscrivant au mieux les avantages et les limites de grands modèles actuels que l'on peut aussi espérer en dépasser les défauts, trouver une voie intermédiaire, qu'elle en soit la synthèse ou simplement un principe plus large ou mieux adapté à notre problème.

3.4.1 Entre onde et corpuscule, numérique et analogique – une histoire qui perdure

Parmi les points communs et antagonismes qui traversent les approches de la forme, il en est un qui nous paraît inévitable, pour ne pas dire invariant, au point où la question se pose de savoir si le problème n'est pas simplement mal posé. Nous avons esquissé cette dualité entre description ondulatoire et description corpusculaire, ou autrement dit, entre description d'objets qui soient intrinsèquement délocalisés, ou parfaitement ponctuels,

sinon parfaitement circonscrits, individualisés. À ce stade, nous tenons à faire une remarque quant aux théories manipulant des objets délocalisés, comme en mécanique quantique. Si l'on parle d'objet délocalisé, c'est bien parce que l'on sous-entend avant tout le caractère ponctuel, particulière de l'objet dont on parle. En effet, l'électron présente une probabilité de présence en les coordonnées spatiales non nulle en une infinité d'endroits - ce qui peut revenir à dire qu'à un instant t , l'électron, objet supposé particulière, voit sa présence en un point de l'espace pondérée d'une amplitude probabilité. Cependant, il est possible d'inverser le point de vue, et ne pas considérer l'électron comme particulière, mais précisément comme une entité d'étendue continue en l'espace, et pouvant se propager ou non, en fonction du temps. En ce cas, l'objet - la forme - de l'électron devient simplement cette onde. C'est en ce sens que les théories de type cordes ou membranes ont évolué : les objets considérés ne sont plus en soi des particules, mais des cordes d'étendues limitées mais non ponctuelles, vibrantes, par exemple.

Il devient tentant, depuis l'avènement du numérique et des succès de la théorie de l'échantillonnage de type Shannon-Weaver, de supposer que cette dualité aux allures de paradoxe, n'a pas forcément lieu d'être : après tout, sous de bonnes hypothèses de cadence adaptée - ou du moins, suffisante - et de signal d'énergie finie, le théorème d'échantillonnage démontre l'équivalence en termes d'information entre une description continue, délocalisée, et une description discrète. C'est également une voie d'abord pour réconcilier théories quantiques et théories relativistes.

Dans le cas de la théorie du signal "classique", et dans le cas des théories quantiques, la transformation de Fourier est centrale, car c'est elle qui permet de lier deux domaines de représentation - par exemple, les domaines temporel et fréquentiel, ou la position et l'impulsion. L'utilisation de cette transformation sous-entend le recours à une base de représentation particu-

lière, en l'occurrence la décomposition sur la base des sinus-cosinus, base de dimension infinie. En théorie du signal, d'autres bases ont depuis été proposées, en partie pour dépasser les limitations de la transformée de Fourier ; on pensera essentiellement aux décompositions de type ondelette. La décomposition en ondelettes utilise un motif de base respectant un certain nombre de contraintes mathématiques, et la base de décomposition est en fait constituée des dilatations-contractions du motif initial. De fait, l'analyse en ondelettes est une analyse de ressemblance (de distance) à un motif particulier et une analyse d'échelle : la ressemblance est testée à toutes les échelles. De manière caricaturale, si le signal à analyser était un signal autosimilaire à toutes les échelles, et que son motif initial était un motif respectant les modalités de définition d'une ondelette, la décomposition sur la base de cette ondelette serait parfaitement adaptée - sa ressemblance serait parfaite à toutes les échelles.

De la sorte, nous voulons attirer l'attention sur l'importance de la base de représentation. Idéalement, il est peut-être possible que toute description continue soit équivalente à une description discrète ; néanmoins, l'efficacité de cette équivalence, le caractère "naturel" de cette équivalence est fortement liée au choix de la base de décomposition. En langage courant, cette problématique revient à dire que pour obtenir la meilleure réponse, il faut savoir poser les bonnes questions.

3.4.2 L'espace plat et le lien probabiliste : la forme ponctuelle

La description que nous qualifierons de quantique, se caractérise par un espace de description de géométrie plate, euclidienne, où la forme "globale" qui nous intéresse est finalement une forme ponctuelle, dans le sens où elle se décompose sur une base propre, invariante. Les particules, à un moment t et en l'absence de mesure, sont certes délocalisées, mais leur forme d'onde

s'exprime comme une combinaison linéaire de vecteurs de la base de description. C'est en ce sens que nous parlons de forme ponctuelle, dans un sens un peu similaire de forme quantifiée. La mécanique régissant l'évolution de ces formes ponctuelles, et des particules, est une mécanique ondulatoire, mais une fois la mesure effectuée, il n'est pas possible d'obtenir un résultat arbitrairement quelconque : le spin sera $1/2$ ou $-1/2$, pas autre chose.

De fait, le coeur de ce modèle, en ce qui nous concerne, est le suivant : les formes observées sont de nature ponctuelles, mais les liens que plusieurs formes entretiennent entre elles est probabiliste, de type corrélational. En effet, pour deux particules, deux formes ayant interagi, étant corrélées l'une à l'autre, l'observation de l'état de l'une déterminera automatiquement l'état de l'autre. L'espace de ce modèle, plat, à états observables bien définis et potentiellement totalement quantifiés, autorise des interactions à distance non médiées par l'étendue, par la localité : la mécanique d'évolution est à la fois ondulatoire et corrélational.

3.4.3 L'espace courbe et le lien local : la forme globale

Le modèle corrélational que nous venons d'aborder s'oppose au modèle dit du "réalisme local", qui veut que toute perturbation d'un élément du système, si cet élément est à même de modifier un autre élément ou une autre partie du système, ne peut qu'être une perturbation se propageant entre les deux parties, c'est-à-dire, modifiant de toute façon tous les points intermédiaires de l'étendue entre ces deux éléments. Ainsi, dans les modèles relativistes, la gravité, son action due à un corps et s'exerçant sur un autre, si le premier corps vient à être modifié, verra son action se modifier également sur l'autre corps, mais pas de manière instantanée ou immédiate – c'est-à-dire, non médiée.

C'est là toute l'illustration des nappes en deux dimensions que l'on mobilise souvent pour représenter l'influence d'un corps grave sur son espace

environnant, en relativité générale : ce corps modifie son voisinage, la géométrie de l'espace, ce qui matérialise son action sur les autres corps pouvant passer à proximité. L'analogie courante est de se représenter une bille lourde sur une nappe fine tendue sur le vide. Si une autre bille vient à passer suffisamment près de la première bille, il est possible qu'elle "tombe" sur la bille creusant la nappe sous l'action de son poids. Par ailleurs, si l'on déplace la bille lourde, la déformation de la nappe va également se déplacer, et la trajectoire des autres billes peut s'en trouver affectée différemment.

Ici, il est donc nullement question de mécanique corrélacionnelle, mais bien de déformations continues. Dans un tel espace, ultimement, il n'est nul besoin d'objets ponctuels, puisque l'on ne s'intéresse qu'aux déformations engendrées par la présence et la distribution d'une information particulière – ici la masse grave. Aussi, la forme d'intérêt serait finalement la forme globale de l'espace, ce qui tend à la tautologie – la forme est la forme. L'élément de description pertinent dans ce type de cas est alors l'élément permettant de mesurer le degré de déformation de l'espace, cette déformation matérialisant l'action d'une information, selon sa nature et son intensité. Cet élément est la métrique de l'espace ; la mécanique d'évolution de la forme dans un tel espace doit donc être une mécanique régissant l'évolution de la métrique. Le modèle que nous qualifierons de relativiste est donc un modèle convoquant à la fois la forme globale et la forme locale, la première se modifiant sous l'action continue et par propagation de la seconde.

3.4.4 L'espace des phases des systèmes dynamiques : la forme comme attracteur

Les théories quantiques et relativistes proposent un abord assez radicalement différent, sinon opposé. La géométrie des premières tient plus d'une approche "spectrale" en espace plat : l'espace sans courbure peut être réduit à un certain nombre d'éléments propres, qui procurent le squelette invariant

selon les observateurs, du système considéré. Les liens entre objets sont essentiellement d'ordre probabiliste, instantané. La géométrie des secondes est une géométrie courbe, variable selon la distribution de l'information massive, et il a même été montré qu'en géométrie Riemannienne, le spectre de ces objets géométriques ne permet pas de les classer exclusivement les uns des autres. L'invariance par observateur est une invariance des lois, non des objets supposés.

Le troisième modèle envisagé tient sur une approche encore différente, et existe aussi bien sous une formulation continue que discrète (que nous n'avons ni n'aborderons en détails dans ces travaux). Le modèle des systèmes dynamiques, où l'on représente un système par son espace dit des phases – une manière de dire que l'on n'est pas contraint par une description "classique" de type position spatiale et temps, mais que l'on peut étendre la description d'inspiration mécanique à un nombre théoriquement arbitraire de paramètres décrivant le système (un angle, une température, un pH, une position spatiale...) – est un modèle qui se veut général, son ambition étant de pouvoir décrire l'évolution temporelle d'un système différentiel général.

Quel objet privilégié de ce type de modèle serait éligible à la notion de forme? Bien entendu, il serait tentant de proposer "simplement" que le système décrit par l'ensemble de ses n paramètres définit une forme. Le gain pour notre quête de définition serait alors relativement pauvre, et surtout, très arbitraire. Notamment, la notion d'invariance, ou à défaut, de stabilité, ne serait pas traitée.

Un des apports majeurs de l'approche par les systèmes dynamiques a été l'établissement de techniques et de résultats généraux permettant d'étudier et de caractériser d'une part la stabilité "simple" de certaines solutions du système, et d'autre part la stabilité structurelle de ces solutions. C'est ainsi que l'on peut définir d'un côté la notion d'attracteur, en lien avec la notion avec la stabilité dite "simple" – la solution retenue est-elle stable sous l'ac-

tion d'une petite perturbation du système ? Attire-t-elle vers sa région de l'espace des phases les trajectoires du système ? – et d'un autre côté l'aspect topologique de ces attracteurs, en lien avec la stabilité structurelle – sous quelles conditions plus qualitatives, structurelles, les solutions stables identifiées préservent la même forme, la même dimension ?

De fait, sous ce type d'approche, il nous semble plus pertinent de considérer la notion de forme du point de vue des attracteurs du système dans un premier temps, et du point de vue des conditions d'existence et d'évolution de ces attracteurs – ce qui est le rôle de l'analyse de la stabilité structurelle. D'une certaine façon, cette approche double ressemble à un compromis entre une approche quantique et une approche relativiste : si un système dynamique présente plusieurs attracteurs, on peut être tenté de rapprocher la description spectrale quantique d'une description en termes d'attracteurs différents – la base de décomposition en théorie quantique devient une base de décomposition en termes d'attracteurs différents ; cependant, là où en mécanique quantique, cette base a pour vocation d'être universelle (le spin d'un fermion ou d'un boson a des valeurs fixées pour tous), la base des attracteurs peut être soumise à modification, voire extinction, selon la stabilité structurelle du système. En effet, même si certaines théories s'intéressent à la possibilité que les constantes fondamentales de la physique ne soient pas toujours constantes, la vaste majorité des approches considèrent que ces constantes sont bel et bien invariantes. La constante de Planck h ne sera pas un paramètre libre du système quantique étudié. En revanche, dans un système dynamique, il est possible que les paramètres "informatiques" au sens d'un système de Turing, les paramètres définissant la structure du problème, puissent varier. Leur variation peut alors entraîner des modifications topologiques, des catastrophes. En ce sens, la topologie et donc la géométrie des solutions, des attracteurs, peuvent être amenées à changer : on se retrouve davantage dans une configuration relativiste.

Cette approche duale a l'intérêt de nous rapprocher d'une définition entre forme et fonction : si la forme est un attracteur, matérialisant par là la propension d'un système à se rassembler autour d'un motif particulier et isolé, alors les paramètres du système réglant sa stabilité structurelle tiennent davantage lieu de paramètres fonctionnels, sinon causaux de l'attracteur, donc de la forme observée. L'étude des transformations portant sur les paramètres du système peut alors également déboucher sur l'identification d'invariances par transformation.

3.4.5 Comparaison des 3 modèles

Puisque nous venons d'aborder successivement les trois grands types de modèles dont nous voulons nous inspirer, nous proposons ici une synthèse récapitulative et comparative de ces trois modèles.

TABLE 3.2 – Tableau comparatif des trois grands modèles théoriques pouvant servir de base pour une théorie de la forme.

Théories quantiques	Théories relativistes	Systèmes dynamiques
Existence d'un produit scalaire, permettant l'obtention de probabilité de mesure, et des motifs interférentiels, donc non linéaires. Techniques projectives sur des sous espaces vectoriels; la projection donne la mesure.	Existence d'un pseudo produit scalaire (métrique), permettant l'écriture d'un intervalle d'espace infinitésimal. Cet intervalle est invariant, non linéaire, et fonde la causalité.	Pas d'utilisation d'un produit scalaire explicite (sinon : espace euclidien), mais il a été montré que l'on pouvait "plonger" les systèmes différentiels dynamiques dans un espace de géométrie Riemannienne.
Existence d'une indétermination pré-mesure, mais également d'une incertitude irréductible et intrinsèque pour les variables conjuguées (eg : position et vitesse). Ce les relations d'Heisenberg.	Les "possibles" sont définis par des trajectoires appelées géodésiques; les trajectoires possibles dépendent de la positivité de l'intervalle d'espace et d'un principe de moindre action (minimisation).	Le temps est absolu <i>a priori</i>
Les états ou les groupes sont séparés, quantifiés. On passe de l'un à l'autre, par saut, non par processus continu.	Le temps et les mesures sont relatives.	Les trajectoires sont remplacées par des orbites, qui sont des courbes paramétrées en le temps, dans l'espace des phases.
Théories reposant sur l'analyse spectrale, l'invariance étant l'invariance de la base des états propres du système. Le temps y est absolu.	Les lois structurant l'espace sont elles invariantes.	La notion d'incertitude se traduit davantage par les notions de variabilité et de sensibilité aux conditions initiales, et d'imprédictibilité, de comportement ou transition chaotique.
La causalité est de type probabiliste (ou corrélationalle).	Pas de place explicite pour l'indétermination ou l'incertitude; elles peuvent être cherchées du côté des comportements non linéaires de la métrique.	Existence d'états stables, confinés à une région isolée de l'espace des phases, de dimension et de structure potentiellement fractales.
Pas de trajectoire, mais des intégrales de chemin.		Ces états sont liés par des trajectoires continues, et non par des sauts. Néanmoins, les variations brusque, les bifurcations et "catastrophes" peuvent, à une échelle d'observation donnée mimer un saut d'état.
Formalisme compatible avec la construction d'un Hamiltonien.	Formalisme compatible avec la construction d'un Hamiltonien.	Analyse spectrale pour la stabilité. Formalisme compatible avec la construction d'un Hamiltonien.

La remarque finale présente pour chaque type de modèle, la possibilité d'un Hamiltonien, est centrale pour la construction de notre théorie, ainsi que nous le verrons dans le chapitre dédié aux fondations de notre approche de la morphodynamique.

Si l'on compare les possibilités qu'offrent ces trois cadres théoriques au cadre théorique des RCT, on s'aperçoit, que les notions de temps et de causalité sont traitées totalement différemment. En effet, dans le cas des RCT, à moins éventuellement d'un design de type avant-après, la temps en tant que chronologie n'apparaît pas explicitement, du moins dans le mécanisme établissant la causalité. Le temps intervient uniquement comme arrière-plan nécessaire et lié aux contingences physiologiques, qui vont permettre d'exprimer une différence de résultat (il faut un certain temps pour qu'un traitement agisse). Mais hormis les études dites de survie, le facteur temps n'est pas utile, voire même est un facteur gênant car ne pouvant être comprimé ou contourné, aboli en pratique, à la démonstration de l'efficacité. Le RCT compare deux groupes, censément et statistiquement comparables en tout point en dehors du facteur de test qui constitue l'intervention. Il n'existe ici aucun principe dynamique, ni même de recherche d'un effet influant sur des caractères multidimensionnels, mais uniquement la recherche d'une différence que l'on estime indépendante, et générée par l'intervention.

Nous venons de passer en revue et de comparer les cadres théoriques, conformément à notre souhait d'assoir notre théorie sur des bases solides, éprouvées par ailleurs. En accord également avec ce que nous avons annoncé, nous proposons à la suite de passer en revue les aspects techniques et pratiques associés à ces trois modèles : ils sont l'assurance que nous disposons déjà d'outils pour tester dans un avenir proche, sous réserve d'adaptations *ad hoc*, nos hypothèses.

3.5 Les techniques associées aux 3 grands modèles

Des trois modèles retenus, tous issus de développements théoriques vus précédemment, il est possible d'extraire ou de relier des techniques de traitement des données particulières. À l'approche ponctuelle, linéaire et spectrale de la mécanique quantique, il est possible d'associer les méthodes linéaires et spectrales des analyses en composantes principales et assimilées ; à l'approche géométrique, topologique et locale des théories relativistes, il est possible d'associer toute une nouvelle classe de techniques, dites d'apprentissage des variétés (*manifold learning techniques*) ; enfin, à l'approche structurelle, dynamique, déterministe et chaotique des systèmes dynamiques, il est possible d'associer des théorèmes et techniques encore en développement de reconstruction des attracteurs à partir des données. Nous terminerons cette partie en mentionnant de récents résultats visant à quantifier et reconstruire des espaces inspirés de l'*epigenetic landscape* de Waddington.

3.5.1 Techniques projectives : ACP/PCA et PCA-likes

L'analyse en composantes principales (ACP, *PCA – Principal Component Analysis* – en anglais) est une des plus anciennes techniques de statistique exploratoire connues. C'est une méthode spectrale, et projective. Son principe repose sur des idées énoncées par Pearson en 1901, et formalisées par Hotelling en 1930. On parle également de transformée de Karhunen-Loève (transformée KL). [154, 133, 134, 136]

Pour un ensemble d'observations donné, de dimension n , l'analyse en composantes principales va travailler sur une matrice $n \times n$ de corrélations linéaires entre ces variables, ou sur une matrice dite de variance-covariance si les variables d'observation n'ont pas été réduites. Les matrices obtenues sont dans les deux cas carrées, symétriques réelles et donc diagonalisables (théorème spectral).

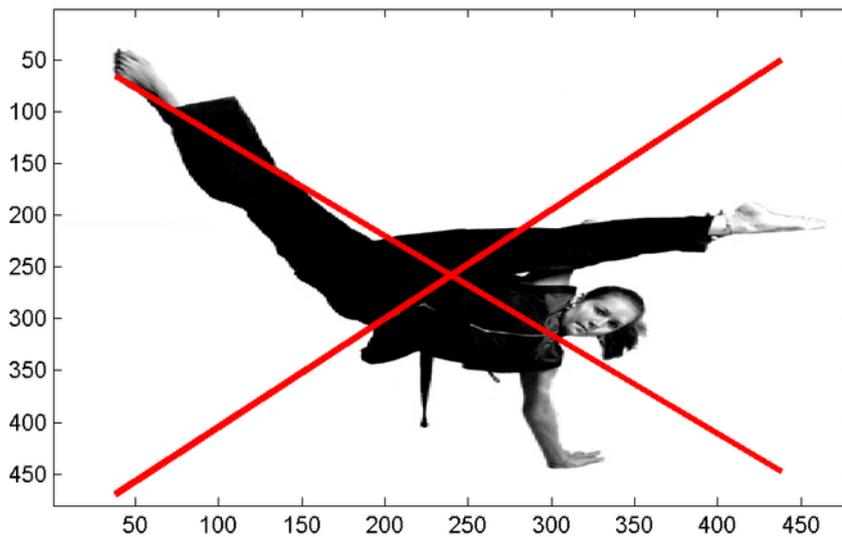
L'idée force ici est donc de chercher la base d'observations dans laquelle les corrélations autres que les auto-corrélations sont nulles ; cela revient à chercher la base dans laquelle les observations ne sont pas, ou sont le moins corrélées entre elles.

La nouvelle base obtenue est ainsi constituée des axes propres des données, au sens de la corrélation. Les données sont alors projetées sur ces nouveaux axes, que l'on appelle également axes principaux, ou axes d'inertie. Cette transformation est optimale au sens de la maximisation de la variance : les axes d'inertie sont classés par ordre croissant des valeurs propres associées, et chaque axe maximise la variance (de façon équivalente : l'inertie) expliquée par l'axe en question.

En pratique, s'il existe n observations, la matrice de corrélations va être carrée, de dimension $n \times n$, et la base obtenue sera également de dimension n . Après le réarrangement de la nouvelle base selon l'ordre croissant des valeurs propres, il est considéré que les derniers axes sont peu représentatifs des données, car expliquent peu de la variance de celles-ci, et peuvent être assimilés à du bruit. L'ACP est ainsi utilisée comme technique de réduction (ou de compression) de la dimension. Il n'existe cependant pas de critère théorique, autre qu'empirique, permettant de déterminer la valeur propre de "coupure", celle au-delà de laquelle il n'est pas nécessaire de décrire les données.

Si l'on considère que la (l'auto) corrélation est une représentation possible de l'énergie des données observées, prises en tant que signal, alors cette approche s'apparente à chercher les valeurs propres de l'énergie et les états stationnaires associés, de la même façon que les valeurs propres de l'Hamiltonien donnent les niveaux d'énergie du système en mécanique quantique, et les vecteurs propres les états stationnaires du système.

FIGURE 3.13 – Les deux axes d’inertie selon l’analyse en composantes principales (Wikipedia, licence CC)



Ainsi formulée, nous constatons que l’ACP est une méthode totalement linéaire, où l’on cherche le meilleur espace de représentation en termes de dimensions les moins corrélées (linéairement), et permettant des projections sur des sous-espaces maximisant sur chacun de ses axes l’inertie expliquée. Cet espace peut être un espace de plus petite dimension que l’espace initial. Une illustration pratique est donnée par la figure 3.13 : pour cette image en deux dimensions, les axes principaux d’inertie sont au nombre de deux également, et représentés par les lignes rouges.

L’espace obtenu sert fréquemment d’espace exploratoire, en particulier, d’exploration visuelle : on y traque les formes les plus homogènes ou concentrées possibles ; cependant, au-delà de deux ou trois dimensions, il commence à devenir compliqué de visualiser correctement la réalité de ces formes. Ce qui peut paraître une forme bien délimitée dans une représentation en deux dimensions – par exemple les deux premiers axes d’inertie – peut très bien ne

plus avoir autant de substance évidente lorsque l'on en visualise les éléments constitutifs selon les axes suivants. Enfin, le pouvoir explicatif, au sens causal du terme, est assez faible : il est essentiellement descriptif.

Cette première technique, l'ACP, a été conçue pour les données continues. Il en existe une formulation pour les données catégorielles, discrètes. En lieu et place de travailler sur une matrice de variance-covariance, il est proposé de travailler sur une table de contingences. Cette table de contingences peut à son tour être diagonalisée ; on parle d'analyse en correspondances multiples (ACM).

De manière plus générale, l'ACP et autres techniques assimilées appartiennent à la même famille de l'analyse factorielle. Ces techniques partagent toutes le même principe de diagonalisation d'une matrice de variance-covariance ou de contingence, mais cette matrice n'est pas constituée des mêmes éléments, selon la méthode utilisée et la nature des données.

Il est notable de relever que, en particulier dans la tradition de l'école française, et plus spécifiquement encore en sciences sociales et économiques, il existe un attelage quasi-systématique de deux techniques : l'utilisation d'une ACP (ou d'une ACM), avec ou sans réduction de la dimension, puis application d'une technique de *clustering* afin d'identifier des objets dans les données.

Initialement, l'une des justifications admissibles à ce systématisme était que l'ACM s'utilisait comme une transformation conservative des données catégorielles en données continues, avec le meilleur système d'axes de représentation au sens de l'inertie. En effet, juste après l'application de l'ACM, on appliquait une méthode de *clustering*, véritable objectif de l'analyse ; or, la majorité des techniques utilisées en pratique courante est conçue pour

s'appliquer sur des données continues, là où les données observées sont plus fréquemment catégorielles. Cette justification reste partiellement convaincante : en réalité, et ce, dès les premiers algorithmes de *clustering*, il existe des techniques, certes moins nombreuses mais tout aussi efficaces et surtout adaptées, conçues pour s'appliquer soit sur des données catégorielles, soit même sur des données mixtes (catégorielles et continues), telles que l'algorithme PAM (*Partitioning Around Medoids*), bâti sur un principe similaire aux k moyennes. [137]

Cette pratique, sans être inintéressante ou injustifiable dans des cas précis et motivés, soulève le problème de la conservation effective de l'information. Comme nous l'avons montré, par construction, l'ACP conserve et ordonne l'information à partir de considérations basées sur les corrélations linéaires entre observations. D'une part, dans l'hypothèse où la dimension finale après ACP soit la même que la dimension initiale, l'optimalité de cette transformation au sens de l'inertie ne garantit en aucun cas l'optimalité au sens d'autres critères, par exemple, un critère de préservation de la topologie ou de la complexité des relations entre données, si celles-ci ne sont pas d'une nature purement statistiques, probabilistes, sans autre loi sous-jacente que celle de leurs distributions respectives ; d'autre part, si la dimension retenue après application ACP est inférieure (s'il y a compression des données) à celle initiale, il est un peu rapide de conclure que toute l'information pertinente est bien présente dans les axes retenus, et que ce qui a été éliminé ne correspond qu'à du bruit, supposé additif qui plus est.

Cette question de préservation de l'information et de la complexité, de reconstruction d'un espace de représentation plus adapté à la recherche des formes, des invariants, nous ramène une fois de plus à la question de la géométrie.

3.5.2 Techniques d'apprentissage des variétés – NLDR

L'ACP, et plus généralement, les techniques d'analyse factorielle et leurs variantes à noyau (à *kernel*), ont et continuent de donner dans un certain nombre de cas de très bons résultats. Elles ont pour elles, entre autres, l'avantage de la simplicité théorique et d'une rapidité d'exécution non négligeable. L'interprétation des données – les axes obtenus sont des combinaisons linéaires des observables de départ – peut elle déjà poser plus de problèmes, en ce sens où les axes d'inertie n'ont pas nécessairement un sens aussi évident que pouvaient sembler avoir les observables initiales : que signifie un axe fait de 40% de sexe féminin, 27% d'âge moyen, 12% de pointure 50, etc. ?

Devant le monopole de l'approche type "maximisation de l'inertie sur la base des corrélations linéaires" et les insuffisances de cette approche dans plusieurs cas, comme celui de la classification des expressions faciales, par exemple, en très basse dimension, d'autres techniques ont été proposées.

Les premières datent de 2000 et sont à ce titre relativement récentes. Elles sont classées sous le terme générique de techniques non linéaires de réduction de la dimension(nalité) (*nonlinear dimensionality reduction techniques* – NLDR). Le prototype en est l'algorithme ISOMAP, pour *ISometric MAPping*. [155, 156]

ISOMAP met à profit dans sa seconde partie une méthode déjà connue et utilisée par ailleurs, depuis les années 30, à savoir le *multidimensional scaling* – MDS, elle-même linéaire par construction. [157] L'originalité de la méthode tient dans sa première partie, qui se base sur des considérations topologiques. En effet, l'idée est de déterminer les voisinages de chaque point, par exemple, les personnes les plus directement similaires d'une personne donnée, et ce pour chaque personne, puis de déterminer les distances point

à point, personne à personne, les plus courtes; autrement dit, les distances géodésiques entre personnes.

La méthode proposée pour ce faire est simple et élégante, elle nécessite la construction d'une matrice de distances point à point, laquelle est donc par définition carrée, de dimension $n \times n$, où n est le nombre de personnes.

Ces distances sont calculées de deux façons différentes, et passent par l'utilisation d'un graphe. Pour une personne donnée, la distance à ses k plus proches voisins est calculée par la distance euclidienne, ce qui correspond à une approximation d'espace plat localement, autour de la personne. Cette opération est répétée pour chaque personne. Pour les personnes au-delà de ces k plus proches voisins, on cherche la plus courte distance reliant deux personnes. Cette plus courte distance est calculée comme le plus court chemin dans un graphe étiqueté, graphe représentant les relations de voisinage : chaque personne est un nœud du graphe, et chaque arête est étiquetée de la distance euclidienne entre les deux nœuds. La distance entre deux nœuds non directement voisins est alors la plus courte distance reliant ces deux nœuds. Le calcul des chemins les plus courts est assuré par un algorithme de type Dijkstra. Les géodésiques sont donc estimées à partir de la somme des distances euclidiennes, suivant le plus court chemin : la courbe globale reliant deux personnes éloignées est approximée à partir de segments calculés localement. De cette façon, il est possible de dire que l'algorithme ISOMAP est un algorithme préservant la topologie des relations inter personnes ou inter observations.

La matrice des distances point à point est ensuite passée comme entrée d'un algorithme classique type MDS, auquel est donnée la dimension finale visée.

Bien sûr, ce genre de technique ne pallie pas toutes les insuffisances des approches linéaires et projectives, et souffre également de limitations plus ou moins importantes. La première, inhérente à cette méthode particulière, tient dans la nécessité de devoir spécifier le type de voisinage, c'est-à-dire, le nombre de voisins à chercher autour d'une personne. D'une façon ou d'une autre, ce choix peut influencer la topologie rendue. C'est notamment le cas s'il existe un sous échantillonnage du système, ou plus grave, un échantillonnage non adapté à certaines parties du système ; dans tous les cas, en l'absence d'algorithme adaptatif en la densité informationnelle et de connaissance précise sur la qualité de l'échantillonnage, le meilleur voisinage à retenir sera celui respectant les parties les moins finement échantillonnées, afin de ne pas aller chercher des voisins proches qui seraient en réalité éloignés dans les parties faiblement peuplées de l'échantillon. Dans le cas où l'on serait dans une partie plus densément peuplée, chercher moins de proches voisins qu'il serait adapté de le faire *a priori* portera moins ou pas à conséquence.

Dans les parties peu densément peuplées, le risque majeur sera ainsi de relier deux points qui ne devraient pas l'être, en autorisant par exemple des "sauts" topologiques théoriquement non valables. De ce fait, la stabilité topologique de l'algorithme ISOMAP avait été questionnée, et il avait été conclu à un comportement relativement robuste en les variations de k , k étant le nombre de plus proches voisins. [158]

La seconde limitation tient davantage à la nécessité de spécifier la dimension visée, et donc estimer la dimension intrinsèque des données, si l'on est dans une optique de préserver la géométrie et la topologie intrinsèques des données. Nous reviendrons sur cette problématique un peu plus loin, et proposerons un ensemble de solutions à ce problème.

Hormis l'algorithme ISOMAP, d'autres techniques, d'autres approches, certaines basées également sur l'exploitation d'un graphe, ont été proposées. L'un des packages logiciels réunissant ces techniques en dénombreait presque une quarantaine il y a encore peu. [159, 160] Néanmoins, beaucoup sont des variantes autour de quelques idées originales. Nous en citons ici quelques unes, et proposons d'en comparer brièvement les performances sur des exemples synthétiques.

Nous pouvons recenser les approches "physiques", basées par exemple sur l'utilisation du Laplacien (méthode spectrale, où l'on cherche les valeurs propres du Laplacien) ou d'une équation de diffusion type équation de la chaleur (techniques dites de cartes de diffusion – *diffusion maps*) ou les approches géométriques locales, qui cherche une approximation linéaire optimale localement, c'est-à-dire, la meilleure base locale pour représenter les points. La première version de l'algorithme LLE (*Locally Linear Embedding*) fut publiée simultanément à ISOMAP, [161, 162] et une version améliorée fut proposée par la suite, en lui adjoignant un critère hessien (HLLE, pour *Hessian LLE*). [163] L'une des idées les plus simples et élégantes géométriquement fut proposé plus tard, sous le nom de LTSA, pour *Local Tangent Space Alignment*. [164] Le principe en est très simple : il s'agit là encore d'approcher localement les données par leur (hyper)plan tangent, et de considérer que si les données suivent une géométrie courbe, alors il est possible d'en respecter la topologie globale en "décourbant" l'espace, c'est-à-dire en alignant les espaces tangents les uns avec les autres.

Bien entendu, comme pour un problème de cartographie, où l'on sait qu'il n'existe pas de projection parfaite du globe terrestre, c'est-à-dire, préservant à la fois les relations angulaires et les relations métriques, [165] tel algorithme NLDR favorisera un déploiement de l'espace respectant les relations métriques (c'est le cas d'ISOMAP, par exemple), et tel algorithme favorisera

quant à lui un déploiement respectant les relations angulaires.

De la même façon que les algorithmes de *clustering* nécessitent dans leur vaste majorité l'introduction d'un paramètre à fixer, le plus courant étant le nombre de groupes à déterminer, la majorité des algorithmes NLDR requièrent en entrée un, voire deux paramètres, ce qui en limite l'utilisation "en aveugle" ou automatique. L'équivalent du nombre de groupes est ici la dimension à viser. Dans le cas d'ISOMAP, nous avons vu qu'il fallait spécifier la dimension à viser, ainsi que le type de voisinage (le nombre k des k plus proches voisins).

Outre leur bonne capacité à restituer la géométrie des données, les différents algorithmes peuvent également être jugés sur leur complexité algorithmique et leur rapidité d'exécution. Ce dernier critère fait que si ISOMAP est parmi les des plus performants en termes de préservation de la topologie, reconnu pour être topologiquement stable et robuste au bruit, il est beaucoup plus critiquable en termes de rapidité ou de scalabilité. En effet, même si l'algorithme de détermination de plus court chemin dans le graphe des distances est rapide est efficace, la seconde étape est beaucoup plus lente – c'est un problème connu de MDS. En regard, un algorithme comme LTSA a l'avantage de se comporter de manière équivalente à ISOMAP en termes de fidélité à la géométrie des données, et d'être beaucoup plus rapide.

Nous comparons ici les performances de six de ces algorithmes, et des approches linéaires classiques, type ACP et MDS. Nous voyons que, sans surprise, les techniques linéaires et projectives, comme annoncé, projettent les données sur les axes d'inertie retenus. Ainsi, si les données ne s'organisent pas selon un principe purement probabiliste, et sont structurées selon une géométrie plus ou moins complexe et courbe, les techniques projectives sont dépassées, et ne préservent pas la topologie intrinsèque. Les résultats sont

consignés dans les figures 3.14, 3.15, 3.16 et 3.17.

Actuellement, sur des données non synthétiques, mais dont on connaît la vérité terrain, les résultats comparatifs des techniques linéaires d'un côté et des techniques non linéaires de l'autre ne sont pas aussi tranchés, et méritent éclaircissement. Il est possible qu'*in fine*, les techniques NLDR fonctionnant correctement soient plus efficaces en général, mais soient encore aujourd'hui mal utilisées, faute de démarche un peu systématique, notamment en ce qui concerne la recherche de leurs paramètres optimaux en fonction des données et de la qualité de l'échantillonnage.

Il est néanmoins tout-à-fait possible qu'aucune des deux approches soit systématiquement meilleure que l'autre, la structure prédominante des données à étudier pouvant éventuellement être régie par des principes totalement différents. Cette constatation nous renvoie au dilemme de la dualité, à savoir qu'il pourrait bien coexister, dans l'attente d'un possible principe unificateur, deux modes de représentation de la réalité, notamment en fonction de l'échelle étudiée, comme c'est le cas en physique.

3.5.3 Techniques de reconstruction des attracteurs

Nous abordons ici le troisième type de technique d'identification de formes et de reconstruction d'espace à partir d'observations. Ce type est directement hérité de la vision des systèmes dynamiques, puisqu'il s'agit ici de chercher à établir l'existence d'attracteurs dans des données, et le cas échéant, de les reconstruire. En effet, si la théorie des systèmes dynamiques est bien assise et documentée, et l'existence d'attracteurs engendrés par la structure même des systèmes d'équations régissant le système à décrire prouvée, la question inverse se pose, à partir d'observations d'un système physique un tant soit eu compliqué, difficilement modélisable exhaustivement et non résoluble analytiquement, de savoir si ces observations recèlent un attracteur.

FIGURE 3.14 – Trois ensembles simples de données, et leur déploiement en deux dimensions

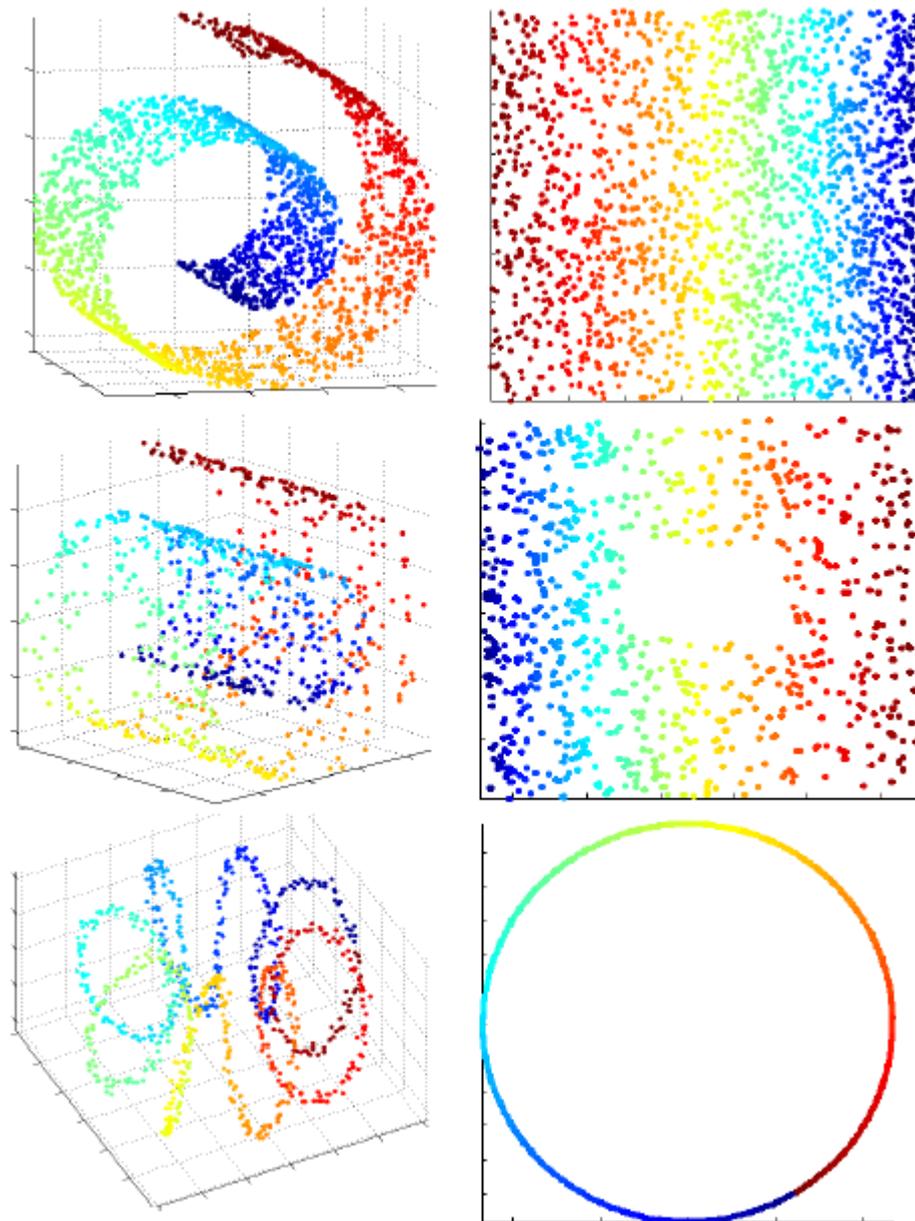


FIGURE 3.15 – 8 algorithmes de déploiement de variétés appliqués à l'ensemble rouleur suisse

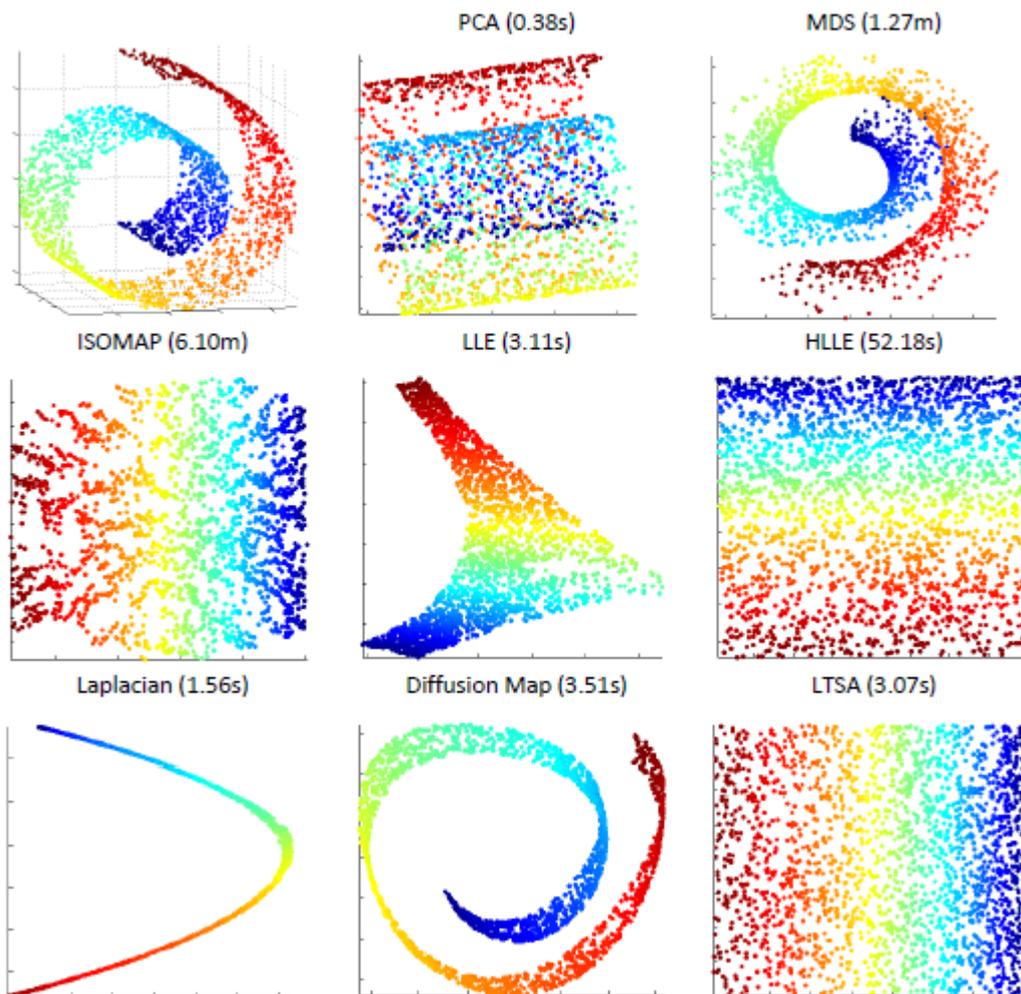


FIGURE 3.16 – 8 algorithmes de déploiement de variétés appliqués à l'ensemble rouleur suisse percé

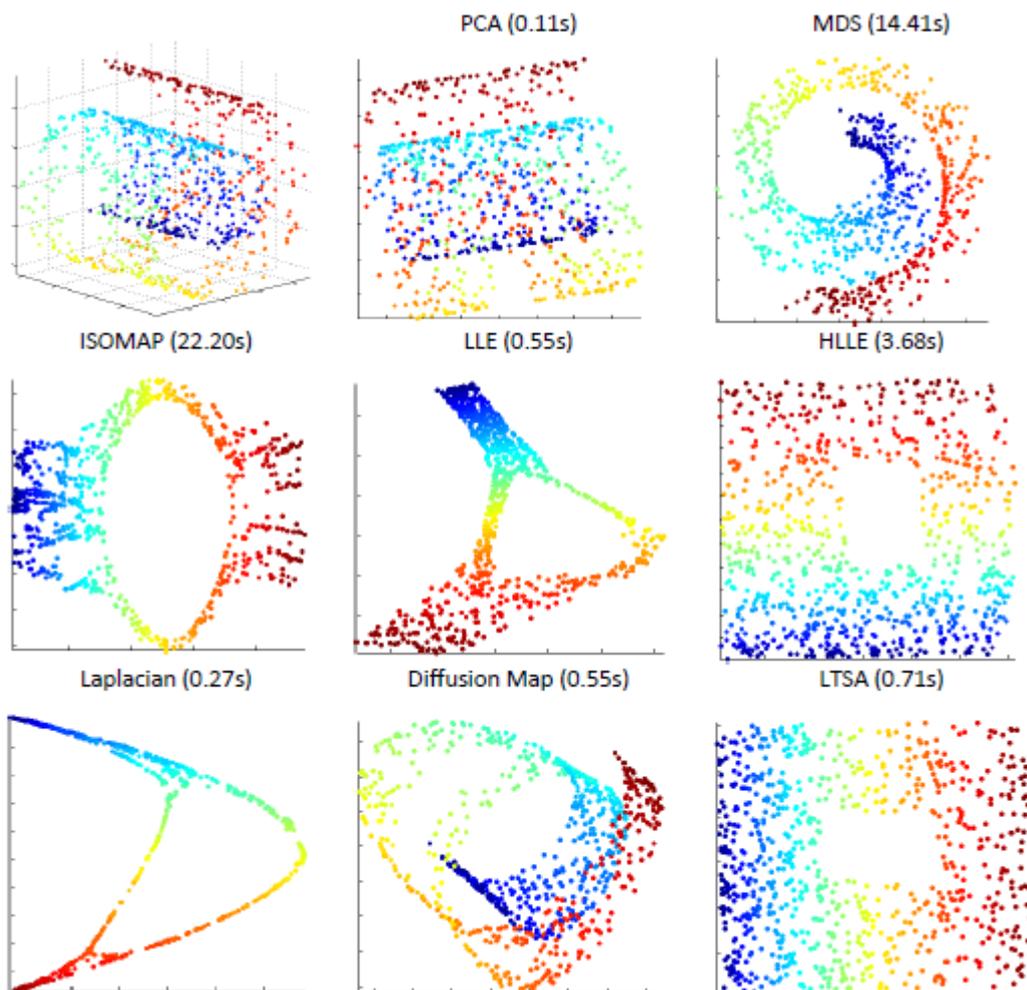
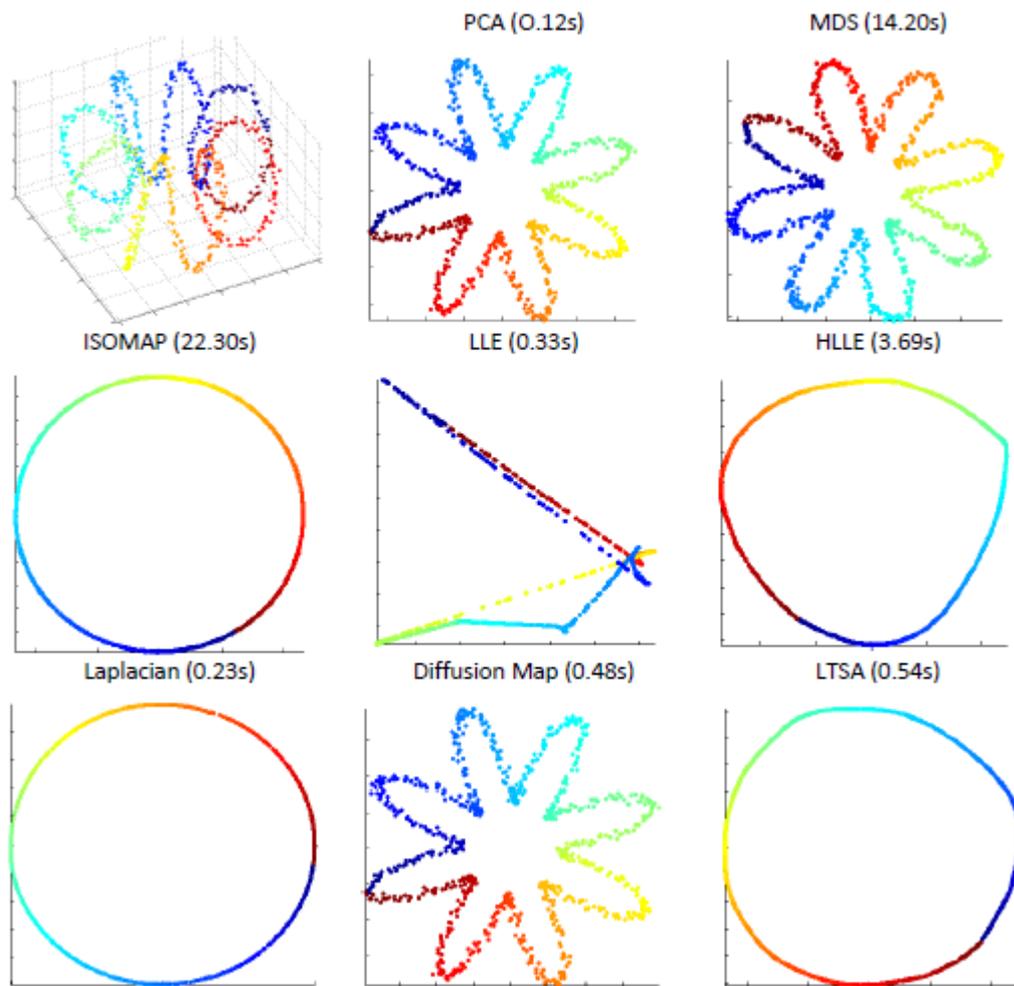


FIGURE 3.17 – 8 algorithmes de déploiement de variétés appliqués à l'ensemble hélice toroïdale



Les résultats généraux concernant ces approches sont eux aussi relativement récents, puisqu'il faut attendre le théorème dit de Takens, pour disposer de la preuve formelle qu'il est possible d'établir l'existence d'un attracteur dans des données si celui-ci existe. [166, 167]

Plus récemment encore, d'autres théorèmes ont été proposés concernant l'existence, mais surtout la reconstruction de la forme de ces attracteurs à partir des données observées. [168, 169] Les techniques proposées ont-elles aussi, à l'instar des techniques précédemment exposées (ACP et techniques NLDR), leurs limitations. L'équivalent du problème de la détermination du nombre de groupes pour le *clustering* et du nombre de voisins pour ISO-MAP est ici le problème du délai, de l'écart de temps entre deux mesures temporelles des observations. En effet, il existe une sensibilité des résultats de reconstruction de l'attracteur quant à ce paramètre. Ce paramètre peut ici encore être comparé à un paramètre de définition du voisinage ou de l'échantillonnage.

3.5.4 Problème de la dimensionnalité *vs* complexité

Nous venons de voir trois types de techniques différentes, toutes visant à la reconstruction d'un espace de représentation minimal. Chacun est le représentant d'une approche différente. Toutes ces techniques posent le même double problème, sans nécessairement l'explicitier : celui de la dimension intrinsèque des observations utilisées, et de la complexité de ces données – ou du moins, la façon dont cette complexité est gérée.

Complexité et dimensionnalité : une autre approche et implications de "*la malédiction de la dimensionnalité*".

Il a été montré que les opérations classiques, impliquant l'utilisation d'une distance (par exemple, une distance euclidienne), peuvent se comporter de

manière contre intuitive lorsque on les applique à des espaces de haute dimensionnalité. Par haute dimensionnalité, nous n'entendons pas nécessairement des espaces de 100 variables, mais des espaces regroupant, une vingtaine de variables. Il semblerait même que les problèmes puissent apparaître dès une dizaine de variables.

Il existe deux problèmes liés à la haute dimensionnalité des espaces, potentiellement liés l'un à l'autre, mais jusque lors présentés comme indépendants l'un de l'autre. Le premier découle d'un problème de combinatoire très simple, et peut expliquer au moins partiellement pourquoi les études *genome-wide* sont difficilement reproductibles, ou fournissent des résultats peu probants et inconsistants. Ce premier problème est nommé "phénomène de l'espace vide", [136, 170] et découle simplement de ce qu'il existe des combinaisons de variables parmi toutes celles enregistrées, qui sont sous échantillonnées, ou même plus directement, inexistantes. En génétique, par exemple, on imagine mal comment peupler suffisamment l'ensemble des combinaisons concernant une cinquantaine de gènes et leurs possibles différents allèles et mutations, même avec une population d'étude large. En outre, il paraît peu vraisemblable que toutes les combinaisons soient équiprobables voire même existantes. L'espace obtenu est ainsi un espace très creux, sur lesquels on applique encore trop souvent, si ce n'est systématiquement, des techniques statistiques classiques inadaptées. La principale concession à la spécificité – non pas spécifiquement la haute dimensionnalité, mais la taille brute des données : le nombre de points, en quelque sorte – de ce type d'études a été de reconsidérer les seuils de significativité des tests statistiques - on parle d'adaptation de la valeur du petit p pour les tests multiples.

Le second problème est appelé "malédiction de la dimensionnalité" (*curse of dimensionality*), ainsi que l'a nommé Bellman lorsqu'il s'est rendu compte du comportement étrange des distributions normales en haute dimension. [136, 170, 171] Il s'est intéressé à la possibilité de discriminer deux indivi-

dus qui seraient différents, et dont les caractéristiques répondraient d'une loi normale multidimensionnelle. Étonnamment, plus le nombre de caractéristiques augmente, c'est-à-dire, plus la dimensionnalité de l'espace croît, plus il devient difficile de distinguer deux individus différents. En effet, la norme de leurs caractéristiques tend vers une norme moyenne, commune à tous les individus de l'espace, ce qui les rend très peu différents, de ce point de vue, les uns des autres. Formulé autrement, on constate une concentration croissante des individus en périphérie de l'hypervolume constitué des différentes variables mesurées. Tous les individus se retrouvent dans une fine couronne (dans le cas d'une hypersphère) dont l'épaisseur tend vers 0 avec la dimension de l'espace. Ainsi, alors que l'on espère mieux discriminer deux individus en enregistrant davantage d'informations – de nature différente – à leur propos, et moins il devient aisé de les discriminer dans cet espace plus "complet", paradoxalement.

La question vient alors de savoir s'il existe globalement ou localement, une dimension intrinsèque de l'espace de haute dimensionnalité, que ce soit en termes de réduction de la redondance de l'information, ou en termes d'espace minimal et suffisant à adapter un modèle performant expliquant ou prédisant un certain nombre de variables d'intérêt. Il existe différentes méthodes proposant d'estimer cette dimension, laquelle est par ailleurs en général non entière – ce caractère est à relier à la possible nature fractale de l'espace ou du phénomène étudiés. En revanche, il n'existe pas à l'heure actuelle et à notre connaissance, de consensus quant à savoir quel estimateur de dimension est à appliquer dans une situation donnée plutôt qu'une autre. Une revue des différents estimateurs a été proposée, qui tend à les catégoriser, notamment en approches locales ou globales, ou en fractale, non fractale. [172]

Des tests sur des ensembles générés par ordinateur pour estimer la di-

mension intrinsèque de différentes variétés semblent montrer le meilleur comportement d'un estimateur particulier, hérité de la théorie des systèmes dynamiques, à savoir l'estimateur du maximum de vraisemblance. [173] Nous suggérons d'éviter, en revanche, l'estimateur basé sur l'ACP, qui lui rend une dimension entière quoi qu'il en soit de la réalité sous-jacente, pouvant donc se tromper d'au moins une unité.

3.5.5 Quantification de *l'epigenetic landscape*

Nous finissons cette partie et ce chapitre sur une approche un peu hybride vis-à-vis des techniques que nous venons de discuter. Nous avons déjà parlé de *l'epigenetic landscape* de Waddington, lequel est resté assez longtemps un principe non éprouvé, car dépourvu de formalisation technique opérationnelle ; aucune méthode explicite et pratique n'avait été proposée afin de tester ou de construire ce paysage.

Nous ne rentrerons pas dans le formalisme développé afin de le rendre plus accessible aux expérimentateurs, et de proposer un outil de reconstruction à partir des mesures effectuées, mais nous désirons porter à la connaissance du lecteur des travaux récents, dont un en particulier, qui doivent retenir l'attention. Un papier s'est intéressé à ce que les auteurs ont appelé la quantification de *l'epigenetic landscape*. Leur entreprise doit pouvoir servir, à n'en pas douter, de base solide pour l'exploration d'ensemble de données, selon le cadre explicatif de Waddington. [120]

Chapitre 4

Capturer, intégrer et traiter l'information : exemples et frontières actuelles des approches par la forme

Sommaire

4.1	Prédire ou expliquer	210
4.1.1	Modèles explicatifs	211
4.1.2	Modèles prédictifs	212
4.1.3	Limites des approches synthétiques en statistique : revues et méta-analyses	214
4.2	La géométrie des infections comme démarche prototype	216
4.2.1	Une brève histoire des maladies infectieuses comme pathologie archétypique	217
4.2.2	La place particulière du VIH-SIDA	220
4.2.3	Panorama actuel et concepts en maladies infectieuses	233
4.2.4	Vers une géométrie des maladies infectieuses : to- pologie de la causalité	238
4.2.5	Au-delà de la pathologie infectieuse : l'exemple du cancer et de son diagnostic	240
4.3	Modéliser un phénomène infectieux	242
4.3.1	Le modèle compartimental et les épidémies	243

4.3.2	Approche des infections ostéo-articulaires : un cas typique	245
4.3.3	Modélisation du système immunitaire	246
4.3.4	Le modèle de réponse immunitaire non spécifique en 3 vagues de l'infection pulmonaire	247
4.3.5	Modélisation de l'homéostasie osseuse	251
4.3.6	Modèle d'IOA : hypothèses de couplage des systèmes	255
4.3.7	Intégration de la stratégie thérapeutique au modèle d'infection ostéoarticulaire	257
4.3.8	Complexité du modèle et analyse de sa stabilité structurelle : limites de l'approche	260
4.4	Typologies et dynamiques sociales	261
4.4.1	Existe-t-il des formes sociales? Premières tentatives en sociologie	262
4.4.2	Exemple en épidémiologie sociale : une typologie de recours aux soins en France	265
4.4.3	Opérationnalité des classes et recherche interventionnelle	269
4.4.4	Entre sociologie des interactions et architecture urbaine : de la taille des villes	272
4.5	Au-delà de la pathologie : les facteurs de risque, les syndromes et autres objets de transition	273
4.5.1	L'exemple des maladies chroniques	278
4.5.2	La pathologie ne saurait plus se baser uniquement sur une donnée uniquement physiologique	280
4.5.3	Polypathologies chroniques : quelles dimensions?	283
4.6	Santé mentale	286
4.6.1	Psychanalyse et physique, une histoire commune : Mach, Freud et Einstein	287
4.6.2	Bref historique de l'objet et de ses classifications en psychiatrie	288
4.6.3	La tentation de la génétique et des neurosciences : insuffisances et impasses	290
4.6.4	Un cas d'école de <i>big data</i> : approche géométrique, intégrative et dynamique basée sur l'observation	293
4.6.5	Principe de modélisation d'un objet psychiatrique : le modèle de Lorenz et la bipolarité	301
4.7	Intégration et traitement des données : les <i>big data</i> en pratique	311
4.7.1	De l'autre côté du spectre des RCT, la fouille de données (le <i>data mining</i>)	311
4.7.2	Capture de routine et entrepôts de données	314
4.7.3	Concept de CDC	315
4.7.4	SPOT-RISC	318
4.8	Retour sur les IOA	329
4.8.1	Modélisation ABM pour la spatialisation	329

4.8.2	Intégrations multi-échelle, multimodale, patient-spécifique	332
4.9	Ce que l'on en retient	333

4.1 Prédire ou expliquer

Nous abordons ce nouveau chapitre dédié aux frontières actuelles, entre approches classiques et approches par la forme, par une distinction qui nous semble importante, et qui, si elle n'est pas faite, peut engendrer confusion et déception. Lorsque l'on parle de modèle ou de modélisation, il n'est pas toujours évident, car peut-être trop souvent implicite, de savoir quel genre d'opérationnalité l'on attend du modèle : est-on dans le domaine du calcul et du calculable, pour atteindre des résultats de manière algorithmique, sûre au sens de la valeur de vérité du résultat délivré, ou bien est-on dans le domaine du modèle conceptuel, du modèle qui illustre et explique une partie du monde ? Certes, il existe des exemples célèbres, historiques et synthétiques qui combinent au final assez bien les deux qualités, mais il semble que cela devienne de moins en moins vrai.

Autrement dit, la question première en modélisation est la suivante : prédire, ou expliquer ? C'est-à-dire, aux extrêmes, un outil le plus efficace possible à la tâche qui lui est dévolue, mais dont le fonctionnement en routine nous est totalement incompréhensible, ou bien une représentation simple qui permette le raisonnement sur une partie de la réalité, que notre intelligence peut manipuler conceptuellement, sans qu'il soit capable en l'état soit d'être calculable, soit de prédire quoi que ce soit correctement ? Les deux types d'approche peuvent être d'excellents supports d'aide à la décision, mais jouant dans des registres totalement différents. Le premier sera notamment roi dans le domaine du quantitatif pur, le second dans le domaine du qualitatif et de l'inductif. En effet, "prédire n'est pas expliquer".[174]

4.1.1 Modèles explicatifs

Nous ne manquons pas de modèles explicatifs. Dans un certain sens, un modèle explicatif, voire éducatif, pourrait être totalement faux ou inopérant en termes de calculabilité, mais parfaitement opérant en ce sens où il permet à la personne d'en faire jouer les mécanismes, et de s'en servir comme support à penser pour décider et agir. En ce sens, un modèle explicatif pur, si cela a un sens, est une histoire, un mythe que l'on se raconte, du registre de l'histoire, voire de la parabole. Il est cependant opérant dans le domaine de la communication intra et interpersonnelle, car s'appuie sur un langage et des éléments de réalité.

Nous pouvons penser au modèle de la personnalité de Freud (deuxième topique), qui structure et hiérarchise la psyché en ça-moi-surmoi : c'est un modèle relativement simple qui tend à donner une explication du fonctionnement d'un objet inaccessible, voire inlocalisable – où vit la psyché ? Pensons également aux modèles successifs de la matière en termes d'atomes : la notion d'atome comme élément constitutif de la réalité, comme grain irréductible, remonte à l'antiquité, se retrouve chez les latins comme chez les grecs. Ils permettent le discours, et même *in fine* l'observation d'un aspect de la réalité, puisque l'on reconnaît à l'atome une réalité objective et mesurable. Néanmoins, les différents modèles, même satisfaisant de plus en plus aux critères de scientificité, ont montré leurs limites prédictives successives, par confrontation à l'expérience (réelle ou de pensée) : le modèle de Bohr et la catastrophe ultraviolette ont ainsi ouvert la porte à la représentation quantique des particules – il existe des quantas dans la nature, mais les particules ne sont pas aussi simplement décrites qu'une planète bien localisée possédant des satellites orbitant autour d'elle. Les électrons sont délocalisés, manifestent des propriétés ondulatoires et non granulaires, mais agissent également comme des particules ponctuelles. Il demeure cependant encore plus simple d'expliquer le modèle planétaire de l'atome que celui de Schrödinger,

où les orbitales remplacent les orbites.

4.1.2 Modèles prédictifs

L'exemple quantique offre une bonne transition pour discuter la différence modèles explicatifs et modèles prédictifs : là où le modèle quantique pour les particules a remplacé un modèle plus figuratif, l'on a gagné en précision de prédiction sur plusieurs ordres de grandeurs, mais ceci au sacrifice d'une intelligibilité incomparable des principes sous-jacents.

Bien sûr, par morceaux, l'on comprend les principes, pris individuellement les uns des autres, de même que les postulats fondant la mécanique quantique. Mais l'ensemble, ainsi que son interprétation – s'il n'y en a qu'une de valable – demeurent insaisissables. Cet état de fait nourrit par ailleurs des difficultés imprévues en termes de projections, de "sens physique", car aussi performant le système soit-il pour prédire avec exactitude un certain nombre de résultats, l'intuition concernant la théorie, le raisonnement inductif en ont souffert. Or, en sciences comme en d'autres domaines, il est important d'avoir l'idée, l'intuition d'un résultat, puis ensuite d'en démontrer rigoureusement le chemin qui mène d'un point à un autre, plutôt que de procéder de façon inverse : tâtonner de façon systématique pour produire un résultats inattendu mais directement intelligible. En outre, d'un point de vue strictement calculatoire et de complexité, il est très rarement rentable, si ce n'est impossible, de procéder de façon exhaustive : d'une part la combinatoire peut être extrêmement importante et croître de manière exponentielle, d'autre part, la complexité intrinsèque du problème peut imposer de recourir à des heuristiques, des approximations. A nouveau, il est difficile de créer de telles approches sans avoir une sensibilité particulière du problème.

Cependant, il existe des domaines où finalement, peu importe si l'exécution *in concreto* de la méthode n'est pas intelligible, si celle-ci remplit au moins deux conditions : elle est efficace et précise à prédire un objet

particulier, et le résultat qu'elle délivre est fiable. Ainsi, en apprentissage automatique, un domaine où l'on cherche à faire effectuer sans supervision à une machine, pour des gains de temps, de productivité ou autres, une ou plusieurs tâches particulières, l'on peut parfaitement recourir à de telles méthodes, parfaitement prédictives, et de valeur explicative nulle. Ce peut être vrai parce qu'elles opèrent sur des masses de données, en volume ou en variété, telles que l'esprit ne saurait les manipuler aussi habilement ou aussi rapidement – mais dans l'idéal, l'on sait contrôler la qualité de la prédiction de la méthode. C'est par exemple le cas des réseaux de neurones, ou des séparatrices à vastes marges (SVM) : le principe de base est connu – fonctionnement du neurone artificiel comme fonction de transfert non linéaire, utilisation d'un noyau pour projeter en haute dimension un problème de plus basse dimension et recherche de l'hyperplan séparateur des données – mais la mise en pratique, effective, n'est pas ou très peu compréhensible. Pour autant, ces deux approches permettent de traiter avec d'excellentes performances différentes classes de problèmes. Ce sont de très bons modèles prédictifs.

Il existe une classe intermédiaire entre ces deux types d'exemples, explicatifs *versus* prédictifs, caractérisée un peu différemment. Il s'agit des modèles de systèmes dynamiques exhibant un comportement chaotique, de type chaos déterministe. Le modèle spécifié est intelligible *a priori*, son pouvoir prédictif à court terme éventuellement très bon, mais il peut rencontrer des configurations où son pouvoir prédictif devient nul, sans que cela soit un problème de défaut d'information, mais par sa nature même : les trajectoires deviennent chaotiques, et par définition imprévisibles, même selon de toutes petites variations autour d'un état connu.

Enfin, ainsi que le révèlent le climat actuel de crise méthodologique en biomédecine, la critique des stratégies de l'industrie pharmaceutique en mal

d'innovation thérapeutique réelle – et qui consiste à davantage soit élargir le panel de personnes susceptibles d'être traitées pour une pathologie particulière : abaissement des seuils diagnostiques, soit créer de nouvelles entités nosologiques pour lesquelles des prescriptions deviendraient nécessaires – certaines confusions demeurent et entretiennent le flou entre pouvoir prédictif et pouvoir explicatif. C'est particulièrement le cas pour les facteurs de risques cardiovasculaires, comme l'hypertension ou l'hypercholestérolémie (excès de cholestérol dans le sang). Traiter systématiquement par le médicament l'hypercholestérolémie n'a pas de sens en soi, même s'il s'agit d'un facteur de risque cardiovasculaire. En effet, le cholestérol n'a pas de rôle causal dans le développement de la pathologie ; il est en revanche de manière quasi-systématique, un bon prédicteur de la mortalité cardiovasculaire. Ainsi, en traitant ce facteur de risque, on n'abaisse pas nécessairement le développement d'une pathologie, seulement un marqueur possible d'une mortalité future et prématurée. Voire pire, il est possible qu'en abaissant le taux de cholestérol, par ailleurs vital, l'on entrave le bon fonctionnement physiologique de l'organisme. [175, 102]. En revanche, traiter les personnes présentant une hypercholestérolémie de manière systématique mènerait à traiter des millions, sinon des milliards de personnes.

4.1.3 Limites des approches synthétiques en statistique : revues et méta-analyses

Nous avons vu dans le chapitre précédent comment l'on produisait de manière générale de la preuve en sciences, notamment biomédicales, avec le plus haut degré de certitude ou de confiance. La méthode expérimentale reste reine, et aura vraisemblablement toujours sa place d'une façon ou d'une autre ; l'implémentation pratique de cette méthode prend la forme de l'essai randomisé contrôlé (RCT). Un argument supplémentaire, venant renforcer la confiance envers la nature causale plus particulièrement des résultats pro-

duits, tient dans la reproductibilité de l'expérience, soit à l'identique, soit avec de petites variations (par exemple, une population un peu différente, afin d'étendre la validité des premiers résultats à une base plus large d'applications).

Nous l'avons également souligné, un RCT est une expérience très fréquemment si ce n'est systématiquement par nature, éloignée de la réalité, en situations ou en diversité – la diversité étant soit écartée, soit formalisée comme un écart tolérable à la moyenne, à la norme (la diversité devient distorsion). Un RCT, une expérience contrôlée est intrinsèquement une étude analytique : elle cherche à isoler un agent causal unique et un effet unique (un antibiotique et la disparition d'une bactérie par exemple).

Force est de constater que les techniques de synthèse, c'est-à-dire, à partir des preuves analytiques, issues d'un raisonnement réductionniste, sont limitées en quantité et en variété. Actuellement, les techniques de synthèse reconnues comme possédant un haut degré de preuve sont les méta-analyses (ou analyses poolées : on agrège les résultats de différentes études portant sur le même sujet), ou leur équivalent côté qualitatif, les revues systématiques. Les deux méthodes possèdent leurs propres guidelines : CONSORT pour les RCT, et PRISMA pour les revues systématiques et les méta analyses. [176]

De la sorte, les preuves établies en sciences biomédicales sont essentiellement des preuves ponctuelles, statiques, peu réalistes ou difficilement transposables (on ne connaît pas la transformation complexe ou élémentaire pour passer d'une situation à une autre), et limitées dans les mesures effectuelles.

Ces méthodes, plus spécifiquement les méta-analyses souffrent par ailleurs de défauts difficilement contournables, et qui entament sévèrement la portée et la crédibilité : le biais le plus important, systématique et non compensable est le biais de publication. En effet, ne sont publiés, ou très majoritairement, que les études ayant mis en évidence des résultats "positifs", c'est-à-dire, montrant un effet "significatif" en termes statistiques. Les

résultats ne mettant aucun effet en évidence ne sont pas ou difficilement publiables. Ainsi, lors des méta-analyses, les effets positifs sont mécaniquement favorisés, aussi petits en intensité soient ils, mais également aussi peu nombreux soient-ils relativement aux potentiels résultats négatifs.

Par les exemples suivants, nous examinons différents domaines pour lesquels nous cherchons des approches synthétiques différentes de celles monopolisant la scène scientifique actuelle.

4.2 La géométrie des infections comme démarche prototype

Les maladies infectieuses ont été et demeurent un modèle en pathologie, tant par leur mécanisme élémentaire et archétypique, que pour les interventions qui en découlent pour les enrayer, que ce soit au niveau de l'individu ou de la population. Très longtemps, la pathologie infectieuse a représenté, aux côtés de la traumatologie et différemment, de la folie, la maladie typique voire unique. Invariant historique – elle est présente jusque les premiers récits humains – elle a jusque ici maintenu son rang de pathologie par excellence, de modèle à étudier et à traduire vers d'autres affections, a confirmé et affiné ses relations causales grâce à la technologie et à la découverte majeure des anti infectieux, dont les antibiotiques.

Cependant, comme nous le verrons, ce modèle, sans être faux – nous parlerons d'approximation locale – vit depuis une trentaine d'année une révolution sans précédent, qui devrait lui ouvrir de nouvelles dimensions, qui demandent encore à être intégrées. Ce renouveau passe certainement par la découverte du VIH, de son lien au SIDA, des moyens tant diagnostiques que thérapeutiques qui lui sont associés, mais également de ses dimensions sociales et épidémiques.

Nous verrons ensuite comment, en prenant pour point de départ les inter-

actions entre agents infectieux et hôtes, le concept de pathologie infectieuse peut être revisité ou étendu : nous en donnerons les nouvelles dimensions et limites, les nouveaux points de vue.

À partir de ce nouvel espace, nous proposerons un début de réponse possible à la question de Foucault, à savoir, de déterminer la géométrie des processus infectieux, tout en respectant une approche centrale au modèle infectieux, à savoir le lien de causalité. Enfin, nous verrons qu'une fois de plus, cette tentative de reformulation de la pathologie infectieuse est certainement transférable, ou à tout le moins similaire à la situation d'autres pathologies d'importance, en l'occurrence le cancer.

4.2.1 Une brève histoire des maladies infectieuses comme pathologie archétypique

Toute histoire de la médecine consacre et consacrera une partie conséquente à la pathologie infectieuse, à ses représentations, ses différentes formes, les croyances les accompagnant, ses hécatombes, ses répercussions ou encore les révolutions pratiques et conceptuelles qu'elle a pu traverser. Elle demeure une préoccupation universelle majeure. [177]

La maladie infectieuse est rapportée sous trois angles, selon trois échelles différentes, selon que l'on s'occupe de l'agent causal, de l'individu infecté, ou bien de l'épidémie. Avant que les pays industrialisés connaissent la double transition démographique et épidémiologique, la pathologie infectieuse marquait tant le quotidien que l'histoire : la traumatologie mise à part, la pathologie infectieuse était la grande pourvoyeuse de décès prématurés – et elle le demeure dans un certain nombre de pays, où sévissent par exemple le paludisme, la tuberculose et infections virales à tropisme hépatique ou gastro-intestinal. La variole tuait abondamment, la pneumopathie infectieuse abattait une personne atteinte sur trois. La pathologie infectieuse a également marqué par ses grandes vagues épidémiques, les vagues de peste étant

parmi les plus notoires, ainsi que les épidémies – les pandémies – de grippe. La maladie infectieuse est donc un fléau individuel et collectif, au titre de la maladie transmissible, que l'on en connaisse clairement les modalités ou vecteurs de transmission, divins ou très rationnels.

En pathologie infectieuse, le paradigme central est celui de la relation de causalité directe entre un agent extérieur, vivant ou peu s'en faut (certains considèrent le virus comme faisant partie du règne du vivant, d'autres non), et l'état pathologique, dont les symptômes peuvent varier d'une pathologie à l'autre, d'une circonstance à l'autre, d'un individu à l'autre, en qualité (présence ou non) et en quantité (intensité des symptômes). Il existe néanmoins, de manière assez constante, la présence de ce que l'on appelle un syndrome inflammatoire, doublé de signes laissant présager que cet état inflammatoire est dû à la présence d'un agent vivant extérieur, en réaction à sa présence, sa prolifération, et éventuellement son action / ses interactions avec l'organisme hôte. S'il existe des signes dits cliniques : une fièvre, une accélération du rythme cardiaque accompagnant cette fièvre, la présence d'une augmentation de volume des ganglions dans une région de drainage donnée et compatible avec le site infecté, une chaleur, une rougeur dans cette même région concernée, il existe également des signes biologiques, témoins d'une activité réactionnelle du système de l'immunité, que cette activité soit innée ou spécifique – que la réaction soit une réaction de première barrière, ou une réaction activée par la mémoire d'une intrusion précédente semblable.

La responsabilité d'un agent extérieur dans le déclenchement d'une séquence de symptômes observables ou subjectifs (comme la douleur) doit être éprouvée, et il a été proposé, par exemple, des critères de causalité, comme les critères de Bradford-Hill (voir tableau 4.1). Ils viennent suppléer le fait que la recherche statistique établit essentiellement des corrélations, et qu'une corrélation entre deux observations (linéaire par ailleurs), n'est pas synonyme de causalité.

TABLE 4.1 – Critères de Bradford-Hill pour la causalité.

Critères	Exemples
Force de l'association	Estimation des risques relatifs ou des odd-ratios
Cohérence	Reproduction des observations
Spécificité	À une cause, on observe <i>a priori</i> un type d'effet bien particulier
Relation temporelle	La cause précède la conséquence
Plausibilité	Mécanismes biologiques, physiques
Preuve expérimentale	Études randomisées
Analogie	Existe-t-il des explications alternatives?

Avec l'invention du microscope, et la découverte du monde microbiologique, la quête incessante de l'école pasteurienne à identifier, pour chaque pathologie infectieuse, l'agent responsable, de visu, de la réaction immunitaire, des couples ont été établis systématiquement, entre agent A et pathologie P. La pathologie infectieuse est devenu une affaire d'investigation systématique, liant la force de la preuve expérimentale à celle de l'observation *in vivo* via ses effets. Elle est également devenue le modèle élémentaire équivalent aux modèles des sciences plus dures, comme la physique, reliant directement cause et effet, mécaniquement.

Ce modèle est puissant par sa simplicité et son efficacité, en ce sens où, ayant identifié le responsable *via* une relation de causalité directe, il devient possible de lutter aussi directement contre les effets, nuisibles et indésirables – la maladie, et souvent, la mort – en supprimant la cause. Après la chasse à l'identification des agents, fut ouverte et demeure ouverte la chasse à l'agent purement et simplement : soit en l'évitant, soit en empêchant sa trop grande prolifération, soit en l'éliminant lorsque la contamination s'est produite.

Dans le premier cas, la stratégie d'évitement, il s'agit par exemple d'éviter de boire de l'eau souillée et éviter le péril fécal, afin de ne pas être contaminé par le virus de l'hépatite A, ou d'éviter les zones à eaux stagnantes et basses altitudes, où les moustiques (les anophèles) vivent, et peuvent transmettre le paludisme (transmission du *plasmodium*, *falciparum* ou *ovale*, par exemple). Il est également possible d'éviter la prolifération, voire les conditions d'exis-

tence des agents infectieux ou de leur vecteur. Dans les deux cas précédents, il s'agit de mesures d'hygiène et d'assainissement, ou de campagne de largage d'insecticide visant les moustiques. Enfin, le dernier cas de figure est davantage curatif que préventif, et monopolise tous les mécanismes possibles visant au ralentissement de la prolifération intra hôte de l'agent ou à sa destruction directe. C'est là l'usage des différents anti-infectieux : antiparasitaire dans le cas du paludisme, antibiotiques dans le cas des infections bactériennes, etc.

Ainsi, la transition épidémiologique a grandement été soutenue par les différentes mesures préventives : les mesures d'hygiène supprimant une grande partie des réservoirs ou vecteurs de contamination, et l'apparition de différentes classes d'anti-infectieux ont parachevé de diminuer et mieux contrôler les dynamiques infectieuses : moins de contamination, plus de moyens d'en contrôler et éliminer les effets délétères, et concomitamment, de meilleurs états nutritifs et immunitaires chez les sujets sains susceptibles d'être infectés. La meilleure compréhension du système immunitaire, sa composante innée et sa composante spécifique, a également abouti à la mise au point et à la généralisation des vaccins contre une grande partie de pathologies fréquentes, notamment infantiles. Notons à ce propos, que pour toutes ces avancées, les deux échelles sont chaque fois concernées et impactées : l'échelle sociale et l'échelle individuelle. Ce sont déjà de grandes politiques de santé qui s'établissent ; par exemple, établissement du calendrier vaccinal.

4.2.2 La place particulière du VIH-SIDA

Avant tout autre propos, nous tenons à souligner ici qu'il n'est pas notre but d'infirmer ou de nier un lien entre VIH et SIDA ; le but est au contraire, d'examiner même brièvement, les différentes dimensions de cette équation fondamentale VIH-SIDA, qui ne sauraient se limiter à la simple virologie, et d'en montrer les implications sur la validité générale de cette équation comme fondation du paradigme de la pathologie infectieuse. Toute décou-

verte scientifique a une histoire, est insérée, comme nous l'avons déjà discuté, dans des contextes sociaux, et reflètent des observations obtenues de points de vue et de capacités de mesures pouvant évoluer. L'histoire du VIH-SIDA est une histoire récente, et une histoire tumultueuse, pour laquelle, certainement, il n'existe pas encore de solution réellement acceptable dans toutes ses dimensions, que ce soit en raison de freins conceptuels que de l'inertie sociale. Cependant, le fait est là : tout le monde, au cours des 30 dernières années, a été surexposé à cette entité, sous une forme ou une autre, bien plus que pour d'autres pathologies. À notre sens, le modèle VIH-SIDA est une excellente opportunité pour reconsidérer les paradigmes actuels, les faire évoluer, et l'examiner à la lumière de la forme, et non de l'observation ponctuelle.

Jusque là, le modèle de la pathologie infectieuse est demeuré simple, et présentait un exemple parfait de relation causale permettant l'action directement suivie d'effets. Néanmoins, la composante épidémique posait conceptuellement déjà plus de problèmes, car les modélisations peinent à expliquer, et surtout prédire, l'apparition et la propagation d'une épidémie, son maintien, et sa disparition. On s'en contentait plus ou moins, en partant du principe que soit le meilleur moyen était de toute façon de vaincre l'infection au niveau individuel, soit que la prévention au niveau social, par la stratégie vaccinale, était suffisante. En bref, on visait l'éradication simple des causes, ou son maintien au plus bas, dans le doute et la méconnaissance fine des mécanismes sous-tendant les phénomènes de propagation. Il semble qu'il y ait eu un flottement entre une sensation de terrain conquis et compris, ou en passe de l'être, et le désir de découvrir autre chose.

C'est dans ce contexte qu'est apparue une nouvelle expression de la pathologie infectieuse, par la découverte du VIH, nom finalement retenu pour le rétrovirus humain. Le VIH est le virus de l'immunodéficience humaine. L'histoire – une histoire – du SIDA, Syndrome d'Immuno-Déficience Acquise, commence non pas par l'observation d'une nouvelle pathologie en

soi, d'un diagnostic nouveau, mais par un ensemble d'occurrences jugées comme successivement trop nombreuses et rapprochées dans le temps, et visant une catégorie sociale particulière, sinon deux lorsque l'association n'était pas systématiquement faite entre elles deux : les homosexuels masculins, et les usagers de drogues. Les premiers signes d'alerte furent donnés devant un nombre jugé anormal de sarcomes de Kaposi, se manifestant avant tout par une atteinte cutanée, voire cutanéomuqueuse, à type initialement de macule, évoluant vers une plaque ulcérovégétante, et se compliquant d'une atteinte viscérale, notamment pulmonaire – ce qui en signe la gravité. Plus formellement, la reconnaissance du SIDA, non du VIH, comme forme d'un problème nouveau et nécessitant diagnostic séparé, peut être datée à partir des communications du CDC d'Atlanta (Centers for Disease Control) aux Etats-Unis, en 1981, à la suite des observations de médecins de New York et San Francisco. [178, 179, 180, 181, 182] Il avait été alors proposé de baptiser cette entité GRID, pour *Gay Related Immune Deficiency*. Cette étape illustre bien l'histoire naturelle que peut suivre la découverte d'une nouvelle forme pathologique, bâtie par ailleurs sur des signes connus, et dont seules l'association, la récurrence spatio-temporelle et ici, liée par biais d'observation ou non, à une ou deux caractéristiques comportementales, hautement discriminantes socialement, apparaissent nouvelles. Il faut bien cerner que si le sarcome de Kaposi, décrit par ailleurs pour la première fois un siècle plus tôt, est un signe intrigant car rare jusque là, les symptômes entrant dans le diagnostic à cette époque sont des signes très aspécifiques : asthénie (fatigue), perte de poids, et possible lymphadénopathie (ganglions). En outre, la toxicomanie et l'homosexualité étaient encore, peu s'en faut, également des diagnostics de pathologie.

L'étape de diagnostic étiologique s'est alors focalisée premièrement sur l'utilisation de drogues, notamment de dérivés nitrés (les *poppers*). Il est

TABLE 4.2 – Critères cliniques (stades) du SIDA, selon l’OMS.

Stades	Symptômes et pathologies
Stade 1	Patient asymptomatique Adénopathies persistances généralisées et fièvre
Stade 2	Perte de poids < 10% du poids corporel Manifestations cutanéomuqueuses mineures Zona dans les 5 dernières années Infections récidivantes des voies respiratoires supérieures
Stade 3	Perte de poids > 10 % du poids corporel Diarrhée chronique inexpliquée pendant >1 mois Fièvre prolongée inexpliquée pendant > 1 mois Candidose buccale Leucoplasie chevelue buccale Tuberculose pulmonaire dans l’année précédente Infections bactériennes sévères
Stade 4	Pneumocystose Toxoplasmose cérébrale Maladie de Kaposi Lymphome Mycobactériose atypique généralisée

convenu et connu qu'une grande partie des drogues, suivant notamment un usage fréquent, a un impact sur l'état nutritif, et sur le fonctionnement du système immunitaire. L'hypothèse infectieuse a été envisagée secondairement, et différents noms ont été proposés pour l'agent responsable, de la maladie, néanmoins appelée syndrome : LAV (*Lymphadenopathy Associated Virus*), ou HTLV-III, par exemple. La question centrale, ici, était et est d'établir un lien de causalité directe, sur le modèle hérité de la tradition pasteurienne : un agent infectieux, une expression pathologique associée et spécifique. L'établissement de ce lien nécessite comme préalable l'isolement de l'agent, son identification habituellement directe, son identification *in vivo* chez les sujets supposés atteints de la pathologie en cause, et de tester l'effet de la suppression de cet agent sur l'expression de la pathologie.

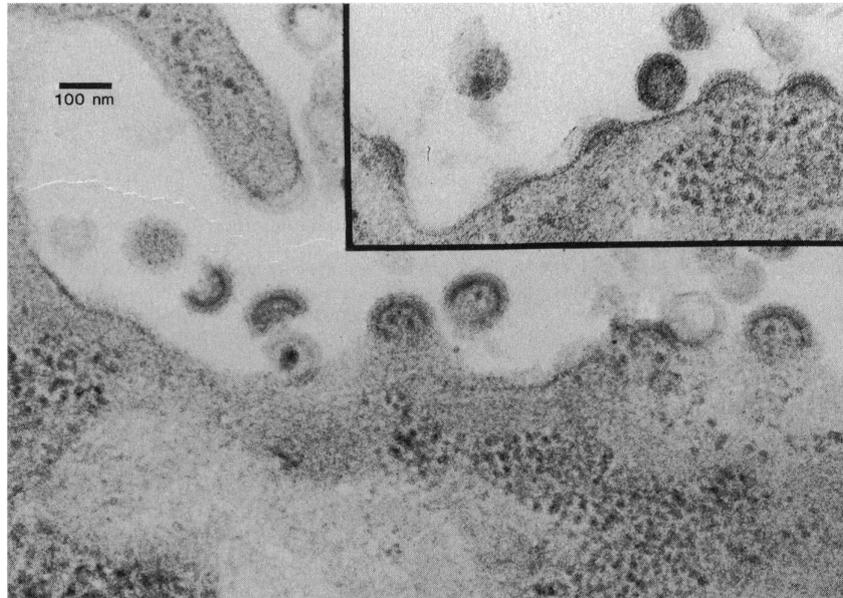
La démarche peut paraître simple et ne nécessitait que du temps, éventuellement des moyens supplémentaires pour parvenir à ses fins. Toutefois, un regard plus attentif montrera toute la difficulté et le périlleux de l'entreprise. Premièrement, même si le SIDA, la pathologie en cause, est bel et bien considéré comme une maladie, il a été nommé premièrement syndrome, ce qui par ailleurs demeure certainement encore pertinent, et plus rigoureux que de parler de pathologie. Ce basculement, au sein de la pathologie infectieuse, semble infime, mais annonce ou devrait annoncer une évolution, si ce n'est une révolution dans la manière de considérer les fondements de la discipline, c'est-à-dire, la fameuse relation agent-pathologie.

Le SIDA est une pathologie à expression variable, même si certains des signes – en fait : certaines des pathologies infectieuses, ou cancers – sont d'apparition temporelle corrélée à l'évolution du syndrome dans le temps – ce qui est discutable, comme nous en reparlerons à propos du rôle du temps comme critère de causalité : plus que le temps lui-même, c'est certainement la conjonction d'un terrain donné et d'événements occurrents à un certain rythme qui est le reflet le plus direct de l'expression différentielle des patholo-

gies associées au SIDA. Aussi, la notion même de pathologie spécifique d'un agent unique, en relation causale directe est mise à mal. Ceci est d'autant plus vrai dans des pays comme l'Afrique, où il existe de base, une prévalence de pathologies pouvant présenter les mêmes signes, voire font partie des pathologies entrant dans le diagnostic du SIDA. La tuberculose, par exemple, est un exemple particulièrement illustratif d'observation interférente : la pathologie est chronique, d'évolution possiblement latente, provoque asthénie et perte de poids, et influe évidemment sur l'état du système immunitaire. Le problème du SIDA comme pathologie infectieuse est donc double, et concerne les deux points clés du paradigme : l'identification de l'agent unique directement causal, et le diagnostic d'une entité observable spécifique. De fait, par définition ou expression différentielle, ou encore par biais d'observation, le SIDA des pays occidentaux n'a que peu à voir avec le SIDA des pays africains, par exemple. L'estimation des prévalences en Afrique est un exercice difficile, si ce n'est impossible – impossible au sens de fournir des chiffres de précision pertinente pour être utilisé dans des modèles de prédiction – comme l'histoire des prédictions démographiques de plusieurs pays africains a pu le montrer (certains pays auraient tout simplement dû disparaître il y a déjà quelques années, selon les modèles prédictifs). Les prévalences ont ainsi été revues à la baisse quasi systématiquement, et de façon importante, au cours des années.

De l'autre côté du problème, se situe la question de l'identification de l'agent causal, puis de son rôle central et nécessaire dans le déclenchement de la pathologie. La quête de cet agent a été violente, et a notamment mis en concurrence les équipes françaises et américaines, menant jusque plusieurs procès et enquêtes pour fraude scientifique, condamnation de Robert Gallo aux Etats-Unis, puis réhabilitation, tractations institutionnelles entre l'Institut Pasteur et le NIH aux Etats Unis, essentiellement autour de questions de brevets concernant le test diagnostic – du VIH, pas du SIDA – et même

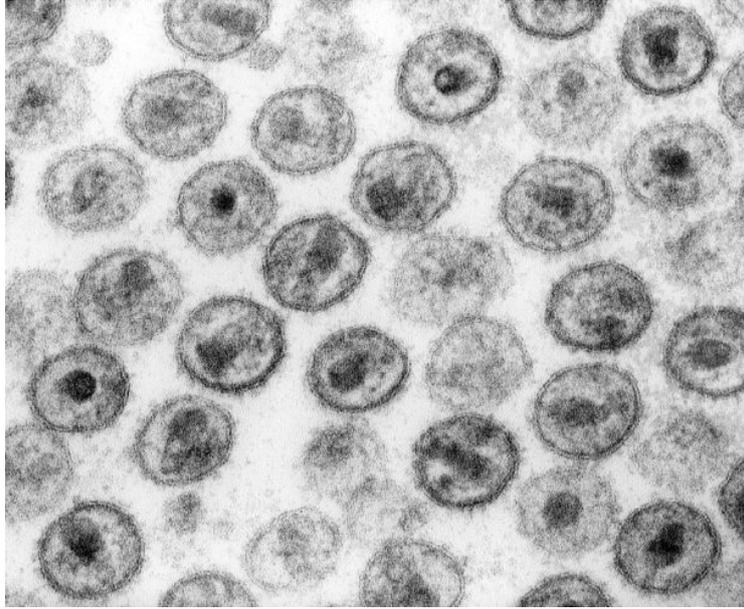
FIGURE 4.1 – Rétrovirus VIH au microscope électronique, comme communiqué en 1983, dans *Science*



intervention et arbitrage politique entre les deux états.

Le diagnostic d'infection par le VIH – disons de sa présence dans l'hôte, et de son insertion dans son fonctionnement – est un diagnostic indirect, et rejoint en cela tous les nouveaux diagnostics basés notamment sur les biomarqueurs, dont la médecine interne et la rhumatologie sont particulièrement friandes ; c'est la caractérisation biologique, immunohistochimique des maladies dites systémiques, comme le lupus, la polyarthrite rhumatoïde, le syndrome sec, etc. Le VIH n'est pas isolé, on en détecte indirectement la présence et son "activité". C'est le cas des tests dits de "dépistages" – présentant par ailleurs un très grand nombre de faux positifs possibles, dus par exemple à la grossesse chez la femme multipare, ou la présence de pathologies infectieuses, type tuberculose, paludisme, de pathologies systémiques, type lupus ou polyarthrite rhumatoïde, ou certains vaccins – et dont l'usage et la détermination du caractère positif ou non est arrêté par un texte de loi, du moins en France. En outre, la question du test de dépistage – et non diagnostic,

FIGURE 4.2 – Rétrovirus VIH au microscope électronique. CDC PHIL



ce qui est là une différence fondamentale – est épineuse, tant la performance d'un test effectué "en aveugle" chute drastiquement. En effet, les performances d'un test dépendent de la prévalence – de la présence de la pathologie – dans la population testée. Plus la prévalence est faible, ce qui est le cas pour le VIH *a priori*, plus le test doit être extrêmement discriminant, séparant les faux des vrais positifs et négatifs. Il doit pouvoir être extrêmement sensible, ou extrêmement spécifique. Quand on parle d'extrêmement, nous entendons des performances de l'ordre de 99.999%, et non 99.0%. Evidemment, un même test effectué cette fois dans une population plus restreinte, plus considérée à risque, donc considérée comme présentant une prévalence plus importante, donnera de meilleurs résultats. Nous sommes là encore face à un problème de reconnaissance de forme, et de résolution : il est nécessaire de posséder la meilleure définition possible de la forme à identifier, son espace de représentation le plus spécifique, et de l'outil le mieux adapté pour l'isoler d'autres formes qui potentiellement, partagent un recouvrement partiel avec

la forme d'intérêt.

Il est donc possible de détecter les anticorps, plus ou moins spécifiques de la présence du VIH, mais pas uniquement. Il est également possible d'identifier la présence, par exemple, d'un antigène, l'antigène p24, qui est une protéine associée au virus. [183]

Au-delà, il est également possible de rechercher la présence de l'ARN viral dans le sang, qui est le reflet d'une insertion du virus dans le fonctionnement de l'organisme hôte, puisque il y serait répliqué activement. La technique utilisée pour la détection est la PCR (*Polymerase Chain Reaction*), une technique d'amplification de séquences nucléotidiques présents dans un échantillon, même en très faible quantité – cette technique vise donc à augmenter le rapport signal sur bruit de telle façon que l'on puisse conclure à la présence d'une forme ou non.

Encore actuellement, la question du rôle précis du rétrovirus dans la morphogenèse du SIDA demeure au moins partiellement débattue ou irrésolue. L'hypothèse de base, aux origines de la découverte, était celle d'un agent infectieux prenant pour cible les lymphocytes T CD4, et les détruisant : c'est un des aspects que l'on constate sur une simple prise de sang, une décroissance de ce type de population de cellules immunitaires.[182, 184] Néanmoins, la latence de la maladie – c'est-à-dire, de l'infection par le VIH et l'observation de ces effets sur le système immunitaire – est surprenante : si le VIH détruit certains types de cellules immunitaires, il ne semble pas le faire de manière soit systématique ou déterministe, soit avec une lenteur qu'il n'est pas évident d'expliquer, du moins directement.

Des études récentes tendent à prouver de manière plus solide, le rôle destructif du VIH concernant les lymphocytes. Il reste néanmoins du chemin à parcourir pour élucider la chaîne causale complète. En outre, cette chaîne causale est d'autant plus difficile à établir avec force que les échelles

de temps viennent brouiller les rôles des différents facteurs pouvant intervenir : le passage de VIH à SIDA peut prendre des années, voire même ne jamais se réaliser – comme pour beaucoup, sinon toutes ? de pathologies infectieuses et notamment virales, il semble toujours exister des porteurs dits "sains" : des personnes portant en eux l'agent, sans pour autant déclarer la pathologie correspondante. Cette autre constatation est là aussi, en soi, source d'interrogations sur la "pureté" de l'équation fondamentale de la pathologie infectieuse, et suggère une plus grande dimensionnalité du problème que sa réduction à deux dimensions (hôte / agent), ou à trois dimensions (hôte / agent / populations).

Un autre point intéressant pour l'explicitation historique et sociologique de la présentation du VIH comme rétrovirus causal du SIDA, tient dans le contexte, la chronologie dans lesquels cette présentation a émergé. Il faut se replacer dans une dynamique de la découverte ou de la construction successive et rapprochée de l'ADN, vu comme base de l'information fondamentale des organismes vivants, et de la théorie de la communication (en général, dite de l'information) de Shannon-Weaver. Les cellules deviennent réellement de potentiels automates au sens de Turing ou de Von Neumann, avec leur ruban d'information décomposable en bases (un alphabet très restreint, de 4 symboles), et leurs unités de traitement, de traduction et transcription de l'information. La théorie du rétrovirus, s'insère parfaitement dans les modèles théoriques qui ont eu le temps de diffuser, dans leurs principes et leurs termes, et offraient un cadre séduisant, mais sans exemple concret pour le valoriser. La notion de virus et de rétrovirus a été une belle opportunité de mettre à l'épreuve une théorie de la communication, essentiellement intracellulaire, avec l'insertion du code viral dans le matériel informationnel de l'organisme hôte. Cette insertion, par ailleurs, pose de nouveau la question de savoir si les formes initiales : hôte et agent, deviennent réellement

dissociables, en pratique, ou si le résultat n'est pas simplement l'obtention d'une troisième forme, résultante supérieure à la simple addition des deux premières, puisqu'elle induit un ensemble de comportements émergents.

Quelle que soit l'orientation prise pour l'étude du VIH-SIDA, tout porte à croire que nous sommes face à une situation atypique en pathologie infectieuse. Cette atypie peut notamment résulter d'un cadre théorique relativement nouveau – celui de la génétique vue comme une théorie de l'information – et d'un agent causal dont les effets présentent des particularités : mécanisme d'action, latence de passage au pathologique... et de manière tout-à-fait flagrante, un agent premier menant à une pathologie elle-même purement composée et caractérisée par d'autres pathologies, infectieuses ou cancéreuses, d'étiologie autre (soit due usuellement à un autre agent pathogène bien identifié, soit multifactoriel dans le cas du cancer, même si des origines infectieuses demeurent suspectées et dans certains cas, établies). De fait, malgré le glissement vers un cadre explicatif plus sophistiqué, et au demeurant très théorique (on dispose plus d'éléments théoriques que d'observations directes, expérimentales), il apparaît que le maintien de l'équation classique a prévalu : en effet, en nommant syndrome ce que l'on désire voir comme une pathologie, le SIDA, qui est décrit comme un ensemble de pathologies et de signes cliniques et biologiques, on a maintenu une entité permettant d'entretenir une relation un agent, un effet pathologique, VIH-SIDA. Ce maintien contribue à rendre la situation relativement inconfortable : il existe par exemple un amalgame entre VIH et SIDA dans l'esprit de beaucoup, les contributions théoriques d'un cadre riche en sont gommées au premier abord, et il ne rend pas compte de la complexité de l'identification de la forme – pas de diagnostic direct et ponctuel.

Au final, l'exemple du VIH-SIDA fournit un prototype riche et fortement multidimensionnel dans sa présentation, avec notamment des dimensions so-

ciales, politiques et des éléments éclairants quant à la production de faits scientifiques, le tout étalé dans le temps. Il faut par exemple, pour se rendre compte à quel point cette histoire a pu être tumultueuse, se remémorer les différentes passes d'arme par revues prestigieuses interposées, du type *Nature*, *Science*, *Lancet* ou *New England Journal of Medicine*. Pour exemple, on peut citer la déclaration de Durban. [185, 186] Il demeure difficile de pouvoir discuter du VIH-SIDA en dehors de toute passion partisane.

D'un point de vue causal, outre la focalisation sur la chaîne directe entre le virus et la pathologie/syndrome, de nombreuses hypothèses ont pu être avancées, et ce dès le début de la quête de la recherche d'un agent causal, certains réfutant jusque la possibilité d'aucun lien entre virus et SIDA. Parmi les hypothèses de la première heure, il y eut la possibilité des drogues ou du comportement sexuel particulier, notamment homosexuel masculin. Concernant la létalité du SIDA, il a également été avancé avec force que les premiers traitements étaient les principaux responsables de la rapidité avec laquelle les patients mouraient. C'était le cas de l'azidothymidine ou communément appelée AZT, dont l'histoire, en particulier aux États-Unis, peut apparaître relativement troublante, notamment quant au comportement et de l'industrie pharmaceutique, et de l'agence de régulation, la FDA (*Food and Drugs Administration*). L'histoire de l'AZT a été assez largement décrite et commentée, notamment concernant son origine. L'AZT avait en effet été désapprouvée initialement comme traitement contre le cancer, d'une part pour des effets peu probants, d'autre part pour d'importants effets secondaires indésirables – dont une importante mortalité des animaux tests. Il a ainsi connu une seconde vie comme traitement anti-VIH. Son efficacité en tant que monothérapie, et à hautes doses a été réfutée ultérieurement.

D'autres hypothèses furent formulées, et soit maintenues ou revenant périodiquement. Parmi ces hypothèses, le rôle prépondérant de l'état nutritionnel, mais également des anti-oxydants. Ces deux arguments ont même

été avancés et promus par le "découvreur" du virus lui-même, L. Montagnier. [187] Ses dernières interventions sur le sujet plaident pour un élargissement du panorama explicatif de la relation VIH-SIDA, et le chercheur estime que si le virus a une influence et un rôle indéniable dans le développement d'un syndrome d'immunodéficience, il serait difficile de l'estimer seul responsable. Selon lui, le maintien d'un état nutritionnel correct – et par là même, d'une façon ou d'une autre, une compétence immunitaire de base suffisante – resterait le meilleur traitement contre la progression ou l'expression du VIH. Cependant, il est frappant de constater à quel point ces retours en arrière, ou ces tentatives d'explorer plus loin que le lien basique VIH-SIDA, par son découvreur célébré, n'a que peu de retentissement sur la communauté scientifique. Celle-ci s'est emparé de son objet, sans pour autant, comme Latour pourrait le dire, l'avoir purifié : en effet, la démarche semble plutôt avoir été opposée ; il a été supposé *a priori* une relation purifiée de cause à effet, un virus-une maladie (de la même façon que beaucoup rêvent encore à un gène-une maladie), puis, par complexité intrinsèque de cet objet et par absence de remise en question du paradigme initial, l'histoire et la représentation du VIH-SIDA comme objet scientifique sont devenues floues, avec une accumulation d'études aux intérêts incertains, la seule certitude demeurant et fédérant une majorité bruyante étant celle qu'il faille éradiquer cet objet. Or, par définition, par son mécanisme même, l'idée d'une éradication d'un rétrovirus paraît au minimum hasardeuse. Le futur nous renseignera sur l'évolution de ce paradigme, de même sur celle de l'épidémie. Quoi qu'il en soit, les CDC ont récemment déclassé l'épidémie de VIH de son statut de menace prioritaire pesant sur la santé publique des américains (CDC list of communicable diseases of public health significance, depuis le 4 janvier 2010).

Nous réaffirmons, à ce stade, que nous ne réfutons pas l'existence du SIDA, ou du VIH en tant que tel – et, finalement, notre question n'est pas tant celle-là, que de se permettre de mettre en perspective cet objet très particulier, presque symbolique. Nous suggérons au lecteur de se reporter, armé de son sens critique et de précautions importantes, à certains ouvrages, polémiques ou orthodoxes, [188, 189, 190] historiques pour la plupart, [191, 192] et en particulier à la déclaration de Durban – et à sa controverse. [185, 186, 193]

Enfin, une collection récente d'articles, notamment publiée par le *Lancet* et le *New England Journal of Medicine*, a souligné de nouveau le statut paradoxal du VIH-SIDA, dans son rôle dual atypique/typique : il est un exemple censé suivre le paradigme infectieux, et dans le même temps, il est le précurseur de bien des concepts. Cette collection a insisté sur l'impulsion qu'auraient procuré la recherche et la réponse des systèmes de soins engagés dans la lutte contre la pandémie, inaugurant une "santé mondialisée" (*globalisation of health*). [194, 195] Ainsi, plutôt que d'élargir le cadre conceptuel permettant de rendre compte des relations complexes, l'histoire montre davantage l'extension des effets imputés au VIH sur différents domaines. Il nous semble ici que ce que nous rappelle Latour s'applique bien, ou doit nous interroger : le VIH-SIDA est-il un objet en soi, ou bien un objet de laboratoire, un objet au moins partiellement socialement fabriqué ? [18]

4.2.3 Panorama actuel et concepts en maladies infectieuses

L'équation fondamentale de la pathologie infectieuse, agent causal-pathologie spécifique tient donc difficilement si l'on étend un peu plus largement le champ de vision ; l'exemple historique et récent du VIH-SIDA nous en apporte des éléments illustratifs conséquents. La question qui se pose est alors : quelles dimensions pour l'espace de (re)définition de la pathologie infectieuse, et quelles relations opèrent-elles dans cet espace ? Ce nouvel espace, et ces

nouvelles relations doivent inclure, pour être satisfaisantes, la formulation de l'équation fondamentale de la pathologie infectieuse, et l'étendre dans les cas où l'approximation locale ne suffit plus à comprendre, prédire et agir sur la réalité.

La définition de l'espace pertinent passe certainement par un certain nombre d'observations factuelles, et de nouvelles hypothèses, des changements de points de vue et d'échelles, qui permettraient de concevoir l'agent infectieux autrement qu'uniquement selon la dimension de l'ennemi monomorphe, ponctuel, à éliminer. Nous proposons ici de montrer que : l'agent infectieux ne peut pas être considéré comme un point dans l'espace, éventuellement répété n fois à l'identique, mais doit être vu comme un élément particulier d'une population non nécessairement monoclonale ; nous montrons aussi qu'elle est capable d'aménager son propre milieu de vie local. Au-delà de ces premières caractéristiques, nous présentons la population bactérienne comme un possible organe à part entière, et donc faisant partie intégrante de notre physiologie. Nous rappelons aussi que les agents infectieux, de natures différentes, sont néanmoins capables de communiquer, d'être en concurrence ou coopérer entre eux. Ils peuvent également participer de notre constitution en tant qu'espèce, et ce sur différentes échelles de temps. Enfin, nous terminons sur la place du système immunitaire.

Nous prenons l'exemple des bactéries. Une des erreurs dont nous nous relevons actuellement, a été de penser une infection bactérienne comme le fait d'une invasion de millions de clones du même individu, là où l'on parle bel et bien d'êtres vivants, régis en partie par un code génétique propre, mais sujet au même type de variabilité que le nôtre. Nous sommes donc bien confrontés à des populations entières, au sens démographique du terme, et ce en termes de millions ou de milliards d'individus, lesquels peuvent interagir entre eux et avec notre organisme. Une bactérie n'est rien moins qu'une

forme de cellule particulière et globalement autonome ; toutefois, là où nous sommes constitués d'environ 10^{14} cellules à l'état adulte, nous abritons en moyenne également, à l'état "normal", environ 10^{15} bactéries, réparties entre 500 et 100 000 espèces. Aussi, nous sommes en nous-même minoritaires, sur le plan des effectifs.

En outre, l'étude des infections ostéo-articulaires, leur difficile caractérisation, leur mauvaise compréhension, a eu le mérite de mettre l'accent sur certaines réalités des colonies bactériennes, dont deux plus particulièrement. La première est que pour une colonie donnée, provenant d'un inoculum, voire d'une souche bien spécifique, il n'en demeure pas moins que la population développée *in vivo* n'est pas monovariée, mais présente bien des variations de germes ; on appelle ces autres colonies, souvent plus quiescentes ou d'expansion moins fulgurante au moment où on les observe, des SCV – des *Small Colony Variants*. [196, 197] En outre, on sait aussi que sous la pression de sélection, en particulier face à un agent anti infectieux plus ou moins efficace, des mutants sont privilégiés, mieux armés contre les agressions similaires. La population bactérienne n'est donc pas homogène en termes d'individus dont elle est constituée, même pour un lieu et une espèce donnés. La seconde est que les bactéries ne sont forcément, voire peut-être jamais, des points disposés sur un tissu, plus ou moins isolément ou sans relation forte entre eux. Les bactéries participent à la formation de ce que l'on appelle désormais le biofilm, qui peut être vu comme la constitution d'un milieu de vie favorisant le développement des bactéries. [198, 199, 200, 201] Il a tout d'un écosystème complexe, celui-ci servant autant de réseau d'alimentation, de protection contre le milieu "extérieur" à la colonie (notamment, les antibiotiques n'y pénètrent pour ainsi dire pas, ou en quantité infime), ou encore de réseau de communication inter agents. Ce biofilm présente aussi des caractéristiques physico-chimiques qui le rendent difficilement éliminable.

La dimension populationnelle et hétérogène ne saurait donc être ignorée, au risque de favoriser le modèle marginal plutôt que la règle. Nous parlons néanmoins ici d'une cohabitation plutôt de conflit, entre hôte et agent infectieux. Il existe d'autres situations, en redéfinition elles aussi, qui ne sont pas si tranchées, que ce soit en faveur d'interactions favorables qu'en faveur d'un rôle pathogène marqué. L'exemple le plus connu, et actuellement de plus en plus étudié, est ce qu'il est désormais convenu d'appeler le microbiote intestinal – ce qu'on appelait, finalement, la flore intestinale. [202, 203, 204, 205, 206, 207] Si l'on suspectait très fortement le rôle des bactéries vivant dans le tube digestif comme participant de la digestion, du catabolisme ou du métabolisme d'un certain nombre de substances d'origine alimentaire, nous en sommes arrivés à l'envisager différemment encore. Il semble être mis l'accent sur l'intégration plus large de la fonction du microbiote au sein de son hôte, tendant finalement à lui conférer la place d'un organe à part, même si celui-ci n'est pas issu, *stricto sensu*, du programme morphogénétique basé sur le matériel génétique des parents. Cependant, aucun organisme n'est conçu pour vivre isolément. Les liens entre microbiote et obésité, ou encore des maladies auto-immunes, comme le diabète de type 1, paraissent gagner en crédibilité chaque jour. Plus étonnant encore, il apparaît qu'il puisse exister un axe "cerveau-tube digestif", notamment par la synthèse bactérienne de substances à tropisme cérébral; plusieurs études mettent en lumière les liens entre pathologies mentales – comme l'autisme, l'anxiété, la dépression – et les différences entre microbiotes intestinaux. Ce n'est par ailleurs pas le seul lien qui a été déjà pu être fait entre agent infectieux et comportement : il en a déjà été question pour le parasite de la toxoplasmose, et l'impulsivité, le comportement suicidaire. Enfin, nous connaissons le rapport entre certains agents infectieux, et le développement de pathologies prolifératives, donc d'un dérèglement de la physiologie cellulaire, et peut-être plus systémique : c'est par exemple le cas pour la bactérie

Helicobacter pylori et le cancer gastrique ou le lymphome du MALT, ou encore le HTLV-1 et certaines leucémies ou lymphomes. [208]

Les agents infectieux participeraient donc largement de la physiologie d'un organisme, que ce soit de manière souhaitable ou non. Mentionnant également la place particulière des agent viraux, qui entretiennent, l'on commence à le comprendre, des relations particulières avec les bactéries, de la même façon qu'ils interagissent avec notre matériel génétique cellulaire : notre code génétique serait en partie hérité de nombreuses inclusions de code viral. Enfin, un autre organe, très particulier dans sa présentation, intervient également ; il s'agit du système immunitaire, lequel est quasi ubiquitaire, distribué et circulant, en perpétuel remaniement, et en interactions incessantes avec nos propres cellules et les agents reconnus comme extérieurs à notre organisme. Une des particularités de cet organe est par ailleurs de présenter une mémoire dynamique : au fil des contacts, le système immunitaire apprend de ses rencontres, et est capable de déclencher une réaction plus rapide et adaptée en cas de nouvelle rencontre avec l'agent extérieur. L'on sait en outre que cet organe, sa capacité à discriminer le soi du non soi, n'a rien d'absolu, et qu'il doit certainement une partie de sa plasticité à cette imprécision de reconnaissance. C'est là encore un argument qui plaide en faveur de l'impossibilité de totalement démêler des formes complexes, du moins statiquement. Sans doute est-il possible de définir la physiologie comme un l'articulation entre différentes formes : en ce sens, le système immunitaire, l'existence de relations imbriquées entre microbiote et le reste de l'organisme hôte, incitent à considérer la notion de forme fonctionnelle. Autrement dit, la forme se distingue par deux types d'information : morphologique, ou statique, par la "visualisation" de ses éléments constitutifs, ceux-ci n'étant pas totalement spécifiques ni pertinents quand isolés, et l'information fonctionnelle, équivalent du mouvement pour la forme, laquelle renseigne sur les différentes fonctions qu'elle est capable de produire, selon les variations relatives aux-

quelles elle peut être soumise. Au fond, il serait nécessaire de comprendre comment circule l'information entre composants de la forme.

4.2.4 Vers une géométrie des maladies infectieuses : topologie de la causalité

Ainsi que l'on vient de le montrer, la relation élémentaire et centrale, paradigmatique de la maladie infectieuse, consistant en la relation causale directe entre un agent "monoclonal" et une pathologie à la définition claire (et pour cause : elle est définie par la présence de sa cause) demeure pertinente, mais revêt les aspects d'une relation linéaire localement valide, insuffisante à décrire l'ensemble complexe des relations entre agent extérieur et organisme hôte. Il s'agit en effet d'un cas extrême et isolé, assez réducteur, mais certes important du fait de ses implications, très longtemps mortelles.

La présence et l'identification, indirecte, d'un agent unique ne suffisant plus à caractériser dans son ensemble la causalité d'une pathologie, qui elle aussi, ne doit plus nécessairement son existence à un tableau clair, isolé des autres tableaux diagnostiques ; les relations entre agents infectieux et hôtes n'étant finalement pas réduites à une relation binaire et manichéenne, mais participant bel et bien d'un fonctionnement global, il devient important de dresser le paysage, dynamique, qui rend compte de tous ces modes de définitions et d'interactions.

Les limites de l'équation fondamentale de la pathologie infectieuse ne doivent pas être vécues comme un échec, ou un acte totalement inopérant, mais plutôt comme une approximation linéaire localement valide : elle ne prend en compte que deux éléments, bien souvent étudiés isolément, et dont on a réduit la relation à celle que l'on a bien voulu étudier, toutes choses égales par ailleurs. Le cas du VIH-SIDA devrait pouvoir nous aider, tout naturellement, tant par ses mécanismes encore largement étudiés, que par l'histoire de sa construction et de sa caractérisation au travers de multiples

dimensions – virologique, thérapeutique, diagnostique, sociale, politique...- et nous permettre de tester un modèle plus large, un modèle qui rende compte des articulations entre systèmes ou facteurs pensés isolément.

Quelle serait la géométrie de la pathologie infectieuse, en ce cas ? Nous montrons qu'elle est loin d'évoluer dans un espace restreint, bi ou tridimensionnel. Nous n'en sommes pas encore à en caractériser la structure fine ; néanmoins, puisque le lien de causalité est un fondamental fort et participant de la définition de la pathologie infectieuse, sans doute pouvons-nous le conserver, en le plongeant dans un cadre formel plus puissant. Il nous semble que le VIH-SIDA, en particulier, par la mise à mal des critères de causalité directs, par la distension temporelle par rapport aux critères habituels entre infection et déclaration de la maladie, enfin par son diagnostic aux frontières et aux expressions floues et variables, doit pouvoir se dépeindre fonctionnellement au moyen de réseaux de causalité, type réseaux bayésiens. La question de l'observation y prend tout son sens, ainsi que l'imprécision, probabilisée, et c'est la circulation de l'information entre les différents éléments associés au VIH qui nous intéressent. Par la constitution de ce réseau, nous aurons fait un premier pas vers la détermination d'une géométrie, car nous en connaissons la topologie, en termes d'information : nous saurons les relations de proximité causale, et non temporelle, entre observations. Il nous restera à en extraire une géométrie plus aboutie, ce qui nous occupera dans le chapitre 5, formellement. Peut-être le VIH sera une arête unique, un goulot d'étranglement incontournable pour mener au stade SIDA ; peut-être verrions-nous apparaître des chemins concurrents ; peut-être enfin, serait-il possible de délimiter un diagnostic syndromique, et donc également une stratégie diagnostique, plus pertinent, que la classification actuelle.

4.2.5 Au-delà de la pathologie infectieuse : l'exemple du cancer et de son diagnostic

Comme nous l'avons mentionné en introduction de cette section, la pathologie infectieuse a tout de la pathologie archétypique : elle se base, ou se basait, sur une relation causale directe entre deux éléments simples, identifiables, et mêmes visibles, pourvu que le microscope ait été inventé. La notion de diagnostic étiologique, et de traitement des causes plutôt que des symptômes tient cet archétype : il ne suffit pas d'identifier un syndrome, une forme syndromique, c'est-à-dire, une forme récurrente, invariante (peu variante) d'une personne à l'autre et d'un observateur à l'autre, pour établir un diagnostic utile. Il est nécessaire de lui adjoindre, s'il n'y a pas de correspondance biunivoque évidente et connue, un diagnostic étiologique, un diagnostic de la cause. Ainsi, le traitement pourra certes s'intéresser à une partie des symptômes – soulager une douleur par exemple – mais il devra avant tout s'attaquer à la cause, afin d'en faire disparaître les conséquences, c'est-à-dire la pathologie diagnostiquée.

Cette traduction depuis la pathologie infectieuse vers les autres types de pathologies n'a *in fine* rien de direct ou d'évident. Cette constatation rejoint, par ailleurs, nos considérations déjà établies sur la dichotomie entre normal et pathologique, entre deux états supposés bien séparés, un état sain et un état morbide, celui-ci étant par ailleurs le seul caractérisé par un diagnostic et des critères s'y rapportant.

Il est intéressant de constater qu'à l'autre bout du spectre des pathologies, du côtés des pathologies chroniques ou liées en partie à l'âge, ce paradigme semble sur le point de se fissurer. Un exemple particulier en serait la pathologie cancéreuse. Une grande partie des cancers sont liés à l'âge, d'une façon ou d'une autre, certainement par l'intermédiaire d'un mécanisme de tolérance ou de réparation des erreurs génétiques de plus en plus défailant avec l'avancement de la sénescence, et donc un risque statistique accru de

voir apparaître la fameuse cellule "immortelle", qui ne reconnaît plus ni ses consœurs, ni de limite particulière.

Un autre fait marquant est que ce type de pathologie, à l'instar de la pathologie infectieuse, fait intervenir des problèmes de signalisation, de communication et finalement, le système immunitaire. Le système immunitaire peut être vu comme un organe particulier, distribué, et avec mémoire, plastique et adaptatif. Il est une sorte de forme du soi, et en complémentaire, de ce qui n'est pas soi. Il régule ainsi les écarts et déséquilibres jusque dans certaines proportions, et pourvu 1) qu'il sache reconnaître correctement la forme du problème, 2) qu'il puisse agir correctement dessus.

Le problème du cancer présente également la composante du diagnostic, et de la classification. Qu'est-ce qu'un cancer, et qu'est-ce qu'une cellule cancéreuse ? Nous disposons à l'heure actuelle d'un certain nombre de projections, de mesure, de ce qui est commun dans plusieurs cas de figures qui pourraient être regroupés dans un même type d'entité – les cancers. Mais il est difficile de s'avancer plus sur la caractérisation formelle et complète de ce qu'est la normalité et le cancer, de façon radicalement séparée. Notamment, la question du dépistage en population générale, en aveugle, systématique et multimodal a été et continue d'être fréquemment proposé comme un dispositif ultime de prévention de la maladie. Il est également établi que cette démarche n'est pas pertinente, et vouée à l'échec par nature, même dans le cas où il y aurait effectivement séparation nette et absolue entre normalité et pathologie. [209, 208, 210, 211, 212, 213, 214, 215, 216]

Sans être allés jusque cette notion de dépistage global et systématique, l'augmentation du nombre d'exams complémentaires et de la fréquence à laquelle ils sont réalisés, ont mis en lumière des formes qui *a priori*, ne pouvaient être dites "normales", et par conséquent, ne pouvaient être que pathologiques et donc, cancéreuses – avec cette sortie de secours tout de même, de la lésion bénigne. Evidemment, le surdiagnostic a engendré et

engendre encore le surtraitement.

Très récemment, il a été proposé, non seulement de rationaliser la prescription des examens complémentaires et des traitements, mais également d'envisager que le terme de cancer ne soit pas la seule alternative à la normalité, caractérisée par l'absence de processus tumoral. Il y aurait donc bien des états intermédiaires, non nécessairement menant à la malignité, voire un continuum entre "normalité" et "cancer". [209, 217, 218]

4.3 Modéliser un phénomène infectieux

Après avoir esquissé ce que pourrait être une approche géométrique, ou du moins à ce stade, une approche topologique, des phénomènes infectieux, nous nous proposons d'étudier plus précisément la forme que peut prendre la modélisation d'un processus infectieux complexe, cette fois à l'échelle de l'individu, sans composante épidémique, mais en couplage avec deux systèmes physiologiques. Nous montrons d'abord comment cette approche dérive d'une approche de modélisation historique des épidémies, *via* les modèles compartimentaux. Nous nous intéresserons ensuite à la modélisation des infections ostéo-articulaires (IOA) comme cas d'étude. Les IOA sont des processus complexes, mal connus. Nous les approcherons en plusieurs étapes : la première sera de présenter un modèle de réponse immunitaire, la deuxième, de suggérer l'emploi d'un modèle d'homéostasie osseuse ; nous nous attaquerons après à l'équation régissant la croissance bactérienne en relation avec ces deux systèmes ; enfin, nous proposerons 4 modèles d'IOA qui correspondent à autant d'hypothèses de couplages entre systèmes. Nous conclurons par des commentaires sur la complexité des modèles obtenus, et les limites de l'approche par l'analyse structurelle des modèles retenus.

4.3.1 Le modèle compartimental et les épidémies

Les modèles les plus connus en modélisation des épidémies sont les modèles dits compartimentaux, dont les modèles SIR (sain, infecté, remis) et ses dérivés sont représentatifs. Ces modèles sont eux-mêmes hérités de modèles démographiques, servant de bilan de populations, entre naissances, décès, immigration et émigration, par exemple.

Nous parlons de compartiments, car l'on considère le problème d'un point de vue populationnel, population divisée en compartiments selon une caractéristique particulière : le compartiment de la population saine (S), le compartiment de la population infectée (I), etc.

Les variations temporelles de chaque compartiment sont données par des équations différentielles ordinaires – l'on fait l'hypothèse de grands nombres pour les populations, et d'une différentiabilité des effectifs des populations. [219, 220] Les modèles SIR peuvent ainsi être décrits comme suit.

Le raisonnement présidant à la construction des modèles SIR est celui présidant à la modélisation de n'importe quel phénomène dont on suppose pouvoir écrire mathématiquement les raisons de la variation d'une variable donnée entre deux instants t et $t + dt$. Autrement dit, en ce qui concerne la population de personnes saines, par exemple, comment explique-t-on dans notre modèle la variation $dS = S(t + dt) - S(t)$, par unité de temps dt ? On remarquera notamment que si $dS < 0$, alors le bilan est négatif, ce qui signifie qu'il y a eu décroissance du nombre de personnes saines sur la période dt . En particulier, si l'on considère la variable I des personnes infectées, alors dI/dt représente l'incidence de la maladie d'intérêt puisqu'elle donne le nombre de nouveaux cas d'infections de la maladie sur la période dt .

Dans le modèle élémentaire SIR, une personne donnée fait partie à $t = 0$ des personnes saines, donc de S ; elle peut être infectée, et quitter S pour

rejoindre et accroître les effectifs de I ; elle peut ensuite guérir avec une "cicatrice immunitaire" si la maladie est immunisante, et passer de I à R . Il convient ensuite de poser des hypothèses concernant les taux d'accroissement des divers compartiments. Notamment, on posera p le taux d'infection. De la sorte, on suppose également que la variation sur dt des infectés est proportionnelle au nombre de personnes saines et au nombre de personnes déjà infectées – il faut une personne infectée pour contaminer une personne saine. Il vient alors :

$$\frac{dS}{dt} = -p.S.I$$

Si maintenant l'on note α le taux de personnes infectées qui se remettent de l'infection, on peut écrire de la même façon :

$$\frac{dR}{dt} = \alpha.I$$

Il vient alors tout naturellement que la variation du nombre de personnes infectées s'écrit :

$$\frac{dI}{dt} = p.S.I - \alpha.I$$

Ceci est vrai si l'on fonctionne à population "fermée", c'est-à-dire que toutes les personnes présentes initialement persistent dans le temps, et sont nécessairement soit S , soit I soit R , et il n'y a ni naissance ni mort, ni immigration ni émigration. Cette hypothèse vérifie bien que : $\frac{dS}{dt} + \frac{dI}{dt} + \frac{dR}{dt} = 0$.

Comme nous l'avons dit, il est possible de raffiner à l'envi ce type de modèle. Par exemple, et pour finir, on peut très bien ajouter un état intermédiaire, entre sain et infecté, l'état "exposé" E . Le passage d'exposé à infecté est donné par le taux β , et permet d'écrire :

$$\frac{dE}{dt} = p.S.I - \beta.I$$

et

$$\frac{dI}{dt} = \beta.E - \alpha.I$$

Ces équations sont relativement simples, et il est possible de les raffiner à volonté : l'on peut introduire d'autres compartiments, des états intermédiaires, l'on peut également introduire une structure d'âge dans la population, une caractérisation de l'hétérogénéité des populations, etc. Le même genre d'approche, dans son principe, peut être proposée pour modéliser les IOA.

Globalement, et plus généralement, les modèles mobilisés, et la forme des fonctions que nous utiliserons dans la suite viennent initialement des considérations de bilans démographiques. Le lecteur pourra ainsi bénéficier d'un plus grand recul en consultant par exemple [221].

4.3.2 Approche des infections ostéo-articulaires : un cas typique

Une IOA, comme son nom l'indique, est une infection touchant une articulation ou une structure osseuse. Elle désigne indifféremment l'atteinte par un agent infectieux de n'importe quelle couche de la structure osseuse, c'est-à-dire aussi bien sa surface (la corticale) que son cœur (la moelle osseuse).

L'histoire naturelle de l'IOA est emplies de suppositions plausibles plus que de preuves objectives et reproduites clairement. Comment peut-on infecter un os ou une articulation ? Il faut la présence d'un germe, *a priori* une bactérie, au niveau du site concerné. Ce germe peut arriver à cet endroit soit par inoculation, par exemple après une blessure profonde ou une fracture ouverte, lors d'une intervention chirurgicale orthopédique, soit par voie hématogène, et donc à partir d'un foyer ou d'un point d'inoculation distant,

et transport par le sang jusque le tissu osseux. [222, 223]

Les IOA sont des infections fréquentes, graves, souvent tendant vers le chronique, mais posent deux types de problèmes pour être analysées simplement par les approches statistiques usuelles : elles ne sont pas assez fréquentes dans un laps de temps raisonnable, et les personnes présentant des IOA présentent également une large variété dans leurs caractéristiques : délai d'apparition de l'IOA, présence de matériel prothétique ou non, si oui, de quel type, schéma thérapeutique reçu... Autant de facteurs peu évidents à déterminer et mesurer, et trop nombreux pour être correctement incorporés dans un modèle statistique classique multivarié. Une alternative est donc d'approcher les IOA par la modélisation : le manque d'information dont nous souffrons dans la réalité sera pallié par des hypothèses bien choisies quant aux mécanismes sous-tendant les IOA.

4.3.3 Modélisation du système immunitaire

Il existe un certain nombre de modèles décrivant tout ou partie du système immunitaire, que cela concerne la réponse immunitaire innée ou spécifique. Dans ces deux cas, les acteurs concernés ne sont pas les mêmes, ni leur efficacité sur les agents pathogènes à éliminer ou freiner. Néanmoins, beaucoup de modèles ciblent une petite partie du système immunitaire, voire s'intéressent spécifiquement au comportement d'un acteur particulier, par exemple le lymphocyte B, ou le lymphocyte T.

Ici, nous nous intéresserons avant tout à la réponse non spécifique, c'est-à-dire, la première barrière face à toute effraction de l'intégrité d'un organisme : la réponse enclenchée face à la reconnaissance d'un corps étranger, reconnu comme « non soi ».

4.3.4 Le modèle de réponse immunitaire non spécifique en 3 vagues de l'infection pulmonaire

Nous décrivons dans cette partie un modèle de réponse immunitaire non spécifique, dont une modélisation en 3 temps a été proposée par A. Smith dans le contexte de l'infection bronchopulmonaire à pneumocoque. [224] Il nous servira de modèle de base concernant la réponse du système immunitaire face à une stimulation bactérienne. Nous en décrivons d'abord les spécifications générales telles que formulées originalement pour le poumon, en exposons le comportement qualitatif et quantitatif, puis proposons des adaptations pour le modèle d'IOA.

Descriptions

Nous reprenons ici une description en 3 temps du modèle de réponse immunitaire non spécifique :

- une première phase impliquant les macrophages (dans le cas du poumon : macrophages alvéolaires),
- une deuxième phase impliquant la sécrétion par les cellules du tissu agressé de cytokines, avec recrutement de neutrophiles,
- une troisième phase où les neutrophiles, sécrétant également certaines cytokines, vont recruter des macrophages dérivés des monocytes, provenant de la circulation sanguine et se déplaçant jusque le site infecté ; cette phase est donc retardée par rapport aux deux autres.

La première phase concerne donc la réponse des macrophages résidents du tissu face à l'agression bactérienne. Le point de vue est populationnel, c'est-à-dire que l'on écrit des équations, des relations réglant l'évolution de populations (ou de compartiments, comme on le rencontre dans la littérature) en fonction du temps. Ce qui nous intéresse ici est l'évolution de la population bactérienne au cours du temps et selon l'action des macrophages. On note

B le nombre de bactéries à un temps t plutôt que $B(t)$ pour plus de lisibilité dans la suite. M_a désigne la population de macrophages résidants.

L'évolution de la population bactérienne est réglée par une balance entre une fonction de croissance bactérienne, et une fonction de répression de cette croissance, ici représentée par la phagocytose bactérienne exercée par les macrophages. On considère ici classiquement que la croissance bactérienne suit une croissance logistique (modèle de Verhulst), selon un taux de croissance r et limitée par la capacité (le "volume" maximal) du tissu à hébergé de façon viable des bactéries K_p . On a donc l'équation d'évolution de la forme suivante :

$$\frac{dB}{dt} = \dot{B} = f_{croissance}(B) - f_{repression}(B, M_a) \quad (4.1)$$

où l'on a

$$f_{croissance}(B) = r \left(1 - \frac{B}{K_p}\right) B \quad (4.2)$$

$$f_{repression}(B, M_a) = \gamma_{M_a} f(B, M_a) M_a B \quad (4.3)$$

avec

$$f(B, M_a) = \frac{n^x M_a}{B^x + n^x M_a} \quad (4.4)$$

De fait, l'évolution de B s'écrit donc en totalité :

$$\frac{dB}{dt} = \dot{B} = r \left(1 - \frac{B}{K_p}\right) B - \gamma_{M_a} f(B, M_a) M_a B \quad (4.5)$$

La deuxième vague implique donc le recrutement de neutrophiles *via* la sécrétion de cytokines, notamment de la part des cellules agressées et en contact avec les bactéries ; dans le cas du poumon, les cellules épithéliales

jouent ce rôle. En outre, les macrophages résidants sont eux aussi capables de sécréter des cytokines recruteuses de neutrophiles, et l'on va intégrer cette réalité en séparant en deux sous populations M_R et M_C la population initiales de macrophages M_a . M_R désigne la part des macrophages ne "inactifs" au sens de la sécrétion de cytokines ; ils peuvent être activés, et alors sécréter des cytokines. M_C désigne *a contrario* la part des macrophages à être activés, et donc à sécréter des cytokines. On a évidemment la relation de conservation :

$$M_a = M_R + M_C \quad (4.6)$$

On introduit par ailleurs d'autres populations, dont les cellules tissulaires – épithéliales dans le cas pulmonaires. On les note T_i et T_a respectivement pour les cellules tissulaires inactivées et activées, de la même façon que pour les macrophages. Les cellules sont activées si elles entrent en contact avec une bactérie, leur nombre est donc proportionnel au nombre de bactéries. Sur la dynamique considérée en première ligne, ces cellules ne sont pas remplacées en cas de destruction.

On note C la population des cytokines. Elle est directement proportionnelle à la quantité de cellules tissulaires activées et de macrophages activés.

On note N la population des neutrophiles recrutés. De la même façon que pour les bactéries, on modélise la "croissance" des neutrophiles comme une fonction pseudo logistique, et soumise à une capacité tissulaire maximale N_{max} . Pseudo logistique, car la croissance des neutrophiles, en particulier initialement, est conditionnée par la population de cytokines, ce qui transparaît dans la formule. Les termes de puits qui font disparaître une partie des neutrophiles sont liés à la mortalité naturelle des cellules, et liés à l'activité cytolytique des bactéries.

On introduit enfin une population "débris", notée D , et consistant en l'ensemble des dégâts causés par la réaction : débris macrophagiques, débris neutrophiliques et enfin débris de dégradation tissulaire. Ces nouvelles populations mises à part, 2 grandes modifications sont à comptabiliser dans notre modèles initial, toutes deux concernant le terme en $f_{repression}(B, M_a)$, qui devient $f_{repression}(B, M_a, N, D)$. D'une part, l'apparition de débris D fait que les macrophages M_a doivent désormais partager leur temps de phagocytose entre bactériophagie et absorption des débris. Apparaît donc un terme d'inhibition de la bactériophagie prenant en compte les débris phagocytés. D'autre part, il convient en effet d'ajouter un terme de répression bactérienne, en raison de l'action bactéricide des neutrophiles.

De fait, le système initial devient :

$$\frac{dB}{dt} = r \left(1 - \frac{B}{K_p}\right) B - \frac{\gamma_{M_a} f(B, M_a)}{1 + k_d D M_a} M_a B - \gamma_N N B \quad (4.7)$$

$$\frac{dT_i}{dt} = -\omega B T_i \quad (4.8)$$

$$\frac{dT_a}{dt} = \omega B T_i - d_T T_a \quad (4.9)$$

$$\frac{dC}{dt} = \alpha \frac{T_a}{1 + k_n N} + \nu \frac{\theta_M M_a B}{(d + \kappa + \theta_M B)(1 + k_n N)} - d_C C \quad (4.10)$$

$$\frac{dN}{dt} = \eta \left(1 - \frac{N}{N_{max}}\right) C - d_{NB} N B - d_N N \quad (4.11)$$

$$\frac{dD}{dt} = \rho_1 d_{NB} N B + \rho_2 d_N N + \rho_3 d_T T_a - d_D D M_a \quad (4.12)$$

La troisième vague de la réponse immunitaire implique le recrutement, retardé, de macrophages dérivés des monocytes, lesquels proviennent du compartiment sanguin, et sont donc chimio attractés jusque le site infectieux. Ces macrophages ne sont pas polynucléés comme les macrophages résidants, mais mononucléés. Nous notons M cette nouvelle espèce de macrophage "tardif". L'introduction de cette population supplémentaire perturbe le système précédent par l'adjonction d'un terme de puits supplémentaire dans la fonction de répression bactérienne, au niveau de l'équation d'évolution de B . Il convient évidemment d'adjoindre au système une dimension supplémentaire, puisqu'il est nécessaire de connaître l'équation d'évolution de M .

Les macrophages dérivés des monocytes sont recrutés essentiellement *via* les neutrophiles. La fonction de croissance de M est donc construite sur un modèle similaire à celui régissant C , suivant une progression pseudo-logistique, proportionnelle en N , et plus précisément en $N(t - \tau)$, puisque relativement aux autres populations considérées comme directement mobilisables, ces macrophages arrivent sur site avec un retard τ . De fait, nous ne récrivons pas à la suite l'intégralité du système, mais seulement l'équation d'évolution de B et celle, nouvelle, de M , le reste demeurant inchangé. On a :

$$\frac{dB}{dt} = r \left(1 - \frac{B}{K_p} \right) B - \frac{\gamma_{M_a} f(B, M_a)}{1 + k_d D M_a} M_a B - \gamma_N N B - \gamma_M M B \quad (4.13)$$

$$\frac{dM}{dt} = \xi \left(1 - \frac{M}{M_{max}} \right) N(t - \tau) - d_M M \quad (4.14)$$

4.3.5 Modélisation de l'homéostasie osseuse

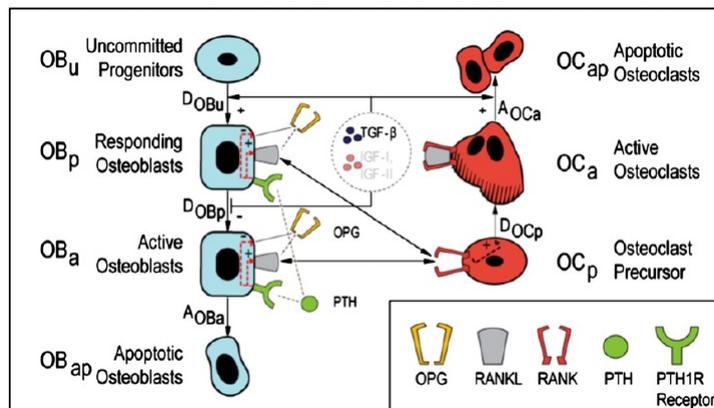
Nous venons d'exposer et de décrire une modèle en 3 vagues successives de réponses immunitaire non spécifique, suffisamment générique pour être

adaptable et adaptée à notre cas de l'IOA. Nous avons appliqué des méthodes issues des systèmes dynamiques afin de mieux en qualifier le comportement, éventuellement en simplifier localement la modélisation. Dans cette partie, nous allons traiter similairement le cas de l'os, qui se substituera en quelque sorte au tissu pulmonaire du modèle précédent.

Ce qui nous intéresse en l'espèce est de formuler un modèle qui régit l'homéostasie osseuse, puisque l'os est un tissu vivant, et en constant remaniement dynamique. En outre, l'activation d'une réponse immunitaire en la présence d'une agression bactérienne peut venir, *via* des couplages, des fonctions différentes assumées par des acteurs identiques, perturber cet équilibre dynamique. Il est à l'heure actuelle acquis que l'équilibre dynamique osseux met en jeu la balance entre ostéoclastes et ostéoblastes, issus tous deux de deux lignées cellulaires bien différentes. [225, 226, 227, 228]

Modèle homéostatique de l'os

Les acteurs retenus dans notre modèle homéostatique de l'os sont rapportés sur le schéma suivant.



Nous distinguons donc 7 types de cellules osseuses dans notre modèle, réparties en 2 catégories, à savoir les ostéoblastes et les ostéoclastes (précurseurs inclus par lignée).

- les progéniteurs ostéoblastiques : OB_u

- les ostéoblastes répondant : OB_p
- les ostéoblastes activés : OB_a
- les ostéoblastes en apoptose : OB_{ap}
- les précurseurs ostéoclastiques : OC_p
- les ostéoclastes activés : OC_a
- les ostéoclastes en apoptose : OC_{ap}

Les acteurs non cellulaires sont par conséquent ici la PTH, l'OPG, le TGF- β et RANKL/RANK.

Nous décrivons dans la suite les équations régissant l'homéostasie de l'os.

$$\frac{dOB_p}{dt} = D_{OB_u} \cdot \pi_{act,OB_u}^{TGF-\beta} - D_{OB_p} \cdot OB_p \cdot \pi_{rep,OB_p}^{TGF-\beta} \quad (4.15)$$

$$\frac{dOB_a}{dt} = D_{OB_p} \cdot OB_p \cdot \pi_{rep,OB_p}^{TGF-\beta} - A_{OB_a} \cdot OB_a \quad (4.16)$$

$$\frac{dOC_a}{dt} = D_{OC_p} \cdot OC_p \cdot \pi_{act,OC_p}^{RANKL} - A_{OC_a} \cdot OC_a \cdot \pi_{act,OC_p}^{TGF-\beta} \quad (4.17)$$

où les grandeurs D_{OB_u} , D_{OB_p} et D_{OC_p} sont les taux de différenciation des populations concernées respectives. Les grandeurs A_{OB_a} et A_{OC_a} sont les taux d'apoptose des ostéoblastes et ostéoclastes activés. Les grandeurs du type $\pi_{act,OC_p}^{TGF-\beta}$ sont les fonctions d'activation (indice en *act*), respectivement de répression (indice en *rep*) liées à la capture de TGF- β soit par les ostéoblastes soit par les ostéoclastes. Nous exprimons plus loin la forme des fonctions d'activation/répression, qui sont là aussi des fonctions de Hill. Le principe est le même pour les exposants en RANKL.

Le volume osseux total VO se comporte dynamiquement de la façon suivante :

$$\frac{dVO}{dt} = -k_{res} \cdot OC_a + k_{form} \cdot OB_a \quad (4.18)$$

Donnons ici l'expression des fonctions d'activation/répression impliquant TGF- β :

$$\pi_{act,OB_u}^{TGF-\beta} = \frac{TGF - \beta}{K_{D1,TGF-\beta} + TGF - \beta} \quad (4.19)$$

$$\pi_{rep,OB_p}^{TGF-\beta} = \frac{1}{1 + \frac{TGF-\beta}{K_{D2,TGF-\beta}}} \quad (4.20)$$

$$\pi_{act,OC_p}^{TGF-\beta} = \frac{TGF - \beta}{K_{D3,TGF-\beta} + TGF - \beta} \quad (4.21)$$

Enfin, nous donnons la forme de la fonction d'activation concernant RANKL :

$$\pi_{act,OC_p}^{RANKL} = \frac{RANKL}{K_{D8,RANKL} + RANKL} \quad (4.22)$$

À ce point, la connaissance de RANKL et de TGF- β est nécessaire. Il est possible d'en donner une expression :

$$RANKL = (\beta_{RANKL} - D_{RANKL}).t \quad (4.23)$$

$$TGF - \beta = \frac{\alpha.k_{res}OC_a + S_{TGF-\beta}}{D_{TGF-\beta}} \quad (4.24)$$

Il est possible, afin de diminuer la complexité du modèle tout en en conservant le minimum nécessaire, de se replier sur le modèle initial proposé par Komarova, [229] où il n'est conservé en première approche que les fonctions autocrines et paracrines des ostéoclastes et ostéoblastes.

$$\frac{dOC}{dt} = \alpha_{OC}.OC^{g11}.OB^{g21} - \beta_{OC}.OC \quad (4.25)$$

$$\frac{dOB}{dt} = \alpha_{OB} \cdot OC^{g12} \cdot OB^{g22} - \beta_{OB} \cdot OB \quad (4.26)$$

4.3.6 Modèle d'IOA : hypothèses de couplage des systèmes

L'ostéoclaste est une cellule osseuse qui a migré jusque ce tissu particulier – il s'agit en effet d'un macrophage, spécialisé, mais dérivant donc de la lignée blanche. De fait, il existe des liens non encore totalement élucidés entre système immunitaire et tissu osseux. Certains auteurs supposent ainsi qu'il existe une reconnaissance par les macrophages de l'environnement osseux ou de certaines de ses cellules constitutives ; les monocytes attaquent notamment l'os minéralisé. De fait, la réponse immunitaire non spécifique pourrait impliquer une attaque partiellement dirigée vers l'agent pathogène, et vers le milieu infecté lors du recrutement des monocytes. [230, 231, 232, 233, 234, 235]

En outre, il est connu que le tissu osseux est le milieu privilégié où est localisé le TGF- β , nécessaire à l'homéostasie osseuse. Cependant, il a été également été montré que le TGF- β exerce une double et contradictoire régulation sur les macrophages. Plus particulièrement, il active les monocytes et induit un recrutement de ce type de cellules, amplifiant ou modulant potentiellement la réponse immunitaire non spécifique une fois amorcée, tandis qu'il désactive les macrophages, pénalisant cette fois la réponse immunitaire une fois celle-ci amorcée. [236, 237, 225, 238, 239, 240]

Ces comportements étant moins spécifiquement ou précisément décrits, en particulier en termes de modélisation, nous proposons de les modéliser en laissant leurs paramètres spécifiques flottant, et d'examiner la stabilité structurelle formelle du système impliquée par les valeurs possibles de ces facteurs.

- les monocytes attaquent les ostéocytes lors de la réponse inflammatoire

H1

- le TGF- β recrute les monocytes H2
- le TGF- β désactive les macrophages H3
- les bactéries peuvent être incluses par les ostéoblastes H4

La traduction de ces 4 hypothèses se fait de la manière suivante :

Hypothèse 1 : un terme d'inhibition freinant l'activité des M_o dans l'équation régissant B , et un terme de puits dans les équations régissant les OB :

$$\frac{1}{1 + K_{OB,M_o} \cdot OB \cdot M_o} \quad (4.27)$$

$$-\gamma_{OB,M_o} \cdot M_o \cdot OB \quad (4.28)$$

Hypothèse 2 : un terme additif de recrutement des M_o dans l'équation régissant les M_o

$$+\gamma_{TGF-\beta,M_o} \cdot TGF - \beta \quad (4.29)$$

Hypothèse 3 : un terme de source et de puits dans les équations régissant les M_r et les M_c

$$+/- \gamma_{TGF-\beta,M_a} \cdot M_c \quad (4.30)$$

Hypothèse 4 : un terme de puits dans l'équation régissant les B , et un terme d'inhibition dans les équations régissant les OB

$$\frac{1}{1 + K_{OB,B} \cdot OB \cdot B} \quad (4.31)$$

$$-\gamma_{B,OB} \cdot OB \cdot B \quad (4.32)$$

Ces 4 hypothèses successives montrent comment il est possible, à partir de modèles non couplés, chacun décrivant une homéostasie isolée – l’homéostasie osseuse, le système immunitaire face à un agent infectieux – de composer différents modèles rendant compte d’interactions entre ces modèles d’homéostasie. En effet, le simple fait que des messagers cellulaires puissent être non spécifiques d’une cellule particulière, ou même d’un système particulier, implique que l’on peut relier les systèmes entre eux *via* certains de leurs composants. En outre, ces interactions élémentaires sont relativement simples, mais généralement non linéaires, et introduisent des rétroactions à différents niveaux, de sorte que le résultat n’est pas nécessairement la superposition triviale des solutions des systèmes pris isolément.

En ce qui nous concerne, il est important de constater que les facteurs intervenant dans la régulation osseuse interviennent également au niveau de la régulation – ambivalente – des composants du système immunitaire.

4.3.7 Intégration de la stratégie thérapeutique au modèle d’infection ostéoarticulaire

En adjoignant au modèle de remodelage osseux une composante immunitaire, et une agression bactérienne, nous avons donné le comportement en "roue libre" de l’organisme, sans intervention extérieure ou toute autre perturbation que celle de l’introduction d’un inoculum bactérien. Il s’avère que dans certains cas, cas pour lesquels par ailleurs il est possible qu’aucun signe clinique patent soit observable dans la vie quotidienne, la croissance et la persistance bactérienne puissent être contrecarrées, et que B soient menées à l’extinction. Dans d’autres cas, le système de base est dépassé, et la croissance bactérienne part à saturation, maximise son occupation tissulaire.

Lorsque l’organisme n’est plus à même d’assurer sa propre régulation, son maintien dans un domaine viable pour lui, l’intervention thérapeutique prend

un sens. Reste la double question de savoir laquelle, et suivant quel schéma "optimal". Ici, la question principale qui nous intéresse est celle de l'efficacité, ou à défaut de l'action de l'antibiothérapie seule sur une IOA, en dehors de toute technique chirurgicale. Pour être parfaitement général, nous parlons de chimiothérapie, plutôt que de se cantonner à l'antibiothérapie, prise dans le sens de toute thérapie agissant par le biais d'une molécule.

Modèles pharmacocinétiques et pharmacodynamiques

Les modèles pharmacocinétiques diffèrent selon plusieurs paramètres, notamment selon la voie d'administration – par intraveineuse directe (IVD) ou par voie orale (*per os*, PO) par exemple. Nous ne qualifierons pas ici tout le cycle ADME – Absorption, Disponibilité, Métabolisation, Élimination – mais il faut garder à l'esprit que des modèles sont disponibles pour chacune des étapes de ce cycle idéal, et qu'il est relativement simple de les adjoindre si nécessaire. En outre, nous faisons l'hypothèse d'une distribution sanguine du produit actif, avec passage hépatique et en circulation générale.

De fait, nous pouvons simplifier le problème de la chimiothérapie en retenant les caractéristiques suivantes :

- la voie d'administration
- la vitesse de disparition du principe actif dans le sang
- la biodisponibilité du produit au site, ici, dans l'os
- la notion de "résistance" bactérienne induite par la chimiothérapie
- l'action bactéricide de la chimiothérapie

Modèles retenus

L'action bactéricide du traitement peut se modéliser de la manière suivante :

$$\frac{k_{max}ATB^\gamma}{EC_{50} + ATB^\gamma}B \quad (4.33)$$

où ATB désigne la concentration en principe antibiotique, γ le facteur de Hill, B la concentration bactérienne, k_{max} l'efficacité maximale de l'agent bactéricide et EC_{50} la concentration en ATB menant correspondant à une efficacité de 50% de l'efficacité maximale.

L'adaptation – entendre, la résistance acquise, quel que soit le moyen – de l'agent infectieux au principe antibiotique peut se modéliser de la manière suivante, en faisant varier EC_{50} avec le temps :

$$EC_{50} = EC_{50t_0} \cdot \alpha \quad (4.34)$$

avec EC_{50t_0} la concentration EC_{50} correspondant au moment initial t_0 d'exposition de l'agent infectieux au principe bactéricide, et α tel que :

$$\alpha = 1 + \beta [1 - e^{-ATB \cdot \tau \cdot t}]$$

où τ est une constante de temps d'adaptation et t le temps.

Enfin, il est possible de modéliser la disponibilité au niveau du tissu cible en fonction de la concentration sanguine et de la perfusion sanguine du tissu. Nous pouvons écrire :

$$V \cdot \frac{dATB_{tissu}}{dt} = Q_{tissu} \cdot \left[ATB_{sang} - \frac{ATB_{tissu}}{R} \right] \quad (4.35)$$

avec ATB_{sang} et ATB_{tissu} les concentrations respectives en principe antibiotique dans le sang et le tissu cible. Pour se rapprocher de l'objet de notre étude, les IOA, il existe une petite fraction d'études pharmacocinétiques et pharmacodynamiques ciblant les propriétés antibiotiques en tissu osseux, notamment en modèles animaux. [241, 242, 243, 244]

Ainsi, il est possible, au-delà des couplages induits par les 4 hypothèses que nous avons formulées plus haut, et issues des données de la littérature, d'ajouter d'autres composantes à notre modèle. Tel quel, nous avons tous les éléments de base pour modéliser les comportements non triviaux de l'homéostasie osseuse sous l'impact d'un agent infectieux unique (au sens de monoclonal), et sous l'impact d'un traitement visant cet agent. Tout serait parfait si nous disposions des valeurs des paramètres réglant chacune des équations. Néanmoins, à ce niveau, nous n'avons pour la plupart de ces paramètres que des conjectures, qu'un minimum d'expériences de laboratoire *in vivo* devrait aider à consolider – ou invalider.

4.3.8 Complexité du modèle et analyse de sa stabilité structurelle : limites de l'approche

Bien entendu, nous pourrions, à partir des données déjà existantes et d'un certain nombre de suppositions, fixer des valeurs pour chacun des paramètres, et faire simuler le modèle numérique sur un ordinateur. Cependant, la qualité de la démonstration demeurerait faible, et très sujet à caution. Il faut bien voir que l'analyse systématique de la littérature concernant de tels modèles ne remonte finalement que peu de données dont on peut supposer qu'elles sont fiables – même pour les modèles que nous avons retenus, publiés par ailleurs dans des revues de bon niveau, il a été nécessaire de contacter les auteurs afin de récupérer ou corriger des éléments des papiers.

Il y a alors la possibilité de ne pas se préoccuper des valeurs des paramètres, et de procéder à une analyse de la forme des équations, et notamment, à une analyse de la stabilité structurelle des systèmes obtenus. Cependant, là aussi, nous sommes rapidement limités, dans le sens où il devient rapidement compliqué d'analyser des systèmes qui dépassent 4 paramètres inconnus – autrement dit, 4 paramètres dont on étudie l'influence sur la stabilité structurelle.

Enfin, qui dit analyse de la stabilité structurelle, dit recherche des points fixes. L'idéal est alors de procéder à la recherche des points fixes de manière formelle et analytique, ce qui peut ici encore poser problème. La forme des solutions envisagées est au mieux polynomiale de degré entier en la variable d'intérêt – ici, la variable B , la variable de population bactérienne. Or, l'on sait qu'il n'existe pas de méthode systématique d'identification des racines des polynômes de degré supérieur ou égal à 5. Un compromis doit donc être entrepris entre recherche des solutions numériques, par simulation, et nécessitant des valeurs suffisamment fiables des paramètres, et analyse formelle et structurelle, qui nécessite quant à elle un système relativement "compact".

En l'état, les valeurs expérimentales n'étant pas les plus simples à obtenir, et requérant du temps et du tact, une piste raisonnable pour des développements plus avancés sera de reformuler le système complet autour d'hypothèses simplificatrices de sa forme, notamment en examinant les éléments les plus négligables relativement à d'autres, ou encore, de possibles symétries. Il sera également possible, pour commencer, d'étudier des domaines et des voisinages particuliers, plutôt que d'envisager de travailler sur tout l'espace de définition.

4.4 Typologies et dynamiques sociales

Nous avons annoncé que la sociologie présentait des intérêts particuliers en termes d'approches par la forme, et ce pour une raison relativement simple : la sociologie étant désireuse de se définir en tant que science tant par sa méthode que par son objet d'étude, il lui faut se déterminer sur les deux fronts, ce qui peut par ailleurs s'envisager conjointement. La méthode retenue, de la même façon qu'un capteur ne restitue qu'une partie du réel, et qu'un observateur rend compte depuis son point de vue, détermine au minimum l'espace de représentation des objets d'études, ainsi que ses contours, son échelle et

TABLE 4.3 – Typologie du suicide, d’après Durkheim (*Le suicide*).

	Régulation sociale	Intégration sociale
Excès	Suicide fataliste	Suicide altruiste
Défaut	Suicide anémique	Suicide égoïste

son degré de précision, de définition au sens de la résolution.

4.4.1 Existe-t-il des formes sociales ? Premières tentatives en sociologie

Nous ne recenserons pas toutes les tentatives faites en sociologie pour déterminer un objet stable et observable par un ensemble suffisamment important d’observateurs ; nous pouvons cependant souligner deux approches faisant écho à deux des possibilités que nous avons identifiées comme pouvant servir de cadre à une approche par la forme. Le premier exemple est d’autant plus important, qu’il est celui choisi par Durkheim immédiatement après l’exposé de sa méthode fondatrice pour en illustrer la démarche : l’étude du suicide comme fait social, et partant, dégager les formes du suicide comme forme sociale – une typologie du suicide. [9] Il est remarquable de constater que cette typologie remplit bien les conditions de la recherche d’un modèle explicatif, au sens de l’exposé d’un ensemble simple et faisant sens, mais également reposant sur une combinatoire, une algèbre très simple composée de deux éléments. Les types de suicide, nous dit Durkheim, sont au nombre de quatre, et sont chacun l’expression de la combinaison d’un excès ou d’un défaut, soit de la régulation exercée par la société et ses règles sur la personne, soit d’intégration de la personne par la société (cf tableau 4.3).

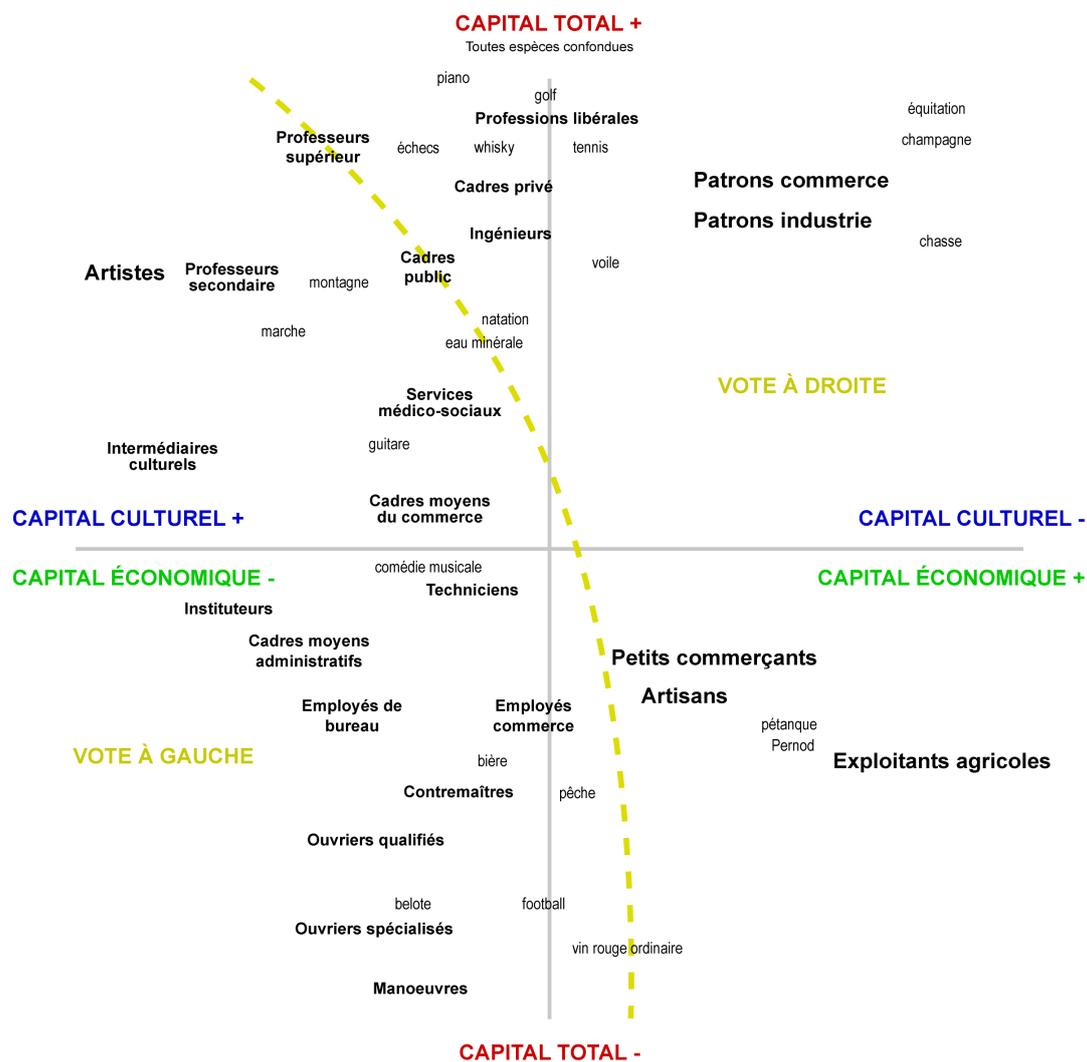
Ainsi, la combinaison aboutissant à un suicide dit fataliste, est un excès de régulation sociale : la personne, écrasée par trop de contraintes, éventuellement perçues comme contradictoires ou dépassant ses capacités d’assimilation, ne voit plus d’autre issue que celle de disparaître en tant que personne.

Le suicide dit anémique, quant à lui, serait dû à un défaut de régulation : la société laisse à penser à la personne que tout y est possible, qu'il n'existe pas d'autre structure qu'une structure amorphe, peu concernée par les personnes, les individualités qui la constituent ; cependant, le principe de réalité rattrape la personne, et entraîne un conflit avec l'apparente omnipotentialité de la personne, et ce qu'elle peut réellement accomplir, matériellement, concrètement.

Cet exemple est historique, sera repris, examiné et affiné par un autre sociologue, M. Halbwachs. Il est également célèbre pour s'appuyer sur des données quantitatives, des statistiques de plusieurs pays en termes de mortalité et suicides. Il est éclairant dans le sens où il propose bien quatre objets de l'échelle des groupes de personnes, de la population, objets qui apparaissent invariants, mais également le résultat d'une dynamique combinatoire. Les classes portent en elles leur construction.

Un autre exemple plus proche de nous, illustre davantage l'approche par la détermination d'un espace de représentation, adossée à des données là aussi, espace dans lequel il devient possible de localiser les personnes. Cet espace n'est pas donné *a priori*, il est déterminé à partir d'un certain nombre de dimensions, qui correspondent à autant de mesures ou questions posées aux personnes enquêtées (directement ou indirectement). Nous parlons ici de la construction de l'espace social dû à P. Bourdieu, exposé dans la Distinction. [15] Le postulat de base de Bourdieu y était que non pas des caractères absolus, mais un ensemble de préférences relatives permettait de distinguer – d'où le titre – des profils, des personnes différentes, et donc une structure sociale. L'avancée conceptuelle tenait également dans le fait qu'il n'a pas été proposé, en dehors de cette hypothèse, et de la sélection d'un certain nombre de "préférences" potentielles à explorer, de structuration *a priori* de l'espace social (cf figure 4.3).

FIGURE 4.3 – Une représentation de l'espace social selon Bourdieu (d'après La Distinction)



Là aussi, la technique, la méthode utilisée a déterminé en partie la construction de l'espace : par cette étude, Bourdieu a popularisé l'emploi de l'analyse par correspondances multiples (ACM), qui n'est autre qu'une version pour donnée catégorielles (ou discrètes, de façon équivalente) de l'ACP. Il s'agit donc d'une technique projective, linéaire, hiérarchisant les axes d'inertie par leur pourcentage de variance expliquée. *Stricto sensu*, le nombre d'axes, donc de dimensions, à retenir n'est pas donné ; il existe différentes techniques, empiriques, afin de déterminer le nombre « suffisant » expliquant les données. L'espace social de Bourdieu est ainsi en général représenté dans le plan. L'analyse qui en a été faite s'est alors essentiellement concentrée sur l'analyse "visuelle", et le fait que les axes d'inertie conservés pour représenter l'espace fournissait un sens assez évident de distinction des personnes dans la société. Il aurait pu y être adjointe une analyse de reconnaissance de formes, afin de déterminer en toute rigueur la présence et la constitution d'une structure en groupes homogènes de personnes partageant des préférences entre eux, et se différenciant suffisamment de celles des autres groupes, plongés dans cet espace social.

Il est également notable qu'une autre technique d'approche par la forme ait été promue par un autre sociologue plutôt que tout autre spécialiste d'un domaine différent : Boudon a introduit en sociologie l'approche par la recherche des classes latentes, qui est également une technique de *clustering* pour données catégorielles, davantage basée sur les statistiques que les autres. [245]

4.4.2 Exemple en épidémiologie sociale : une typologie de recours aux soins en France

Nous avons eu l'occasion de travailler et d'examiner des données à la croisée des données médicales – ou de santé, et des données sociales et économiques, en exploitant les données d'une cohorte francilienne, la cohorte SIRS. On

parle à propos des études visant à explorer les (mal nommés) déterminants sociaux de la santé, d'épidémiologie sociale. Ces travaux ont été menés avec P. Chauvin, de l'équipe ERES (Équipe de Recherche en Épidémiologie Sociale, UMRS Inserm 1136), et ont fait l'objet d'une thèse de médecine. [246] Ces travaux sont encore actuellement en cours, dont une grande partie est soumise à publication dans un autre contexte.

Face à des données nombreuses et de nature très variées, et ne voulant pas introduire d'a priori massif ni tester des hypothèses précises, nous avons préconisé une approche du problème qui nous concernait par la forme. La question était de savoir s'il existait des comportements particuliers de la part des personnes envers le système de soins, officiel tant qu'officieux. Et dans le cas où il existerait des formes bien identifiées de recours aux soins, quels en seraient les facteurs associés ?

Nous avons ainsi étudié une vingtaine de variables qualifiant des recours différents au système de soin (cf tableau 4.4). Nous avons, sur la base de ce premier espace réduit et naïf, utilisé différents algorithmes de *clustering*, en conjonction avec une approche dite par la robustesse des formes ; cette approche permet en effet d'estimer le nombre optimal de groupes dans les données, en se basant sur la fréquence avec laquelle les personnes sont classées en couples identiques, quel que soit le brassage aléatoire des données, le nombre d'itérations de l'algorithme, les différents algorithmes, et les paramètres de ces algorithmes. Si deux personnes sont tout le temps classées ensemble, le couple est robuste. Cette propriété est examinée pour tous les couples possibles.

En l'occurrence, nous avons pu déterminer quatre grands types de recours aux soins, quatre formes distinctes, exceptionnellement robustes (cf figure 4.4). Ces types de recours ont ensuite été analysés au regard de leurs autres caractéristiques, telles que l'âge, le genre, la position socio-économique, leur

TABLE 4.4 – Les variables de recours aux soins utilisées pour la constitution d'une typologie.

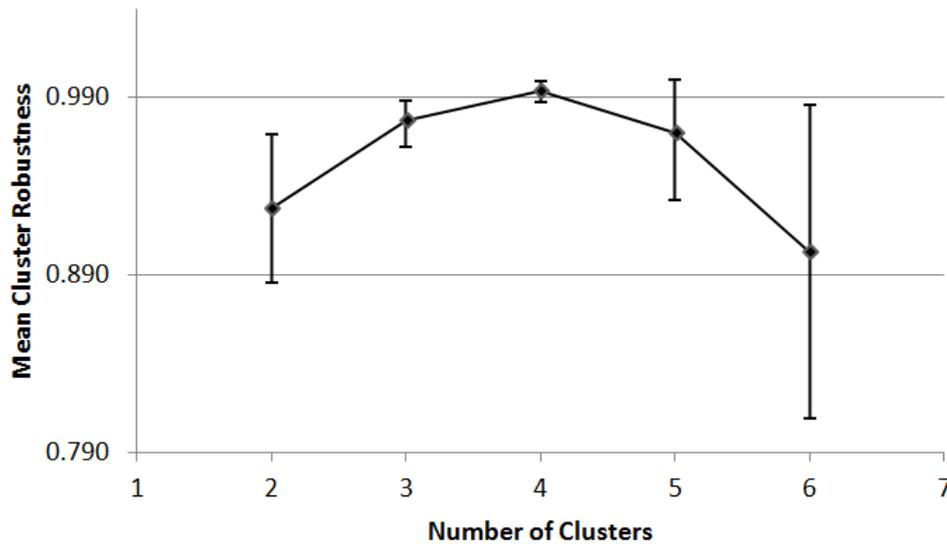
Groupes de variables	Variables	Niveaux
Soins primaires	Date de visite du dentiste	4 niveaux
	A un généraliste	Oui/non
	Spécialiste d'accès direct (AD)	Oui/non
	Fréquence spécialiste AD	5 niveaux
	Bilan centre sécurité sociale	Oui/non
Médecine alternative ou complémentaire	Avis de l'entourage	4 niveaux
	Acupuncteur, ostéopathe	Oui/non
	Autres médecins	Oui/non
Spécialistes d'accès indirect	Spécialistes d'accès indirect (AI)	Oui/non
	Fréquence spécialiste AI	5 niveaux
Lieux de consultation	Généraliste hôpital public	Oui/non
	Spécialiste hôpital public	Oui/non
	Généraliste hôpital privé	Oui/non
	Spécialiste hôpital privé	Oui/non
	Généraliste cabinet	Oui/non
	Spécialiste cabinet	Oui/non
Urgences, soins non programmés	À domicile	Oui/non
	Fréquence à domicile	5 niveaux
	Service d'urgences (SAU)	Oui/non
	Fréquence SAU	5 niveaux

couverture maladie, leur intégration sociale, etc. Nous pouvons ainsi distinguer, qualitativement et quantitativement, les comportements et profils suivants :

- le Français âgé, aisé et possédant une bonne couverture sociale, socialement isolé, inactif et en mauvaise santé. Ce type de personne présentait un recours aux soins intensif, selon toutes les modalités et lieux de recours possibles, et également le plus fréquemment ;
- le jeune homme, fréquemment d'origines étrangères, chômeur et peu aisé, possédant une mauvaise couverture maladie. Il présente néanmoins une bonne santé, et également le recours le moins diversifié et le moins fréquent (par exemple, le généraliste, une fois par an) ;
- le représentant d'une population d'âges divers, souvent d'origines étrangères et au niveau socio-économique faible, inactif et ayant une couverture sociale médiocre, plutôt socialement isolé. Il est en mauvaise santé, a davantage tendance à avoir recours aux généralistes de ville ou aux services des urgences pour des soins non programmés ;
- la jeune Française active, de haut niveau socio-économique, à la très bonne couverture maladie, bien entourée socialement. Elle est en bonne santé, a un recours fréquent aux spécialistes de cabinet et aux soins alternatifs et paramédicaux.

Des discussions plus larges des questions d'épidémiologie sociale, des déterminants sociaux de la santé, et des rapports d'échelle entre personne et environnement, prises sous un aspect plus classique ou sociologique, peuvent être notamment consultées parmi les publications de P. Chauvin et I. Parizot, pour ne citer qu'eux. [247, 248]

FIGURE 4.4 – Robustesse des groupes en fonction du nombre de groupe, pour la typologie de recours aux soins



4.4.3 Opérationnalité des classes et recherche interventionnelle

Les politiques de santé ne sauraient se satisfaire d'une simple description statique, d'un état de fait, qui est par ailleurs dangereux s'il devient institutionnalisé, considéré comme un état « naturel », immuable. La question se pose donc, une fois une typologie déterminée, de son existence, de sa stabilité, et des moyens pour en contrôler les entrées et sorties, ou en contrôler la forme.

Si l'on s'en tient à notre étude précédemment citée sur les types de recours aux soins, nous nous sommes assurés que les types, ou de manière équivalente, les formes que nous avons identifiés n'étaient *a priori* pas artificiels, et qu'ils étaient par ailleurs d'une excellente robustesse sous des contraintes de rééchantillonnages et d'algorithmes différents. La forme globale de ces types ainsi que leur existence semblent, dans les limites de cette étude, rai-

sonnablement acquises. Cependant, maintenant que nous avons obtenu ce qui peut tenir lieu de décompositions en états propres ou fondamentaux des modes de recours aux soins, qu'en faire ? Ce type d'approche, à ce stade, est en effet très descriptive, et nous ne savons d'une part, sous quelles conditions il est éventuellement possible pour une personne de passer d'un type à l'autre, d'autre part, si la structure en 4 types est une structure stable, ou bien s'il existe des facteurs en réglant la stabilité.

Ainsi, il serait intéressant, afin de dépasser le stade descriptif, de pouvoir examiner la stabilité structurelle d'une telle décomposition, ou, en supposant par ailleurs une stabilité acquise autour d'un voisinage raisonnable, quels sont les principes permettant de passer d'un état à un autre. Par extension, si la notion de type de recours aux soins est certes une notion de groupe, d'entité sociale plutôt qu'un comportement uniforme prêté à toutes les personnes appartenant à ce groupe, serait-il sensé de considérer l'état d'une personne donnée, en termes de recours aux soins, comme un état mixte de ces 4 groupes, de la même manière qu'une particule présente, à tout moment, un état mixte, indéterminé, jusque à ce qu'on le mesure ? La question se pose d'autant plus que le recours, tel qu'il est mesuré, est précisément quelque chose de révélé lorsqu'il y a recours, ou non. Dans l'intervalle, et en l'absence d'événements, quels qu'ils soient, modifiant ou incitant à un recours, la situation personnelle instantanée paraît indéterminée. En revanche, on peut trouver certainement un moment donné où l'état n'est pas ambigu, car révélé et mesuré. Dès lors, l'intérêt d'un opérateur d'évolution peut prendre du sens. Néanmoins, la construction d'un tel opérateur, dans le cas présent, nécessiterait de connaître la forme de l'Hamiltonien du système, ce que nous ignorons.

La réponse usuelle à ce type de question est en général cherchée dans les techniques de recherche dite interventionnelle. Nous retombons ici dans les

méthodes assimilées aux RCT, lesquelles sont très délicates sinon impossibles d'applications quand il s'agit de contrôler l'effet d'une politique de santé qui ne se borne pas à prescrire un médicament pour une pathologie bien identifiée. Mais ici, que faudrait-il tester comme paramètres afin de répondre à nos deux questions – stabilité et évolution ? Dans un monde simple ou simpliste, l'on trouverait un paramètre unique qui aurait un impact tel qu'il changerait radicalement la typologie observée. Il en existe certainement (une guerre, une pandémie dévastatrice...) mais de tels paramètres ne sont finalement pas pertinents en l'état. Ils ne nous renseignent pas sur la mécanique interne qui sous-tend les formes identifiées.

On peut, en guise d'illustration et d'exercice de pensée, imaginer le cas suivant : supposons que nous observions 3 groupes, 3 classes différentes, bien séparées et distinctes les unes des autres. Peut-on considérer qu'elles résultent de mécanismes séparés, assurant leur ontologie propre et leur stabilité, et que seules leurs différences relatives les caractérisent ? Ou bien ne peut-on pas imaginer, par exemple, un processus de type système dynamique, très simple, à un ou deux paramètres, tel qu'un des trois groupes soient constitués d'éléments initiaux, dont l'évolution amène à bifurcation et séparation en deux groupes séparés, que seraient nos deux autres groupes ? De fait, dans un tel cas, il existerait bien, et en basse dimension, un processus unique permettant de décrire l'existence de trois groupes séparés. Leurs conditions d'existence seraient également réglées par un système unique et simple. Cette expérience de pensée simple signifie uniquement que l'observation de groupes, aussi bien séparés, stables soient-ils sous certaines opérations, par exemple, des opérations de rééchantillonnage aléatoires, ne doivent pas nous faire conclure hâtivement à un mode d'existence plus qu'un autre. À l'extrême, ce type de raisonnement revient à nous poser la question de la place du modèle dans la production de la preuve.

4.4.4 Entre sociologie des interactions et architecture urbaine : de la taille des villes

Nous avons abordé la question des modèles allométriques, des lois d'échelle, ainsi que de la perspective sociologique, de l'existence de groupes sociaux particuliers eu égard à un sous-espace de description donné, censé représenter la mesure de comportements élémentaires.

Plus généralement, il est possible de se poser la question des relations structurantes à diverses échelles, et ce en relation d'une part avec la notion de *stockage* – la réalité des choses s'ancrant dans une forme de matérialité – et d'autre part avec la notion d'*utilité*, ou d'usage des ressources. En théorie de l'information [23], il est ainsi possible d'examiner et de mettre en relation la fréquence de l'utilisation de mots en fonction de leur rang de fréquence d'utilisation (par exemple, un lien entre le mot le plus fréquemment utilisé, disons 8000 fois, et son rang de fréquence d'utilisation, en l'occurrence, 1).

Shannon avait proposé, dans le même esprit, que le coût de représentation (d'utilisation) d'une information augmente logarithmiquement avec le nombre d'informations. À la suite, en guise de synthèse, Mandelbrot proposa que coût de stockage et coût d'utilité soient égaux. Le lien entre la fréquence d'utilisation et le rang de fréquence d'utilisation devenait alors, pour n le rang :

$$f(n).(a + bn)^c = K,$$

où K est une constante. [23]

Si stockage et usage sont ainsi liés selon une relation d'échelle, il devient possible d'examiner les relations structurantes pour tout type d'information. Récemment, Bettencourt [249, 250] s'est proposé de tester dans quelles proportions il était possible d'expliquer la taille des villes selon les interactions sociales s'y déroulant, quel que soit le type d'interaction considérée (voir

FIGURE 4.5 – Taille et organisation des villes, d'après *Bettencourt*

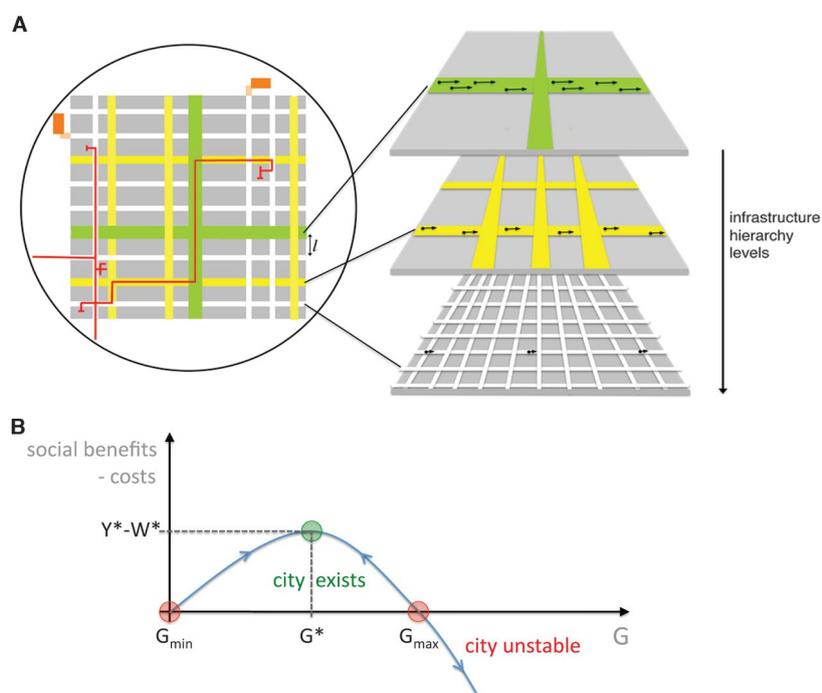


figure 4.5).

Il s'avère que nombre de villes, notamment de grandes tailles, semblent respecter une loi d'échelle mettant en lien interactions sociales et structure de la ville.

4.5 Au-delà de la pathologie : les facteurs de risque, les syndromes et autres objets de transition

Jusque l'histoire relativement récente, le panorama des pathologies était assez simple, même s'il ne manquait pas de variétés, de déclinaisons. Il était dominé par la pathologie aiguë et traumatique, laquelle pouvait être plus facilement létale que de nos jours. En l'absence de protection sociale, également, d'éducation large des populations, et d'infrastructures et conduites favorisant l'hygiène minimale, l'espérance de vie était significativement plus

courte, même il y a à peine un siècle, dans les pays occidentaux. La transition démographique, c'est-à-dire, le passage d'une forte mortalité et d'une forte natalité à un régime de faible mortalité et faible natalité, s'est doublée d'une transition épidémiologique et sanitaire. Cette transition se caractérise par une amélioration de l'hygiène, de l'alimentation et des systèmes de santé et par une mutation des causes de décès, les maladies infectieuses reculant devant les maladies chroniques ou dégénératives.

L'importance des maladies chroniques a été soulignée ces dernières années, voire décennies, l'accent étant mis et sur les coûts engendrés, et sur le changement de paradigme qu'entraîne une prise en charge qui s'étale dans le temps, si ce n'est à vie. Pourtant, il semblerait qu'en réalité, nous ayons jusque présent déployé beaucoup d'effort pour ne faire qu'effleurer une partie de la réalité, tout occupé à nos approches réductionnistes et (sur) spécialisées.

Ainsi, actuellement, il semblerait que nous tenions pour relativement acquis le nombre de pathologies connues et connaissables, la majorité des "nouvelles" pathologies trouvant généralement leur place sans problème dans la nomenclature usuelle, et se présentent comme des variantes ou sous catégories, des raffinements de classifications établies. En marge de ces territoires jugés globalement cartographiés, s'est dessiné tout un continent inférieur, environnant ou annonçant le passage de l'état sain à l'état pathologique, tout un ensemble d'entités qui par définition, ne seraient pas des pathologies, mais des états soit prédisposant à la pathologie de manière statistique, soit à considérer comme des pré-pathologies.

C'est de cette façon, concomitamment à la recherche de causalité entre événements et pathologies, que nous en sommes arrivés au concept de facteur de risque. Un facteur de risque n'est pas en soi une pathologie (il peut en être une, mais dans le sens où cette pathologie exposerait à un risque

accru d'en acquérir une autre), ni une pré-pathologie, même si certaines interprétations ou prises en charge proposées tendent à les assimiler.

Le facteur de risque d'une pathologie est une notion purement statistique. Il s'obtient lorsque l'on compare les occurrences relatives d'une pathologie donnée, dans deux groupes exposés différemment au facteur que l'on suppose comme au moins partiellement lié à la pathologie considérée. Un cas d'école est par exemple de comparer le nombre de cancers bronchopulmonaires parmi les fumeurs de tabac, et parmi les non fumeurs de tabac. Ici, le facteur d'exposition serait bien entendu le tabac, et la pathologie le cancer bronchopulmonaire. On considèrera que le tabac est un facteur de risque du cancer bronchopulmonaire si, entre autres, la fréquence des cancers est significativement (au sens statistique) plus élevée parmi les fumeurs comparativement au non fumeurs. Dans l'idéal, pour se rapprocher d'un lien de causalité, il faudrait par exemple confirmer cette différence de fréquence par une expérience contrôlée, où les deux groupes seraient égaux par ailleurs, statistiquement, en dehors de la consommation ou non de tabac.

Cette notion étant statistique, cela signifie également qu'il demeure tout à fait possible qu'une personne fume toute sa vie sans déclarer aucun cancer bronchopulmonaire, et *a contrario*, qu'un non fumeur déclare un tel cancer.

Dans l'histoire récente, en particulier dans le domaine du cardiovasculaire – un domaine qui demeure en avance sur le plan des recherches des causes et de la physiopathologie – deux exemples sont intéressants à mentionner, montrant les amalgames qui peuvent être faits, et les implications qui en découlent.

Il y a l'exemple de l'hypertension artérielle (HTA), qui, en France en particulier, est considérée à la fois comme pathologie et comme facteur de risque cardiovasculaire. Néanmoins, la majorité des cas d'HTA sont des HTA dites primaires ou primitives, entendre par là qu'elles n'ont pas de cause

bien identifiée en dehors des mécanismes que l'on considère liés à l'âge et au temps. De fait, si l'on exclut les minoritaires mais bien existants cas d'HTA secondaires – liés à une cause bien identifiée – en quoi l'HTA est-elle une pathologie plutôt qu'un syndrome, défini par des mesures de tensions artérielles systoliques et diastoliques particulières ? Ce d'autant que ce n'est pas tant l'HTA qui est associée causalement et directement au sens eclassique du terme avec des événements de type mortalité, mais davantage comme facteur de risque de pathologies menant à ces événements – tels que les accidents vasculaires cérébraux.

Cette année (2014), de nouveaux seuils de définition pour l'HTA ont été proposés, basés sur les degrés de preuve dont on dispose pour soutenir l'intérêt d'un traitement ou non de cette HTA, par exemple en fonction de l'âge, des pathologies associées (maladie rénale) ou évidemment des chiffres tensionnels. Les recommandations sont actuellement âprement discutées. [251]

De fait, récemment (2011), la Haute Autorité de Santé (HAS), en France, a décidé que l'HTA (sévère) n'était pas une pathologie mais un facteur de risque, et qu'à ce titre, elle n'avait pas à figurer parmi la liste des Affections Longue Durée (ALD) ouvrant droit à une prise en charge à 100 % des frais médicaux par l'assurance maladie.

Un second exemple est la place du cholestérol dans le paysage des pathologies et facteurs de risque cardiovasculaires, encore une fois. Jusque présent, il semblait acquis que le cholestérol devait figurer en bonne place des facteurs de risque cardiovasculaires – c'est-à-dire qu'un taux élevé de cholestérol dit LDL (*Low Density Lipoproteins*) est supposé être lié à un taux accru d'infarctus du myocarde et d'accidents vasculaires cérébraux (AVC). La réalité n'a jamais, semble-t-il été celle-là, et le rôle causal direct du LDL dans l'infarctus du myocarde est très discutable, au mieux. Traiter les personnes présentant ces taux plus élevés seraient donc inutile en soi, puisque sans action sur les infarctus (ce qui n'est cependant pas acquis pour les AVC). Néanmoins, ce

taux est une information pertinente. En effet, et comme nous avons pu le souligner précédemment, s'il n'a pas de rôle causal dans la survenue d'événements indésirables, le LDL se révèle être dans la majorité des modèles un très bon prédicteur de survenue (statistique) de ces événements. Nous voyons donc ici toute l'ambiguïté que peut receler la notion de facteur de risque, telle que nous l'utilisons actuellement. [175, 102]

Un peu différemment, et plus anciennement également, nous assistons à une résurgence et une multiplication des syndromes. Un syndrome est, par définition, un ensemble de signes, qui, s'il a un sens, doit être suffisamment fréquemment observables en tant que tel, donc présenter à la fois une forme de stabilité et de reproductibilité – à tout le moins, des occurrences répétées et observables. Les syndromes peuvent être des ensembles de signes mixtes, aussi bien au sens de signes cliniques (fièvre, perte de poids), que de signes biologiques (anémie, augmentation des leucocytes), ou encore radiologiques – la liste n'est limitée que par les modalités d'observation mobilisables. Les syndromes demeurent conceptuellement bloqués entre la pathologie – théoriquement, un syndrome peut coïncider avec diverses pathologies, il appartient à un niveau de diagnostic inférieur, ou bien encore un syndrome ne se traite pas comme une pathologie, dans le sens où généralement, le syndrome ne permet pas, au contraire de la pathologie, d'y associer une cause biunivoque – et le facteur de risque : on peut ainsi penser au syndrome métabolique, qui serait en soi un facteur de risque cardiovasculaire, sans être une pathologie.

Les notions de facteurs de risque, comme de syndromes et de prépathologies rendent finalement compte d'un vide tant théorique que pratique et d'une tentative de combler ce vide : en supposant une existence catégorielle des pathologies acquise et stable, la recherche d'une causalité, et plus largement, des chemins menant de l'état "sain" ou "normal" à l'état pathologique se heurte à un manque cruel d'outils. L'avènement des statistiques en bio-

médecine a dirigé cette recherche naturellement vers une description type facteur de risque : on étudie des groupes, l'effet indépendant d'un ou plusieurs facteurs sur l'occurrence d'une pathologie, et l'on détermine les risques relatifs de déclarer une pathologie selon une exposition à un événement. La dynamique proposée est donc très fruste, de nature statistique, axée sur des concepts plutôt péjoratifs (le risque demeure une notion sous-entendant la crainte et la non maîtrise d'une situation), et consiste en un saut catégoriel entre sain et pathologique, saut qui est aléatoire. Les syndromes apparaissent comme une manière de chercher à identifier des groupes homogènes de situations qui ne soient ni nécessairement pathologiques, ni synonymes ou définitoires d'une normalité. Ils s'affranchissent parfois des contraintes d'une causalité unique linéaire et ponctuelle, en présentant des objets réalistes et complexes, sans qu'il soit actuellement possible d'en dire beaucoup plus à leur sujet. La seule histoire du SIDA en tant que syndrome/pathologie suffit à montrer comment ces tâtonnements de paradigmes peuvent être douloureux.

Quant aux prépathologies, leur intérêt nous paraît quasi nul en termes scientifiques : elles ne jouent pas de rôle de réelle redéfinition, font office d'ajustement de seuils parfois douteux, et les événements importants qui pourraient leur être associés (mortalité, morbidité, incapacités fonctionnelles...) sont mal étudiés.

Nous voyons donc ici comment le champs de la pathologie s'est sensiblement métamorphosé, sans s'être encore stabilisé, entre tradition anatomo-clinique pure et définition large de la santé, comme état "ne pouvant se définir que comme l'absence de maladie" (définition OMS).

4.5.1 L'exemple des maladies chroniques

Le premier malaise concernant la transition épidémiologique et la survenue au premier plan de la pathologie chronique arrive dès la définition même de

pathologie chronique : il s'agit, historiquement, d'une définition, voire pluri-définition, par défaut. Est pathologie chronique ce qui n'est pas pathologie aiguë, aigu étant entendu comme court et transitoire, avec en général guérison avec ou sans intervention visant à éradiquer ou diminuer les causes de la pathologie. Cependant, le caractère aigu, et partant le caractère chronique, s'il possède bien cette dimension temporelle essentielle, n'est pas le même selon la spécialité à laquelle l'on s'intéressera, pour la simple raison que les pathologies aiguës pulmonaires ne durent pas nécessairement la même durée qu'une pathologie aiguë ORL, par exemple.

Il est toutefois possible de considérer qu'une pathologie est dite chronique à partir du moment où elle dépasse au minimum un mois de durée, plus fréquemment deux ou trois ; l'OMS va même jusqu'à proposer de qualifier de chronique une affection d'au moins six mois, entraînant par ailleurs un retentissement important sur la vie de la personne concernée. Ainsi, seront classiquement cataloguées chroniques les pathologies telles que certaines maladies infectieuses (VIH, tuberculose, hépatites chroniques, paludisme), certaines pathologies cardiaques comme l'insuffisance cardiaque, pulmonaires, comme la broncho-pneumopathie obstructive chronique ou l'asthme, endocriniennes comme les diabètes sucrés, ou encore la majeure partie des cancers. A ce stade, on comprend bien qu'il peut y avoir un sens à regrouper ces pathologies sous la même catégorie générique, mais on identifie rapidement également une forte hétérogénéité associée inter pathologies. Certaines entretiennent très probablement ou de manière avérée des relations particulières, notamment physiopathologiques (maladies cardiovasculaires et diabète, diabète et surpoids...), tandis que d'autres sont connues pour s'associer plus fréquemment, sans que leur substrat commun relève nécessairement d'un mécanisme physiopathologique élémentaire partagé ; on sait notamment que les personnes souffrant de pathologie mentale ont plus souvent une autre pathologie chronique somatique associée que ceux n'en souffrant pas. Les

explications courantes en sont par exemple, la prise de médicaments favorisant l'apparition d'autres troubles, ou bien des comportements de santé conduisant à davantage de pathologies. [252]

Le caractère prolongé ou permanent, répété, ou bien le caractère grave, nécessitant des traitements ou des suivis coûteux et lourds, font que les personnes atteintes de pathologies chroniques, et *a fortiori* multiples, doivent supporter un poids financier important. Ce poids peut s'avérer d'autant plus difficile à assumer que ces pathologies peuvent être discriminantes, ou impliquent une baisse du travail, donc de la rémunération.

De telle sorte qu'une autre vue, donc une autre forme de la polypathologie chronique émerge du point de vue non de la personne malade, mais du système de soins qui doit la traiter, et du système de financement des soins. En France, par exemple, la pathologie chronique est définie par le système de soins par ce que l'on nomme les ALD (Affections Longue Durée), au nombre de trente, auxquelles s'ajoutent deux autres un peu à part. La polypathologie chronique, elle, n'est pas réellement définie au niveau du système de prise en charge. C'est cependant un défi important, sachant que les ALD pèsent un poids prédominant dans les dépenses de l'assurance maladie.[253, 254, 255]

4.5.2 La pathologie ne saurait plus se baser uniquement sur une donnée uniquement physiologique

Nous avons eu l'occasion d'entreprendre des recherches aux côtés de R. Gagnayre, au sein de son Laboratoire de Pédagogie de la Santé (Paris 13, SMBH, EA3412), et notamment d'y mener une revue systématique sur le concept de polypathologie chronique, et des facteurs lui étant associés (mortalité, déterminants, incapacités fonctionnelles, qualité de vie...). Ces travaux sont également en cours de publication. Nous venons de donner quelques éléments démontrant la nécessité d'approcher la polypathologie chronique, qui recouvre intuitivement une réalité certes multiforme, mais suffisamment pré-

occupante pour que divers acteurs s'y attèlent, autrement que par un point de vue unique, ou à tout le moins selon une seule dimension d'approche, que serait par exemple, l'approche physiopathologique unique.

Cette remarque est d'autant plus importante, que les frontières entre facteurs de risque, syndromes et pathologies deviennent chaque jour plus floues. Une approche naïve mais potentiellement efficace en première approximation, serait de définir la polypathologie chronique comme la cooccurrence d'au moins deux pathologies chroniques, à recenser parmi une liste de pathologies candidates. Dans les faits, cette approche a été jusque ici la plus répandue, avec des succès variables. En effet, tout le monde n'est pas d'accord sur la liste des pathologies candidates ; notamment, un certain nombre d'études éliminent de cette liste les pathologies ostéo-articulaires, dont on sait qu'elles sont très répandues, et non sans conséquences sur l'ensemble de la population. [256]

Par ailleurs, certaines autres études ont pris le parti d'incorporer ce qui peut être considéré comme des facteurs de risques, ou des syndromes dans cette liste, voire même se focaliser sur une association pathologie plus facteur de risque (par exemple : dyslipidémie ou hypertension artérielle), ou pathologie plus syndrome (par exemple, syndrome métabolique). Existe également tout le pan des études ciblant une diade particulière, du type pathologie chronique somatique (diabète ou asthme) et pathologie chronique psychiatrique (dépression ou anxiété). Au mieux, ces études s'intéressent au retentissement bidirectionnel de l'une sur l'autre, mais en l'état actuel des connaissances, en dehors de constater une prévalence accrue de ces associations particulières, nous ne disposons pas de d'hypothèse physiopathologique forte. [257]

Il est aussi important de décrire les rapports que peuvent entretenir ce que l'on appelle depuis plus longue date les comorbidités et la polypathologie chronique. Lorsque nous parlons désormais de polypathologie chronique,

nous faisons l'hypothèse d'une entité à part, indépendante, éventuellement englobant la notion de comorbidité, qui selon nous, devrait être amenée à disparaître.

La comorbidité, par définition, se rapporte à une morbidité qui serait principale, et à laquelle s'ajouteraient d'autres pathologies au sens large. [258, 259, 260] Toutefois, comment définir le caractère principal de cette pathologie ? Dans les faits, elle est soit celle qui semble la plus grave, voire la plus urgente à traiter, ou bien plus simplement en dehors de toute grande urgence, comme la pathologie à laquelle un spécialiste, à un temps t et un lieu donné, va s'intéresser. Ainsi, le diabétique cardiopathe sera vu comme tel par l'endocrinologue ou le diabétologue – le patient présente avant tout un diabète, et des problèmes cardiovasculaires comme comorbidités – et vu comme un cardiopathe diabétique par le cardiologue – pour qui le patient est un cardiopathe, présentant comme comorbidité, voire comme facteur de risque, un diabète. Au vu du système de soins, il est vu comme un diabétique d'un œil, et comme un cardiopathe de l'autre, ALD diabète, et ALD cardiopathie. Cette vision d'un cas simple et courant illustre bien l'approche réductionniste et cloisonnée actuelle.

Une autre question, mettant possiblement en lumière une différence conceptuelle entre comorbidité et polypathologie, elle-même récurrente de nos approches par la forme, est celle de la dimension temporelle – vu comme flèche temporelle, donc succession d'événements, et non comme mesure de durée, définitoire du caractère chronique. La comorbidité peut prendre son sens si l'on considère que la pathologie index est la première à avoir été déclarée. De même, des profils de polypathologies pourraient se baser sur des trajectoires distinctes et récurrentes, plutôt que sur des associations tout temps confondu. Cependant, il faudra se garder, le cas échéant, d'y voir la preuve indéniable et trompeuse d'un caractère physiopathologique, causal, fausse-

ment évident et représenté par des successions temporelles distinctes.

4.5.3 Polypathologies chroniques : quelles dimensions ?

Ainsi, pour se doter d'une définition objective de la polypathologie chronique, il est nécessaire de se poser la double question : quel(s) point(s) de vue adopter pour qualifier l'objet – question de la classe d'observateurs – et quelles dimensions retenir pour l'espace de représentation ? Et de fait, quel instrument de mesure pour la définir et la mesurer ? Le fait de ne pas disposer de substrat physiopathologique évident et commun, c'est-à-dire, finalement, d'hypothèse causale plus ou moins simple mais dont on dériverait chacune des entités possibles de polypathologie chronique, est en soi une chance d'aborder le problème de la causalité comme mécanisme de définition d'un objet. Si l'on tient à la relation cause – effet, doit-on nécessairement sous-entendre qu'il existe un délai de propagation, un temps irréductible structurant ces deux éléments, ou bien peut-on considérer que le simple lien logique entre les deux peut suffire ? Dit d'une manière un peu différente mais proche, la dimension temporelle linéaire est-elle seule garante ou seule structure possible pour expliciter la relation causale, ou bien existe-t-il, tout simplement, une causalité formelle, une causalité vue comme structuration de la forme, sans mention explicite ou nécessaire du temps linéaire commun ?

Voyons les deux approches privilégiées actuellement dans le domaine. D'une part, nous avons pu identifier quatre grands procédés de mesure, et donc de définition, utilisés pour caractériser la polypathologie chronique ; sachant que pour chaque catégorie, il existe des variantes quantitatives importantes. D'autre part, de manière peut-être plus fondamentale dans son abord initial, beaucoup d'études se sont davantage intéressées aux conséquences supposées de la polypathologie : donc, à considérer des définitions qui soient construites en lien avec une mesure bien particulière, qui n'est pas directement la mesure

de la polypathologie.

La revue de littérature à laquelle nous nous sommes livré à mis en évidence 4 grands abords pour caractériser la polypathologie chronique (que nous nommerons ici MCC - *Multiple Chronic Conditions*). Nous les présentons ici brièvement :

- le décompte et l'addition pure et simple de pathologies chroniques, prises dans une liste restreinte et arbitraire de pathologies candidates ; de fait, cette approche suit un modèle dans la tradition directe du réductionnisme pur et du modèle biomédical. Il est à noter que d'une étude à l'autre, les listes ne sont pas les mêmes, ignorant assez fréquemment des pathologies à forte prévalence, comme les affections rhumatologiques ; [252]
- le dénombrement des diades et triades les plus fréquentes, ce qui correspond davantage à une approche réaliste-holistique ; on se base sur une analyse des co-occurrences les plus représentées, donc existantes dans les faits ; [261]
- les approches basées sur la construction d'un index, lequel est bâti selon un objectif de meilleure corrélation possible à un type d'événements ou de mesure – par exemple, une forte consommation médicamenteuse durant l'année passée, ou prédisant au mieux la consommation médicamenteuse dans l'année qui suit ; [262]
- les approches basées sur l'identification de groupes homogènes de personnes, sur la base de leurs pathologies et de caractéristiques associées ; ce sont donc des approches utilisant des techniques de *clustering*, ou d'ACP. Celles-ci sont largement minoritaires. [263]

Outre le fait que les approches différentes que nous venons de citer relèvent de choix parfois radicalement opposés, du moins dans leur démarche, le principal problème engendré par cette multiplicité est la non reproductibi-

TABLE 4.5 – Prévalences de la polypathologie chronique en fonction des définitions.

Prévalence (%)	Mesure	Population	Référence
14	Addition	99 997 adultes anglais (généralistes)	Salisbury 2011
56	Index (ACG)	99 997 adultes anglais (généralistes)	Salisbury 2011
23.2	Addition	1 751 841 Écossais (généralistes)	Barnett 2012
71.2	Addition	28 millions d'Américains hospitalisés	Steiner 2013
24.5	Addition	198 670 Espagnols (généralistes)	Garcia-Olmos 2012
29.7	Addition	60 857 Néerlandais (généralistes)	v.d. Akker 1998
37.1	Index (CIRS)	3 398 Australiens (généralistes)	Britt 2008
19.3	Addition	13 806 Hollandais (généralistes)	Westert 2001

lité des études et des résultats obtenus. Formulé autrement, si l'on se doute qu'il existe quelques similarités partagées par les différentes approches, il n'en demeure pas moins que les mesures effectuées sur la base de ces définitions varient drastiquement selon les cas. Il est inutile de chercher à comparer des mesures subtiles, l'examen pur et simple des prévalences obtenues dans des populations *a priori* relativement comparables présentent des valeurs pouvant être d'un rapport 10 entre elles. Définir un objet d'étude dans ces conditions paraît illusoire (cf tableau 4.5).

Tout aussi troublant, il est étonnant de constater que si la question de savoir quelles pourraient être les causes communes à certains profils de MCC a été posée finalement précocément, aucune n'étude ne semble s'en préoccuper frontalement, et plus étonnant encore, certaines s'orientent à l'opposé : on cherche à définir la meilleure mesure de la MCC selon les conséquences qu'on peut lui attribuer. En outre, les conséquences les plus fréquemment étudiées demeurent des mesures de consommation ou de coût de prise en charge, ce qui nous éloigne assez des préoccupations physiopathologiques.

Ainsi que nous avons parlé précédemment des objets transitionnels en santé, comme les facteurs de risque ou les syndromes, il est frappant de constater que dans le cadre de la polypathologie, un groupe de médecins,

constitué autour de cette question spécifique – la polypathologie chronique – a cherché à analyser les données de la littérature portant sur les définitions de la MCC, ce dans le but de proposer une définition synthétique et de validité large. Leur conclusion tient dans une définition précisément mixte, à savoir que la MCC serait, par exemple, définie comme la présence d’au moins une pathologie chronique et d’un facteur de risque. Nous sommes en plein dans une recherche de définition d’objet, où les frontières se brouillent entre pathologie au substrat organique bien isolé, et notion d’association statistique, à peine probabiliste, à la survenue d’un événement indésirable ou à un groupe diagnostique particulier.

Le mieux que nous semblions être en mesure d’accomplir en ce qui concerne ici la MCC reste l’identification de motifs récurrents, et stables au travers des populations et des études. Nous approchons alors une description qui tienne certes plus de la description syndromique que de la définition physiopathologique, mais c’est un premier pas. L’étude des facteurs associés à ces motifs, dépassant la simple interprétation biochimique inévitable, doit fournir les éléments suivants à la compréhension de ces motifs. Quelles sont les relations qu’entretiennent ces pathologies jusque lors considérées comme isolées ? Selon quelle dynamique, quelle synergie ? Un premier pas vient d’être franchi, pouvant sensibiliser davantage les cliniciens et chercheurs, éventuellement décideurs publiques, par la publication d’une revue systématique des études portant sur l’identification, par méthodes de *clustering*, des motifs récurrents en MCC. Ces études sont encore trop peu nombreuses, et de qualité méthodologique hétérogène, ou faible. L’effort est à poursuivre. [264]

4.6 Santé mentale

Nous avons esquissé en introduction, un bref historique des sciences prenant la psyché ou la santé mentale comme objet d’étude, ainsi que les difficultés intrinsèques qu’elles rencontrent depuis leur naissance. Nous prenons ici

le temps de décrire plus précisément les enjeux de ce que nous appellerons de façon générique la psychologie, prise en tant que science. Nous rappellerons avant tout que la genèse de la psychologie s'est faite en grande partie culturellement, historiquement et même géographiquement dans un contexte commun aux révolutions physiques que sont la relativité et la mécanique quantique, puis l'informatique et la théorie de l'information. Après avoir souligné les insuffisances des classifications et cadres nosologiques actuels, ainsi que des perspectives dominantes, nous exposerons comment aborder différemment, théoriquement et pragmatiquement, le problème de la caractérisation des personnes dans un espace psychologique. Enfin, nous terminerons par un exemple théorique plus concret et plus ciblé : la modélisation d'un trouble bipolaire au sens large à partir d'un système dynamique simple.

4.6.1 Psychanalyse et physique, une histoire commune : Mach, Freud et Einstein

Il est peut-être assez peu connu ou reconnu que la psychanalyse, en tant que domaine autonome, qu'on lui accorde le statut de science ou non, possède des racines communes avec le développement d'autres sciences modernes, telles que la mécanique quantique, la relativité générale ou l'informatique. L'un des dénominateurs communs à la poignées de personnes ayant contribué au premier plan à l'érection de ces sciences est certainement Ernst Mach, grand physicien de la fin du XIXe, père spirituel de beaucoup de scientifiques de l'époque, dont Einstein. Mach était physicien, mathématicien, et a travaillé avec des physiologistes et des médecins, comme Joseph Breuer, notamment parce qu'il s'intéressait aux sensations et à leurs liens à la réalité physique. Le principe de Mach est resté comme l'idée d'un relativisme des masses inertielles, s'inscrivant contre l'idée d'une action absolue et isolable du reste du monde.

4.6.2 Bref historique de l'objet et de ses classifications en psychiatrie

Si l'on considère que l'objet de la psychologie, plus que celui de la psychanalyse, est le fonctionnement cognitif et relationnel d'une personne, et partant, dans une perspective de séparation du monde en normal et pathologique, alors la plus grande partie du travail des psychologues – au sens large là aussi : psychologues, psychiatres... – a bel et bien été de catégoriser les personnes en entités pathologiques. La classification traditionnelle séparait classiquement la pathologie en névroses et psychoses, classification qui a laissé sa place à celle plus internationale reconnaissant notamment davantage les troubles de l'humeur et les troubles psychotiques (troubles de la relation au réel). Il existe à l'heure actuelle deux classifications internationales concurrentes, non totalement superposables : la CIM, à sa version 10 (classification internationale des maladies, non exclusivement dédiée à la pathologie mentale) que l'on doit à l'OMS, et le DSM, tout juste arrivé à sa version 5 (manuel diagnostique et statistique des troubles mentaux) que l'on doit à l'APA (*American Psychiatric Association*). Si l'on a déjà souligné la difficulté de distinguer normal et pathologique, la norme communément admise en biomédecine étant plus largement celle de la norme biologique, qu'en est-il de la psychologie, où les entités nosologiques sont essentiellement portées par des symptômes, des signes subjectifs et non des mesures biologiques ? L'histoire du DSM est à ce titre riche en enseignement, lui qui caractérisait il n'y a pas si longtemps encore l'homosexualité comme pathologique (jusqu'en 1973, DSM II), ou qui, le temps aidant, s'enrichit tant et plus de nouveaux diagnostics. Le "statistique" de DSM peut induire en erreur quant à son processus de production.

Les catégories diagnostiques sont des catégories conçues comme étanches, et même s'il n'est pas formellement impossible d'être schizophrène et dé-

primé, cet état sera caractérisé comme un état de superposition, une addition de deux conditions bien définies. Les diagnostics sont fondés sur des critères diagnostiques, souvent répertoriés en majeurs et mineurs : il faut par exemple deux critères majeurs parmi quatre, et deux mineurs parmi une dizaine pour être catégorisé dans tel ou tel groupe. Le "statistique" ne va pas plus loin que refléter le mode de production des preuves étayant les catégories : les catégories s'appuient sur des études statistiques. Néanmoins, comme nous avons pu le souligner précédemment, ces études ne correspondent qu'à une étape analytique, voire auto-confirmatoire, basée sur des groupes déterminés *a priori* ; la synthèse des preuves conduisant à la fixation des critères diagnostiques se fait par avis d'experts, réunis en groupes, chaque groupe ne s'occupant que d'une entité pré-déterminée. Nous sommes donc rapidement confrontés à un problème de logique circulaire dont la survivance peut paraître un peu étrange. Quel expert aurait intérêt à voir disparaître son objet d'expertise ?

Chaque nouvelle édition du DSM amène son lot de protestations, de critiques plus ou moins virulentes quant à la pertinence peu avérée des entités nosologiques, et de plus en plus fréquemment son coordinateur à écrire tribunes sur tribunes pour dénoncer l'entreprise. [265] Pourtant, rien ne change, ou presque. Parmi les faits les plus frappants, il est de considérer les quelques études s'intéressant à la stabilité temporelle des catégories diagnostiques pour une personne donnée : concernant les troubles bipolaires (dépression et manie) et la schizophrénie, il a été montré que sur une dizaine d'années, près de la moitié des personnes diagnostiquées dans une des catégories passaient dans l'autre. Les psychologues ont proposé dès lors deux possibilités pour améliorer la situation de leur science : abandonner les signes subjectifs, et chercher les substrats les plus somatiques possibles à la pathologie mentale, à commencer par les gènes et l'architecture, morphologique ou fonctionnelle, cérébrale.

4.6.3 La tentation de la génétique et des neurosciences : insuffisances et impasses

Il convient évidemment de ne pas pousser à la caricature, et de ne pas d'une part, mettre tous les psychologues dans la balance pro génétique inconditionnelle, ni d'autre part considérer que les pro génétiques pensent uniformément que seul le gène explique la pathologie – même s'il resterait à définir ce qui est gène normal et gène pathologique – ils parlent en effet en général de paradigme gène-environnement, pour signifier que le gène forme la personne, plus ou moins sous l'influence d'interactions avec des événements non génétiques. Néanmoins, un certain nombre estime que l'examen exhaustif du génome doit pouvoir fournir de manière quasi bi univoque les clés de la pathologie. Le premier mouvement consistant à chercher dans le génome le support aux formes nosologiques pré définies a échoué, ce qui ne devrait pourtant que surprendre qu'à moitié : si les classes existantes sont mal définies et non exclusives, comment espérer qu'il existe une réalité matérielle sous-jacente qui les départage parfaitement ? Les études récentes ont montré que pour les cinq principaux troubles (bipolaires, psychotiques, autistiques, retard mental et trouble hyperkinétique avec déficit de l'attention), il n'existait aucun cluster bien séparé des autres pour lesquels on établirait une correspondance avec ces entités. Pire, chaque entité partage avec les autres les mêmes gènes candidats. Notons bien que ces études sont exécutées sur ce qu'on appelle des études *genome-wide* (ou études GWAS, pour *genome wide associations study*) : on y examine l'intégralité du génome à la recherche d'associations entre une pathologie, un trait, et un gène. Une limite connexe à celle-ci tiendrait dans le fait qu'il semble qu'à ce jour, aucune étude génétique de ce genre n'a pu être reproduite dans ses résultats. Nous avons vu en chapitre 3 quelques raisons pour lesquelles il est probable qu'avec les outils classiques, voire même de façon intrinsèque, il n'est pas possible d'obtenir des résultats cohérents (cf malédiction de la dimension, et phénomène de l'espace vide).

Les psychologues, après un divorce d'avec les neurologues – il existait, et en existe encore quelques uns, des neuropsychiatres, avant que neurologues et psychiatres soient séparés – en reviennent aux neurosciences, en particulier par le biais de l'imagerie, qu'elle soit morphologique ou fonctionnelle, avec une préférence semble-t-il pour cette dernière modalité. L'idée est apparemment assez simple : on cherche à étudier par IRM fonctionnelle par exemple, l'activité cérébrale et ses variations selon la catégorie diagnostique à laquelle l'on est censé appartenir. Des études récentes s'intéressent même à chercher dans ces images les marqueurs infra cliniques, précoces, de pathologie mentale. Avec cette limitation *a priori* indépassable qu'il n'existe pas de juge, de *gold standard* pour juger du bon classement ou non par les algorithmes scrutant les images fonctionnelles : puisque l'on se trompe dans près de 50% des cas de diagnostic, quand celui-ci est par ailleurs d'expression patente et non latente, quelle valeur attribuer à ce type d'approche ? C'est enfin sans compter sur le traitement, en général statistique, réservé aux effectifs de ces études : la majorité des conclusions s'appuient sur des séries, de vingt, trente patients pour les bonnes études. Là aussi, la majorité reste impossible à reproduire dans leurs résultats. [96] A ce titre, il est intéressant de noter que très récemment, parmi les techniques reconnues par les psychiatres comme prometteuses pour ouvrir vers une plus grande scientificité, sont mentionnées les techniques de *clustering*. [77] Sauf que, malheureusement, il semblerait qu'ils aient préféré leur utilisation en tant que techniques d'apprentissage, et l'on en revient au problème du gold standard impossible, et non en tant que techniques de reconnaissance de forme / découverte de groupes homogènes.

Les versions du DSM se suivent, se ressemblent assez largement, et sont depuis les dernières itérations, sans cesse critiquées, de plus en plus avant même leur publication, et fait en passe de devenir une constante, désavouées par leur coordinateur principal – ce qui pose question.

Par son fonctionnement, mais également par une certaine inertie de la communauté – un retard? – des spécialistes en santé mentale, le DSM valorise, ou promeut de manière exclusive, un modèle catégoriel, étanche, des pathologies mentales. Il existe pourtant quelques publications, avant le DSM 5, ayant tenté une approche dite "dimensionnelle", c'est-à-dire, continue dans son approche des symptômes, et au-delà, éventuellement des tableaux nosologiques. Malgré ces quelques tentatives, intéressantes, peu d'écho. [115, 266, 267, 268]

Toutefois, deux mouvements récents méritent notre attention. D'une part, la sortie du DSM 5 a vu la publication par les ténors de la psychiatrie, voire les promoteurs de la nouvelle mouture, d'articles suggérant que l'approche catégorielle n'est peut-être pas idéale, et qu'une approche dimensionnelle serait plus adaptée, notamment pour les cinq grands tableaux nosologiques (dont le retard mental, l'autisme, les troubles psychotiques et les troubles bipolaires). Cependant, il est remarquable de constater la pauvreté conceptuelle de ce qui a été proposé : l'abord dimensionnel tiendrait à représenter les cinq grands types de troubles le long d'un axe continu – ce qui, de la sorte, ordonne les pathologies les unes par rapport aux autres. Dans un tel cas, alors qu'une approche catégorielle permettait encore de passer d'une catégorie à une autre par un "saut" diagnostique, que penser d'une représentation unidimensionnelle ? Il ne serait ainsi pas possible de présenter à la fois un retard mental et un trouble psychotique ? Comment placer les personnes dans un tel espace, si le mot a encore un sens ici ?

L'autre mouvement s'avère plus étonnant, et doit être surveillé. Il s'est amorcé dans une discrétion plutôt singulière, alors que l'organisme à son origine n'a rien de confidentiel. En effet, depuis quelques années, le NIMH, *National Institute for Mental Health*, l'organisme états-unien pour la recherche en santé mentale, propose la construction de ce qui est nommé RDoC, pour *Research Domain Criteria*. L'approche se veut matricielle, et basée sur les

observations, regroupées en systèmes et sous-systèmes, (presque) indépendamment des entités diagnostiques actuelles. Elle entend de la sorte explorer les fonctionnements, notamment au niveau des circuits neuronaux, mais également biomoléculaires, en lien avec un comportement normal ou "anormal", différents intermédiaires étant possibles entre les deux. À ce jour, cette initiative qui tranche avec le système DSM, se pose comme la seule alternative un peu conséquente et crédible, en dépit du fait qu'il n'apparaît pas évident, en pratique, de déterminer et la forme et ce qui en sortira.

À cette initiative, nous devons ajouter et associer les travaux, les idées de T. Insel, qui souligne régulièrement les problèmes inhérents au système DSM, et promeut davantage une approche basée sur les faits et les observations. [269, 270, 81, 271, 272, 273]

4.6.4 Un cas d'école de *big data* : approche géométrique, intégrative et dynamique basée sur l'observation

Tant que l'on ne se donnera pas les moyens de redéfinir, d'explorer la pertinence avec le moins d'a priori possible des entités nosologiques actuelles, il semble que nous soyons condamnés à une impasse conceptuelle et pratique. Abandonner les définitions subjectives paraît en outre critiquable, dans le sens où même en pathologie somatique, le symptôme exprimé a sa place, et peut servir d'entrée diagnostique. Chercher des signes objectifs ne paraît pas dénué d'intérêt, et l'on peut penser qu'ils apporteraient une stabilité accrue des catégories. Cependant, et rejoignant ici partiellement l'approche dimensionnelle et non catégorielle, il ne nous paraît pas nécessaire de postuler *a priori*, l'existence d'entités bien séparées. Rien ne dit qu'après examen rigoureux des données, nous ne retombions pas *in fine* sur des catégories, aux frontières sensiblement différentes que celles existantes actuellement – mais cette hypothèse n'est pas nécessaire et est même un obstacle à la conduite correcte de la construction d'un cadre nosologique plus réaliste. Nous donnons dans

TABLE 4.6 – Méthode étape par étape de construction d’un paysage psychologique.

- | |
|---|
| <ol style="list-style-type: none"> 1. Champs et niveaux des questions et mesures retenues. Création du protocole : questionnaires, examens complémentaires éventuels (imagerie, génétique, dosages sanguins) 2. Échantillonnage de la population générale 3. Reconstruction de l’espace minimal de description à partir de l’échantillonnage : détermination de la dimension intrinsèque des données (estimateurs fractals), et préservation de la topologie de l’espace (techniques NLDR) 4. Analyses intermédiaires des dimensions de l’espace : facteurs principaux associés aux dimensions de l’espace, notamment en vue de la production d’un questionnaire général minimal (optique plus orientée soins primaires et tests sensibles) 5. Partitionnement de l’espace minimal de description, par des méthodes de <i>clustering</i> associées à des techniques assurant la stabilité du partitionnement. Les régions retenues sont à analyser et comparer aux catégories classiques 6. Recherche des géodésiques (distances) entre régions, et analyses des paramètres critiques entraînant des changements brusques de direction (stabilité structurelle) 7. Recherche alternative des chemins menant d’une région à une autre, par réseaux de causalité (réseaux bayésiens). Possibilité d’identifier des critères critiques menant d’une région à une autre 8. Par régions identifiées : recherche d’attracteurs au sens des systèmes dynamiques, et caractérisation de la forme de ces attracteurs 9. Par régions identifiées : possibilité d’analyse locale, et de détermination de la dimension intrinsèque locale (possibilité de questionnaires minimaux spécifiques de la région d’intérêt : optique de tests diagnostiques spécifiques) |
|---|

le tableau 4.6 une méthode pas à pas permettant de tester l’approche que nous préconisons.

Nous avons souligné l’importance de la dimension sociale dans la définition bien articulée des pathologies, notamment celles qui s’inscrivent dans le temps, et réciproquement, l’importance de la dimension psychologique, individuelle dans la constitution de groupes sociaux faisant sens. Nous partons également de catégories ignorant très largement tout signe objectif, qu’il soit issu de la biologie ou de l’imagerie.

Nous proposons de repartir de la base, de fonder une éventuelle classification, si c’est ce qui en émerge, non pas basée exclusivement sur des syndromes de signes subjectifs, mais sur un volume de données multidimension-

nel, davantage structuré selon le modèle biopsychosocial que biomédical pur. L'analyse adaptée et rigoureuse de ces volumes de données doit permettre de constituer une cartographie basée sur les faits, tant objectifs que subjectifs, faits eux-mêmes recueillis en population générale, et non exclusivement sur des populations ciblées et sélectionnées par leur supposée classe diagnostique.

Ces données doivent aussi bien concerner la personne directement, que son écologie directe – son entourage, sa famille, son milieu professionnel – et son milieu plus macroscopique : zone d'habitat, culture ou niveau économique. Concernant la personne, les données doivent concerner ses états rapportés – ses symptômes tels qu'il les vit –, ramenés à une base la plus élémentaire et athéorique possible, et caractérisant son fonctionnement ainsi que ses états intériorisés (tristesse, joie, colère, détachement, anxiété...). Le recueil des données devra se faire d'au moins deux manières complémentaires : par administration d'un auto questionnaire (la personne seule face à elle-même et ses ressentis), et par entretien (la personne en situation d'inter subjectivité).

Les données objectives personnelles peuvent concerner des variables de type biologique, voire moléculaire, ainsi que des données fonctionnelles, quel que soit leur mode d'acquisition. Notons que parmi les modulateurs de comportement, l'hypothèse infectieuse est envisagée : c'est par exemple le cas de *Toxoplasma Gondii*, parasite responsable de la toxoplasmose, dont on suspecte le rôle dans les accidents de la route, l'impulsivité ou certaines formes de suicide, ou bien du rôle des bactéries dans ce que l'on commence à désigner comme l'axe *brain-guts* – les bactéries de la flore intestinale (le microbiote), par leur diversité dans la population, secréteraient des molécules modulant partiellement l'humeur ou le comportement. L'hypothèse virale est également suggérée, du temps de la personne, ou comme reliquat inséré dans le matériel génétique et transmis.

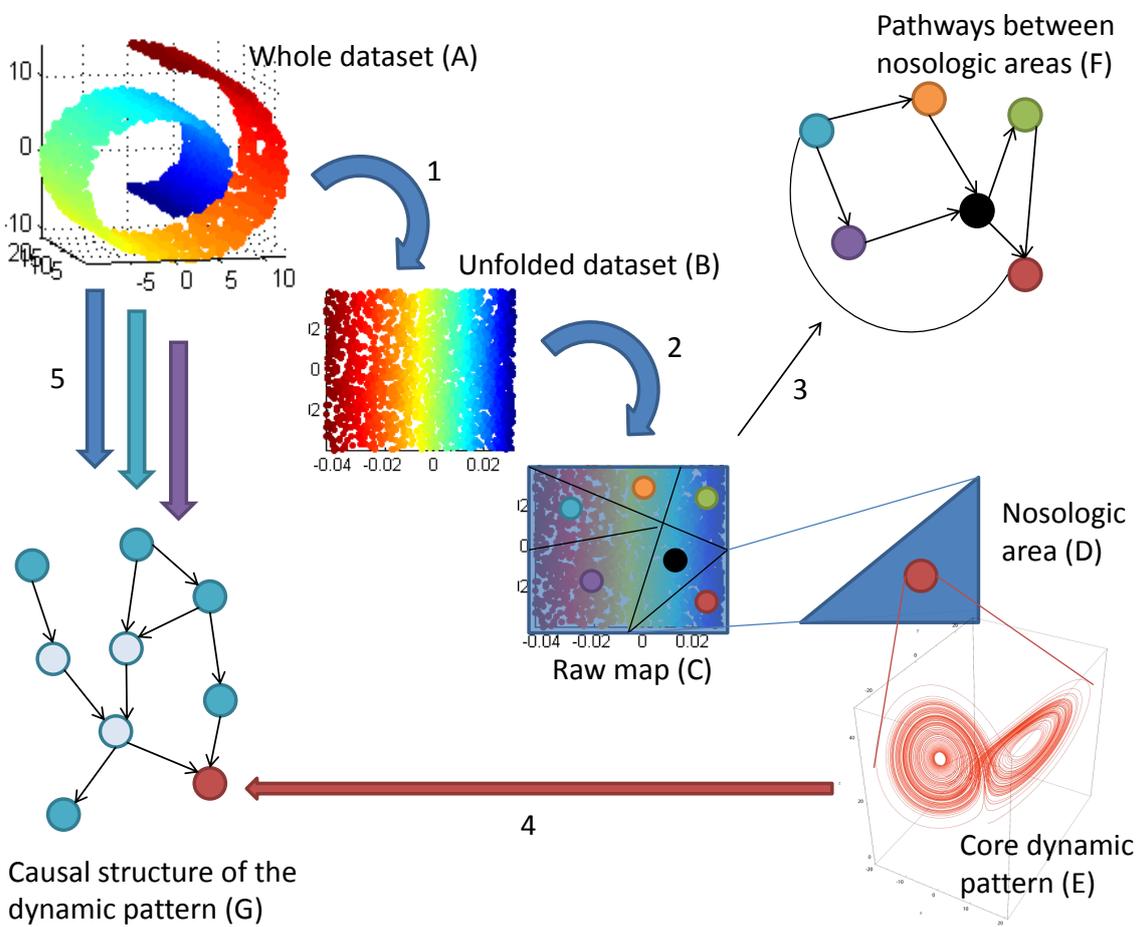
La haute dimensionnalité, le volume ainsi que l'hétérogénéité de ces données en font un excellent cas d'école de ce qu'on appelle les *big data*. La médecine, entre autres domaines, n'est clairement pas équipée actuellement pour relever ce type de défi, même si elle en a récupéré le terme assez récemment, tout en amputant régulièrement l'intégralité du sens ; nous y reviendrons dans la section suivante (section 7). Plusieurs démarches pourraient être envisagées pour aborder ce volume de données, et y chercher la cartographie spécifique des profils récurrents. Nous en proposons ici une qui a le mérite d'être directement praticable techniquement, pourvu que les données soient accessibles et utilisables.

Il s'agit d'une stratégie inspirée de la vision de Waddington et de son paysage épigénétique, bien que sensiblement différent à l'arrivée. Cela consiste à "tailler" ou à modeler dans le volume de données la carte de plus faible dimension possible, tout en respectant la topologie initiale au maximum : au-delà de toute notion de normalité ou de pathologique, ce qu'il nous intéresse d'approcher tient dans les relations de proximités entre personnes ; proximité au sens de ressemblance selon différentes variables, dont les profils subjectifs et objectifs. Un schéma de principe de toute la démarche est donné en figure 4.6.

La première étape est ainsi de procéder à l'estimation de la dimensionnalité intrinsèque globale de la carte recherchée, par exemple grâce aux estimateurs dont nous avons parlé en chapitre 3 (l'estimateur du maximum de vraisemblance donne en général de bons résultats). Cette dimensionnalité obtenue, il sera appliqué au volume de données un algorithme de réduction non linéaire de la dimensionnalité, type LTSA, choisi pour sa rapidité par rapport à ISOMAP, par exemple.

Une première carte brute et réduite est alors obtenue ; à partir de celle-ci, en fonction de la dimensionnalité résiduelle, si elle n'est pas trop haute

FIGURE 4.6 – Schéma de principe pour la construction d'un paysage psychologique



– de préférence, inférieure ou égale à 10, éventuellement 15, maximum 20
– nous appliquerons différents algorithmes de reconnaissance de forme, en conjonction avec des techniques de détermination du nombre optimal de classes à retrouver. Nous aurons une préférence initiale pour une approche par la robustesse des groupes, afin de déterminer des régions les plus stables. Plusieurs techniques sont utilisées de telle sorte qu'un consensus soit atteint. Il s'agit là de la deuxième étape, qui vise à partitionner la carte brute en régions homogènes, en régions que l'on nommera nosologiques.

La troisième étape est une étape de détermination plus fine des formes dynamiques structurant vraisemblablement les régions nosologiques : par région, des algorithmes de recherche et reconstruction des attracteurs seront utilisés, et c'est ici que l'on introduit un paradigme différent des classes diagnostiques statiques et ponctuelles, isolées. Nous postulons que plutôt que des classes rigides, définies exclusivement comme pathologiques, il est possible d'identifier des profils, sans qu'il soit nécessaire de les qualifier *a priori* de normal ou pathologique, et que ces profils correspondent à des formes dynamiques, des attracteurs au sens des systèmes dynamiques.

De fait, les chemins menant aux attracteurs sont des chemins continus, que la transition soit brutale ou non d'une trajectoire à l'autre, et l'appartenance à une région, un profil, est une trajectoire de variabilité intrinsèque, unique et personnelle, non statique. L'analyse d'une part des chemins menant aux attracteurs nous enseignera les facteurs critiques associés aux situations indésirables, d'autre part, l'analyse des facteurs structurant l'attracteur lui-même doit permettre d'en estimer la stabilité structurelle au sens large, et le cas échéant, comment le détruire ou le modifier, s'en échapper.

Les régions nosologiques, les attracteurs y siégeant, peuvent être d'une dimension intrinsèque différente de celle de la carte globale – que l'on suppose supérieure ou égale à celle des régions. Ainsi, après avoir sélectionné les

personnes concernées par un attracteur donné, il est possible d'une part d'estimer la dimensionnalité intrinsèque locale en estimant celle de l'attracteur à partir de la carte réduite, d'autre part d'estimer cette même dimensionnalité intrinsèque locale en revenant au volume de données initial; ce, afin de vérifier la cohérence de nos analyses avant et après réduction du volume initial.

L'analyse en termes de facteurs critiques ou "causaux" des chemins et des attracteurs peut se mener par une analyse type réseaux bayésiens, en réutilisant le volume de données initial là aussi, et en cherchant les relations non nécessairement linéaires associant l'appartenance à un attracteur donné et d'autres facteurs sous-jacents, subjectifs comme objectifs. Enfin, nous donnons dans le tableau une comparaison point par point des principales différences entre une approche classique catégorielle type DSM et une approche type paysage psychologique telle que celle que nous proposons.

TABLE 4.7: Tableau comparatif des deux descriptions : actuelle selon le DSM, et selon une hypothèse type paysage psychologique

	Description type DSM	Description type paysage psychologique
<i>Espace de description</i>	Catégoriel, isolé (diagnostic) <i>a priori</i> Patients labellisés par un diagnostic	Continu, sans catégorie <i>a priori</i> Personnes placées dans l'espace les unes par rapport aux autres
<i>Construction de l'espace</i>	Catégories <i>a priori</i> , historiques	Construction à partir d'un échantillonnage de la population, et reconstruction sans <i>a priori</i>
<i>Dimensions de l'espace</i>	Par axes, et par critères principaux et secondaires (diagnostic)	Dimension intrinsèque globale de l'espace de description. Dimension intrinsèque locale selon les régions d'intérêt
<i>Description des diagnostics</i>	Selon des critères, de seuils arbitraires et universels	Partitionnement de l'espace de description sans <i>a priori</i> autre que la notion de voisinage et de position relative des personnes Définition d'un diagnostic par la notion d'un attracteur, qui assure le caractère stable de la région
<i>Définition d'une normalité ?</i>	Par défaut, si une personne n'entre pas dans une catégorie diagnostique Stratification sous-jacente de	Pas de normalité, mais des positions uniques de personnes dans un espace continu. Pas de position absolue, mais relative.

<i>Transition d'un état à l'autre</i>	<p>l'espace de normalité par la distribution des facteurs de risque</p> <p>De la normalité à une catégorie diagnostique : par saut, avec préexistence de facteurs de risque ou non.</p> <p>D'une catégorie diagnostiques à une autre : des facteurs de risque ?</p>	<p>Passage continu d'une position à une autre, selon une trajectoire unique par personne.</p> <p>Possibilité de passer d'un état stable à un autre selon l'action de différents facteurs critiques (modifiables ou non), d'un attracteur à un autre.</p>
<i>Variabilité individuelle</i>	<p>Caractérisée par un écart à la moyenne (à la norme), ou la présence ou non de critères diagnostiques secondaires</p>	<p>Position unique et relative dans l'espace de description.</p> <p>Variabilité exprimée en termes d'orbites autour d'un attracteur.</p>
<i>Caractérisation de l'incertitude</i>	<p>Incertitude statistique (écart à la moyenne).</p> <p>Incertitude de mesure (due à l'observabilité des critères classants).</p> <p>Incertitude de méconnaissance (informations encore non connues de la science)</p>	<p>Transitions chaotiques possibles.</p> <p>Variabilité intrinsèque d'un système dynamique, assurant la stabilité de certaines régions.</p> <p>Incertitude potentiellement due à des trajectoires chaotiques par endroit, donc imprédictibles à long terme.</p> <p>Incertitude de méconnaissance</p>
<i>Production de connaissance</i>	<p>Association d'un diagnostic à facteur causal ou de risque indépendant des autres (association point à point)</p>	<p>Analyses en réseaux de causalité sans <i>a priori</i> (réseaux bayésiens)</p> <p>Recherche des géodésiques (chemins naturels) entre points caractéristiques</p> <p>Identification et forme des attracteurs</p> <p>Étude des paramètres critiques influençant les trajectoires</p>
<i>Méthodes de construction des espaces</i>	<p>Diagnostics <i>a priori</i></p> <p>Études d'associations linéaires et statistiques entre un diagnostic et une mesure</p> <p>Comparaisons statistiques inter diagnostics</p>	<p>Construction par intégration nativement multivariée de toutes les dimensions retenues, préservant les relations non linéaires</p>

Une autre possibilité peut s'offrir à nous à ce stade, concurrente ou complémentaire, à savoir la modélisation de la structure de l'attracteur par équations structurelles. Nous proposons par ailleurs de développer cette approche sur un cas plus ciblé et concret dans la sous-section suivante.

4.6.5 Principe de modélisation d'un objet psychiatrique : le modèle de Lorenz et la bipolarité

Nous illustrons ici un cas plus précis d'étude d'un attracteur rendant compte d'une forme psychologique. Nous choisissons pour des raisons de simplicité de modéliser un trouble bipolaire – qui correspondrait par exemple, dans le cadre nosologique actuel, à un trouble bipolaire alternant phases dépressives et phases maniaques. Il est à noter que l'on ne se place pas dans une situation nécessairement contradictoire avec ce que nous venons d'exposer, à savoir que les catégories de pathologie actuellement recensées présentent d'importantes insuffisances : nous proposons un modèle d'une part dynamique et non statique et ponctuel, affranchi des notions de norme, et par ailleurs, que l'on peut atteindre par un chemin continu, et non par saut d'une catégorie à l'autre. Enfin, par bipolaire, nous pouvons entendre une catégorie générique plus large que celle à laquelle la psychiatrie actuelle fait référence : nous invoquons un motif dynamiquement stable sous certaines conditions, notamment structurelles, qui correspond à une alternance peu ou non exactement prédictible entre deux pôles d'un espace de représentation donné. À ce stade, nous ne faisons pas d'hypothèse sur la dimensionnalité de l'espace de représentation, ni sur les grandeurs d'une part mesurables pour qualifier le trouble, et d'autre part structurante de l'attracteur.

Une connaissance minimale des systèmes dynamiques classiques et des attracteurs fait inévitablement penser à un attracteur type attracteur étrange de Lorenz lorsque l'on évoque un attracteur à deux pôles. Notre idée est donc ici de tenter de modéliser un trouble bipolaire par un système de Lorenz.

L'idée peut sembler étonnante à première vue : pourquoi un objet complexe comme un trouble psychologique, exhibant par ailleurs visiblement une composante double mais peu prédictible, de trajectoires inter personnelles et de variabilité intra personnelle très importantes, pourquoi un tel objet

devrait-il se plier à une modélisation simple, en particulier sous la forme d'un système de trois équations ?

Montrons que cette première objection n'a pas réellement lieu d'être, et surtout, que l'approche n'est en fait en rien réductrice, si l'on s'y prend correctement.

Pour commencer, il s'agit de renverser la conception classique qui soutient ce type d'objection : que l'observation d'un comportement complexe et peu prédictible doit nécessairement être représenté par un modèle lui-même très complexe, au sens de son intelligibilité : un système avec beaucoup de variables, liées les unes aux autres de manières non évidentes et multiples. Or, s'il est bien un enseignement fondamental que l'on doit aux systèmes dynamiques, c'est bien qu'un système *a priori* très simple dans sa formalisation – par exemple, un pendule : une boule pesante au bout d'un fil, que l'on lâche en dehors de son point de repos – peut exhiber des comportements complexes, et non compliqués, jusque devenir imprévisible et chaotique dans ses trajectoires. Inscrit en faux contre cette idée de "conservation de la complexité" entre modèle et réalité, apparaît également la question de la dimensionnalité du problème : il n'est pas nécessaire d'étudier un espace de très haute dimensionnalité pour obtenir un comportement complexe. Il semblerait même que les systèmes, au moins localement, présentent des phénomènes de transition intéressants, critiques, pour de petites dimensionnalités (cf par exemple la théorie des catastrophes de Thom). Les espaces de très haute dimensionnalité étant par ailleurs soit structurés davantage de manière "probabiliste" (de type corrélation), soit étant possiblement uniquement de haute dimensionnalité en apparence ou par construction, et non par dimension intrinsèque.

Ensuite, examinons la critique de la contrainte formelle : pourquoi le phénomène serait-il, sous prétexte qu'il soit oscillant autour de deux pôles distincts, nécessairement de la forme d'un système de Lorenz ? Deux réponses

différentes à cette question : d'une part, effectivement, il est possible que d'autres systèmes d'équations, *a priori*, fournissent un attracteur bipolaire différent de celui de Lorenz ; d'autre part, ce qui nous intéresse en l'espèce, n'est pas que l'entièreté de la personne présentant ce type de trouble, soit modélisable par trois équations particulières, mais bien que l'on puisse rendre compte au moins partiellement des observations faites par cette forme de système. C'est bien la structure du modèle, sa qualité et non son aspect rigide ou quantitatif qui nous intéresse ici. En termes plus pratiques et techniques, il suffit d'examiner si une partie des observations peuvent être exprimées par des équations types Lorenz, liées entre elles, et de mesurer l'erreur résiduelle que n'explique pas le modèle. Ceci implique donc le recours à une méthode statistique d'interprétation des données.

Néanmoins, nous introduisons la contrainte – ou l'hypothèse – selon laquelle trois dimensions suffisent à représenter un trouble bipolaire. Evidemment, ces trois dimensions n'ont cependant aucune nécessité de correspondre directement aux items retenus par les classifications, par exemple : nous considérons simplement que trois dimensions, liées tel que le système de Lorenz le spécifie, peuvent rendre compte de cette forme.

Afin de tester cette hypothèse, nous distinguons deux parties au problème : 1) la spécification du système, qui est donc donné dans sa structure (système de Lorenz), et 2) l'identification du système à partir de données observées. Pour cette seconde partie, nous préconisons de recourir à la modélisation par équations structurelles (ou SEM : *Structural Equation Modeling*), et estimation des paramètres par approche bayésienne.

Cette technique présente dans notre cas précis plusieurs avantages. Elle permet d'expliquer par une méthode statistique une partie des données, et d'estimer l'erreur résiduelle – autrement dit, la partie non expliquée par la forme de notre modèle. Elle permet également de recourir à ce que l'on

appelle des variables latentes, à savoir, des variables représentant un concept complexe, non mesurable directement, mais modélisable par un ensemble de grandeurs qui sont elles mesurables. Un exemple illustratif de variable latente est l'intelligence. Si l'on admet que l'on ne puisse assimiler intelligence et quotient intellectuel, on peut alors considérer l'intelligence comme une variable latente, mais modélisable par un ensemble d'observations, telles que justement différents tests de QI, des tests d'intelligence émotionnelle, des tests d'aptitudes créatrices, etc.

De la même façon, nous pouvons explorer des concepts, au nombre de trois, qui seraient liés selon le système de Lorenz, et estimés par des variables elles observées, notamment par des échelles validées et reproductibles. Ces concepts pourraient être des potentiels émotionnels : équivalents d'une énergie potentielle physique, sorte de réserve pouvant être convertie sous une autre forme. Potentiel de tristesse, potentiel maniaque, par exemple. Ceci est également d'importance, car cela diminue la contrainte induite par le choix d'un espace de dimension trois pour représenter la forme : il est bel et bien possible de mesurer davantage de variables que trois, certaines participant structurellement d'une des trois variables latentes utilisées par le système.

Un autre avantage décisif de cette technique de SEM est de permettre une modélisation non linéaire, et donc bien de représenter correctement et formellement le système de Lorenz.

Si la phase d'identification des paramètres du modèle n'aboutit pas à des paramètres nuls, il sera ainsi possible de dire que le trouble peut s'expliquer en partie par un système de Lorenz. Dans le même mouvement, nous aurons une construction des trois dimensions du trouble à partir de mesures. Un tel résultat ne serait pas important que d'un point de vue purement identificatoire, mais surtout conceptuel : l'on passerait d'une conception de la pathologie statique et étanche, vue comme une déviation excessive à la

norme d'un certain nombre d'items, et dont les variations inter et intra individuelles seraient essentiellement vues comme des effets aléatoires, à une représentation sans norme, mais avec une infinité de trajectoires déterministes mais imprévisibles et variables, respectant une forme globale et dynamique. L'imprédictibilité partielle ou totale serait intrinsèque à ce trouble, et non l'effet d'un manque d'information, que d'aucun serait tenté d'aller chercher dans les mutations génétiques, ou un phénomène épigénétique exclusif.

Il pourra nous être objecté un autre problème à cette entreprise de modélisation : l'aspect temporel du modèle de Lorenz. A cela, nous pouvons remédier de deux manières au moins. Soit l'on parvient à suivre les mêmes personnes au cours du temps, et les observations sont datées, soit échantillonnées à pas fixes, soit lorsque survient un événement significatif. Soit l'on a davantage accès à des observations de genre étude transversale, recueillies une ou deux fois auprès des mêmes personnes, mais concernant davantage d'individus que dans le premier cas. Dans ce second cas de figure, il est également nécessairement de dater les événements. Nous ferions alors l'équivalent d'une hypothèse de type ergodique, c'est-à-dire que l'observation moyenne temporelle est assimilable à l'observation moyenne "spatiale" – ici, le spatial est l'ensemble des individus concernés. La formulation différentielle du système de Lorenz serait approchée par une variable normalisée mesurant l'écart de deux états rapportée à l'écart de temps séparant ces deux états.

Pour tester cette hypothèse de modèle, il serait ainsi nécessaire de pouvoir accéder à une étude type cohorte de patients, à défaut étiquetés "bipolaires de type 1 ou 2".

Formalisme de base de la modélisation par équations structurelles

Les modèles à équations structurelles servent à représenter et estimer des concepts ou des grandeurs complexes, réputées non directement observables.

L'exemple le plus fréquemment donné est l'exemple de l'intelligence, dont on peut supposer qu'on peut en approcher le concept par différents types de mesures, comme des tests de QI, mais non directement. Dans ce cas, l'intelligence serait une variable latente, quand les résultats des tests fourniraient des variables mesurées.

Un modèle à équations structurelles est alors constitué de deux types d'équations : une équation dite de mesures, et une équation dite structurelle. Les deux lient les variables mesurées et les variables latentes, l'équation structurelle étant supposée donner la forme du concept en fonction de différentes variables latentes, tandis que l'équation de mesures lie les mesures aux variables latentes. Pour reprendre le cas de l'intelligence, on pourrait vouloir (caricaturalement) diviser ce concept en une intelligence procédurale, cognitive (IC), et une intelligence émotionnelle (IE), les deux pouvant évidemment entretenir des relations entre elles. L'IC et l'IE seraient ainsi des variables latentes, non mesurables telles qu'elles directement, et constitutives de l'intelligence. Les mesures permettant d'approcher des dimensions de l'IC et de l'IE seraient par exemple les tests de QI, les tests de personnalités, etc. Ces derniers, ces tests sont les mesures. De fait, l'équation de mesures donnera les liens entre variables latentes IC et IE et les mesures (les tests), et l'équation structurelle donnera la forme des relations entre l'intelligence, et les variables latentes que sont l'IC et l'IE. Dans chaque équation, il existe un terme d'erreur, de mesure dans l'équation de mesure, et de modélisation dans l'équation structurelle.

Si les premiers modèles étaient linéaires, il a rapidement émergé des versions non linéaires, étendant ainsi la portée de cet outil. Enfin contrairement à beaucoup de modèles statistiques, le caractère normal ou gaussien des va-

riables n'est pas nécessaire à la validité du modèle.

Nous donnons ici la forme de l'équation de mesure en général :

$$y = \mu + \Lambda\omega + \epsilon \quad (4.36)$$

Avec y de dimension n , le nombre de variables mesurées, μ de dimension n , vu comme intercept, Λ une matrice de taille $p \times q$, ω de dimension q variables latentes et ϵ un vecteur d'erreur de dimension n .

De même, nous donnons la forme générale de l'équation structurelle, dans sa version non linéaire :

$$\eta = \Pi\eta + \Gamma F(\xi) + \delta \quad (4.37)$$

Avec η un vecteur de variables latentes de dimension q_1 , et ξ de dimension q_2 , tel que $q = q_1 + q_2$. δ est un vecteur de dimension q_1 , représentant l'erreur résiduelle du modèle. F est un vecteur de p fonctions linéairement indépendantes et différentiables avec $p \geq q_2$. Enfin, nous avons la partition des variables latentes suivantes : $\omega = (\eta, \xi)$.

Un modèle pour les phénomènes de type bipolarité

On pose X une forme de potentiel interne à la personne, Y une forme de potentiel maniaque et Z une forme de potentiel dépressif. On suppose par ailleurs que les troubles que l'on nomme actuellement bipolaires, et présentant soit l'une soit les deux caractéristiques de manie et de dépression dans le temps, peut s'écrire comme une combinaison de ces 3 grandeurs : 2 grandeurs potentielles, et 1 grandeur à voir comme une variable interne, *a priori* conservative, au moins dans un voisinage temporel raisonnable. Nous nous référons par exemple aux critères diagnostiques actuels, issus du DSM 5 (cf tableau 4.8).

Sous forme infinitésimale, la première équation du système de Lorenz que

TABLE 4.8 – Critères diagnostiques pour les troubles bipolaires, selon le DSM 5 : 1 des 2 critères principaux (manie ou dépression) et 3 symptômes dits "mixtes".

Critères principaux	Manie ou hypomanie	Dépression majeure
et 3 symptômes parmi :	Dysphorie excessive ou humeur dépressive Diminution de l'intérêt ou du plaisir Ralentissement psychomoteur Fatigue ou perte d'énergie Dévalorisation ou culpabilité excessive Pensées de mort récurrentes, idées suicidaires	Humeur élevée, expansive Augmentation de l'estime de soi, idées de grandeur Plus grande communicabilité Fuite des idées Augmentation de l'énergie ou de l'activité Engagement excessif dans des activités à potentiel élevé de conséquences dommageables Réduction du besoin de sommeil

l'on rappelle ci-dessous :

$$\begin{cases} \frac{dX}{dt} = \sigma(Y - X) \\ \frac{dY}{dt} = \rho X - Y - XZ \\ \frac{dZ}{dt} = XY - \beta Z \end{cases}$$

peut s'écrire :

$$dX = \sigma(Y - X).dt$$

Ainsi, on note qu'il existe une relation "élastique" entre Y et Z , entre potentiels maniaque et potentiel interne, que représente ici X . De façon très grossière, le profil temporel d'une personne souffrant d'un trouble bipolaire pourrait être approximé par l'intégrale sur le temps de X , la normothymie (ou euthymie) étant alors définie comme une grandeur moyenne sur le temps, et valant $\int_t dX = \langle X \rangle_t = 0$ (euthymie au premier ordre).

On note également ici la relation privilégiée, et asymétrique pour Z , entre X et Y . En effet, nonobstant le terme en X du second membre de la deuxième équation du système, la deuxième et troisième équations du système sont quasi-symétriques, avec chacun un terme liant la variable "op-

posée" (i.e. Y pour Z et inversement, pour manie et dépression) et un terme linéaire en la variable d'intérêt (i.e. Y pour l'équation en $\frac{dY}{dt}$).

De fait, l'accroissement du potentiel maniaque Y sera réglé par l'intensité du potentiel interne X , et freiné par rétroaction par l'intensité de Y , et l'intensité d'interaction entre les potentiels X et Z , soit XZ . L'intensité de l'accroissement maniaque est donc réglée d'une part par l'intensité du potentiel interne, et freinée par la tendance maniaque augmentant, ainsi que par la propension elle aussi croissante à la dépression.

La troisième équation en Z nous renseigne sur la manière qu'a le potentiel dépressif à croître alors que l'intensité d'interaction XY , c'est-à-dire, le potentiel interne (la "réserve" à changer d'état que possède la personne) avec le potentiel maniaque, croît elle aussi. Cet accroissement est d'autant plus important localement que l'écart entre interaction XY et le potentiel dépressif Z est grand. En d'autres termes, plus forte est la crise maniaque, et plus grande est la capacité de la personne à pouvoir changer d'état, et plus grand sera l'accroissement du potentiel dépressif.

Pour plus de simplicité pour la suite, nous renommons les variables X , Y et Z respectivement comme suit : Fo , M et T . Fo pour fonctionnement, M pour manie et T pour tristesse (de l'humeur). La modélisation par équations structurelles impliquent une séparation des variables en 2 types : les variables mesurées et les variables latentes (non observées ou complexes).

Pour les variables mesurées et observables, il existe des échelles de thymie, plus ou moins spécialisées, validées, qui permettent de mesurer à un temps t l'intensité de la thymie plutôt maniaque ou dépressive. Il existe de la même façon des échelles de fonctionnement, spécialisées ou non, validées

et rendant compte du fonctionnement d'une personne, qu'il soit personnel, interpersonnel, social, etc.. Les variables mesurées seront donc Fo , M et T . Puisque ces mesures seront effectuées à aux moins 2 temps différents, on construira les variables qui leur sont respectivement associées et standardisées (en écart et en intervalle de temps, elles sont vues comme des variables aléatoires mesurées de manière répétée), que l'on notera : Fo^N , M^N et T^N .

Les variables latentes vont nous permettre d'atteindre 2 buts : obtenir une expression au premier ordre de la variable BP , modélisant le trouble bipolaire, comme annoncé, et vérifier si la structure empruntée au système de Lorenz est satisfaisante pour expliquer les données mesurées (ce qui s'évaluera tant sur les indices d'adéquation, de *fitness* du modèle, que sur les termes d'erreur résiduelle de modélisation). On notera alors ΔM , ΔT et ΔFo les variables latentes équivalentes à dM , dT et dFo du système différentiel déterministe.

De la sorte, nous avons le système mesuré comme suit :

$$\begin{cases} Fo^N = \mu_1 + \lambda_{1,1} \cdot \Delta Fo + \lambda_{1,2} \cdot Fo + \lambda_{1,3} \cdot BP + \epsilon_1 \\ M^N = \mu_2 + \lambda_{2,1} \cdot \Delta M + \lambda_{2,2} \cdot M + \lambda_{2,3} \cdot BP + \epsilon_2 \\ T^N = \mu_3 + \lambda_{3,1} \cdot \Delta T + \lambda_{3,2} \cdot T + \lambda_{3,3} \cdot BP + \epsilon_3 \end{cases}$$

Par ailleurs, nous avons l'équation structurelle suivante, d'après notre hypothèse suivant le modèle de Lorenz :

$$\begin{cases} BP = \gamma_{1,1} \cdot \Delta Fo + \gamma_{1,2} \cdot \Delta M + \gamma_{1,3} \cdot \Delta T + \gamma_{1,4} \cdot Fo + \gamma_{1,5} \cdot M + \gamma_{1,6} \cdot T + \pi_1 \cdot BP + \delta_1 \\ \Delta M = \gamma_{2,1} \cdot \rho Fo - \gamma_{2,2} \cdot M - \gamma_{2,3} \cdot FoT + \delta_2 \\ \Delta T = \gamma_{3,1} \cdot FoM - \gamma_{3,2} \cdot \beta T + \delta_3 \\ \Delta Fo = \gamma_{4,1} \cdot \sigma M + \gamma_{4,2} \cdot \sigma T + \delta_4 \end{cases}$$

On remarquera que la première équation, qui est classiquement l'équation structurelle principale, celle servant à modéliser le phénomène recherché (ici, le trouble bipolaire), telle qu'elle est écrite, est analogue à un développement au premier ordre de Taylor de BP , si l'on considère que BP s'écrit comme une fonction f telle que $f=f(Fo, M, T)$, et que ΔX s'assimile à la différentielle du premier ordre de X .

4.7 Intégration et traitement des données : les *big data* en pratique

Nous avons effleuré ci-dessus le concept à la mode des *big data*, concept par ailleurs bien souvent maltraité, et utilisé pour dire simplement qu'il y a beaucoup de *bits* à traiter, sans présumer d'autres qualités spécifiques de ces données. Nous expliquons ici, au-delà du terme à la mode, la réalité des *big data*, leur place possible en sciences, ainsi qu'un système de traitement des données complet, à vocation première hospitalière, mais très certainement transposable à d'autres domaines. En particulier, les exemples que nous avons jusque ici traités ont mis en lumière un problème récurrent, pour aussi peu nouveau qu'il soit, celui de l'existence, puis de l'accessibilité, de la disponibilité et enfin de l'utilisation des données, dès lors que l'on ne veut pas se cantonner à des modèles purement hypothétiques, sans chance d'être éprouvés, corrigés, confirmés ou abandonnés. Autrement dit, impossible de s'intéresser au traitement de l'information médicale sans parler et qualifier l'information à traiter, donc les données médicales.

4.7.1 De l'autre côté du spectre des RCT, la fouille de données (le *data mining*)

À la suite de ce que nous avons déjà débattu précédemment, et qui a montré que le modèle prédominant quant à la production de preuves en

sciences est jusque ici un modèle basé sur le modèle expérimental, et les essais randomisés contrôlés, nous examinons l'autre terme de l'alternative, à savoir les études dites *data-driven*.

Les plans d'expériences peuvent être plus ou moins complexes et recherchés, ils visent globalement tous les mêmes buts : équilibrer les effectifs de sorte qu'ils soient comparables, à la différence d'un paramètre d'intérêt. Ces approches sont essentiellement des tests d'hypothèse. Toute l'expérience est montée de telle façon à mettre d'une part en lumière une différence entre deux groupes, et d'autre part, que cette différence soit uniquement attribuable au paramètre d'intérêt.

Cependant, avec l'accumulation de données, la difficulté accrue de formuler une hypothèse univariée ou de contrôler les expériences, un autre type d'approche s'est fait jour et tend à se développer de plus en plus : la fouille de données, ou *data mining*. On n'y cherche pas en premier recours à tester ou vérifier une hypothèse, mais davantage à mettre à jour des corrélations entre observations, ou à découvrir des motifs, des formes liant ou structurées par plusieurs variables.

L'augmentation des flux de données et des capacités de stockage ont amené certains à parler de *big data*. On attribue par exemple au META group, devenu depuis Gartner, d'avoir les premiers utilisé en particulier le terme de *big data*, en les définissant de la sorte : les *big data* se fondent sur les trois V, pour Volume, Variété et Vitesse. Certains ajoutent un quatrième V pour Vérité. Les *big data* concernent donc des gros volumes de données, mais également des données de natures hétérogènes, qu'il faut savoir traiter rapidement. En recherche, les *big data* sont souvent entendues comme la manipulation de gros volumes, ce qui n'est ni réellement nouveau, ni pertinent. En revanche, que l'on ait à traiter des volumes de haute dimensionnalité, de sources variées, et non structurées est un défi assez récent, rendu possible techniquement en termes d'accès à ces données. Leur manipulation correcte

est un autre aspect du problème.

Nous sommes donc arrivés, en sciences, à ce point où le spectre des moyens de production de savoir occupe tout un spectre de possibilités, les extrêmes en étant l'expérience totalement contrôlée à vocation de vérification d'hypothèse unique, et la masse de données amorphes accumulées sans *a priori*, dans laquelle il doit être possible d'explorer des relations sous-jacentes entre observations. Là où l'expérience contrôlée a une vocation unique : elle ne sert à démontrer si probante qu'une seule chose à la fois – l'expérience est dédiée à un résultat unique – et peut ainsi nécessiter des coûts considérables à sa mise en place pour ne plus ensuite servir, les *big data* sont censés être éventuellement coûteux en termes d'infrastructure initiale, mais servir à la découverte de différents résultats.

Il existe, ou existait de façon transitoire, une forme intermédiaire, représentée par les grandes cohortes de personnes, où il s'agit d'enregistrer un grand nombre de paramètres, mais dans un périmètre d'investigation relativement restreint – par exemple : les déterminants sociaux de la santé parmi les salariés d'une grande entreprise (cohorte GAZEL). En restant raisonnable, et en utilisant correctement certains outils statistiques – l'analyse multivariée en particulier – il y est possible de se rapprocher des conditions quasi expérimentales, en se servant des données comme variables d'ajustement, c'est-à-dire, substituer les techniques statistiques plus sophistiquées à ce que le design de l'expérience contrôlée contribuait à réaliser : équilibrer les sources de variabilité par tirage aléatoire. Cependant, ces techniques ont un certain nombre de limitations, et sont assez fréquemment mésusées en biomédecine, en lien avec une méconnaissance des effets induits par une haute dimensionnalité, ou un sur ajustement des modèles sur un trop grand nombre de variables. Par ailleurs, outre les limites des modèles de régression appliqués aux *big data*, nous ne savons pas mieux faire actuellement que d'identifier des corrélations linéaires : on ne sait pas détecter de façon automatique

des corrélations de plus hauts degrés, ou de degré fractionnaire. Il est alors nécessaire de se doter d'autres types d'approches, lesquelles, sans être forcément capables d'identifier automatiquement des relations complexes entre variables, doit pouvoir au moins les conserver. [274, 275, 276]

4.7.2 Capture de routine et entrepôts de données

Nous venons de dire que les *big data* pouvaient, comme désavantage, engendrer d'importants coûts afin d'en établir l'infrastructure nécessaire à la collecte des informations. Ce n'est pas toujours le cas, notamment là où, de manière quasi automatique ou routinière, des informations sont collectées à chaque instant, et alimentent diverses bases de données, plus ou moins structurées, de qualité et de nature variables.

Dans le cas du milieu hospitalier, il existe une informatisation croissante à tous les niveaux, et les établissements sont équipés de systèmes d'information plus ou moins élaborés. Il n'en reste pas moins vrai que la numérisation des données, que ce soit de l'image (imagerie conventionnelle, nucléaire, IRM...), des résultats quantifiés (biologie, microbiologie...) ou encore du texte (comptes rendus divers), gagne du terrain considérablement. Evidemment, cette numérisation se faisant sur différents fronts, de manière asynchrone, non standardisée, et concernant des informations de types variés, coexistent et se développent indépendamment plusieurs bases de données au sein du même système, le tout devant respecter des conditions de sécurité en théorie drastiques, afin de protéger le secret médical et la confidentialité des données de santé nominatives.

Du côté de la communauté des épidémiologistes, essentiellement, l'intérêt est grandissant pour l'exploitation de vastes quantités de données. Cependant, l'angle d'attaque apparaît sensiblement différent de celui relativement fondamental qui pose les limites intrinsèques à l'exploitation de ces ensembles

de données (par exemple, le phénomène d'espace vide). Les problématiques abordées sont notamment d'au moins deux grands types : la possibilité de conduire des tests – statistiques, essentiellement et toujours – sur des données distribuées, éparpillées dans diverses bases de données, et la valeur des résultats dits *data-driven*, c'est-à-dire issus de données recueillies en dehors de tout contexte prospectif ou de design conçu comme pour une expérience. Ces questions sont évidemment légitimes, mais manquent sans doute en partie le fond du problème, notamment en s'interdisant de recourir à d'autres méthodes plutôt qu'à des extensions des méthodes déjà utilisées, et critiquées pour leur utilisation dans des contextes pourtant plus favorables. À noter l'exception d'un auteur particulier, qui s'est récemment prononcé sur le sens que l'on peut donner à l'exploitation de ce type de données, selon le traitement que l'on en fait. [275]

4.7.3 Concept de CDC

Autre question intervenant, celle de la qualité des données, à partir du moment où elles ne bénéficient pas d'une attention particulière et calibrée pour un emploi bien spécifié. Se surajoute le problème lié au codage si l'information doit être entrée selon une certaine nomenclature (par exemple, le codage des actes et des diagnostics).

Il n'est pas possible, *a priori* et en pratique courante actuelle, d'assimiler la qualité des données capturées en nombre et temps limité dans le cadre d'un essai contrôlé dédié à une question précise et celle des données capturées et enregistrées sans utilité immédiate ou différée ressentie par les professionnels de santé. Au-delà de la qualité des données saisies, dont on sait qu'elle peut et devrait s'améliorer, notamment parce qu'il existe une marge de progression probablement substantielle (le codage du PMSI s'est ainsi amélioré de façon drastique sur quelques années d'existence, ce qui doit rassurer, même si son périmètre est beaucoup plus restreint, et son utilité encore davantage), se

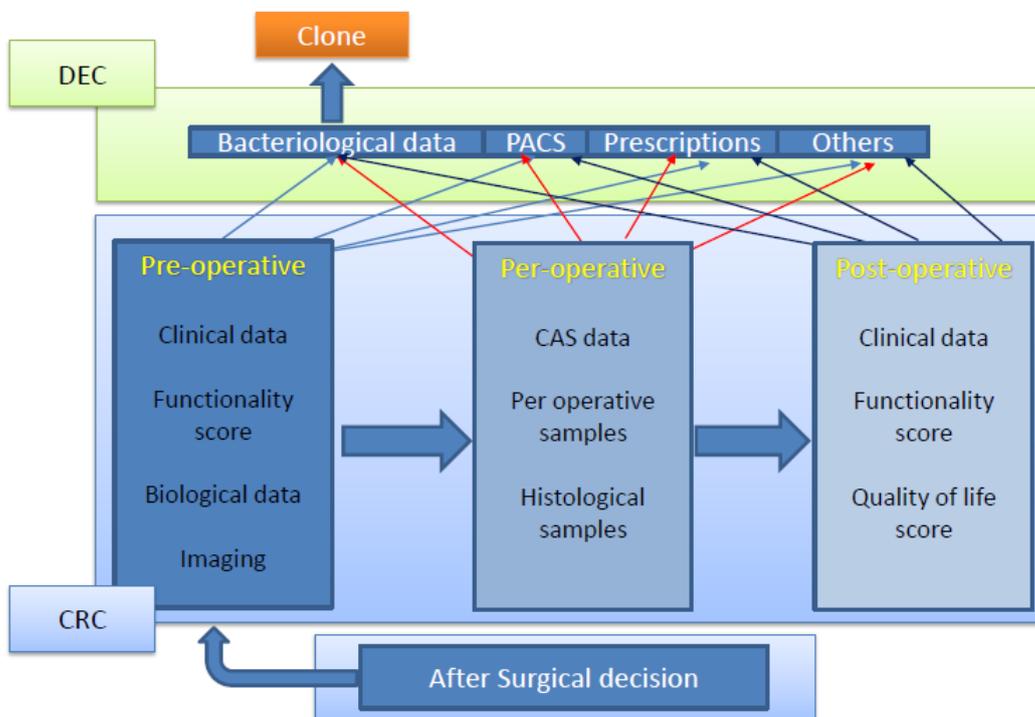
pose la question du manque de spécificité de certaines données, de l'absence de données dans les bases qu'il serait pourtant pertinent de recueillir dans le cadre de certaines activités, mais que la pratique usuelle ne prévoit pas de manière systématique ou bien définie.

Le concept de CDC – initialement, pour Centre de Données Cliniques – incorpore, en sus de et appuyé sur l'exploitation secondaire de données, la mise en place de filières spécialisées, ces filières étant dans l'idéal distribuées, à ressources constantes ou peu s'en faut, et ne matérialisant rien d'autre que la constitution de flots de patients articulés autour d'une prise en charge globale, multidisciplinaire, mais centrée sur une ou un ensemble de pathologies faisant sens. C'est notamment le cas pour la prise en charge des patients de chirurgie orthopédique, et présentant un profil complexe, par exemple en raison de complications infectieuses ; on retrouve ici notre abord par les IOA.

La filière spécialisée tient alors en des recueils de données enrichis, ciblés sur des signes spécifiques et pertinents, si possibles standardisés, des rendez-vous dédiés à la prise en charge globale et non fractionnée, et l'intervention de personnel formé pour ces recueils et prise en charge. Outre l'amélioration espérée immédiate de la prise en charge des patients, on espère pouvoir améliorer la qualité des informations entrant dans le système d'information, de même que l'information sera enrichie de données spécifiques pour ces patients précis. Une idée centrale au CDC est de renforcer l'attention et la cohérence des examens proposés aux patients répondant d'un certain profil – pour ce qui nous intéresse ici, présentant des IOA. Le concept de CDC est résumé, en pratique, dans la figure 4.7.

L'autre pilier du CDC consiste à se doter d'un système d'exploitation secondaire des données, tant pour le patient, que pour la recherche.

FIGURE 4.7 – Concepts de CDC et de SPOT RISC en pratique clinique de routine



4.7.4 SPOT-RISC

Le projet et système SPOT-RISC relève de ce que l'on nomme informatique médicale, et plus spécialement, des systèmes d'exploitation secondaire des données de santé. Ce système se base sur la création d'un entrepôt de données, c'est-à-dire une base de données plus ou moins structurées, sur laquelle il est possible d'effectuer des traitements.

L'idée sous-jacente est simple : avec la numérisation croissante et programmée des données patients en milieu hospitalier dans un premier temps, et en ambulatoire plus difficilement, des informations sont enregistrées à chaque instant, concernant les patients entrant dans le système de soins. Ces données sont pour certaines formatées, codées (par exemple, codage des diagnostics et des actes), d'autres sont davantage amorphes, sous forme de texte.

Le flux de patients passant par les hôpitaux est comparable à la constitution d'une très grande cohorte de patients, ouverte. Moyennant la caractérisation et le formatage minimal des données saisies et disponibles (essentiellement, leur qualité), nous disposons donc d'un observatoire unique en son genre, et d'une grande richesse. Bien entendu, nous sommes à l'opposé du principe d'application des RCT, mais nous disposons pour analyser et modéliser les relations entre données d'un très grand nombre d'outils, dont ceux déjà utilisés en épidémiologie et études observationnelles. Il existe un certain nombre de difficultés à mener ce genre de projet, les difficultés techniques étant bien présentes, mais loin d'être les plus insurmontables, comparées aux contraintes juridiques, et aux obstacles politiques à des échelles variées.

En pratique, dans ce projet, il existe trois grands défis techniques :

1. celui lié au morcellement des données en différentes bases de données (biologiques, anatomopathologiques, d'imageries...),
2. au caractère amorphe encore d'une grande partie des données, non

structurées,

3. la mise en commun des informations provenant de différents centres hospitaliers, afin de se placer, de manière transparente pour l'utilisateur, dans un contexte multicentrique, mais d'accès unique.

Le concept de SPOT-RISC est un concept par couche et modulaire, présentant une triple destination, du moins en ce qui nous intéresse ici (il existe en fait une quatrième destination : l'utilisation des données à but médico-administratif, au-delà du PMSI).

Cette triple destination est la suivante : le système doit pouvoir servir à la fouille de données et au prototypage de RCT (aussi étonnant que cela puisse paraître, il n'est pas rare de ne pas disposer d'informations suffisantes, voire d'information tout court, pour renseigner sommairement un problème scientifique : pas de données de prévalence, par exemple), à la mise en place d'une médecine prédictive personnalisée, et enfin à la surveillance épidémiologique. Ce système a été initialement pensé pour servir à l'étude des infections ostéo-articulaires, mais le principe ne lui est en rien spécifique. L'importante dimensionnalité des données se rapportant à un patient donné doit permettre, avec un certain nombre de limites évidemment, non seulement de permettre à continuer d'utiliser, à but d'analyses localement valides, les approches usuelles, mais également d'approcher les problèmes par d'autres moyens, dont la caractérisation de formes directement mesurées, et non synthétisées par l'esprit, après analyse de la littérature et des données isolées. Dans l'idéal, ce système se doit d'être pédagogique et accessible, afin de sensibiliser le public des chercheurs et des praticiens à des approches différentes de celles qu'ils connaissent, notamment en en reliant l'utilisation à leurs milieux et problèmes quotidiens.

Architecturalement, le système repose sur l'implantation des mêmes logiciels sur plusieurs centres (à fin décembre 2013, le déploiement est prévu et

financé pour 3 centres hospitalo-universitaires, dont Brest, Rennes et Tours ; l'année 2015-2016 doit normalement voir la dissémination du système sur les CHU d'Angers, Nantes et Poitiers), afin qu'il soit possible pour un centre d'accéder à ses propres données fédérées, mais également, à moyen terme, à l'ensemble des données des centres concernés, par un accès unique, et vu comme une seule base de données. SPOT-RISC s'appuie sur 4 couches fonctionnelles, superposées : une couche d'accès aux données, une couche de structuration des données, une couche de traitement des données puis une couche d'interfaçage homme-machine. Nous détaillons brièvement ces 4 couches dans ce qui suit, en appuyant un peu plus sur la structuration des données, et sur la couche de traitement – la première détermine l'espace initial dans lequel nous pourrions travailler, et la seconde, des outils nous permettant de le modéliser et de le manipuler. L'architecture et les concepts derrière SPOT-RISC sont représentés dans les figures 4.8, 4.9 et 4.10. D'un point de vue technique, ce projet s'appuie sur l'expertise et la participation de deux partenaires industriels : TEKLIKO et ENOVACOM.

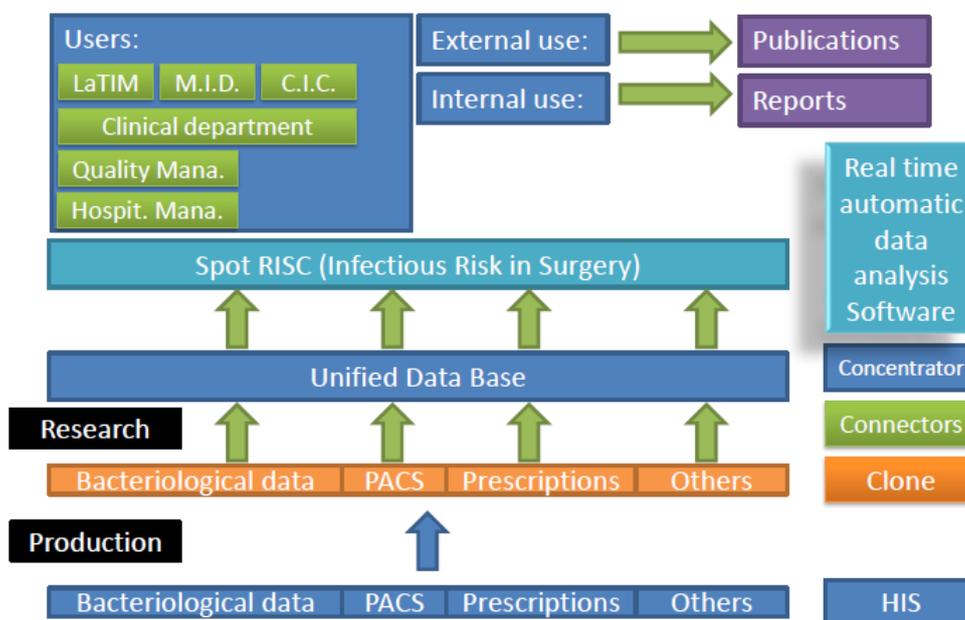
Présentation des 4 couches logiques et architecturales : ETL – structuration – traitement – IHM.

La couche d'ETL (*extraction – transport – loading*) est conçue pour la récupération des données dans les différentes bases du système d'information hospitalier : bases de données biologique, microbiologique, d'imageries, cliniques, génétique, pharmacologiques. En effet, les bases où sont stockées les différentes informations et les résultats d'examen sont souvent séparées, organisées distinctement, et la première étape lorsque l'on désire "reconstituer" un patient ou un profil, est d'aller, de manière transparente à l'utilisateur, récupérer les informations éparpillées un peu partout dans le système d'information. Globalement, les verrous techniques sont ici peu importants, et les principaux problèmes rencontrés sont de deux ordres : la politique de sécurité

informatique, locale et négociable, et le caractère propriétaire des logiciels structurant ou permettant l'exploitation de chacune des bases de données à interroger. En effet, même si nous n'avons besoin que de nous connecter aux bases, encore faut-il bénéficier des droits et de la documentation nous permettant de le faire en toute confiance, légalité et conformité. Les industriels en charge des différents logiciels sont autant d'interlocuteurs à atteindre, et à convaincre (bien que l'interopérabilité des logiciels, et plus encore des progiciels, soient théoriquement une obligation – ce fut le principal motif retenu pour l'amende infligée à Microsoft par l'Union Européenne).

Couche de structuration et de constitution de l'entrepôt de données. La couche de structuration des données repose sur deux types d'éléments : le format de la structure elle-même, à déterminer et standardiser, et le moteur permettant d'alimenter ce format en données observées. Le moteur peut être découplé du reste tant qu'il respecte les besoins et formats d'entrée et sortie des données. Quant à la structure des données, il s'agit purement et simplement de proposer une sorte de "grille de lecture" des observations d'un patient, ou autrement dit, côté médical, donner l'équivalent d'un dossier médical standardisé. Le format de cette structure est à préciser selon les différentes utilisations qui peuvent être faites de l'outil, mais également selon les données disponibles ; il est nécessaire de s'orienter vers un format établi de façon concertée et consensuelle, pour une interopérabilité de base correcte, par exemple entre centres hospitaliers. Nous avons, au niveau du laboratoire et plus précisément de l'équipe s'intéressant aux risques infectieux en milieu chirurgical, proposé deux premiers ensembles de variables à récupérer de manière systématique. Ces formats seront à confronter à l'expérience de leur utilisation, mais également des utilisations désirées par d'autres équipes. Le premier ensemble est l'ensemble que nous avons jugé comme médicalement et techniquement minimal, et a un but général, non spécifique d'une disci-

FIGURE 4.8 – Architecture générale de SPOT RISC



plaine particulière. D'une certaine façon il recoupe les observations routinières effectuées en milieu hospitalier pour caractériser le point de fonctionnement d'une personne, physiologiquement parlant. Parmi les données structurantes, certaines ne sont pas médicales, mais à vocation d'identification, notamment temporelle, des mesures effectuées.

Les données recueillies sont les suivantes :

Données d'identifications temporelle et personnelle : Id unique par patient, et des time stamp doubles, comportant la date et l'heure d'une mesure ou d'une observation. Les dates d'hospitalisation peuvent être retenues, mais sont relativement peu fiables, et davantage contextuelles.

- Données cliniques : date de naissance ou âge au moment de l'examen, température en degrés Celsius, pressions artérielles systolique et diastolique en mmHg, fréquence cardiaque en battements par minute, la

FIGURE 4.9 – Architecture et départements en charge de SPOT RISC

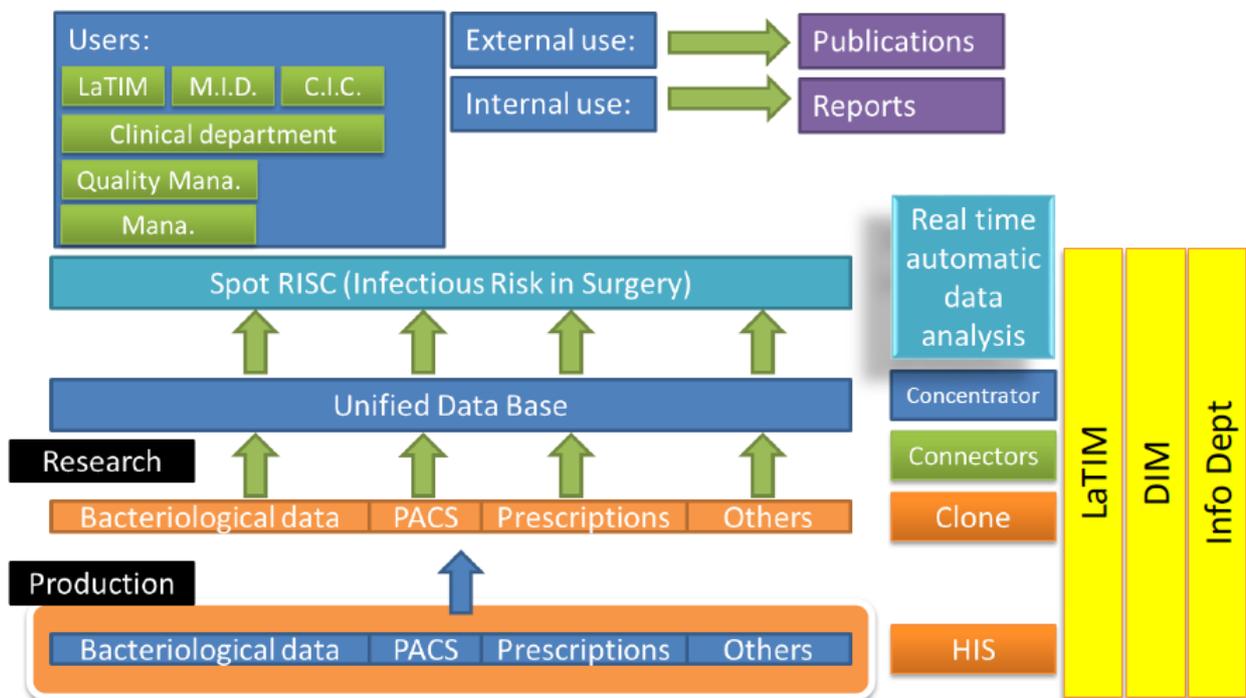
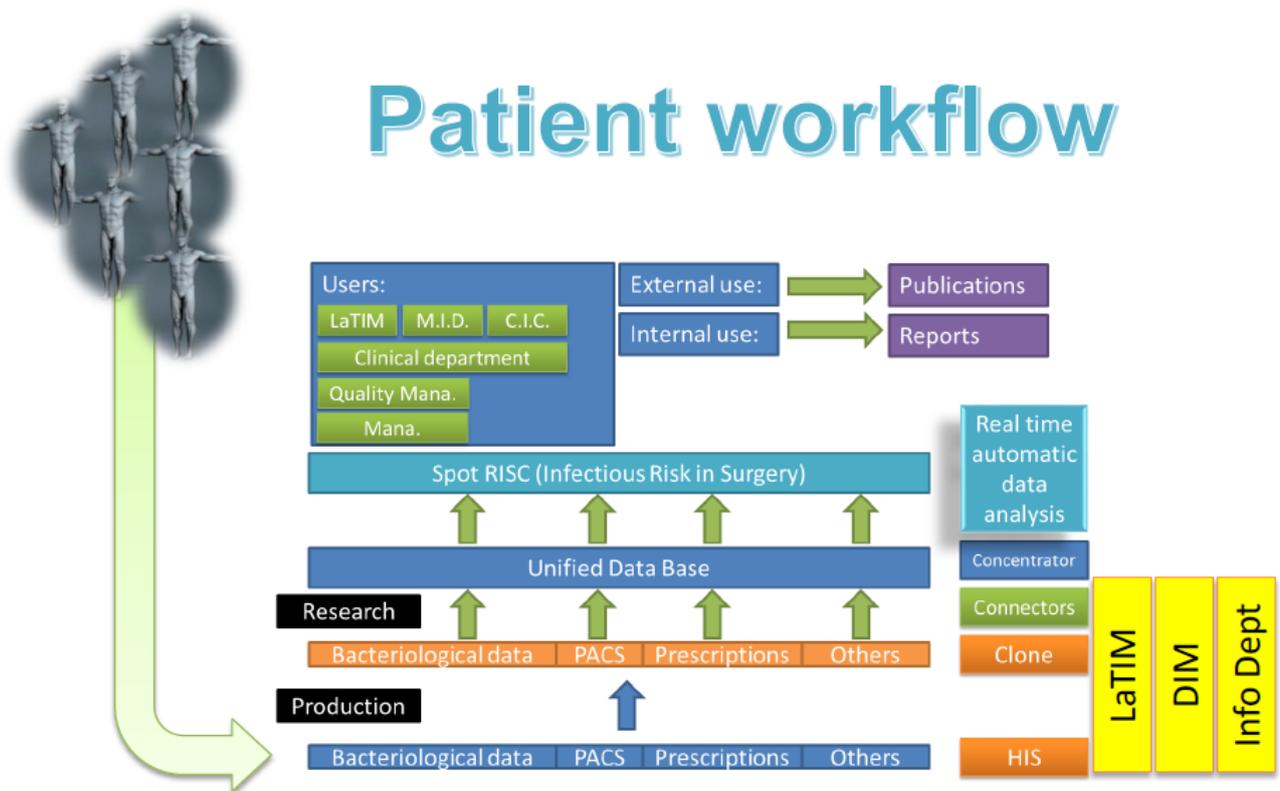


FIGURE 4.10 – Principe d'alimentation en données de SPOT RISC



saturation en oxygène en pourcents, la glycémie en grammes par litre, le poids en kilogrammes et la taille en centimètres.

- Données diagnostiques : antécédents, codés en CIM-10, diagnostiques contemporains de l'hospitalisation ou de l'examen, primaires et secondaires si existants. Allergies connues : oui/non et mention des allergisants.
- Données thérapeutiques : traitements médicamenteux (classes) à l'entrée, prescrits pendant l'examen ou l'hospitalisation – les posologies seront codées dans des versions ultérieures ; traitements non médicamenteux (chirurgie oui/non, et type de chirurgie) ; pose de dispositif intracorporel (oui/non, type d'implant).
- Données d'examen clinique : présence d'une douleur (oui/non, cotation, localisation), présence d'une limitation fonctionnelle (impotence fonctionnelle : oui/non, localisation), état de conscience (score de Glasgow, ou présence d'items permettant le score de Glasgow).
- Données biologiques : données récupérées de la numération de la formule sanguine (hématies, leucocytes, et les différentes mesures et morphologies), albuminémie, urémie, créatininémie, CRP, VS, transaminases hépatiques, Gamma-GT, phosphatases alcalines, bilirubinémie, données de gaz du sang, troponines, D-dimères, Beta-HCG, données du ionogramme sanguin (calcémie, natrémie, chlorémie, kaliémie, bicarbonatémie), données de bilan de coagulation (TP, TCA), dosages thyroïdiens (TSH-US, T4, T3), fer, ferritinémie, vitamine B12, Hba1c.
- Données d'imagerie : données issus du compte rendu (type d'examen, présence d'anomalies ou de mesures standardisées – par exemple, la fraction d'éjection ventriculaire gauche, FEVG, en échographie cardiaque).
- Données de microbiologie : type d'examen (ECBU, ECBC, écouvillon, ablation de matériel...), présence et type d'un agent infectieux (bacté-

rien, fongique, viral, parasitaire – nom d'espèce), quantité normalisée de l'agent présent (par exemple, en CFU), données de l'antibiogramme associé (sensibilité aux anti infectieux par type de germe).

- Données sérologiques : VIH 1 et 2, antigénémie p24, TPHA-VDRL, antigènes et anticorps virus de l'hépatite B et C, PCR (virus de la grippe).

Le second ensemble est plus spécifique de l'activité et de l'organisation de l'activité concernant la maîtrise des risques infectieux en milieu chirurgical, et plus spécifiquement encore, des IOA. Il comprend en plus des données minimales générales, la mention des mesures effectuées dans le cadre du CDC.

- Données de CDC IOA : amplitudes articulaires (localisation de l'articulation concernée, type d'amplitude mesurée, mesure de l'amplitude), matériel prothétique (présence oui/non, type de matériel), mesure de qualité de vie.
- Données de prise en charge anti infectieuse : traitement anti-infectieux (voie d'administration, posologie, heure et vitesse d'administration).
- Données biologiques : bilan phosphocalcique, dosage de la parathormone, interleukines.

Certaines de ces données sont dès l'origine structurées, ou du moins codées, ce qui permet de les récupérer sans traitement intermédiaire ou minime, et de les affecter de façon univoque au dossier reconstitué. D'autres sont accessibles et récupérables au niveau de données textes, essentiellement des données de compte rendu (compte rendu d'hospitalisation, d'imagerie...). De fait, en l'état actuel, il est possible de récupérer l'information si l'on dispose d'un moteur de fouille de texte (*text mining*). Les termes souhaités sont repérés, et leur présence matérialisée par un codage adapté au niveau du dossier reconstitué. Par exemple, les symptômes dits subjectifs, comme la

douleur, peuvent être mentionnés, voire même quantifiés (*via* des échelles), mais n'apparaîtront au mieux que dans le texte; il est ainsi possible de coder la variable douleur en O/N s'il est trouvé mention d'une douleur, et d'accompagner cette variable d'une possible cotation de 0 à 10, matérialisant son intensité. Le même traitement peut être réservé à l'évaluation des déficits sensitivo-moteurs, à la présence de signes inflammatoires, d'une rougeur, etc.

Couche de traitement des données.

Les moteurs de traitement de l'information sont essentiellement basés sur les packages logiciels libres d'utilisation de la plateforme R. Les moteurs sont modulaires, même si un noyau de base sera conçu et nécessaire à l'exécution correcte des modules additionnels (il existe des interdépendances entre modules, afin de ne pas recoder et alourdir les modules en code redondant).

- Module d'extraction et d'échantillonnage, d'exportation des données. Pas d'exhaustivité recherchée, mais du quasi-temps réel, en décharge du réseau primaire de soins, sans accès en écriture et clonage des informations. Approche par échantillonnage aléatoire suffisant et représentatif. Accès web, et intégration de moteurs de traitement de l'information libres de droit.
- Module de pré traitement des données en haute dimensionnalité : Techniques de traitement de données de haute dimensionnalité : techniques de réduction, techniques bayésiennes.
- Module de traitement statistique classique et de tests d'hypothèses : Techniques classiques d'épidémiologie : comparaisons de moyenne et proportion, modèles de régression et modèles de survie. Modèles multiniveaux, calculs d'effectifs et de précision.
- Module de détermination de modèles prédictifs : Techniques pour la médecine prédictive personnalisée : techniques d'apprentissage supervisées ou automatiques, assises sur la connaissance engendrée par la

pratique mono et multisite.

- Module de détection d'anomalies et d'événements rares : Techniques de surveillance d'effets indésirables ou d'événements rares : séries temporelles, détection d'anomalies.

Nous glissons donc d'approches classiques – conservées, notamment dans une perspective de prototypage et d'aide à la conception d'études ciblées et contrôlées, autrement basées que sur la seule littérature ou intuition – vers des approches par la forme : détection de motifs dans les données, sans hypothèse *a priori*, mobilisation de techniques hautement prédictives et non explicatives.

Couche d'interface homme machine. Construite selon les cas d'usage envisagés, et selon les profils d'utilisateurs, aux droits et types d'usage prévus différents.

Actuellement, et à notre connaissance, il n'existe qu'un seul véritable concurrent à ce projet, déjà disponible et utilisé en différents endroits du monde. Ce projet est le projet i2b2, pour *Informatics for Integrating Biology and the Bedside* ; il est soutenu par le NIH, et a démarré aux États-Unis en 2004, et piloté par Harvard. Le logiciel est téléchargeable gratuitement, et le déploiement du système mène à de plus en plus de publications, que ce soit pour expliciter de nouveaux composants logiciels ou traitements, ou pour exposer le résultat de recherches basées sur l'exploitation secondaire des données enregistrées en routine. [277, 278, 279, 280, 281, 282] Il a été montré entre autres l'intérêt d'un tel système pour identifier et constituer des cohortes ou prototyper des essais. [283]. Actuellement, les efforts se portent notamment sur l'utilisation de données temporelles, à partir d'informations amorphes, ce qui est effectivement un défi, ou sur la surveillance d'indicateurs de santé, à partir de données distribuées. [284, 285]

4.8 Retour sur les IOA

L'une des idées qui ont présidé à la conception, et à la réalisation, qui débute, d'un système tel que SPOT-RISC, était que certains événements étaient suffisamment courants ou graves médicalement parlant pour ne pas être ignorés, mais difficilement enregistrables ou étudiables *via* les outils statistiques classiques ; c'était notamment le cas des IOA, qui présentent des cas de figure à la variabilité inter et intra individus très importants, notamment par rapport au nombre de cas. On peut difficilement se retrouver dans pire situation en statistiques.

La capture et l'utilisation secondaire de données enregistrées en routine doit permettre l'étude plus systématique et multimodale de phénomènes du type des IOA. Nous avons proposé un abord des IOA par les systèmes dynamiques, approche qui, si elle n'est pas dépourvue d'intérêt, demeure actuellement encore délicate à calibrer. Nous proposons, dans un futur proche, et comme démarche prototypique d'autres phénomènes, d'utiliser une approche hybride, entre modélisation et exploitation des données réelles, entre explication et prédiction. Pour ce faire, nous proposons notamment l'utilisation des modèles dits multi-agents, ou encore *Agent Based Models* – ABM. [286]

4.8.1 Modélisation ABM pour la spatialisation

Si l'on se souvient et se réfère à la description initialement faite en termes de systèmes dynamiques pour les IOA, on comprendra comment il est possible d'utiliser la modélisation ABM pour mieux comprendre les IOA. Si l'on se place à un niveau cellulaire – les cellules osseuses, de manière équivalente, les cellules de l'immunité ou les bactéries – les agents, autrement dits, les éléments génériques à prendre en compte sont précisément les cellules, lesquelles peuvent interagir *via* des règles à spécifier, des messages particuliers.

La modélisation par agents tient de la programmation objet, en informatique : on définit un agent, un objet donné (une cellule osseuse type),

lequel possède des attributs généraux (une taille, un type de déplacement, des messages à envoyer, des messages à recevoir, un métabolisme, une durée de vie), ainsi qu'un comportement qui peut être décrit, cette fois encore, par un automate. Tous les agents à étudier sont donc définis de manière générale.

L'étape suivante tient alors de la simulation. Il est instancié autant d'agents que l'on désire, d'une classe donnée, et dont les caractéristiques particulières peuvent être tirées selon une loi de probabilité – par exemple, la taille peut être tirée selon une loi normale, de moyenne m et de variance σ . La simulation elle-même permet alors d'observer l'évolution du système, en le laissant "vivre" selon les règles d'interaction qui ont été spécifiées à une échelle locale. On voit ici que le parallèle, ou la transition, entre système dynamique et ABM se fait assez aisément en se référant aux modèles plus spécifiques de réaction-diffusion de Turing : des règles chimiques sont spécifiées localement, et un processus de diffusion tient lieu de loi de déplacement. La modélisation ABM permet bien sûr d'implanter ce type de modèle, à la Turing, mais également de s'affranchir des contraintes d'une formulation analytique ou continue.

Ainsi, un des points forts de l'approche ABM est de pouvoir prendre en compte la spatialisation, dans des configurations à la fois aléatoires, et précises, spécifiques, ce qui est en général plus délicat en système dynamique. Bien entendu, comme pour toute simulation, on pourra objecter au fait que l'on observe "ce que l'on a introduit". Cette remarque demeure toujours aussi limitée, dans le sens où, par hypothèse, tout modèle définit un certain nombre d'hypothèses et de mécanismes. Cependant, il n'est stipulé ici que des règles locales, et il devient difficile de faire le procès direct de l'observation "de ce que l'on voulait voir", s'il émerge des comportements collectifs ou des motifs à une échelle autre que l'échelle de spécification, et qui ne soit pas spécifiquement encodés à la base.

Un autre écueil est celui de la particularité, du "minimum local" d'une

simulation donnée. Il est évident qu'il n'est pas possible de se limiter à une itération unique, aussi représentative soit-elle – ou pense-t-on qu'elle soit. Il est possible de répéter les simulations autant de fois que voulu, et d'enregistrer les paramètres de chaque agent, notamment en fonction du temps, ou de surveiller la survenue d'événements particuliers. De la même façon, comme on le ferait pour une analyse de stabilité structurelle en système dynamique, il est possible de procéder à diverses itérations du même système, mais en faisant varier des paramètres de celui-ci – par exemple, le nombre initial de bactéries, leur virulence, la porosité d'un matériau, etc.

Ces approches sont en plein essor, et il demeure un certain nombre de problèmes pratiques à résoudre. Les plus prégnants sont les suivants : la prise en compte native des trois dimensions (la majorité des plateformes ne gèrent réellement que la simulation en deux dimensions), la gestion de larges quantités de données générées, et l'exploration, la représentation et la synthèse des résultats recueillis. [287, 288, 289, 290]

Pour ces dernières questions, les approches déployées en *big data*, en fouille de données, autrement dit, en exploitation de données observées restent tout-à-fait mobilisables, et ne le sont que peu en pratique, jusque présent. On peut notamment penser à utiliser des algorithmes de recherche et de reconstruction des attracteurs d'un système : si les simulations sont correctement enregistrées, rien ne s'oppose à ce que l'on y recherche de tels objets.

Il existe un certain nombre de plateformes de programmation ABM. Pour la poursuite des travaux esquissés ici, qui donneront lieu à un doctorat au sein de l'équipe, les deux plateformes envisagées pour la modélisation ABM des IOA seront NetLogo,[288] qui n'a plus à faire ses preuves, mais nécessite de repousser deux limites (le nombre d'agents modélisables, et la prise en compte de la 3D), ou Flame, qui est un projet open source beaucoup plus récent, issu d'un appel à projet européen FP. [291] Ce dernier gère native-

ment la 3D et semble permettre d’instancier un nombre beaucoup important d’agents simultanément. Cependant, il ne bénéficie pas de l’ergonomie, de la rapidité de programmation ni de l’appui et de la reconnaissance de NetLogo.

4.8.2 Intégrations multi-échelle, multimodale, patient-spécifique

Le but ultime de l’utilisation de SPOT-RISC pour l’enregistrement et l’analyse des IOA, et de l’approche ABM, est de proposer l’intégration d’un moteur ABM dans SPOT-RISC, bâti pour la simulation et la prédiction de l’évolution des IOA, et ce de manière multi-échelle, multimodale et basée tant sur l’expérience acquise des patients connus que sur les données spécifiques du patient à traiter.

Le point de départ de la modélisation ABM est le niveau cellulaire, ou à tout le moins, d’ensembles cellulaires, et des tissus les hébergeant. Nous sommes donc à une échelle relativement basse. Afin d’étudier les répercussions biomécaniques, mais également la progression des bactéries dans un milieu respectant la géométrie osseuse, un partenariat est en cours avec une université portugaise, leader européen en *tissue engineering*, et pouvant nous fournir et des images à l’échelle nécessaire des tissus, et les caractéristiques biomécaniques de ce tissu, à diverses étapes de la colonisation bactérienne. Ces résultats auront pour but de nous aider à construire notre premier modèle d’IOA.

Nous remonterons alors en échelle, une fois que l’on aura maîtrisé le premier niveau, jusque parvenir à l’échelle de l’os en tant qu’objet, au moins partiel, mais macroscopique. Nous utiliserons alors les paramètres patients, notamment les données d’imagerie afin d’injecter au modèle la géométrie de l’organe spécifique, mais également les données microbiologiques, entre autres. Le but étant ainsi d’estimer au mieux la stratégie thérapeutique.

4.9 Ce que l'on en retient

Le modèle expérimental a ses intérêts, mais présente des limites indépensables, notamment en termes de production de savoir réaliste et pratique, efficient. L'approche causale isolée et linéaire, ponctuelle, ne saurait suffire. Il existe d'autres approches possibles, basées sur l'analyse de données, et surtout, la caractérisation dimensionnelle adéquate du problème à étudier. Il existera un compromis entre volonté de prédire et volonté d'expliquer, entre complexité et efficacité. Les approches doivent également être pensées comme intégratives au minimum, dès le départ : même les approches non linéaires visant à reconstruire les dynamiques complexes par articulations de sous systèmes présentent rapidement une forte complexité, et sont certainement très fragiles.

Chapitre 5

Cadre théorique pour une morphodynamique basée sur la géométrie et l'information

Sommaire

5.1	Quels principes transversaux ?	335
5.1.1	Plan général de la démarche	335
5.1.2	Principes de Fermat et de Maupertuis	339
5.1.3	Principe de moindre action, formulation lagrangienne	341
5.1.4	Principes dynamiques	343
5.1.5	Principe de relativité	344
5.2	Causalité et temps – forme fonctionnelle	345
5.2.1	Structure causale d'un ensemble d'observations échantillonnées	347
5.2.2	Temps interne et mesure des événements causaux	348
5.2.3	Utilisation d'automates informationnels – analyse bayésienne	349
5.3	Géométrie informationnelle	350
5.3.1	Formulation du problème général et enjeux	350
5.3.2	Approche relativiste	351
5.3.3	Forme générale du problème	352
5.3.4	Définition du système observé	352
5.3.5	Définition d'un GAD, type réseau Bayésien	352

5.3.6	Définition de l'espace informationnel associé au GAD	357
5.3.7	Définition de la métrique et de l'intervalle infinitésimal	361
5.3.8	Courbure de l'espace informationnel et origine de la dynamique géométrique	364
5.3.9	Hypothèse du lien entre courbure de l'espace et contenu de l'espace	367
5.3.10	Une autre vision : de l'information locale vers l'Hamiltonien du système	369
5.3.11	Hamiltonien multi-événements ou multi-temps	373
5.4	Aspects opérationnels et problèmes ouverts	374
5.4.1	Reconstruire l'espace informationnel : une approche algorithmique	374
5.4.2	Formulations alternatives basées sur l'action	379
5.4.3	Intégration de l'échelle	380
5.4.4	Quelle interprétation du concept relativiste dans le cadre de la géométrie informationnelle?	381
5.4.5	Le temps physiologique de West	382

5.1 Quels principes transversaux ?

5.1.1 Plan général de la démarche

De l'intérêt de rechercher un Hamiltonien caractérisant un ensemble d'observations (resp. de données) : nous avons précédemment passé en revue 3 grands cadres théoriques et leurs formulations mathématiques associées. Actuellement, il n'existe aucun cadre réconciliant totalement, en un seul formalisme, ces 3 approches.

La géométrie de la relativité générale n'est pas compatible avec la structure d'espace plat de la mécanique quantique. La relativité générale s'accommode mal de la quantification de la mécanique quantique. Quant au chaos émergent des systèmes dynamiques, il n'existe pas à proprement parler en mécanique quantique.

Parmi les points communs, ou les ponts que l'on s'efforce de construire pour rapprocher ces théories, on peut citer : la généralisation de grandeurs formelles, comme l'action (intégrales de chemin en mécanique quantique, par exemple, introduites par Feynman, et qui a donné lieu à la théorie quantique des champs, puis par extension aux théories de type théorie des cordes ou théorie M), les grandeurs héritées de la mécanique analytique, telle que le Lagrangien ou l'Hamiltonien. La géométrisation des théories est également un biais d'approche, la relativité étant considérée comme principalement une théorie géométrique – la géométrie différentielle est devenue depuis une approche théorique qui se généralise pour la mécanique classique. La quantification pose également problème, et selon le schéma persistant d'une dualité onde corpuscule, revisité à l'aulne de la théorie de la communication, il a été proposé qu'une bonne candidate à l'unification des théories passe par le postulat d'une réalité qui soit aussi bien continue que discrète, *via* un possible théorème d'échantillonnage généralisé – et des grandeurs limites, type longueur de Planck pour la plus petite échelle du réel. [292, 293, 294, 295, 296, 297] Un autre point commun de ces 3 approches repose dans la non linéarité des relations. La métrique de la relativité générale est réglée par la distribution de masses, et évolue dynamiquement avec cette distribution, en interaction. Les relations liant les éléments infinitésimaux de l'espace n'ont pas à être linéaires en toute généralité. La mécanique quantique repose certes sur un formalisme à base d'espaces vectoriels, de méthodes projectives et d'algèbre linéaire – de type analyse spectrale – il existe cependant un saut qualitatif dans le concept apparemment simple d'interférences, comme en optique : les observations réelles sont basées sur des relations d'ordre 2 en les amplitudes de probabilités. Là aussi, le phénomène même d'intrication est un phénomène débordant le principe de simple superposition et est non linéaire. Enfin, la richesse des systèmes dynamiques repose sur le couplage des équations, et la formulation non linéaire de ces

équations.

Néanmoins, quelle que soit la théorie, il a été montré qu'un Hamiltonien pouvait être construit selon des hypothèses plus ou moins strictes dans chaque cas. En premier lieu, la formulation de la mécanique quantique repose grandement sur l'Hamiltonien du système. Celui-ci apparaît directement dans l'équation de Schrödinger, et règle pour ainsi dire l'évolution possible d'un système. La mécanique hamiltonienne a également prouvé sa capacité à exhiber des comportements chaotiques, et est une formalisation particulière des équations différentielles régissant le système. [298, 299] Inversement, il a été montré qu'il était possible, pour un système dynamique quelconque, de lui associer une métrique type Riemannienne, et donc de plonger les équations du système dans un espace Riemannien.

Le problème qui survient alors est de déterminer, d'une manière ou une autre, la forme de l'Hamiltonien d'un système. En effet, en dehors de systèmes aux interactions bien définies, et en général relativement simples, il est difficile sinon impossible de donner une expression analytique exacte de l'Hamiltonien du système. Si nous disposons d'un moyen général de dégager l'Hamiltonien d'un système quelconque, alors certaines indéterminations quant à sa forme plus précise peuvent être levées en choisissant préférentiellement un cadre théorique plutôt qu'un autre – par exemple, choisir un cadre quantique plutôt que relativiste.

La préoccupation d'une science est de dégager des objets identifiables et reproductibles selon des conditions plus ou moins bien identifiées. Il lui est donc demandé de savoir dégager des invariants sous un certain nombre de transformations. Le principe de relativité, notamment, ne stipule pas naïvement que tout peut être vrai, et que tout dépend de l'observateur – de la subjectivité – mais dérive bel et bien de la nécessité pour une loi "universelle" de présenter la même forme pour tout observateur. Les conséquences

de cette universalité amènent justement la relativité des mesures effectuées, et non des principes régissant le possible.

Notre démarche répond de ces préoccupations, et rend compte d'une contrainte supplémentaire, à savoir le mode de représentation du temps et de la causalité. Supposant qu'il n'existe *a priori* pas de temps universel, mais des fonctionnements universels et reproductibles, capables d'enregistrer et de poser des actions selon des séquences élémentaires possibles ou probables et d'autres qui ne le sont pas, nous prenons le parti de caractériser un ensemble d'observations par les relations qu'entretiennent ces observations, des relations de type causal et possible. Le temps de ce système est ainsi un temps subjectif, propre aux liens unissant les observations. Il est alors postulé que le schéma global caractérisant ces observations est un invariant du système – en d'autres termes, une extension du principe de loi physique classique. Ce schéma invariant est un automate informationnel et causal.

Nous proposons donc de créer ou de composer avec un Hamiltonien rendant compte de cet automate. Dans un cadre davantage orienté systèmes dynamiques et relativistes, nous associons un espace géométrique à cet automate. La métrique générale de cet espace est déterminée par l'automate. La nature de cet espace géométrique est au minimum une variété différentielle, voire pour la construction rigoureuse d'un Hamiltonien multi-temps, un fibré de jets $J^1(T, M)$, T et M étant eux-mêmes des variétés différentielles.

Dans un cadre relativiste, la détermination complète et dynamique de cette métrique est liée à la distribution de l'information. Cette information peut être présente ou non, et à des cadences relatives différentes. L'automate règle la possibilité d'utiliser cette information, ainsi que les conséquences en découlant. De manière plus générale, et selon un cadre d'interprétation à la Turing, l'espace informationnel associé à l'automate peut être vu comme un espace pour lequel, en tout point, il existe un automate ponctuel lisant de l'information et pouvant en écrire ; son état courant peut changer selon

l'information reçue.

La formulation qui consiste en un système multi causal, plutôt que réglé selon un temps unique et uniforme, implique la prise en compte de plusieurs variables tenant lieu de temps – des variables d'événements. Les théories classiques ne prennent en compte qu'une variable temporelle, et c'est également le cas de la relativité générale. Mathématiquement parlant, il est donc nécessaire de se doter d'outils permettant de rendre compte de plusieurs dynamiques de type temps. Le calcul multifractal ou calcul fractionnaire aborde cette problématique selon l'angle d'échelles de temps multiples. D'un point de vue formel, des développements récents quant à la géométrie différentielle et centrée sur les formulations lagrangiennes ou hamiltoniennes indiquent qu'il est possible d'écrire un Hamiltonien qui soit explicitement multi-temps. L'extension des cadres relativistes classiques peut ainsi passer par ces techniques, et nous servir de fondations pour l'écriture d'un Hamiltonien quelconque, mais multi-temps.

Enfin, une interprétation essentiellement géométrique de l'espace construit et muni d'un Hamiltonien ou assimilé, passe par l'application d'un principe de moindre action. L'application d'un principe de moindre action permet de dériver les équations du mouvement, ou autrement dit, d'explicitier les variations possibles d'information en fonction des valeurs des objets géométriques considérés.

5.1.2 Principes de Fermat et de Maupertuis

Principe de Fermat

Le principe de moindre action est un principe d'extrémalité qui traverse la physique et le temps avec un beau succès. Il a été rencontré dans divers domaines plus spécifiques, avant que ne lui soit donné sa forme générale, variationnelle, portant sur la grandeur que l'on appelle action. C'est donc un principe d'économie, qui veut que parmi toutes les possibilités, la nature

choisirait celle qui coûte le moins, la plus directe.

En optique géométrique, par exemple, on peut rencontrer le principe de Fermat, lequel affirme que, dans un milieu isotrope, le chemin suivi par la lumière entre deux points fixes P et Q est celui qui rend minimal le temps de parcours, c'est-à-dire qui rend extrémale l'intégrale

$$\int_{P \rightarrow Q} dt = \int_{P \rightarrow Q} \frac{ds}{c},$$

où ds est un élément infinitésimal, une infime portion du chemin allant de P à Q , c est la célérité de la lumière, et $\frac{ds}{c} = dt$ est l'intervalle infinitésimal de temps nécessaire pour parcourir l'intervalle ds .

Le principe variationnel s'écrit alors en assurant qu'une petite variation de cette intégrale est nulle, la valeur retenue étant extrémale :

$$\delta \int_{P \rightarrow Q} \frac{ds}{c} = 0$$

où δ signifie ici une variation très petite autour du trajet retenu pour relier P à Q .

Principe de Maupertuis

Plus généralement, on peut rencontrer en mécanique analytique le principe de Maupertuis, lequel est valable pour tous les systèmes dits conservatifs ; autrement dit, les systèmes dont l'énergie totale $E = T + V$ ne varie pas dans le temps.

Cette fois, le principe d'extrémalité portera sur l'énergie cinétique T au cours du temps, et permettra de déduire les équations du mouvement du système, sa trajectoire en fonction du temps. Ainsi, pour nos deux points P et Q , la trajectoire respectant ce principe d'extrémalité est celle qui rend extrémale l'intégrale curviligne :

$$\int_{P \rightarrow Q} 2T dt = \int_{P \rightarrow Q} m v ds,$$

où ds est un élément infinitésimal de trajet, $2T = m.v^2$ et $v = \frac{ds}{dt}$.

Néanmoins, cette formulation est théorique, car il est rare de pouvoir écrire simplement et explicitement, voire analytiquement la vitesse v en fonction de la coordonnée s ou du temps t , pour un système quelconque.

On peut écrire l'énergie totale du système $E = T + V$, où $V = V(s)$ représente l'énergie potentielle du système, et ainsi déduire $mv = \sqrt{2m(E - V)}$. L'intégrable extrémale devient :

$$\int_{P \rightarrow Q} m v ds = \int_{P \rightarrow Q} \sqrt{2m(E - V)} ds$$

Le principe de Maupertuis, de manière analogue au principe de Fermat, s'écrit alors $\delta \int_{P \rightarrow Q} \sqrt{2m(E - V)} ds = 0$.

5.1.3 Principe de moindre action, formulation lagrangienne

Considérons pour simplifier un point matériel décrit, dans un repère galiléen, par un seul degré de liberté, noté $q(t)$ à l'instant t . Le Lagrangien L est la différence entre l'énergie cinétique (du point matériel) et l'énergie potentielle :

$$L = T - V$$

$$L(q, \dot{q}, t) = \frac{1}{2} m \dot{q}^2 - V(q, \dot{q}, t)$$

où $\frac{1}{2} m \dot{q}^2$ est l'énergie cinétique du système et $V(q, \dot{q}, t)$ est l'énergie potentielle, qui en général ne dépend pas explicitement de \dot{q} .

L'action de la trajectoire, étant la somme totale de la différence entre

l'énergie cinétique et l'énergie potentielle durant la trajectoire, est définie par :

$$S = \int_{t_i}^{t_f} L(q(t), \dot{q}(t), t) dt$$

où t_i et t_f désignent respectivement l'instant initial et l'instant final.

Équations d'Euler-Lagrange Les conditions d'extrémalités d'une fonctionnelle peuvent se formuler selon ce que l'on appelle les conditions ou les équations d'Euler-Lagrange. Concernant les conditions d'extrémalité, appliquée à la notion d'action, il est intéressant de comprendre que :

- l'obtention d'un chemin minimisant localement l'action implique que tout autre chemin, de mêmes conditions initiale et finale, exhibera une action de plus grande valeur ;
- étant données certaines conditions initiales ou finales, il est possible qu'il n'existe pas de chemin minimisant localement l'action. Cependant, si ce chemin existe, alors il est unique, car déterminé par des conditions initiales non équivoques et le caractère continu du mouvement.

Notons par ailleurs que le Lagrangien n'est pas défini de manière unique. Il est possible de lui adjoindre une fonction $f(q, \dot{q}, t)$ telle que $f(q, \dot{q}, t) = \frac{dF}{dt}$. Dans ce cas, il est ajouté à l'action une fonction $F(t_{final}) - F(t_{initial})$. Cette fonction ne dépend que des extrémités ; par conséquent, en termes relatifs, elle s'annule quand on fait varier l'action par rapport au chemin.

Le chemin $q(t)$ effectivement suivi par le point matériel entre les instants t_i et t_f fixés est un extremum de l'action. Une petite variation de l'action est donc nulle autour de ce chemin, ce qui se traduit par :

$$\delta S[q] = 0$$

L'équation d'Euler-Lagrange correspondant à cette condition est alors la suivante :

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{q}}(q, \dot{q}, t) - \frac{\partial L}{\partial q}(q, \dot{q}, t) = 0$$

Cette équation fournit donc en toute généralité un moyen formel de définir un principe de moindre action à partir d'un Lagrangien.

5.1.4 Principes dynamiques

L'application d'un principe d'extrémalité sur une grandeur appelée action, cette action s'écrivant en fonction du Lagrangien, amène à l'écriture des équations d'Euler-Lagrange. Nous avons dit que l'application d'un tel principe permettait de dériver les équations du mouvement – autrement dit, qu'il permettait de dégager un principe dynamique à partir d'une grandeur identifiée. Nous en montrons une illustration dans un cas classique.

Supposons ici que l'énergie potentielle ne dépende explicitement que de q et de t : $V(q, \dot{q}, t) = V(q, t)$. On obtient :

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{q}}(q, \dot{q}, t) = m \cdot \ddot{q}$$

Les équations d'Euler-Lagrange s'écrivent alors :

$$m \cdot \ddot{q} = - \frac{\partial V}{\partial q}$$

On retrouve ainsi les lois du mouvement de Newton avec :

$$\Sigma (\text{forces extérieures}) = - \frac{\partial V}{\partial q}$$

Cette réécriture des lois de Newton selon le Lagrangien et le principe de moindre action, telle que proposée ici, ne fonctionne bien entendu *a priori* que si toutes les forces appliquées au système dérivent d'un potentiel V . L'application à d'autres théories de la même démarche permet de dériver de manière générale les équations du mouvement, au sens large. C'est no-

tamment ce qui permet de déduire l'équation des géodésiques en relativité générale.

5.1.5 Principe de relativité

À la base des théories relativistes, notamment de la relativité générale, existent les principes d'équivalence et de relativité. Le principe d'équivalence stipule que l'effet de la gravitation peut être localement représentée de manière équivalente par une distorsion de l'espace, donc un effet géométrique. Le principe de relativité, qui donne à entendre en sociologie ou autres sciences humaines et sociales des significations qui prêtent à contre-sens, met en relation observateurs et lois physiques. Plus précisément, il postule que les lois physiques, pour avoir un statut légitime de lois de la nature, doivent avoir la même forme pour tout le monde, et en particulier dans des référentiels dits inertiels. La notion de référentiel renvoie directement à une conception particulière de l'observateur : l'observateur a pour lui un système d'axes, un instrument (une règle) de mesure des distances spatiales, et un système de mesure du temps (une horloge).

Sous cette hypothèse, les lois physiques doivent conserver la même forme que l'on passe d'un observateur à un autre, d'un référentiel inertiel à un autre. Il faut donc que les relations de passage, les transformations permettant de déduire les mesures de l'un à partir de celles de l'autre respectent cette invariance de forme. C'est cette exigence qui a mené à la transformation de Lorentz que nous avons déjà évoquée. Les conséquences pour les mesures des différents observateurs en sont les transformations de dilatation-contraction du temps et des longueurs.

Il est nécessaire que ces notions soient claires, car le principe de relativité, rendu populaire par la formule "tout est relatif", n'est pas ce que l'on appelle régulièrement un relativisme – relativisme sévèrement attaqué en général. Le principe de relativité ne consiste pas à dire qu'il n'existe aucun objet stable

et possiblement consensuel, que tout est possible, une assertion comme son contraire, selon l'observateur. C'est à l'opposé l'affirmation qu'il existe des fonctionnements universels, et c'est la contrainte formelle engendrée par ce postulat qui amène la relativité de mesures élémentaires au niveau d'observateurs différents. Ces différences de mesures sont précisément les garantes de l'invariance du principe dynamique sous-jacent.

Plus généralement, la version formelle et élargie du principe de relativité est le principe de covariance : il implique que les relations de passage d'une base de représentation à une autre conservent la forme des équations fondamentales régissant les interactions.

Dans un tel cadre, il n'est ainsi absolument pas contradictoire de vouloir chercher à identifier des fonctions invariantes pour un certain nombre de classes d'observateurs, ceux-ci devant être définis, et à la fois de constater que l'identification de ces fonctions invariantes puisse impliquer des variations des mesures relatives. C'est justement l'existence et la forme de ces variations induites que nous désirons repérer sur la base d'un principe d'invariance fonctionnelle.

5.2 Causalité et temps – forme fonctionnelle

Selon nous, la recherche d'invariants fonctionnels nous ramène directement aux notions de causalité, de temps, et *in fine* à la définition de la forme. Au fond, ce qu'il nous intéresse de poursuivre, scientifiquement parlant, est un double but : nous désirons pouvoir nommer et identifier, ainsi qu'agir sur ce que nous nommons. Notre relation au monde est un aller-retour permanent entre ce que nous mesurons, comprenons selon notre référentiel personnel et social, et les actions que nous entreprenons en accord avec ces mesures et cette compréhension.

Ainsi, de la même façon que parler d'information dans l'absolu, au sens de théorie de l'information, n'a pas grand sens, contrairement à parler soit

de communication, d'interaction ou éventuellement d'information relative, il paraît difficile de parler d'objet ou de forme dans l'absolu sans parler des actions ou des facteurs que nous pensons associés à l'apparition, au maintien ou à la modification de cet objet ou de cette forme.

Cette présomption de lien intrinsèque entre forme et fonction est particulièrement présente dans un domaine comme l'anatomie, ou la mécanique. L'invention de la roue autorise de pouvoir rouler ; un cube peut certes dévaler une pente, mais il paraît moins adapté à cette fonction de manière pérenne ou "naturelle" ; sa forme ne semble pas à même de pouvoir minimiser le coût de ce type d'utilisation. Une articulation anatomique présente un minimum de congruence géométrique, afin qu'elle puisse jouer son rôle ; un des objectifs de l'orthopédie étant d'ailleurs de restaurer, initialement, au mieux la géométrie des os ou articulations sous l'hypothèse que cela rétablira au mieux, par transitivité, la fonction.

En un certain sens, l'existence (ce qui regroupe ici : l'apparition, le maintien et la transformation) d'une forme est liée à l'existence de sa fonction et inversement, ce lien tenant lieu, finalement, de lien de causalité : la forme peut causer une variation de fonction, tandis qu'une variation de fonction peut causer une variation de forme. Evidemment, à cette échelle, celle de la forme elle-même, cette relation de causalité circulaire mène à une voie sans issue. Il s'agit donc de se déplacer en échelle, et prendre le parti, par exemple, de considérer que la forme globale, de même que la fonction, peut se construire à partir de relations élémentaires entre mesures élémentaires ; cette construction pouvant être de nature linéaire ou non linéaire. Pour reprendre l'analogie relativiste, l'espace-temps peut être euclidien – plat, comme le globe terrestre - localement, et non euclidien – courbe, comme la géométrie globale de la terre – globalement. C'est par ailleurs là un principe de base du calcul différentiel.

De cette sorte, nous ouvrons la voie à une définition possible de ce que

nous appellerons une forme fonctionnelle.

5.2.1 Structure causale d'un ensemble d'observations échantillonnées

Nous étendons ici la définition d'un observateur et de référentiel à un nombre arbitraire de mesures, lesquelles n'ont pas à être de type spatial ou temporel au sens usuel du terme. Un observateur demeure ainsi caractérisé par un ensemble d'appareils de mesure au sens large, lesquels doivent délivrer, bien entendu, des mesures identiques toutes choses égales par ailleurs.

Considérant qu'il semble acquis qu'il existe une limite d'échelle indépassable aux phénomènes physiques, nous prenons le parti de considérer que toutes ces mesures sont des échantillons (dans l'idéal, bien entendu, récolté à une cadence adaptée). Nous considérons également qu'un observateur est entièrement caractérisé par ce qu'il est en état de mesurer : son espace se résume donc aux événements qu'il est capable d'enregistrer.

La question que nous posons alors, est la suivante : pour tous les observateurs de la même classe (i.e définis par les mêmes instruments de mesure), il existe des relations de type causal entre les différentes observations effectuées – quelle est la structure de ces relations ?

Cette structure définit le fonctionnement interne de ce qui est mesuré, et sert de support à la possibilité d'une ou de plusieurs formes. En effet, le monde d'un tel observateur ne se décrit qu'en termes des mesures enregistrables, à des occurrences ou débits relatifs différents. Pour que ce monde lui soit intelligible, et qu'il ait la possibilité d'y agir, des relations entre types d'observations doivent exister, structurant à la fois sa représentation du monde et réglant ses actions potentielles sur celui-ci.

Notons enfin que la notion d'échantillonnage doit nous permettre de nous affranchir ou d'élargir la conception réductionniste : la réduction en "atomes" d'information ne présume pas de leurs relations complexes ou non, ni du

mode de reconstruction du système à partir des échantillons. Plus rigoureusement, la manière d'échantillonner correctement, à l'heure actuelle, dépend du cadre théorique et donc des transformations prescrites par ce cadre théorique, mais ne se limite pas à une approche purement réductionniste.

5.2.2 Temps interne et mesure des événements causaux

Dans l'optique d'une caractérisation complète du monde d'un observateur par les mesures qu'il est capable de réaliser, à l'instar de l'approche relativiste, cela signifie que le temps est compté, enregistré de manière interne au monde, et que celui-ci permet d'ordonner les séquences d'observations. Si les observations se structurent selon des relations causales entre elles, cela signifie qu'un type d'observation est enregistrée si et seulement si un autre ou d'autres types d'observations – celles qui causent le premier type – sont enregistrées. Un observateur donné va donc pouvoir coder ses observations en séquences non totalement arbitraires ou non totalement aléatoires, puisque ordonnées par des observations causales, que nous nommerons événements plus loin. Ces événements fournissent la cadence des observations.

Un exemple simple serait d'imaginer la situation d'une personne placée dans une pièce telle que toute stimulation de ses sens soient abolie : pas de lumière, pas de son, pas de perception de l'espace. Rapidement, comme certaines expériences ont pu par ailleurs le montrer avec des chambres anéchoïdes et obscures, il devient impossible de s'orienter – voire même de conserver un équilibre. En réalité, il reste les bruits physiologiques comme les battements de coeur qui permettent d'obtenir une forme d'information rythmique. Supposons que ces bruits n'existent pas. Par moments, un signal lumineux ponctuel pourrait être activé. Ce signal deviendrait le seul étalon possible pour recréer un décompte temporel. Or, que ce signal soit émis à un rythme régulier ou un rythme totalement aléatoire, l'évaluation de l'intervalle à l'aide d'autres informations étant impossibles, ce temps sera pris

comme base homogène (i.e dont les intervalles sont par hypothèse réguliers) pour cet observateur privé d'autre repère. Pour un autre observateur, dans une situation quelque peu différente, la base de temps ne serait pas nécessairement la même. Nous faisons jouer le rôle du signal lumineux aux variables causales.

Puisque nous sommes dans une optique de causalité non nécessairement unique ou purement linéaire, il devient possible qu'il existe plusieurs observations de type événements. De fait, la base de temps peut devenir asynchrone et multi-cadence. Ce cas de figure paraît naturel pour n'importe quel système qui présente plusieurs échelles d'organisation : les signaux causaux cellulaires ne s'activent pas nécessairement aux mêmes rythmes que les signaux causaux physiologiques macroscopiques. Bien entendu, il existe pour ces organismes des relations inter échelles entre signaux causaux, et donc entre cadences différentes.

Reste à voir comment nous pouvons reconstruire la structure causale sous-jacente aux observations, et ainsi catégoriser les observations comme événements ou non.

5.2.3 Utilisation d'automates informationnels – analyse bayésienne

Ainsi que nous l'avons déjà mentionné, il est possible de recourir aux concepts de réseaux de causalité, tels que les réseaux bayésiens. En effet, ceux-ci fournissent l'automate informationnel qui règle la circulation de l'information entre observations possibles. De fait, l'observation d'un signal lumineux, vu comme base de temps, peut induire l'observation d'une variation concernant une ou plusieurs autres variables.

Nous proposons donc, pour un ensemble de données qui correspondent à autant d'observations que peut accomplir un type d'observateur, de représenter la structure informationnelle et causale de cet ensemble par un

automate informationnel, par exemple, un réseau bayésien. Quelle que soit l'utilisation que nous ferons de cet automate dans la suite, il servira de base commune.

5.3 Géométrie informationnelle

5.3.1 Formulation du problème général et enjeux

Le travail présenté ici repose sur un certain nombre d'hypothèses. Nous dénombrons de fait 5 grandes hypothèses structurant la démarche.

- i) on observe un système complexe par échantillonnage d'un certain nombre de ses variables. Ces variables présentent donc un caractère discret, et peuvent être représentées si besoin par leur variations relatives.
- ii) il existe un espace dit informationnel, représenté par une variété, et dont la métrique est déterminée par un graphe acyclique directionnel (GAD). Par conséquent, ce GAD fournit un intervalle infinitésimal caractéristique des données observées. Ce GAD est un automate dont on suppose qu'il décrit le comportement informationnel pour n'importe quel élément du système. C'est un invariant.
- iii) les observations se répartissent en 2 sous-espaces, l'un constitué des variables descriptives, l'autre des variables d'événements ou de comptage. Ces dernières sont des variables analogues à des variables de temps, en ce sens où elles mesurent au sein d'un système observé la survenue d'un événement observable. Ces sous-espaces sont des variétés différentielles.
- iv) la variation des probabilités conditionnelles selon l'information observée fournit la distorsion de la métrique.
- v) la métrique et ses distorsions permettent d'écrire la géométrie des données observées en fonction de l'information et de ses variations.

Des géodésiques peuvent donc être explicitées, et fournissent de fait un principe dynamique basé sur la géométrie.

Une des idées est de faire disparaître la notion de temps absolu, et de lui substituer la notion d'événements et de code séquence : la représentation par un automate invariant se prête à cette interprétation, puisque les variables descriptives ne varient que selon l'occurrence d'une variable de comptage. Ces variables représentent le "temps interne" du système. Dans le cas d'une interprétation type relativité générale, les distorsions de la métrique sont mises en relation avec le contenu en information de l'espace. L'expression "locale" de l'automate est alors remplacée par l'équivalent du tenseur d'énergie-impulsion, que nous appellerons tenseur informationnel. Notamment, ce tenseur fournit les débits de modifications des variables descriptives selon les variables de comptage. Cet aspect permet d'envisager une forme de programmation particulière du système : sachant l'impact sur l'espace informationnel d'une modification de son contenu localement, il devient possible, en théorie, de modeler géométriquement ce espace afin d'obtenir la forme locale ou globale désirée.

5.3.2 Approche relativiste

Cette partie présente la démarche construisant l'espace informationnel associé au système observé, partant de l'hypothèse que la forme de cet espace est une variété dont la géométrie, plus précisément la courbure, dépend du contenu local en information. De ce fait, nous serons amenés à utiliser une équation type équation d'Einstein pour la relativité générale.

Globalement, nous nous appuyerons sur deux références générales reconnues pour la majorité des définitions formelles rencontrées ci-après, pour la relativité [65], et pour la physique théorique [300].

5.3.3 Forme générale du problème

Dans l'approche dite relativiste, nous cherchons à obtenir ce que nous appellerons une *équation d'évolution de la métrique de l'espace informationnel*, dans l'hypothèse où l'espace qui nous intéresse se déforme selon son contenu en information. La forme générale de cette équation est alors :

$$f(g_{ij}) = T_{ij} \quad (5.1)$$

où f est une fonctionnelle non nécessairement linéaire, g_{ij} est la métrique de l'espace, et T_{ij} un objet décrivant le contenu local de l'espace en information.

5.3.4 Définition du système observé

Les systèmes ici considérés recouvrent n'importe quel système observable, i.e. dont on peut en mesurer des grandeurs, et ce de façon reproductible. Les systèmes sont des systèmes *a priori* réels, i.e. à énergie finie. Ce système est dénoté dans la suite S . À noter qu'il n'est pas imposé à un tel système d'être monobloc *physiquement* ou *apparemment*, en temps ou en espace. On qualifiera ce système de forme s'il s'en dégage en effet une structure causale entre ses différentes observations et réalisations. Une forme se définit ainsi par l'observation d'un ensemble possédant une structure logique causale. Tout l'intérêt du présent travail est alors de se donner les outils afin de définir et décrire la géométrie de cette forme dans un espace approprié.

5.3.5 Définition d'un GAD, type réseau Bayésien

Nota Bene : pour éviter de fastidieux allers-retours, à ce stade de la construction de notre cadre théorique, nous reprenons à l'identique la présentation qu'a été faite des GAD plus tôt. Il nous semble nécessaire de garder à l'esprit leurs hypothèses pour comprendre les constructions qui leur seront

associées.

L'idée ici est de représenter le système étudié par la distribution de probabilité jointe de l'ensemble des observations du système. La structure minimale de cette distribution est approchable *via* l'hypothèse d'indépendance conditionnelle, et permet de décrire la structure informationnelle du système par un graphe acyclique dirigé (GAD). De fait, les différentes observations faites sur le système, et les liens orientés associant ces variables permet d'interpréter le GAD comme un automate particulier.

Le "moteur" logique qui conduira à la construction de l'automate informationnel tient dans la formulation du théorème de Bayes, que nous rappelons ici :

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)} \quad (5.2)$$

Où A et B correspondent à deux variables aléatoires, et P dénote leur probabilité, $P(.|.)$ leur probabilité conditionnelle.

Données d'un graphe et d'une probabilité jointe décrivant le système

Soit l'ensemble O des observations du système S . O est constitué de n observations suivant p dimensions. Il est possible de considérer chaque p dimension comme une variable aléatoire discrète, avec n tirages. A partir de ces p variables, notées $\{X_i\}_i$, il est possible de représenter l'ensemble de S par la distribution de probabilité jointe $P(X_1, \dots, X_p) = P(S)$. Parmi les techniques permettant d'approcher cette distribution de probabilité, nous pouvons utiliser les techniques bayésiennes, en particulier l'approche des réseaux bayésiens.

Un réseau bayésien RB est la donnée de 2 éléments, à savoir un graphe G et la distribution de probabilité jointe sur S notée $P(S)$. On a donc par définition :

$$RB = \{G, P(S)\} \quad (5.3)$$

Le graphe G peut être dirigé ou non. Dans notre cas, nous nous munirons d'un graphe dirigé. On impose en outre à celui-ci d'être acyclique, c'est-à-dire de ne présenter aucune "boucle" au sein du graphe. Un tel graphe est appelé graphe acyclique dirigé (GAD). En outre, puisque nous construisons un graphe, nous nous dotons de la notation suivante afin de représenter l'arête joignant X et Y , X et Y étant deux variables observées, chacune représentées par un noeud du graphe :

$$X \rightarrow Y$$

Ce qui signifie ici que l'on circule de X vers Y et non l'inverse. Dans un GAD type réseau bayésien, l'orientation de l'arête symbolise le sens de la relation de causalité entre deux variables. Dans l'exemple précédent, il s'agit donc de considérer que la connaissance de X entraîne la connaissance de Y . Tout l'intérêt du réseau bayésien vient alors du fait de "squelettiser" la distribution de probabilité $P(S)$ du système, squelette qui sera donné par l'allure du GAD et ses arêtes. Cette squelettisation est possible grâce à la propriété d'indépendance conditionnelle, laquelle va réduire le nombre de relations point à point entre variables du système.

Notion d'indépendance conditionnelle

La propriété d'indépendance conditionnelle se définit assez simplement. Soient X , Y et Z trois variables du système étudié. On dit que X est indépendant de Z conditionnellement à Y si la distribution de probabilité de

X conditionnellement aux variables Y et Z est égale à la distribution de probabilité de X conditionnellement à Y . Autrement dit :

$$P(X|Y, Z) = P(X|Y) \quad (5.4)$$

On écrit la propriété d'indépendance conditionnelle de X par rapport à Z sachant Y de la façon suivante :

$$X \perp Z|Y \quad (5.5)$$

Par ailleurs, un réseau bayésien est doté de la propriété de Markovianité : chaque variable X du système est conditionnellement indépendante des autres variables qui ne sont pas ses descendants, sachant les parents de X . Voyons comment ces propriétés simples permettent d'écrire la probabilité jointe du système $P(S)$.

Ecriture de la probabilité jointe avec l'hypothèse d'indépendance conditionnelle

Suivant les propriétés d'indépendance conditionnelle et de Markovianité, la probabilité jointe $P(S)$ peut s'écrire :

$$P(S) = P(X_1, \dots, X_p) = \prod_{i=1}^p P(X_i|X_{i-1}, \dots, X_1) \quad (5.6)$$

En effet, X ne dépend que de ses prédécesseurs. Allons une étape plus loin : tous les X_{i-1}, \dots, X_1 ne sont pas nécessaires, si l'on considère que l'on peut s'en tenir qu'à la dépendance aux parents directs de X_i . Nous noterons ici les parents de X_i par $P(C(X_i))$, $C(\cdot)$ pour "cause" de. Nous pouvons alors écrire :

$$P(S) = \prod_{i=1}^p P(X_i|C(X_i)) \quad (5.7)$$

Automatiquement, nous nous apercevons que la structure "utile" de la probabilité $P(S)$ s'allège considérablement, et devient par ailleurs plus lisible et structurée. Cette structuration se fait en termes de circulation de l'information et de causalité locale. Nous allons déterminer un exemple très simple à trois variables qui nous servira tout au long de l'exposé d'exemple compagnon pour illustrer les propriétés dont nous avons besoin.

Présentation d'un exemple compagnon simple : modèle X,Y,Z

Le modèle de réseau bayésien que nous proposons ici est volontairement simple si ce n'est simpliste ; cependant, il est suffisant pour permettre de tester les concepts avancés. Soit un système S_{ex} constitué de trois variables observées binaires 0/1, X , Y et Z . Nous avons donc $S_{ex} = \{X, Y, Z\}$. Le GAD associé à S_{ex} est le suivant :

$$Y \rightarrow X \leftarrow Z$$

Y et Z sont donc les "causes" de X . Un tel modèle est décrit par sa table de probabilité jointe. La table associée à S_{ex} serait ainsi de la forme :

TABLE 5.1 – Distribution de la probabilité jointe de S_{ex}

	$P(Y = 1)$		$P(Y = 0)$	
	$P(Z = 1)$	$P(Z = 0)$	$P(Z = 1)$	$P(Z = 0)$
$P(X = 1)$	X_{111}	X_{110}	X_{101}	X_{100}
$P(X = 0)$	X_{011}	X_{010}	X_{001}	X_{000}

Nous avons introduit les probabilités sous la forme de $X_{X=i,Y=j,Z=k}$, avec $\{i, j, k\} \in \{0, 1\}$ car la variable à décrire est X , qui dépend des variables Y et Z .

Interprétations en termes d'automate

Le modèle S_{ex} peut donc se décrire par son GAD associé, et $P(S_{ex})$ peut se réduire à :

$$P(S_{ex}) = P(X|Y, Z).P(Y).P(Z) \quad (5.8)$$

Aussi, à chaque événement observé concernant Y ou Z , nous gagnons de l'information sur la valeur de X . Notons également que :

$$Y \perp Z$$

Ainsi, un tel automate informationnel permet de décrire le comportement local d'un système pour lequel nous allons observer l'occurrence ou non d'événements particuliers, ici Y ou Z . Il ne s'agit donc pas à proprement parler d'un automate tel que l'on en rencontre en robotique ou en automatique. Plus loin, nous verrons que nous attribuons des rôles particuliers aux variables telles que Y et Z , si l'on considère que les variables dont il nous intéresse de décrire l'évolution sont des variables telles que X .

5.3.6 Définition de l'espace informationnel associé au GAD

Le système observé est donc représentable par un automate informationnel. Nous supposons ici de plus que cet automate est un invariant du système, et qu'il est possible d'établir une correspondance entre cet automate et une variété géométrique. Cette variété constituera le support de l'espace informationnel associé à l'automate. Pour structurer cette variété, nous séparons en 2 sous-espaces D et E l'espace O . D et E sont des variétés différentielles.

Hypothèse de structure de variété sous-jacente comme espace d'évolution – position relative

Dans la suite, nous supposons que l'évolution du système S sera déterminée par l'évolution de l'ensemble des variables observées O . Nous considérons que les informations obtenues en O permettent de localiser un élément particulier dans un espace qui aura la structure générale d'une variété. Nous passons donc d'une structure logique décrite par un GAD, construite sur la base d'observations échantillonnées, à un espace *a priori* continu. La succession d'informations obtenues par un élément de l'espace s'y matérialise comme sa trajectoire, et définit une séquence d'information. Qualitativement, le passage à une variété et à un espace continu nous paraît justifié par i) le théorème général de l'échantillonnage et le simple fait réel que l'on observe des parties de systèmes, des atomes, ii) que ce qui différencie deux éléments régis par le même automate informationnel est leur différence relative en information reçue, donc leur position relative sur la variété, ainsi que la séquence reçue, donc leur trajectoire dans l'espace continu, iii) les outils dont on dispose en géométrie différentielle sont puissants, et s'appliquent naturellement en espace continu.

Il est intéressant de noter qu'une étude s'est intéressée, à titre d'hypothèse, à caractériser les espaces dits "positionnels", selon une métrique relativiste. Ces espaces positionnels se rencontrent en biologie théorique, dont la biologie du développement, et peut, d'une certaine manière, être rapproché d'une conception type Waddington : on désire, dans un espace de configuration décrivant l'ensemble des valeurs possibles des paramètres caractérisant un système, positionner à un instant t , l'état d'un organisme. Les auteurs avaient alors proposé l'utilisation d'une métrique type relativité générale, pour des raisons similaires aux nôtres : étant donné l'équivalence – ou la rétroaction dynamique – entre métrique et contenu de l'espace, une telle mé-

trique permet de reproduire un modèle cybernétique (avec rétroaction, donc, et dynamique), localement, en tout point de l'espace, et variables selon les conditions et la position en espace. [301]

Séparation de l'espace d'observation O en 2 sous-espaces D/E – rôle des variables D et E

Soit un système S . On en observe l'ensemble de variables O . Dans cet ensemble O , certaines variables sont celles dont on aimerait comprendre l'évolution, par exemple le mouvement d'une personne ou une transmission. Nous appellerons ces variables les variables de description, ou si l'on veut, les variables d'espace, par analogie avec la relativité. Nous notons l'ensemble de ces variables D .

Parmi $O \setminus D$, nous nous intéressons aux variables X_i telles $X_i = C(D_i)$, donc les variables qui sont "causes" des variables de D . Nous nommons E cet autre sous-ensemble de O , E pour événement. Nous aurions également pu le nommer T , par analogie avec la relativité, car nous leur ferons jouer le rôle de variables temporelles. En effet, nous nous affranchissons ici de toute dépendance explicite et directe *a priori* avec un temps extérieur au système. Nous comptons et décrivons l'évolution de D selon les occurrences de E . Nous verrons dans la suite comment nous traduisons cette dichotomie et relierons les deux types de variables au travers de la définition d'un intervalle d'espace informationnel.

Lorsque l'on énoncera des propriétés vraies en général, indépendamment de la nature D ou E de la variable observée X_i , nous nous contenterons pour simplifier les notations de nous référer aux X_i .

Définition d'un produit scalaire et traduction des hypothèses d'indépendance conditionnelle et de séparation de l'espace

Afin de structurer l'espace informationnel et d'en dégager la géométrie, nous devons nous munir d'un produit scalaire. Ce produit traduit la propriété d'indépendance conditionnelle naturellement par une propriété d'orthogonalité. Afin de ne pas alourdir les notations, nous proposons d'écrire le produit scalaire ".". Ainsi, si deux variables X et Y sont indépendantes, nous écrivons :

$$(X \perp Y) \Rightarrow (X.Y = 0)$$

Aussi, dans l'exemple S_{ex} , voyons la traduction du produit scalaire d'une combinaison linéaire des variables avec elles-mêmes :

$$\begin{aligned} (aX + bY + cZ).(aX + bY + cZ) &= a^2 X.X + b^2 Y.Y + c^2 Z.Z + \\ &\quad abX.Y + acX.Z + baY.X + bcY.Z + caZ.X + cbZ.Y \\ (aX + bY + cZ).(aX + bY + cZ) &= a^2 f(Y, Z)X^2 + b^2 Y^2 + c^2 Z^2 + \\ &\quad 2abg(Y)X.Y + 2ach(Z)X.Z + 0 \end{aligned}$$

où $f()$, $g()$ et $h()$ sont des fonctions des variables observées. Leur dépendance spécifique en certaines variables dérivent des propriétés du GAD : $f()$ dépend de Y et Z car X dépend uniquement de ces variables ; $g()$ ne dépend que de Y car Y est indépendante de Z et que X est dépendante uniquement de Y et Z – on considère que la projection que réalise le produit scalaire annule la dépendance en Z .

Il est possible également, étant données les dépendances décrites par le GAD, de supposer qu'il existe une dépendance systématique en les variables concernées, en sus d'une dépendance en les parents. Ainsi, il est possible de postuler

l'existence d'une fonction $g(X_i, X_j)$ ou encore écrite g_{ij} telle que :

$$(aX + bY + cZ).(aX + bY + cZ) = a^2g(X, X)X^2 + b^2g(Y, Y)Y^2 + c^2g(Z, Z)Z^2 + 2abg(X, Y)X.Y + 2acg(X, Z)X.Z + 0$$

avec :

$$g(X, X) = g_{XX}(X, Y, Z)$$

$$g(Y, Y) = g_{YY}(Y)$$

$$g(Z, Z) = g_{ZZ}(Z)$$

$$g(X, Y) = g_{XY}(X, Y)$$

$$g(X, Z) = g_{XZ}(X, Z)$$

Ces notations nous amènent insensiblement à l'écriture de la métrique dont nous munirons la variété de l'espace informationnel. Au passage, notons que nous avons fait l'hypothèse de la symétrie des fonctions $g_{ij} : g(X, Y) = g(Y, X)$.

5.3.7 Définition de la métrique et de l'intervalle infinitésimal

Maintenant que nous avons établi la correspondance entre automate et espace informationnel, nous allons expliciter la grandeur centrale qui permet de lier automate et variété, par la définition d'un intervalle infinitésimal construit sur la structure de l'automate. L'écriture de l'intervalle nous amène alors à formuler la métrique de la variété.

Écriture de l'intervalle

En relativité restreinte, et pour un espace bi-dimensionnel, l'intervalle d'espace-temps est donné par :

$$\begin{aligned} ds^2 &= dx^2 + dy^2 - c^2 dt^2 \\ &= g_{xx} dx^2 + g_{yy} dy^2 + g_{tt} (cdt)^2 \end{aligned}$$

avec les g_{ij} les éléments de métrique, ici respectivement égaux à $\{1, 1, -1\}$ pour la triade $\{x, y, ct\}$. Il n'existe pas ici d'éléments de type $dxdt$ ou de type $dx dy$, car ces dimensions sont censées être orthogonales entre elles. Nous avons donc de manière équivalente des $g_{ij} = 0$ pour tout $\{i, j\}$ tels que $i \neq j$. A partir de ce moment, nous allons construire notre intervalle similairement, en respectant les règles suivantes : $g_{ij} \neq 0$ si X_i et X_j ne sont pas conditionnellement indépendants ; $g_{ij} < 0$ si X_i ou $X_j \in E$. Ainsi, pour l'exemple S_{ex} , nous avons $X \in D$ et Y et $Z \in E$. L'intervalle correspond s'écrit :

$$dS^2 = g_{XX} dX^2 + g_{YY} dY^2 + g_{ZZ} dZ^2 + 2g_{XY} dX dY + 2g_{XZ} dX dZ$$

et

$$g_{XX} > 0$$

$$g_{YY} < 0$$

$$g_{ZZ} < 0$$

$$g_{XY} < 0$$

$$g_{XZ} < 0$$

et les autres $g_{ij} = 0$ sinon.

Écriture de la métrique

Dans le cas général associé au GAD de S , et en respectant la subdivision de O observation de S en sous-ensembles D et E , l'intervalle infinitésimal structurant la variété de l'espace informationnel est donné par la formule suivante :

$$dS^2 = g_{ij}dX^i dX^j \tag{5.9}$$

avec les conditions suivantes :

- $g_{ij} = 0$ si l'on a $X_i \perp X_j$
- $g_{ij} < 0$ si l'on a X_i ou X_j de genre E
- $g_{ij} > 0$ si l'on a X_i et X_j de genre D
- $g_{ij} = g_{ij}(X_i, X_j, C(X_i) \cap C(X_j))$

En pratique, on peut choisir d'estimer les valeurs des g_{ij} à partir de la table de $P(S)$, soit estimer les valeurs à partir des données observées de S .

Écriture de l'inverse de la métrique

Dans les grandeurs que l'on utilisera dans la suite, le calcul de l'inverse de la métrique est nécessaire. La métrique et son inverse permettent de transformer notamment les symboles de Christoffel, ou les différentes formes de la courbure. Notamment, on a simplement :

$$g_{ij}g^{ij} = 1 = Id \quad (5.10)$$

De manière plus générale, on peut écrire :

$$g_{ij}g^{jk} = \delta_i^k \quad (5.11)$$

où δ_i^k est le symbole de Kronecker.

Nous définissons également une grandeur liée à la métrique fréquemment utilisée :

$$\det(g_{ij}) = g \quad (5.12)$$

avec \det le déterminant. Nous verrons dans la suite comment l'on utilise la métrique pour faire varier les indices des différentes grandeurs, notamment la courbure.

5.3.8 Courbure de l'espace informationnel et origine de la dynamique géométrique

L'espace informationnel est muni de sa métrique, qui permet de mesurer toute chose sur cet espace. Puisque nous avons supposé que l'automate était un invariant, et que nous lui avons associé un intervalle infinitésimal de fait invariant, nous allons étudier comment évolue cette métrique selon des variations des variables D et E. Ceci nous amène à définir les symboles de Christoffel associés à la métrique, lesquels mesurent la distorsion de cette

dernière. Nous montrons également comment ces distorsions contribuent à façonner l'espace informationnel, et donc à le courber localement : nous écrivons la courbure en fonction des symboles de Christoffel. Enfin, il s'agit de comprendre comment à partir d'arguments purement géométriques, il est possible de définir une dynamique : c'est là ce que nous abordons avec la notion de géodésique.

Déformation de la métrique et écriture des symboles de Christoffel

On définit les *symboles de Christoffel* $\{\Gamma_{ik}^l\}_{i,k,l}$ de la façon suivante :

$$dX_i = \Gamma_{ik}^l X_l dX^k \quad (5.13)$$

ou de manière équivalente :

$$\frac{\partial X_i}{\partial X^k} = \partial_k X_i = \Gamma_{ik}^l X_l \quad (5.14)$$

Selon la structure du GAD et la propriété d'indépendance conditionnelle, il est possible de réduire la dépendance en les $\{X_i\}$ des symboles de Christoffel Γ_{ik}^l . En effet, la variation de X_i s'exprimera uniquement en fonction de lui-même et des fils de X_i : une variation dans la connaissance de X_i répercutera en la variation de la connaissance de ses fils. En outre, on considèrera que les Γ_{ik}^l , qui sont des fonctions des coordonnées $\{X_i\}$, ne dépendront que des parents de X_i et de X_i lui-même : en effet, X_i ne s'écrit qu'en fonction de lui-même et de ses parents.

Les symboles de Christoffel permettent aussi d'écrire de façon condensée la différentielle du tenseur métrique :

$$dg^{ij} = d(X_i \cdot X_j) = dX_i \cdot X_j + X_i \cdot dX_j$$

d'où :

$$\partial_k g_{ij} = g_{il} \Gamma_{jk}^l + g_{lj} \Gamma_{ik}^l \quad (5.15)$$

Lien entre symboles de Christoffel et courbure de l'espace

Inversement, il est possible d'écrire les symboles de Christoffel uniquement en fonction de la métrique, comme suit :

$$\Gamma_{jk}^i = \frac{1}{2} g^{ih} (\partial_j g_{kh} + \partial_k g_{jh} - \partial_h g_{jk})$$

Il peut être utile également de faire varier la hauteur d'indices des symboles de Christoffel, comme suit :

$$\Gamma_{jk}^i = g^{il} \Gamma_{ljk}$$

Nous avons défini la grandeur structurante de l'espace – la métrique, nous avons décrit comment expliciter la dépendance d'une déformation des variables en fonctions des autres variables ; il ne nous reste plus qu'à décrire la grandeur qui va donner sa forme à l'espace, à savoir la courbure. C'est la grandeur qui va différencier un espace plat d'un espace non-euclidien.

La courbure peut s'écrire de différentes façons :

$$R_{jkl}^i = \partial_k \Gamma_{jl}^i + \Gamma_{hk}^i \Gamma_{jl}^h - \partial_l \Gamma_{jk}^i - \Gamma_{hl}^i \Gamma_{jk}^h \quad (5.16)$$

ou aussi :

$$R_{ijkl} = \frac{1}{2} [\partial_{jk} g_{il} + \partial_{il} g_{jk} - \partial_{ik} g_{jl} - \partial_{jl} g_{ik}] + g_{rh} (\Gamma_{il}^r \Gamma_{jk}^h - \Gamma_{ik}^r \Gamma_{jl}^h) \quad (5.17)$$

Ici aussi, les variations de hauteur d'indice concernant les différentes courbures peuvent s'exprimer à l'aide de la métrique. Notamment, nous aurons besoin plus loin de la courbure scalaire R et de la courbure de Riemann R_{ij} .

Ces deux courbures sont définies comme suit :

$$R = g^{ij} R_{ij} = R_i^i$$

et

$$R_{ij} = R_{ikj}^k$$

De façon générale, ce type d'opérations visant à diminuer l'ordre d'un tenseur sont appelées des opérations de contractions.

Principe de la géodésique comme la courbe de moindre action : dynamique

Le principe de moindre action amène à la formulation de ce qu'on appelle une géodésique. Il s'agit de l'équivalent en espace courbe d'une distance usuelle en espace plat. Il existe différentes formulations de l'équation géodésique. Nous en donnons une générale, comme suit :

$$\frac{d^2 X^i}{dS^2} + \Gamma_{jk}^i \frac{dX^j}{dS} \frac{dX^k}{dS} = 0 \quad (5.18)$$

À noter que la géodésique concerne l'ensemble des variables de l'espace, même si elle est généralement présentée comme paramétrée par une variable particulière, le plus souvent le temps t .

5.3.9 Hypothèse du lien entre courbure de l'espace et contenu de l'espace

L'espace informationnel, caractérisé par sa métrique et sa courbure, n'est donc pas nécessairement plat. Ce caractère courbe induit une dynamique. Cependant, comment relier et caractériser la courbure en fonction du contenu informationnel local de l'espace ? Autrement dit, en une région donnée, quelle est la relation entre information, flux d'information et courbure de l'espace,

donc la dynamique engendrée ? Pour ce faire, nous introduisons ici une analogie avec la relativité générale, *via* une équation de métrique du type équation d'Einstein. Ce type d'équation met en relation 3 types d'objets : la métrique, la courbure et une entité appelée tenseur d'énergie-impulsion : nous définirons ici par analogie un tenseur informationnel, décrivant la quantité d'information et son flux localement. La définition d'un tenseur informationnel nous amène enfin à faire le lien inverse avec l'automate informationnel.

Analogie avec la relativité générale : équation d'Einstein

Rappelons ici la forme générale de notre problème de base :

$$f(g_{ij}) = T_{ij}$$

En relativité générale, il est postulé que la courbure de l'espace-temps est régie par le contenu local en matière. Il existe ainsi une équation dite équation de métrique, laquelle donne la relation entre la courbure, la métrique et le contenu local. Nous prendrons pour exemple la forme de l'équation d'Einstein comme forme particulière du problème. Évidemment, nous nous affranchissons des constantes usuelles, pour ne conserver que la structure de l'équation. L'identification des coefficients et éventuelles constantes seront à évaluer à partir des données. L'équation métrique devient :

$$R_{ij} + \alpha g_{ij}R = \beta T_{ij} \tag{5.19}$$

où R_{ij} est le tenseur de Ricci, R la courbure scalaire, et T_{ij} le tenseur d'énergie-impulsion donnant le contenu local en matière et énergie. α et β sont deux constantes à estimer. Contrairement à ce qui peut se voir en cosmologie, nous n'introduisons pas d'équivalent à la constante cosmologique, qui rajouterait donc un terme en Λg_{ij} , Λ étant la constante cosmologique.

Forme et détermination du tenseur informationnel

La relativité générale lie le tenseur d'énergie-impulsion à la courbure, *via* l'équation d'Einstein. Ce tenseur a une forme particulière, que nous retenons dans notre cas, et qui dépend de la présentation des variables retenues. Ici, nous rappelons que nous avons deux sous-ensembles de O que sont D et E , respectivement analogues des variables de type espace et de type temps en relativité. De fait pour un tenseur informationnel, nous avons la matrice suivante :

$$T_{ij} = \begin{pmatrix} D-D & D-E \\ E-D & E-E \end{pmatrix}$$

Ce qui peut encore se voir de la manière suivante :

$$T_{ij} = \left(\begin{array}{c|c} \text{stress morphologique} & \text{densité de moment} \\ \hline \text{flux d'information} & \text{densité d'information} \end{array} \right)$$

5.3.10 Une autre vision : de l'information locale vers l'Hamiltonien du système

L'approche que nous avons suivie jusque-là est celle de la construction de l'espace informationnel *via* sa métrique et ses distorsions, qui induisent la dynamique interne. Il est possible également de caractériser un système par ce que l'on appelle un Hamiltonien. L'Hamiltonien, de même que le Lagrangien, qui entretiennent des liens d'équivalence, sont à issus de la volonté de formaliser la mécanique classique. Ces 2 objets sont intéressants pour leur

compacité et leurs propriétés. L'Hamiltonien est lié à une autre grandeur, que l'on appelle l'action. Dans notre cas, il est possible d'expliciter l'action en fonction du tenseur informationnel. L'action permet d'accéder à ce que l'on appelle l'impulsion généralisée : il s'agit du type de variables dont nous manquons afin de définir totalement l'Hamiltonien du système.

Pourquoi un Hamiltonien ? Relation avec le Lagrangien

On note \mathcal{L} la grandeur appelée *Lagrangien*, et \mathcal{H} la grandeur appelée *Hamiltonien*. On appelle par ailleurs *coordonnées généralisées* les grandeurs notées $\{q_i\}_i$ et vitesses généralisées les grandeurs notées $\{\dot{q}_i\}_i$.

Il est possible, à partir de \mathcal{L} et des $\{\dot{q}_i\}_i$, de définir une troisième grandeur, appelée *impulsions généralisées*, notées $\{p_i\}_i$ et exprimées comme suit :

$$p_i = \frac{\partial \mathcal{L}}{\partial \dot{q}_i} \quad (5.20)$$

L'Hamiltonien est défini comme la transformée de Legendre du Lagrangien :

$$\mathcal{H} = \mathcal{H}(q_i, p_i, t) = \sum_j^p \dot{q}_j p_j - \mathcal{L}(q_i, \dot{q}_i, t) \quad (5.21)$$

En outre, les relations définissant totalement l'Hamiltonien \mathcal{H} sont données par :

$$\dot{q}_i = \frac{\partial \mathcal{H}}{\partial p_i} \quad (5.22)$$

$$\dot{p}_i = -\frac{\partial \mathcal{H}}{\partial q_i} \quad (5.23)$$

Le Lagrangien comme l'Hamiltonien sont des grandeurs utiles pour exprimer et dériver le principe *de moindre action*. Il s'agit d'un formalisme transversal et unificateur nous permettant d'étendre et transposer à notre contexte un certain nombre de concepts.

Par ailleurs, l'Hamiltonien est algébriquement plus simple à manipuler que le Lagrangien (l'Hamiltonien ne fait intervenir explicitement que des dérivées d'ordre 1 en ses paramètres), et intervient dans différentes formulations et théories, à commencer par l'équation de Schrödinger. L'étude des systèmes hamiltoniens ouvre également sur l'étude dynamique des systèmes, et notamment la détermination des régimes chaotiques. Il permet donc une caractérisation concise d'un système. La recherche d'un Hamiltonien ouvre ainsi la possibilité de formuler un principe dynamique du système.

Par ailleurs, la recherche d'un Lagrangien ou d'un Hamiltonien est motivée dans notre cas par un argument simple, faisant référence à notre présentation des 3 cadres théoriques candidats à une théorie morphodynamique (quantique, relativiste, et systèmes dynamiques). Dans chacune de ces théories, il est possible de rapporter, parfois sous certaines hypothèses plus ou moins contraignantes, la dérivation des équations fondamentales de la formulation d'un Lagrangien ou d'un Hamiltonien. Évidemment, en l'absence d'unification de ces théories, ces Hamiltoniens divergent de formulation ou d'interprétation à un moment ou un autre. Néanmoins, leur principe initial demeure commun, ainsi que la dérivation des équations à partir du principe de moindre action. Pour illustration, l'Hamiltonien donne donc l'équation d'évolution en théorie quantique, mais se rencontre également en mécanique générale, et en systèmes dynamiques, où le comportement chaotique peut y être identifié également. [150, 299, 298, 302, 303] Enfin, en théories relativistes, si certains auteurs considèrent acquise l'existence d'une formulation Hamiltonienne de la relativité générale "canonique", cette existence reste controversée. [304, 305]

Principe de l'action. Relation tenseur/ action

Soit A la grandeur appelée *action*. Il existe un lien entre action A , tenseur d'énergie-impulsion T_{ij} et la métrique g_{ij} :

$$T_{ij} = \frac{1}{\sqrt{g}} \frac{\delta \mathcal{A}}{\delta g^{ij}} \quad (5.24)$$

Relation action / impulsion

Il est également possible de relier l'action à la courbure et à la métrique. On a :

$$\mathcal{A} = \frac{1}{K} \int R \sqrt{g} d^p X \quad (5.25)$$

où K est une constante de normalisation, R est la courbure scalaire, g le déterminant de la métrique, p la dimension de O .

D'autre part, il existe une relation entre action, impulsions généralisées et coordonnées généralisées :

$$p_i = \frac{\partial \mathcal{A}}{\partial q_i} \quad (5.26)$$

Ainsi, partant de la connaissance de T_{ij} , c'est-à-dire, la formulation locale de l'automate informationnel dans la variété de l'espace informationnel, et de la métrique g_{ij} , il devient possible d'approcher l'action \mathcal{A} . Sachant l'action, il devient possible de définir une impulsion sans en postuler *a priori* la forme exacte. La connaissance des p_i permet alors de définir un Hamiltonien du système.

Relations de définition de l'Hamiltonien liant variables **D**, impulsions et variables **E** : exemple à une variable temporelle

Ainsi que nous venons de le déterminer plus haut, l'Hamiltonien est entièrement défini dans le cas classique à une variable temporelle unique par

ses relations liant \mathcal{H} aux q_i et aux p_i :

$$\dot{q}_i = \frac{\partial \mathcal{H}}{\partial p_i}, \quad \dot{p}_i = -\frac{\partial \mathcal{H}}{\partial q_i}$$

5.3.11 Hamiltonien multi-événements ou multi-temps

Nous avons proposé la construction d'un Hamiltonien à partir de considérations relativistes. En dehors de systèmes simples, où il n'existerait qu'une variable de comptage, ou dans le cas où les variables de comptages seraient réductibles à une seule (on peut par exemple supposer que toutes sont proportionnelles à une variable temporelle unique), le formalisme classique de l'Hamiltonien nous limite lorsque l'on envisage plusieurs variables de comptage et que l'on veut retomber sur des interprétations directement analogues à celles rencontrées en mécanique classique. Récemment, plusieurs auteurs se sont penchés sur la possibilité de construire un Hamiltonien dit multi-temps, c'est-à-dire, prenant en compte différentes variables temporelles (voir par exemple, [306, 307, 308, 309]).

L'extension au cas multi-temps soit du Lagrangien soit de l'Hamiltonien peut se faire *via* l'écriture des équations d'Euler-Lagrange avec une variable temporelle multidimensionnelle, et transformation de Legendre du Lagrangien pour l'obtention de l'Hamiltonien, soit en s'appuyant directement sur la construction d'un espace particulier, que l'on appelle le fibré de jets du premier ordre, et noté $J^1(T, M)$. Le formalisme, dans ces deux cas, est relativement lourd d'écriture, et technique ; aussi, nous vous renvoyons aux annexes pour plus de précision.

Quoi qu'il en soit, l'intérêt pour ce formalisme, notamment à base de fibré de jets, lequel permet l'expression de grandeurs invariantes naturellement, grandit, alors que ses bases avaient été posées il y a au moins 80 ans, notamment par de Donder [310] puis Dirac. Ce dernier cherchait en effet une formulation générale de la dynamique à partir d'un Hamiltonien. [304]

S'il est possible de proposer, sous certaines hypothèses, une forme particulière de tels Hamiltoniens, nous préférons, dans notre cas et dans une optique de construction de l'opérateur à partir de données observées, de nous baser sur des relations générales et simples, comme nous venons de les présenter en dimension 1 pour le temps. En ce qui nous concerne, nous écririons donc notre espace support $J^1(E, D)$.

5.4 Aspects opérationnels et problèmes ouverts

5.4.1 Reconstruire l'espace informationnel : une approche algorithmique

Le cadre formel et mathématique fourni jusqu'ici a le principal avantage de nous donner des outils manipulables selon des règles valides et d'en cautionner les résultats dérivés. Cependant, la complexité combinatoire impliquée par les développements précédents pourrait s'avérer rapidement rédhibitoire, sans par ailleurs être absolument nécessaire dans chacun des cas. En particulier, si l'on désire essentiellement reconstruire l'espace informationnel à partir d'observations, et d'en identifier les géodésiques, nous proposons une approche inspirée de techniques dites de *manifold learning*, autrement dit, d'apprentissage de variétés. Nous nous appuyerons notamment sur un algorithme princeps de ce genre d'approches de reconstruction non linéaire, à savoir l'algorithme ISOMAP.

Position du problème général et *manifold learning*

Le problème général posé aux algorithmes de *manifold learning* est de recouvrer la variété supposée sous-jacente aux données observées. Le but poursuivi habituellement n'est pas tant de viser une reconstruction de la structure géométrique pour tirer partie de ses propriétés intrinsèques que de s'en servir pour réduire la dimension des données observées. C'est la raison

pour laquelle les mêmes algorithmes sont également appelés *nonlinear dimensionality reduction techniques* – *NLDR*, ou techniques de réduction de la dimension.

Les techniques NLDR se sont développées à partir de 2000, avec la publication de l’algorithme ISOMAP. Il existe actuellement différentes approches, selon notamment que l’on vise à conserver la topologie et donc les relations métriques de voisinage, ou que l’on tend à préserver davantage la structure globale ou les relations angulaires plutôt que métriques. L’algorithme ISOMAP préserve la topologie des données.

De fait, le problème général du *manifold learning* est de se donner un ensemble O d’observations, de dimension p , et d’identifier la variété sous-jacente M respectant la géométrie des données, et de dimension m telle que $m \leq p$. La question de l’estimation de la dimension intrinsèque de M à retrouver à partir de O doit évidemment être traitée, car les algorithmes NLDR nécessitent la dimension visée comme paramètre d’entrée. Dans le cas général, différents estimateurs de la dimension existent, et il n’y pas à ce jour de réel consensus sur l’utilisation préférentielle de l’un plutôt que l’autre. Il apparaît que les estimateurs fractals tel que la *correlation dimension* et l’estimateur selon le maximum de vraisemblance (*maximum likelihood estimator*) sont parmi les plus fiables.

Le prototype de *manifold learning* : algorithme ISOMAP

L’algorithme ISOMAP, pour *ISometric MAPping* est paru en 2000 et fait office de prototype d’algorithme NLDR. Outre sa lenteur relative face à d’autres techniques, il demeure l’un des plus performants tant sur données artificielles que sur données réelles. ISOMAP fonctionne selon 3 étapes, la dernière étant l’application d’une technique linéaire connue par ailleurs, appelée *Multidimensional Scaling* – *MDS*.

Première étape : détermination pour chaque point de O le voisinage, par

l'algorithme des k plus proches voisins ($k - NN$), par exemple.

Deuxième étape : création d'un graphe étiqueté des distances point à point. Les distances des points directement voisins à un point donné A sont évaluées par la distance Euclidienne. Pour deux points A et B qui ne seraient pas voisins directs comme déterminés par la première étape, s'ils sont reliés par un ou plusieurs chemins, voient leur distance évaluée par le plus court chemin, donc par leur géodésique approchée. La détermination de la plus courte distance se fait par l'application d'un algorithme type Dijkstra. Des points non reliés ont une distance fixée à l'infini. Au terme de cette étape, l'on dispose d'une matrice des distances point à point.

Troisième étape : selon la dimension intrinsèque déterminée par un estimateur sur O , on applique à la matrice de distance l'algorithme MDS. M est alors déterminée.

À noter qu'entre O et M , il est possible, et même courant de "perdre" des observations ; cela tient à ce que le graphe construit à l'étape 2 n'est pas nécessairement totalement connecté. Les points ne faisant pas partie de la composante principale du graphe sont donc considérés comme *outliers*.

Algorithme de détermination de l'espace informationnel et des géodésiques

Eu égard au fonctionnement de l'algorithme ISOMAP, il nous est possible de nous en inspirer afin de déterminer l'espace informationnel et ses géodésiques à partir de O . Dans notre cas, il ne nous est pas nécessaire de recourir à un estimateur de la dimension intrinsèque des données, car nous approchons la dimension de l'espace à partir du GAD. En effet, la dimension de notre espace informationnel sera celle donnée par $dim(D) + dim(E)$.

En revanche, le GAD nous fournit la forme de la métrique et l'intervalle infinitésimal associés à l'espace informationnel. Ainsi, la reprise du schéma algorithmique d'ISOMAP, mais en substituant la métrique à la distance Eu-

clidienne utilisée par ISOMAP nous permet de nous mettre dans une situation analogue à celle que serait le passage de la relativité restreinte à la relativité générale : la première étape donne une approximation locale en espace plat par la relation de voisinage de chaque point, et la deuxième étape par parcours du graphe étiqueté, à l'approximation des géodésiques et donc à la courbure de l'espace.

Si l'on souhaite "déplier" l'espace informationnel, il est évidemment possible d'appliquer la troisième étape d'ISOMAP, et donc d'appliquer MDS sur la nouvelle matrice de distance point à point, obtenue à partir de la métrique associée au GAD.

Il existe néanmoins une modification importante à mentionner : dans le cas d'ISOMAP, les distances sont toujours positives ou nulles, ou "infinies". Dans le cas de notre métrique, précisément, certaines portions seront négatives. Pour la détermination des géodésiques pertinentes, il existe donc une étape intermédiaire, consistant à "filtrer" les distances directes entre paires de points, c'est-à-dire, éliminer toutes celles étant négatives (on peut proposer, comme dans ISOMAP, de les placer à l'infini). Les seuls chemins admissibles deviennent ceux se composant de segments positifs ou nuls.

Intérêts d'une analyse en espace informationnel

En reconstruisant l'espace informationnel, c'est-à-dire, en utilisant une métrique reflétant la circulation de l'information et la relation entre variables causales et variables de description, nous disposons directement des trajectoires telles qu'elles sont façonnées par l'information. L'analyse de géodésiques spécifiques peut également informer sur les effets non triviaux, non linéaires des variables causales sur les variables descriptives. En outre, à partir des géodésiques, on peut espérer estimer la forme du tenseur informationnel. Il s'agit d'une technique d'analyse multivariée non linéaire, préservant *a priori* les relations complexes entre variables sans avoir à les expliciter

totalement.

Principe de la représentation locale de l'automate *via* le tenseur informationnel et programmation non linéaire

Nous avons structuré notre métrique à partir de la forme du GAD représentant l'automate informationnel lié au système observé. Nous avons considéré que cet automate représentait les conditions nécessaires et suffisantes, en probabilité, pour l'observation de l'évolution des variables D , selon le "temps interne" au système, matérialisé par les variables E . L'intervalle infinitésimal construit sur l'espace informationnel correspond à la structure informationnelle des données telle que représentée par l'automate : aussi, l'analogie du tenseur énergie-impulsion, que nous avons appelé tenseur informationnel décrit localement les observations D et E . Cet objet représente donc l'équivalent dans l'espace informationnel de l'automate.

Supposons que l'on s'intéresse à un système S spatialisé, i.e. dont une partie des observations est de nature spatiale. On peut penser à n'importe quel système où un mouvement peut être perçu dans un plan ou un espace tri-dimensionnel. Selon notre théorie et l'approche relativiste, nous pouvons reformuler l'évolution de S par l'équation générale, que nous rappelons ici :

$$f(g_{ij}) = T_{ij}$$

Nous avons par ailleurs déterminé la forme générale de T_{ij} , qui est en lien avec la déformation de l'espace local. Ces T_{ij} sont autant d'automates locaux, dont on peut éventuellement programmer le débit et les séquences d'information qu'ils émettent. Si l'on se donne un objectif local ou global pour le système S , on peut le traduire en termes de géométrie du système. Il devient alors possible de procéder à ce que l'on appellera une *programmation non linéaire* ou une *programmation géométrique* : sachant la déformation

engendrée localement par une séquence d'informations, le problème général de programmation revient à trouver la séquence de transformations continues de l'espace afin d'obtenir l'état souhaité, s'il existe. Ce type d'approche s'inscrit dans les démarches que l'on nomme actuellement le *morphogenic engineering* : comment "programmer", en toute généralité, un système complexe ? [311, 312]

5.4.2 Formulations alternatives basées sur l'action

Nous avons souligné que les principales grandeurs nous intéressant, à savoir l'action, la métrique ou l'Hamiltonien, peuvent être reliées les unes aux autres, notamment dans le cas relativiste. De fait, une des voies utilisées pour unifier les théories physiques consiste en la recherche de l'écriture d'une action – afin, finalement, d'en dériver les équations fondamentales selon un principe de moindre action. C'est le cas pour les théories des cordes, des membranes, ou encore la théorie quantique des champs, dérivée de l'invention des intégrales de chemin par Feynman. Ainsi, dans l'optique générale où l'on détermine un automate informationnel à partir des données observées, la forme de l'action qui lui serait associée dépend et peut diverger selon le cadre interprétatif privilégié ; par exemple, de considérer une action type relativité, ou une action type quantique.

Une piste qui nous paraît par ailleurs séduisante est une piste suggérée par De Broglie, dans sa thermodynamique "cachée". [313] Il y met en lien deux grandeurs fondamentales, que sont donc l'action, toujours, mais également l'entropie. Leur mise en relation est de la forme :

$$\frac{A}{h} = -\frac{S}{k},$$

où A est l'action, h la constante de Planck et S l'entropie. L'entropie serait ainsi vue en relation avec l'action et le paramètre d'échelle et de quantification qu'est la constante de Planck. Étant donnée une entropie, par exemple

l'entropie de Shannon, ce type de relation permet de construire encore un autre type d'action – sans que l'on sache à ce stade à quel point cette action différerait des autres envisagées. Nous pensons également à l'analogie avec le temps physiologique tel que West le suppose, lié à l'entropie du système.

5.4.3 Intégration de l'échelle

L'approche que nous proposons ne fait pas intervenir explicitement la notion d'échelle. Il existe à notre sens différentes possibilités pour la faire apparaître ou l'intégrer dans notre formalisme. L'une d'entre elle consisterait à prendre la voie d'une interprétation quantique, dans le sens d'une équation fondamentale de type Schrödinger, à notre Hamiltonien, et de donner crédit à l'équation qu'a dérivé Nottale dans le cadre de sa relativité d'échelle. Ainsi que nous l'avions évoqué, il y fait intervenir explicitement un opérateur fractal, lequel nous permet d'ouvrir la porte à une analyse multi-échelle.

Il serait également possible, si les données sur lesquelles reposent notre Hamiltonien en présentent l'opportunité, de considérer des informations d'échelles différentes, donc des X_i et des X_j d'échelles différentes. Dans ce cas, à la manière des modèles allométriques où nous avons des variables X_{ij} , où i, j indiquent respectivement des individus et des espèces, nous aurions des combinaisons de variables, et donc une métrique, multi échelle, du type $g_{ij}X_iX_j$. Nous aurions ainsi une métrique d'échelle.

Un peu similairement, par l'établissement d'une équivalence entre temps physiologique de type West, et temps propre de type relativité générale, la loi de probabilité déterminant les événements de type temps pour West étant une loi d'échelle, nous aurions directement une relation liant la métrique, déterminée à partir d'un certain niveau d'information, et une relation d'échelle, laquelle lie le temps et la "masse", en l'occurrence, son équivalent informationnel.

Enfin, et certainement de manière pas totalement décorrélée des proposi-

tions précédentes, mais plus généralement, *via* le formalisme de l'analyse fractionnaire, notamment par la recherche de grandeurs canoniquement associées à l'Hamiltonien, si l'on en considère sa version multi-temps : en effet, une interprétation fréquente du multi-temps est une interprétation multi échelle multi fractale (de temps), et le calcul de grandeurs conjuguées usuelles, type position et impulsion (ou par défaut, vitesse), entraîne l'utilisation du formalisme fractionnaire, dans lequel il devient possible de caractériser une dérivée non entière.

5.4.4 Quelle interprétation du concept relativiste dans le cadre de la géométrie informationnelle ?

Dans l'exposé que nous venons de faire, concernant l'écriture d'une métrique associée à un automate informationnel, nous avons clairement favorisé l'interprétation relativiste sur l'interprétation quantique, notamment en cherchant à établir une métrique non linéaire, et à lier contenu de l'espace en information et courbure de cet espace.

Aussi, se pose la question de l'interprétation relativiste, ou disons, du point jusque lequel on peut envisager pousser l'analogie relativiste, de cette approche basée sur l'information et la géométrie. Un des aspects de cette analogie tient à l'existence, et à la possibilité de calculer l'équivalent d'une vitesse limite, prise comme c en général en théorie relativiste. Cette notion est importante, car elle sert de référence aux transformations de type Lorentz. En outre, selon notre vision en termes d'information, et de débit d'information, il serait logique de concevoir que ces débits ou traitements d'information soient soumis à une vitesse indépassable.

Cette vitesse indépassable doit-elle être, cependant, une vitesse "continue", "différentiable", ou de nature probabiliste, et donc, à calculer et interpréter comme une probabilité maximale, indépassable ?

Dans tous les cas, il nous semble cohérent d'assimiler l'existence d'une vi-

tesse de référence, par "canaux", c'est-à-dire, associée à chaque source constitutive du temps interne du système, comme une "fréquence de coupure". Cette question de la vitesse maximale renvoie alors à la question de l'échantillonnage du système observé : sous la condition que l'on puisse connaître cette vitesse maximale, les critères permettant d'établir si le système considéré a été correctement observé (i.e : échantillonné) seront plus simples à établir.

L'optique initiale est de dire que l'on bâtit un automate informationnel sur des observations échantillonnées, donc plutôt de prime abord, des informations de nature continue. Le traitement qui en est fait ensuite sous la forme d'une correspondance avec un espace géométrique et la détermination d'une métrique va en ce sens. Cependant, les automates informationnels, type réseaux bayésiens, ont à l'origine une assise probabiliste, discrète ou continue, mais probabiliste. Pourquoi ne serait-il ainsi pas possible, en maintenant le raisonnement menant à la construction d'une métrique d'espace, et donc d'un temps propre *a priori* différent de la simple addition des temps internes élémentaires, de déterminer les lois stochastiques régissant notre système ?

Enfin, dans le cadre d'une interprétation relativiste type "relativité générale", nous avons déjà explicité l'analogie principale, à savoir la déformation de l'espace en fonction du contenu et du débit en information locaux : équivalents de la courbure de l'espace-temps par la distribution de la masse.

5.4.5 Le temps physiologique de West

La théorie selon laquelle l'activité physiologique des organismes suivrait des relations allométriques, donc des relations d'échelles, a été assez largement discutée, étudiée et développée par BJ West au cours des deux dernières décennies. Le cadre mathématique sur lequel s'appuie cette théorie est celui du calcul fractionnaire et des multifractales. Par définition, les grandeurs issues de séries temporelles y sont exprimées par des fractales multi-temps.

Plus récemment, BJ West a proposé l'hypothèse d'un temps physiologique, que l'auteur avait par ailleurs pressenti (rapport de recherche, non publié) et proposé d'explorer. Ce temps serait à la fois le produit d'une activité propre au sujet vivant, par mise en concours de ses différents rythmes et activités, et une caractéristique fondamentale de ce sujet vivant. Il est à mettre en correspondance avec la notion de métabolisme : à quantités et système identiques, le métabolisme entretenu par un organisme peut agir à des vitesses différentes. Les informations traitées sont les mêmes, et traitées selon les mêmes voies. Néanmoins, dans un cas, un sujet aura vécu davantage d'événements qu'un autre, métabolisant moins vite. Les voies métaboliques, la vitesse de traitement des informations dans le cadre du temps physiologiques sont à mettre en parallèle avec le GAD, la circulation de l'information selon certaines voies privilégiées et leur débit en fonction des variables E .

Rappel : le temps propre en relativité générale et relation à la métrique

En relativité générale, l'intervalle infinitésimal est assimilé *via* deux constantes au *temps propre* du système étudié, par opposition à la variable de temps incorporée dans l'intervalle d'espace. Nous avons ainsi :

$$d\tau = m.c.ds \tag{5.27}$$

où τ est le temps propre, m la masse du système étudié, et c la célérité de la lumière. En général, la formulation du temps propre est donnée de manière normalisée, à savoir :

$$d\tau = ds \tag{5.28}$$

Cependant, dans notre cas, où nous avons une formulation différente des variables D et une dimension de E *a priori* supérieure à 1, il convient de faire

attention à la présence des deux constantes qui seront en fonction du système S observé et de la nature des observations, à différencier par variables.

Une possibilité serait de considérer que l'équivalent de la masse, ainsi que nous l'avons fait jusque ici, est la distribution d'information, et l'on peut ainsi se doter d'un équivalent multidimensionnel X , comprenant les variables d'événements et les variables d'observation, en toute généralité. Sans en préciser la formulation précise à ce stade, on peut également se munir d'un vecteur de même dimension C , représentant les vitesses limites associées à chaque variable d'information. On peut alors se donner l'équivalent suivant :

$$d\tau = mc ds \equiv X^\top G C ds$$

où G est la matrice constituée des éléments g_{ij} de la métrique, et $ds^2 = g_{ij} dX^i dX^j$.

Se pose ensuite la question de la forme de C . De manière générale, si l'on s'en tient à notre exemple S_{ex} compagnon, il peut y avoir plusieurs canaux menant à une seule observation. La question de la vitesse de changement d'une variable d'observation est donc une question de vitesses partielles, dues à des phénomènes concurrents. Nous avons notés $C(X_j)$ les "causes" (ou les parents) de X_j . Puisque nous cherchons à déterminer une vitesse limite, nous pouvons supposer que la vitesse limite C_j associée à la variable X_j , est de la forme :

$$C_j = \max_{X_i \in \{C(X_j)\}} V_{X_j|X_i},$$

avec $V_{X_j|X_i}$ la vitesse associée à X_j , par rapport à la "cause" X_i de X_j .

La question de la vitesse $V_{X_j|X_i}$ dépend alors du type de données et d'encodage disponible, et dans le cas probabiliste, du caractère continu ou discret des variables. Néanmoins, de manière tout-à-fait générale, nous cherchons à quantifier une grandeur maximale, du type :

$$V_{X_j|X_i} = \max P(\delta X_j \neq 0 | \delta X_i \neq 0)$$

Autrement dit, quelle est la probabilité maximum d'observer un changement de valeur de X_j sachant que X_j a aussi changé de valeur ?

Pour obtenir une formulation plus analytique encore de ce temps propre, il serait nécessaire de résoudre la question de l'approximation de la forme et l'estimation des valeurs des coefficients des g_{ij} . Si l'on considère que les fonctions des variables pondérant les éléments $dX^i dX^j$ peuvent s'approcher par des fonctions polynomiales, on pourra se donner une forme générale des coefficients. En reprenant là encore notre exemple S_{ex} compagnon, et pour une dépendance de g_{XZ} en X et en Z , on aurait, à l'ordre 2 :

$$g_{XZ} = g_{XZ}(X, Z) = \alpha + \beta_1 X + \beta_2 Z + \gamma_1 X^2 + \gamma_2 Z^2 + \gamma_3 XZ$$

Il serait étonnant qu'il soit nécessaire, pour un modèle de bonne performance, de devoir nécessairement pousser l'approximation beaucoup plus avant. Quoi qu'il en soit, ce type d'approximation doit également permettre d'une part d'estimer les coefficients α , β_i ou γ_j , mais également de pouvoir faire des calculs plus formels, notamment de dérivation de la métrique totale, par exemple pour inférer la forme du tenseur informationnel, ou encore remonter à l'action ou à l'Hamiltonien.

Temps physiologique et temps propre

Au vu de la définition du temps propre de notre espace informationnel, et de la logique présidant à la détermination d'un temps physiologique, nous pouvons postuler leur équivalence en général.

Autrement dit :

$$\tau^{physio} \equiv \tau^{propre}$$

Plus précisément, ce type d'équivalence, pour qu'elle puisse s'écrire formellement, nous impose soit de faire un choix quant à la nature du membre de droite, à savoir de considérer que l'on a construit une relation métrique sous la forme d'une densité de probabilité jointe de type $f(X_i, X_j, g_{ij})$, soit d'envisager qu'elle induit une forme de relation d'échantillonnage en base de temps physiologique de l'espace informationnel.

Dans le premier cas, nous aurions la relation du type :

$$p(\tau < t < \tau + d\tau).d\tau = f(X_i, X_j, g_{ij}) \cdot \sqrt{dX_i dX_j}$$

Il demeure possible d'imposer la forme de la densité de probabilité comme West le propose, c'est-à-dire, respectant une relation invariante d'échelle ; cette hypothèse n'est cependant pas obligatoire, et la contrainte peut en être relâchée, et supposer une forme plus générale.

Dans le second cas, on peut envisager le raisonnement suivant : plus la probabilité pour le temps physiologique de "tomber" entre τ^{physio} et $\tau^{physio} + d\tau^{physio}$ est grande, plus la densité de points de l'espace informationnel est importante. De fait, on écrit :

$$\tau \approx h(X_i, X_j, g_{ij})$$

et

$$\frac{\tau}{d\tau} \approx \frac{h(X_i, X_j, g_{ij})}{k(X_i, X_j, g_{ij})}$$

peut permettre de quantifier la cadence d'échantillonnage relative de l'espace informationnel.

Chapitre 6

Conclusions et perspectives

Sommaire

6.1	Limites des travaux	388
6.1.1	Autres approches : réseaux booléens, théorie de la viabilité, grammaires génératives	388
6.2	Développement des algorithmes de géométrie informationnelle	390
6.2.1	Une question en suspens : réalisme local <i>vs</i> mécanique quantique	392
6.2.2	Formulation hamiltonienne multi-événements . . .	393
6.2.3	Géométrie non commutative	394
6.3	Tests et interprétations en pratique courante .	396
6.3.1	Observations et nouveautés	396
6.3.2	Simulations, données réelles, et développement des algorithmes.	396
6.4	Publications et communications	397
6.4.1	Publiés ou acceptés	398
6.4.2	Soumis	398
6.4.3	En cours	398
6.4.4	Communications orales	398
6.4.5	Autres publications de l'auteur	399

Par les travaux présentés ici, nous espérons avoir tant soulevé les limitations de méthodes largement employées et encore plébiscitées, notamment en recherche biomédicale, que proposer une voie possible et solidement étayable

pour la fondation d'une morphodynamique basée sur l'information et la géométrie. Cette dernière proposition s'inspire et propose une approche basée sur trois grands cadres théoriques (les théories quantiques, relativistes et des systèmes dynamiques ou de la stabilité structurelle).

Néanmoins, et comme le lecteur s'en sera aperçu, il existe encore un certain nombre de développements nécessaires, ainsi que des tests validant la pertinence de l'approche.

6.1 Limites des travaux

Parler de morphogenèse et en souligner l'intérêt comme approche alternative aux méthodes prédominantes actuellement ne peut se faire, dans un volume raisonnable de temps et de pages, qu'en laissant de côté plusieurs théories ou approches au profit d'autres qui nous ont soit paru plus pertinentes, soit que nous maîtrisions mieux.

6.1.1 Autres approches : réseaux booléens, théorie de la viabilité, grammaires génératives

Parmi les approches de la morphogenèse que nous avons ignorées dans ces travaux, nous pouvons mentionner tous les systèmes basés sur les grammaires dites "génératives". Ceux-ci ont fait preuve d'une efficacité indiscutable pour rendre compte de plusieurs formes. L'un des exemples les plus célèbres étant l'obtention de végétaux, tout ou partie, en particulier de feuilles, à partir de motifs élémentaires, qui seront répétés selon des règles pré-spécifiées (une grammaire). On parle plus spécifiquement dans ce cas de L-système – pour Lindenmayer (voir figure 6.1). Diverses variantes existent, notamment introduisant une dimension stochastique, aléatoire à des grammaires déterministes.

Il existe par ailleurs une théorie nommée théorie constructale, proposée

FIGURE 6.1 – Exemples d’herbes sauvages générées à partir d’un L-système.

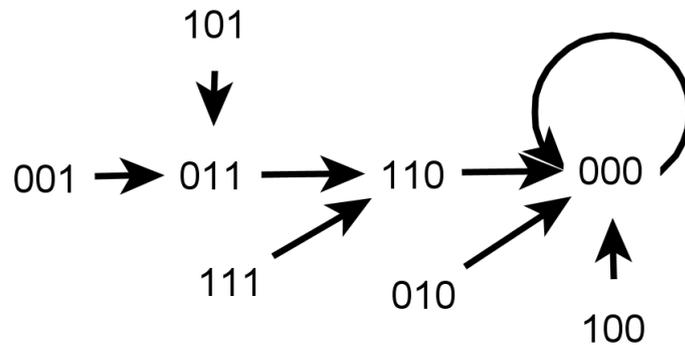


par Bejan, et prenant le contre-pied de l’approche fractale. Partant du principe qu’il existe une limite inférieure aux échelles des phénomènes, du moins, une limite inférieure aux échelles pertinentes en termes de fonctionnalité et de forme caractéristiques, Bejan propose d’assembler à la façon des grammaires génératives des formes élémentaires, selon des contraintes d’optimisation particulière, jusque obtenir une forme complexe, d’échelle supérieure. [314]

Ce type de démarche, pour pertinente et intéressante qu’elle soit, s’intégrait difficilement au type de cadre théorique et formel que nous cherchions – notamment, une théorie basée sur l’information quelconque, sur des observations. Nous avons donc choisi de les écarter, sans en méconnaître l’existence.

Il semble également exister un intérêt croissant pour les réseaux dits booléens, lesquels représentent des automates particuliers, aux états booléens. Des états attractants peuvent y être cherchés et modélisés, et l’on peut étudier l’espace des phases et des transitions de phases (voir figure 6.2). Ces réseaux sont également appelés systèmes dynamiques séquentiels. [315]

FIGURE 6.2 – Un exemple d’espace des phases d’un réseau booléen (Wikipedia, licence CC)



Enfin, nous mentionnerons la théorie dite de la viabilité, due à Aubin, dont nous connaissons l’existence, mais ne maîtrisons suffisamment correctement ni les outils ni le cadre. [316] Notre absence de maîtrise élémentaire de la théorie ainsi que les contraintes de temps et de choix font que nous n’avons pas intégré cette approche à notre étude non plus. Par ailleurs, s’il semble qu’elle commence à être davantage reconnue, il s’agit d’une théorie encore insuffisamment connue.

6.2 Développement des algorithmes de géométrie informationnelle

Le cadre général que nous avons proposé, c’est-à-dire, la construction d’un espace informationnel, d’interprétation essentiellement géométrique, basée sur la structure fonctionnelle et causale des informations, n’est évidemment qu’une étape. En particulier, pour reprendre le cas relativiste, si l’on approche le problème directement, i.e. en cherchant à estimer toutes les inconnues et résoudre chaque équation, il est peu probable que la théorie soit praticable. Cependant, l’arsenal mathématique et formel déployé nous paraissait nécessaire dans le sens de la présentation d’un cadre correctement

étayé.

L'opérabilité de notre théorie passe quant à elle par la recherche de techniques plus "intelligentes" que la simple tentative de résolution directe du problème initial. C'est précisément dans cet esprit que nous avons proposé le principe d'un premier algorithme d'identification des géodésiques en espace informationnel. Cet algorithme a également pour vocation d'inspirer d'autres voies d'explorations permettant d'accéder aux grandeurs pertinentes de la théorie : par exemple, outre les géodésiques, l'estimation de la forme du tenseur informationnel.

L'une des questions les plus importantes à régler reste de déterminer une méthode visant à estimer les coefficients des éléments "infinitésimaux", autrement dit, de la métrique de l'espace. En l'occurrence, la difficulté est certainement plus ou moins grande selon le point de vue préféré : dans une optique d'espace "plat", de type quantique ou relativité restreinte, il n'existe pas de dépendance en les variables d'observations pour les coefficients de la métrique : il s'agit davantage de déterminer au mieux la valeur absolue de ces coefficients, lesquels seront fixes sur l'ensemble de l'espace. En revanche, dans une optique de métrique dynamique, cette dépendance en les variables existe. Bien entendu, un certain nombre d'hypothèses peuvent être formulées, à commencer par des hypothèses usuelles de dépendance linéaire en ces variables des coefficients. On peut encore raffiner, et chercher des solutions sous formes polynomiales.

Le développement de tels algorithmes, outre leur intérêt quant à la possibilité qu'ils offriraient de pouvoir utiliser en routine notre approche, permettrait d'explorer une question autrement fondamentale, à savoir : quel mécanisme principal suit l'information ? L'évolution d'une forme basée sur une structure informationnelle fonctionnelle est-elle davantage de type quantique, ou bien relativiste ? Le réalisme local correspond-il davantage aux mouvements de l'information, ou bien est-ce finalement le caractère probabiliste

et la mécanique des corrélations qui prédominent ?

6.2.1 Une question en suspens : réalisme local *vs* mécanique quantique

Einstein était radicalement contre l'idée d'une mécanique gouvernée par le hasard, et plus encore, par des phénomènes instantanés. C'est notamment cette position qui l'avait mené, avec deux autres protagonistes, à formuler le fameux paradoxe EPR. Néanmoins, la mécanique quantique accumulant tant de succès et permettant des calculs d'une précision telle, force a été de considérer qu'il devait y avoir quelque chose de vrai dans celle-ci, et quelque chose de faux dans la position d'Einstein.

Dans un monde par Einstein, les déformations de l'espace, la propagation de l'information ont des causes et ne peuvent se manifester instantanément, mais bel et bien à vitesse finie. En mécanique quantique, l'état d'une particule 1 intriquée avec une particule 2 peut voir son état être instantanément modifié, sans autre interaction nouvelle avec la particule 1, et ce quelle que soit la distance les séparant, pourvu que l'on effectue une mesure sur 1. Les inégalités de Bell, formulation précise sur laquelle se fonde le paradoxe EPR, ont servi de support à différentes expériences pour démontrer qu'Einstein se trompait. Pour autant, l'interprétation est un peu plus subtile : oui, il est possible d'observer des "interactions instantanées" à des distances arbitraires, mais ceci sans pour autant que le principe de causalité soit violé. Ce qui est possible uniquement parce que l'action de mesurer les états des particules intriquées impliquent en réalité et en pratique, une séquence et un temps déterminés.

Outre ces considérations toute théoriques, il n'en reste pas moins que l'utilisation maintenant ancienne d'algorithmes de réduction de dimension de type projectif / spectral tel que l'ACP d'une part, et l'apparition de nouveaux algorithmes dits non linéaires de réduction de dimension, pour

beaucoup basés sur des considérations géométriques et topologiques d'autre part, ont permis de confronter les résultats que les premiers et les seconds fournissent quand il s'agit de réduire la dimensionnalité d'un ensemble de données. Les algorithmes type ACP ressemblent à une approche quantique, avec identification des vecteurs propres et projection sur une base diagonale, tandis que les approches NLDR s'apparentent bien plus à une approche relativiste, où l'on considère un voisinage localement plat, pour une géométrie globalement et potentiellement courbe.

Les résultats donnés par les deux classes d'algorithmes sont parfois étonnants. Sur des exemples artificiels, comme ceux présentés dans ces travaux, les techniques non linéaires s'en tirent en général parfaitement bien, et surtout, mieux que les approches linéaires. En pratique, des résultats sur données réelles montrent que la réalité est moins simple, et il existe des cas où l'ACP fait mieux que les algorithmes NLDR, et inversement, sans que l'on comprenne exactement pourquoi. Ceci est intrigant, car les hypothèses et les mécanismes sous-jacents sont distincts, sinon opposés, du moins en apparence. De fait, la structure informationnelle des données suivraient plutôt tantôt une structure relativiste, tantôt purement corrélative. [317, 318, 319, 320]

Une hypothèse que nous formulons ici, serait que les techniques projectives décriraient mieux les ensembles de données de haute dimensionnalité intrinsèque, et de faible densité en observations, tandis que les techniques topologiques réussiraient mieux pour des données de dimensionnalité intrinsèque plus faible, et de plus grande densité en observations. Tout ceci reste donc à éclaircir.

6.2.2 Formulation hamiltonienne multi-événements

Un des postulats forts de notre approche tient sur la disparition du temps en tant que tel, pour lui substituer des variables d'événements, internes à

la forme fonctionnelle considérée. Ceci a le tort d'introduire l'équivalent de plusieurs variables temporelles, plutôt qu'une seule. Or, toutes les théories usuelles n'intègrent qu'une seule variable de type temps, quel que soit le rôle qui lui soit dévolu, tandis que les variables de type espace peuvent être multipliées à l'envi.

Les récents développements théoriques en géométrie différentielle, se proposant de déterminer des espaces multi-temps, nous ont procuré un cadre pour appuyer et expliciter notre approche. Cependant, ces développements demeurent jeunes, encore peu pratiqués, et nécessitent du temps pour qu'ils soient consolidés ou corrigés. Plus particulièrement, le passage des formulations des lois usuelles connues pour un Hamiltonien temps unique n'est peut être pas aussi direct qu'il n'y paraît au premier abord, et nécessiterait un examen plus rigoureux. C'est notamment le cas pour les liens entre actions et Hamiltoniens. De même, dans l'optique d'une interprétation quantique, l'opérateur d'évolution est lui aussi temps unique, et fait également intervenir l'Hamiltonien. Quelle devrait-être, en ce cas et en toute rigueur, la formulation correspondante en multi-temps ?

Etant donnée la forme des expressions multi-temps, il serait surprenant que les expressions correspondantes des quantités connues en temps unique soient fondamentalement différentes. La preuve doit cependant en être apportée.

6.2.3 Géométrie non commutative

Il existe une approche dont nous n'avons pas parlé, même si l'auteur l'avait envisagée sous différentes formes à plusieurs reprises – il s'agit d'une approche basée sur les relations de commutation. Si l'on considère un ensemble d'observations, il est possible de considérer, *a priori*, que chaque type d'observation peut être vu comme une mesure d'une observable correspondante.

Les observables sont à la base de la formulation quantique, comme nous

l'avons déjà mentionné. Une manière d'étendre le principe quantique serait de postuler de base des relations de commutations entre les observables principales du système étudié, sans présumer de la transformation, qu'elle soit potentiellement analytique ou non, permettant de lier une observable à une autre dans une relation de non commutation (dans le cas de la mécanique quantique, c'est la transformation de Fourier qui implique la forme et le lien des relations des non commutation : c'est notamment le cas des variables position et vitesse, ou temps et énergie).

Ce biais est utilisé par Alain Connes, qui a développé la géométrie non commutative, dans une tentative de réconciliation des approches quantiques (relations de non commutation) et relativistes (géométriques). [321] L'entreprise est séduisante, mais pour les mêmes raisons que nous n'avons pas abordé d'autres théories comme celle de la viabilité, nous avons préféré écarter dans un premier temps le cadre de cette géométrie. Nous n'en maîtrisons pas suffisamment les outils de base. Elle n'en reste pas moins un cadre théorique à surveiller, ne serait-ce que comme source d'idées pour le développement d'algorithmes.

Enfin, toujours dans le cadre des relations de non commutation, ou de manière équivalente, des relations de conjugaisons entre variables, il resterait à trouver un abord ou un test simple permettant de déterminer quels couples de variables sont en effet conjugués. Dans cette optique, nous pensons que l'approche par réseau bayésien peut là encore fournir des éléments fondamentaux, notamment en considérant que deux variables directement parentes l'une de l'autre, sont liées par une relation de causalité, donc une notion de "déplacement" de l'une sous l'action d'une variation de l'autre. Il est alors possible de postuler que toutes les relations de ce type entre deux variables sont de type "position-vitesse", et donc conjuguées, quelle que soit, là encore, la forme de la transformation reliant ces deux variables.

6.3 Tests et interprétations en pratique courante

6.3.1 Observations et nouveautés

Comme pour la majorité des algorithmes, notre proposition ne dit pas comment gérer la nouveauté, que l'on peut envisager sous différents cas de figure.

On entend par nouveauté, par exemple, la modification de la topologie révélée par l'automate informationnel : quel est l'impact sur la métrique associée, et plus spécifiquement, sur l'estimation des paramètres de cette métrique, de l'ajout ou de la disparition d'une arête (donc, de la modification des relations topologiques) ? De manière plus générale, peut-on, et si oui, comment, gérer l'impact de la création ou de l'annihilation d'arêtes ou de nœuds au sein du graphe associé à l'automate informationnel ? Quel serait le lien entre la dynamique du graphe et la métrique ? Sans pouvoir répondre nécessairement directement à ces questions, en supposant qu'il soit possible d'y répondre simplement, il serait intéressant d'évaluer empiriquement le comportement de la métrique induit par des modifications élémentaires et simples du GAD. C'est en effet ici l'une des limitations classiques des approches non linéaires, par rapport aux approches linéaires, où précisément, il est simple et direct d'ajouter des composantes.

6.3.2 Simulations, données réelles, et développement des algorithmes.

La pertinence de notre famille d'approches, via la construction d'un espace informationnel et de sa géométrie multi-temps associée, indépendamment des préoccupations de complexité algorithmiques, ne pourra être validée qu'en développant des algorithmes visant à la construction d'un tel espace, et à l'estimation des paramètres nécessaires à sa construction.

L'une des possibilités *in silico* serait d'avoir recours à des modèles multi-

agents, notamment dont on connaît par ailleurs la formulation sous forme de systèmes dynamique, et d'enregistrer, à la manière d'une cohorte synthétique de personnes, les données de chaque agent au cours des simulations.

Disposant de ces données générées, nous pourrions déterminer l'automate informationnel décrivant la circulation de l'information entre les différents niveaux, afin de caractériser l'automate correspondant à l'activité collective des agents, au système considéré. De cet automate, il s'agirait d'en tirer la forme de la métrique associée, et selon le point de vue théorique retenu (quantique ou relativiste), estimer les paramètres de la métrique numériquement. Cette métrique connue, l'application de l'algorithme modifié de type ISOMAP nous donnera accès aux géodésiques de l'espace informationnel – ces géodésiques sont les trajectoires naturelles, au sens du principe de moindre d'action, du système étudié. Selon la dimension retenue de notre système (la dimension du GAD associé), nous pourrions également déployer l'espace informationnel, toujours selon l'algorithme modifié type ISOMAP. Des régions spécifiques pourront y être recherchées, comme états attractants ou groupes homogènes. Dans ce dernier cas, nous aurons un avantage sur une approche type clustering usuelle, puisque les groupes identifiés le seront d'après des principes dynamiques. L'approche est bidirectionnelle, car muni de ce principe dynamique : d'un côté, nous synthétisons le comportement du système sous forme d'une écriture type GAD et équation de métrique, d'un autre côté, nous pouvons, via l'estimation du tenseur informationnel, essayer de programmer le système, en nous donnant une distribution informationnelle telle que nous façonnons les trajectoires voulues, vers les états désirés.

6.4 Publications et communications

Les travaux présentés dans ce manuscrit ont fait l'objet de publications, de communications orale ou affichées, et fournissent le matériau pour d'autres publications soit déjà soumises, soit en cours de rédaction, soit enfin en

projet.

6.4.1 Publiés ou acceptés

- Lefèvre T, Lepresle A, Chariot P. "An alternative to current psychiatric classifications : a psychological landscape hypothesis based on an integrative, dynamical and multidimensional approach", en révision pour *Philosophy, Ethics and Humanities in Medicine*.
- Lefèvre T, Stindel E, Ansart S, Roux C. "Mathematics in medicine : beyond iatromathematics", *Lancet* 2014 ;383(9916) :513. doi : 10.1016/S0140-6736(14)60183-5. [lettre]

6.4.2 Soumis

- "What do we mean by multimorbidity ?", soumis à *BMC Health Services Research*.

6.4.3 En cours

- Article Géométrie informationnelle. *Pre-print* pour arxiv, puis développement sur la base d'un exemple pour en illustrer l'intérêt.
- Article Modélisation mathématique des infections ostéoarticulaires. À destination de *J Theor Biol*.
- Article *Big data et système d'exploitation secondaire des données hospitalières – SPOT RISC*. À destination de JAMIA.
- Article : 50 ans de Naissance de la Clinique : quelle est la géométrie des infections ? Sous forme de *Review* ou *Perspective*.

6.4.4 Communications orales

- Journée scientifique du CRIOGO, Brest, octobre 2013 : "modélisation mathématique des infections ostéoarticulaires". Communication orale.

- Séminaire du Laboratoire de Pédagogie de la Santé, Paris, décembre 2013 : "les modes d'existence de la polypathologie chronique". Communication orale.
- 7^{me} journée de l'IFR65, Paris, juin 2013 : "Techniques non linéaires de réduction de la dimension : principes et revue brève". Communication orale.
- Séminaire de l'école de santé publique Pierre Louis, St Malo, octobre 2012 : "Pour la préservation de l'environnement et de la géométrie : techniques non linéaires de réduction de la dimension." Communication affichée.
- Orateur invité au séminaire de l'UMRS 1136, ERES (13 juin 2014) : "Formes et fonctions en science : Einstein, Schrödinger et Lorenz répondent à Foucault". Communication orale.
- 48^{ème} Congrès international de médecine légale, Marseille, 2013 : "Déterminants de l'incapacité totale de travail : une analyse par réseaux bayésiens". Communication orale.

6.4.5 Autres publications de l'auteur

Articles publiés

- Denis C, Lepresle A, Lefèvre T, Chariot P. "Natsal and sexual violence in Britain". *Lancet* 2014 ;383(9920) :866-7. doi : 10.1016/S0140-6736(14)60438-4. [lettre]
- Chariot P, Briffa H, Lepresle A, Lefèvre T, Boraud C. "Fitness for detention in police custody : a practical proposal for improving the format of medical opinion". *J Forensic Leg Med* 2013 ;20(8) :980-5
- Chariot P, Lepresle A, Lefèvre T, Boraud C, Barthès A, Tedlaouti M. "Alcohol and substance screening and brief intervention for detainees kept in police custody. A feasibility study." *Drug Alcohol Depend* 2014 ;134 :235-41

- Briffa H, Lefèvre T, Boraud C, Chariot P "Physicians attending detainees in police custody : an improved format of the doctor's opinion" *Presse Med* 2013 ;42(1) :e9-e15
- Lefèvre T, Briffa H, Thomas G, Chariot P. "Evaluating the functional impairment of assault survivors in a judicial context – a retrospective study". *J Forensic Leg Med* 2012 ;19(4) :215-8
- Lefèvre T, Singh-Manoux A, Stringhini S, et al. "Usefulness of a single-item measure of depression to predict mortality : the GAZEL prospective cohort study" *Eur J Public Health* 2012 ;22(5) :643-7
- Nabi H, Chastang JF, Lefèvre T, et al. "Trajectories of depressive episodes and hypertension over 24 years : the Whitehall II prospective cohort study". *Hypertension* 2011 ;57(4) :710-6
- Visvikis D, Lefèvre T, Lamare F, et al. "Monte carlo based performance assessment of different animal PET architectures using pixellated CZT detectors" *Nuclear Instruments and methods in physics research section A* 2006 ; 569(2) :225-9

Articles soumis

- Lefèvre T, Watier L, Lorin G "Discriminating factors in blunt trauma between deaths with fall from a low height and homicidal blows", soumis à *Forensic Science International*
- Lefèvre T, Chauvin P "Advanced methods in multivariate analysis : nonlinear dimensionality reduction techniques to overcome principal components analysis limitations", soumis à *JAMIA*

Annexe A

Éléments de calcul vectoriel

A.1 Outils de calcul vectoriel

Nous rappelons ici les notions de base sous-tendant l'existence des espaces vectoriels. Nous abordons aussitôt pour des raisons de commodités la convention d'écriture d'Einstein.

A.1.1 Espace vectoriel et espace dual - forme linéaire

Il est nécessaire de savoir parler d'espace vectoriel, et d'avoir assimilé les principes de base, notamment la décomposition d'un vecteur sur une base de l'espace vectoriel. On rappelle ici qu'un espace vectoriel est caractérisé par sa dimension, qui peut être infinie ou finie. La dimension d'un espace vectoriel correspond finalement au nombre minimal, nécessaire et suffisant, d'éléments à partir desquels on pourra décrire exactement l'ensemble de tout l'espace, autrement dit, à partir desquels on peut construire absolument tous les vecteurs de l'espace. Ceci est possible en raison de la structure d'un espace vectoriel, qui assure notamment que toute combinaison linéaire de vecteurs est un vecteur. Nous supposons aussi le lecteur familier des notions de calcul matriciel. Rappelons seulement qu'une matrice n'est qu'une représentation commode des fonctions, c'est-à-dire des transformations, de vecteurs d'un

espace vers un autre, voire d'un espace dans le même – on parle d'endomorphismes.

Rappelons l'écriture d'un vecteur sur une base.

Soit une base $e(e_1 \dots e_n)$ d'un espace vectoriel E , de dimension n , x un vecteur de cet espace. x se décompose sur cette base de la façon suivante :

$$x = \sum_{i=1}^n x^i e_i$$

Dans ce cas, les x^i sont les coordonnées de x dans la base e . Evidemment, on comprend que si l'on se dote d'une autre base, par exemple e' , les coordonnées ne seront plus les mêmes, mais des x'^i . Pour nous, cette constatation basique a son importance, car cela signifie ni plus ni moins que 2 observateurs distincts peuvent être précisément considérés comme 2 systèmes, 2 bases de décompositions différentes (si vous faites face à quelqu'un, votre droite n'est pas la sienne). Il convient donc de savoir passer d'une représentation à une autre, donc d'une base à une autre.

Les vecteurs e_i d'une base de l'espace E étant par définition des vecteurs, il est possible de les décomposer suivant une autre base, par exemple la base e'_I . On écrit alors :

$$e_i = \sum_{i=1}^n a_i^I e'_I$$

où les a_i^I sont les coordonnées des e_i dans la base des e'_I . Pour représenter le passage d'une base à l'autre, il faut donc se donner n^2 a_i^I , ce qui peut se représenter sous forme matricielle, la matrice A . A est la matrice de passage de la base e'_I vers la base e_i . L'inverse de la matrice A , notée A^{-1} , existe et est la matrice de passage dans le sens inverse, soit de e_i vers e'_I .

Les matrices de passage permettent d'écrire un vecteur quelconque x exprimé dans une base donnée dans une autre base, et ainsi de retrouver ses

nouvelles coordonnées. Ainsi, pour x de coordonnées x^i dans la base e_i , on obtient les coordonnées x'^I dans la base e'_I de la façon suivante :

$$x'^I = \sum_{i=1}^n a_i^I x^i$$

ou sous forme matricielle :

$$X' = AX$$

A.1.2 Notation d'Einstein

Manipuler des vecteurs en dimension n peut vite devenir fastidieux à écrire, notamment lorsque l'on effectue des produits, où des sommations apparaissent aussitôt et alourdissent l'écriture. Pour remédier à ce type de problème, nous allons adopter une convention d'écriture dite convention d'Einstein. L'indexation des i et I pour les coordonnées de vecteur n'était pas fortuite, certains se trouvant en exposant, d'autres en indices. L'utilité en est la suivante : par convention, la répétition d'un indice en haut et en bas de deux variables différentes tient lieu de sommation deux à deux sur l'ensemble des valeurs possibles de cet indice. Voici un exemple :

$$x = \sum_{i=1}^n x^i e_i$$

va s'écrire en convention d'Einstein de la façon suivante :

$$x = x^i e_i$$

La sommation est implicite en raison de la répétition de l'indice i en haut puis en bas des 2 variables x et e . Autre exemple, pour la multiplication matricielle :

$$AB = C$$

va s'écrire :

$$c_j^i = a_k^i b_j^k$$

A.2 Outils de géométrie Riemannienne

La remise en cause d'une géométrie unique, valable quelque soit l'objet à décrire, c'est-à-dire, la géométrie dite euclidienne, s'est faite en plusieurs étapes, à partir du moment où il y eut des esprits pour critiquer l'axiomatique, qui revêt un caractère arbitraire et des conséquences importantes, puisque l'axiomatique de base fonde toute la théorie qui en dérive. L'intuition d'une géométrie qui ne soit pas euclidienne, donc qui ne remplisse pas nécessairement tous les axiomes d'Euclide, peut se trouver dans des considérations pratiques.

Ainsi, alors que tout un chacun a appris que la somme des angles d'un triangle vaut 180 degrés, ce qui est vrai en géométrie euclidienne, il s'avère que cette propriété n'est pas vérifiée pour d'autres géométries. Représentez-vous la surface de la terre, en tant que sphère, par exemple. Supposez que vous n'apparteniez à un monde qu'en 2 dimensions, celles de la surface terrestre, courbe. Dessinez un triangle suffisamment grand sur cette surface, mesurez en les angles : leur somme ne vaut pas 180 degrés. Ceci est dû à la courbure de la surface, contrairement à un espace euclidien, qui lui est plat.

Nous nous intéressons donc ici à la géométrie dite riemannienne, de son concepteur, Riemann. Sa définition tourne autour de la notion de courbure de l'espace, et l'on construit l'équivalent des outils habituels en géométrie classique, ce qui nécessite de se poser certaines questions, à propos de manipu-

lations "naturelles", comme le transport parallèle d'un vecteur. Finalement, l'on construit pas à pas une véritable géométrie, littéralement, par analogie avec un arpenteur qui se doterait de moyens de mesures comparables tout au long de son périple sur la surface.

A.2.1 Définition du tenseur métrique

Tout calcul, toute dérivation en géométrie riemannienne implique la construction et la connaissance d'un tenseur fondamental, que l'on appelle le tenseur métrique. C'est lui qui va servir d'outil à l'arpenteur mobile, dont les instruments de mesure vont varier de manière relative lors de ses déplacements, notamment si l'espace est courbe ou déformé : en effet, dans le cas d'une géométrie plate, euclidienne, les instruments sont "rigides". Le tenseur métrique se définit simplement comme le produit scalaire des vecteurs de base, deux à deux, autrement dit :

$$g_{ij} = e_i \cdot e_j$$

Le tenseur métrique structure l'espace intrinsèque, et est défini sur tout l'espace. Il revêt une importance tant théorique que formelle, notamment par le confort de calcul qu'il induit ; il permet, selon sa variance, d'abaisser ou élever, les indices d'une suite tensorielle. Nous donnons donc les définitions formelles des différentes variances du tenseur métrique.

Sous forme contravariante :

$$g^{ij} = e^i \cdot e^j$$

$$g^{ij} = e^i \cdot e^j$$

Sous forme mixte :

$$g_j^i = e^i \cdot e_j$$

$$g_j^i = e^i \cdot e_j$$

A.2.2 Coordonnées curvilignes et symboles de Christoffel

Nous introduisons ici la notion de coordonnées curvilignes, ainsi que les symboles dits de Christoffel, dont l'utilisation sera justifiée dès après, dans la subsection concernant les connexions. Les symboles de Christoffel permettent en outre d'écrire bien des relations. On définit les symboles de Christoffel de la façon suivante :

$$de_i = \Gamma_{ik}^l e_l du^k$$

Ils permettent par exemple d'écrire de façon condensée la différentielle du tenseur métrique :

$$dg^{ij} = d(e_i \cdot e_j) = de_i \cdot e_j + e_i \cdot de_j$$

d'où :

$$\partial_k g_{ij} = g_{il} \Gamma_{jk}^l + g_{lj} \Gamma_{ik}^l$$

A.2.3 Espace tangent

Dans l'hypothèse d'un espace courbe, si l'on veut pouvoir continuer de comparer 2 vecteurs qui ne sont pas "ancrés" au même point de l'espace, il est nécessaire, comme en espace plat, de "transporter" sans "modifier" un des 2 vecteurs au point d'ancrage de l'autre : il s'agit du transport parallèle. Il suppose que l'on puisse mettre pied à pied 2 vecteurs que l'on va alors décomposer dans la même base, afin de pouvoir les comparer, coordonnée par coordonnée. Dans l'espace plat, cela ne pose a priori pas de problème : la structure de l'espace est rigide et la même en tout point.

En espace courbe, le transport parallèle n'a rien d'évident a priori : com-

ment transporter un vecteur vers un autre, pour les comparer, si l'espace ambiant se distord ? Pour ce faire, on introduit la notion d'espace tangent en un des 2 points d'ancrage des vecteurs à comparer ; le second vecteur sera projeté sur ce plan, et c'est dans ce plan que l'on va transporter un vecteur vers l'image de l'autre pour les comparer.

Cependant, un tel procédé nécessite, outre certaines hypothèses ad hoc, de savoir comment la base du point de tangence se transforme en la base au point d'ancrage du vecteur image, dans le plan tangent. Il faut alors introduire le concept de connexion, et en particulier, nous utiliserons une connexion affine.

Connexion, connexion affine

La connexion est donc la transformation qui lie deux bases dans un plan tangent de l'espace courbe. L'idée est comme dans le cas d'un changement de base vers une autre du cas classique, de déterminer la relation de passage. Cette relation, dans le cas d'une connexion affine, fait intervenir évidemment le tenseur métrique, mais aussi les symboles de Christoffel introduits plus hauts.

La connexion affine peut entièrement s'écrire en fonction du tenseur métrique et des symboles de Christoffel, comme suit :

$$\Gamma_{jk}^i = \frac{1}{2}g^{ih}(\partial_j g_{kh} + \partial_k g_{jh} - \partial_h g_{jk})$$

Définition de la courbure

Nous avons défini la grandeur structurante de l'espace – le tenseur métrique, nous avons décrit comment se repérer dans l'espace à l'aide des coordonnées curvilignes ; nous avons expliqué que comparer 2 vecteurs nécessitent le concept de plan tangent et de connexion ; il ne nous reste plus qu'à décrire la grandeur qui va donner sa forme à l'espace, à savoir le tenseur de

courbure, ou courbure. C'est ce tenseur qui va différencier un espace plat d'un espace non-euclidien.

La courbure est une suite tensorielle, elle peut s'écrire de différente façon, en tout point M de l'espace :

$$R_{jkl}^i(M) = \partial_k \Gamma_{jl}^i + \Gamma_{hk}^i \Gamma_{jl}^h - \partial_l \Gamma_{jk}^i - \Gamma_{hl}^i \Gamma_{jk}^h$$

ou aussi :

$$R_{ijkl}(M) = \frac{1}{2}[\partial_{jk}g_{il} + \partial_{il}g_{jk} - \partial_{ik}g_{jl} - \partial_{jl}g_{ik}] + g_{rh}(\Gamma_{il}^r \Gamma_{jk}^h - \Gamma_{ik}^r \Gamma_{jl}^h)$$

Géodésiques en espace Riemannien

Nous donnons ici la formulation générale d'une courbe géodésique, par l'équation paramétrique suivante :

$$\frac{d^2 u^i}{dt^2} + \Gamma_{jk}^i \frac{du^j}{dt} \frac{du^k}{dt} - \frac{1}{2f^2} \frac{d(f^2)}{dt} \frac{du^i}{dt} = 0$$

sachant que f^2 est donnée par :

$$f^2 = g_{ij} \frac{du^i}{dt} \frac{du^j}{dt}$$

A.2.4 Dérivation tensorielle et écriture des opérateurs classiques

L'opérateur divergence d'un champ de vecteurs $V(M)$ s'écrit :

$$div V = \frac{1}{\sqrt{|g|}} \partial_i (\sqrt{|g|} v^i)$$

Le gradient d'un champ scalaire $\varphi(M)$ s'écrit :

$$grad_k(\varphi(M)) = \nabla_k \varphi(M) = \partial_k \varphi(M)$$

$$\nabla_k \varphi(M) = g^{ik} \partial_k \varphi(M)$$

Le Laplacien s'écrit quant à lui :

$$\Delta \varphi = \operatorname{div}(\operatorname{grad} \varphi) = \frac{1}{\sqrt{|g|}} \partial_i (\sqrt{|g|} g^{ij} \partial_j \varphi)$$

soit encore :

$$\Delta \varphi = \nabla_i (g^{ij} \partial_j \varphi)$$

Annexe B

Éléments de géométrie différentielle

B.1 Variétés et fibrés

Variété topologique : Une variété topologique M de dimension n est un espace topologique à base dénombrable, tel que chacun de ses points admet un voisinage homéomorphe à un ouvert de l'espace vectoriel topologique \mathbb{R}^n .

Plus précisément :

$\forall x \in M$, il existe un voisinage ouvert U_x et un homéomorphisme $\varphi_x : U_x \rightarrow \varphi_x(U_x) \subset \mathbb{R}^n$. On dit alors que (U_x, φ_x) est une carte locale de M .

Atlas : Une famille de cartes (U_i, φ_i) qui recouvre (entièrement) M constitue un atlas de la variété M . Un tel atlas est dit de classe C^k , $1 \leq k \leq \infty$, si :

pour tous les indices i, j tels que $U_i \cap U_j \neq \emptyset$, l'application de changement de cartes :

$\varphi_i \circ \varphi_j^{-1} : \varphi_j(U_i \cap U_j) \rightarrow \varphi_i(U_i \cap U_j)$ est un difféomorphisme de classe C^k , c'est-à-dire une bijection de classe C^k dont la bijection réciproque est

aussi de classe C^k .

Compatibilité des atlas : Deux atlas de classe C^k sur une même variété topologique M sont dits compatibles lorsque leur réunion est encore un atlas de classe C^k . La relation de compatibilité ainsi introduite est une relation d'équivalence pour les atlas.

Variété différentielle : Les classes d'équivalence définissent la structure de variété différentielle : une variété différentielle (ou variété différentiable) de classe C^k est une variété topologique munie d'une famille d'atlas de classe C^k tous compatibles avec un atlas donné. Une variété lisse est une variété différentielle de classe C^∞

Dans chaque classe d'équivalence existe un représentant privilégié, l'atlas maximal, obtenu en considérant toutes les cartes compatibles avec l'atlas initial.

Différentiabilité d'une fonction : La différentiabilité d'une fonction f entre deux variétés différentielles M et N se définit en procédant à la lecture de f dans des cartes locales au voisinage de chaque point.

Formellement, on considère deux variétés M et N de classe C^k et un entier j inférieur à k . Une application $f : M \rightarrow N$ est dite de classe C^k quand, pour tout point m de M , on peut trouver une carte locale (U, ϕ) de M centrée en m et une carte locale (V, ψ) de N centrée en $n = f(m)$ telles que :

$f_{\phi, \psi} = \psi \circ f \circ \phi^{-1}$ est bien définie et est de classe C^j de $\phi(U)$ dans $\psi(V)$.

Si on considère de nouvelles cartes locales, la régularité des applications de changement de cartes assure que les fonctions $f_{\phi', \psi'}$ correspondantes seront régulières, ce qui montre la cohérence de la définition.

Il est un peu plus délicat de définir la notion de différentielle d'une application différentiable, puisque cela demande d'introduire au préalable les vecteurs tangents. Cependant, le rang de la différentielle de $f_{\phi, \psi}$ au point m

ne dépend pas des cartes choisies. On parle notamment d'immersion si ce rang est égal en tout point à la dimension de M , de submersion s'il est égal en tout point à la dimension de N .

Vecteur tangent à une variété : On peut définir de façon relativement simple la notion de vecteur tangent à une sous-variété de R^n , et cette définition donne des vecteurs de l'espace ambiant. Cependant, il est utile et important de disposer d'une construction qui ne fasse pas intervenir de plongement spécifique, mais qui soit valable au niveau des variétés abstraites. Un mode de définition possible est basé sur la relation de tangence entre les courbes. On peut, par lecture dans une carte locale, définir la notion de courbes tangentes en un point m de la variété M . Les vecteurs tangents en m à la variété sont alors les classes d'équivalence pour cette relation. Ils forment un espace vectoriel $T_m M$ qui a la même dimension n que M .

Fibré vectoriel La définition d'un fibré vectoriel de rang n nécessite la donnée des éléments suivants :

- deux espaces topologiques : soit X un premier espace appelé *base* et E un second espace, appelé *espace total* ;
- une application continue $\pi : E \rightarrow X$, nommée *projection* ;
- chaque fibre $\pi^{-1}(x)$ est dotée d'une structure d'espace vectoriel.

On emploie ce qu'on appelle une *trivialisation locale*, laquelle consiste en la donnée d'un ouvert U tel que $U \in X$, et d'un homéomorphisme ϕ dit de trivialisation : $U \times R^n \rightarrow \pi^{-1}(U)$ tel que pour tout point x de U :

- $\pi\phi(x, v) = x$ pour tous les vecteurs v de R^n
- l'application $v \rightarrow \phi(x, v)$ induit un isomorphisme de R^n sur $\pi^{-1}(x)$.

Enfin, une dernière propriété est requise pour déterminer un fibré vectoriel : on requiert que chaque élément de la base X appartienne à un voisinage ouvert U de trivialisation.

De la même façon que pour nombre de constructions d'espaces, il est possible d'enrichir la définition de fibré vectoriel par un certain nombre de propriétés additionnelles. On peut ainsi parler de fibré vectoriel différentiable de classe C^k . Dans ce cas, il est nécessaire d'avoir :

- la base X et l'espace total E sont des variétés différentielles de classe C^k ;
- la projection π est de classe C^k aussi ;
- il existe des trivialisations locales de classe C^k .

La définition d'un fibré nous permet alors d'aborder et de construire la notion de fibré tangent à une variété.

Fibré tangent à une variété : La réunion disjointe de tous les espaces tangents $T_m M$ pour m appartenant à M forme une variété TM de dimension $2n$, appelée fibré tangent à la variété. L'espace TM possède une structure de fibré vectoriel de base M . Cependant cette construction engendre une perte de différentiabilité : si M est de classe C^{k+1} , le fibré tangent est de classe C^k . On peut ensuite définir la différentielle d'une application différentiable f entre deux variétés M et N . On note m un point de M et n son image ; alors la différentielle $T_m f$ de f au point m est une application linéaire de l'espace tangent $T_m M$ vers $T_n N$. Pour la définir, on réintroduit les applications $f_{\phi, \psi}$ représentant la fonction f lue dans des cartes locales, et on en prend la différentielle au point $\phi^{-1}(m)$ correspondant à m . Une nouvelle fois, l'objet ainsi construit est indépendant des cartes utilisées.

Fibré tangent à une variété, plus formellement : On définit $T(M)$ en se donnant pour chaque ouvert U de M une trivialisatation locale :

$$\varphi_U \left\{ \begin{array}{l} T(M) \rightarrow U \times V \\ m \mapsto (P_U(m), v_U(m)) \end{array} \right\}$$

où V est un espace vectoriel isomorphe à l'espace tangent à M en n'importe quel $P \in U$ et pour chaque $m \in T(M)$, $v_U(m)$ appartient à l'espace

tangent à M en $P_U(m)$.

Par ailleurs φ_U doit satisfaire à la condition de recollement suivante :

Si $P_0 = P(m) \in U_1 \cap U_2$ où U_1 et U_2 , sont des ouverts associés à des cartes x_1^μ et x_2^μ alors on doit avoir (en notation de coordonnées pour les vecteurs v_{U_1} et v_{U_2}) :

$v_{U_2}^\mu(m) = \frac{\partial x_2^\mu}{\partial x_1^\nu} \Big|_{P_0} v_{U_1}^\nu(m)$ où on a adopté la convention de sommation d'indices répétés d'Einstein.

B.2 Notion de jets

Fibré de jets (*Jet Bundles*) En géométrie différentielle, on appelle fibré de jet un type de construction qui consiste à bâtir un nouveau fibré à partir d'un fibré initial. L'intérêt de cette construction tient dans la possibilité d'écrire des équations différentielles dont le support est alors les sections des fibrés, et dont la formulation sera invariante. Les jets peuvent également être vus comme une version des développements de Taylor, indépendante de la base de coordonnées. Les fibrés de jet sont désormais reconnus comme étant l'objet adapté à une formulation covariante en théorie des champs, de type relativité générale, par exemple.

Jets : définition. Soient M une variété de dimension m et (E, π, M) un fibré. Pour tout p de M , on pose $\Gamma(\pi)$ l'ensemble constitué des sections locales contenant p . Soit $I = (I(1), I(2), \dots, I(m))$ un m -uplet d'entiers ordonnés), on a alors :

$$|I| := \sum_{i=1}^m I(i) \quad \frac{\partial^{|I|}}{\partial x^I} := \prod_{i=1}^m \left(\frac{\partial}{\partial x^i} \right)^{I(i)} .$$

Si $\sigma, \eta \in \Gamma(\pi)$ sont deux sections locales, on dit qu'elles ont le même r -jet en p si :

$$\frac{\partial^{|I|} \sigma^\alpha}{\partial x^I} \Big|_p = \frac{\partial^{|I|} \eta^\alpha}{\partial x^I} \Big|_p, \quad 0 \leq |I| \leq r.$$

Ce type de relation – posséder le même r -jet en un point p – définit une relation d'équivalence. Un r -jet définit donc une classe d'équivalence pour cette relation. On note $j_p^r \sigma$ un tel r -jet. r est l'ordre du jet, p est sa source et $\sigma(p)$ sa cible.

Cas en dimension 1 Soit $f : R \rightarrow R$ une fonction réelle présentant au moins $k + 1$ dérivées successives définies sur un voisinage U du point x_0 . Le théorème de Taylor nous permet d'écrire :

$$f(x) = f(x_0) + f'(x_0)(x - x_0) + \cdots + \frac{f^{(k)}(x_0)}{k!}(x - x_0)^k + \frac{R_{k+1}(x)}{(k+1)!}(x - x_0)^{k+1}$$

où

$$|R_{k+1}(x)| \leq \sup_{x \in U} |f^{(k+1)}(x)|.$$

Dans ce cas le k -jet de f au point x_0 est défini par le polynôme :

$$(J_{x_0}^k f)(z) = f(x_0) + f'(x_0)z + \cdots + \frac{f^{(k)}(x_0)}{k!}z^k.$$

Ainsi, les jets sont souvent considérés comme des polynômes en la variable z , et non comme des fonctions polynomiales de cette variable. Autrement dit, z est vue comme une variable indéterminée, ce qui laisse un degré de liberté permettant diverses opérations algébriques sur les jets. La véritable variable dont dépendent les jets, en tant que fonction, est en fait le point x_0 . Si l'on fait varier ce point x_0 , la définition du jet en ce point entraîne l'existence de polynômes d'ordre au moins k en chacun de ces points. Ce détail a son importance, dans le sens où, habituellement, les développements de Taylor sont davantage utilisés pour leur dépendance fonctionnelle en leur variable z , et non en leur point de développement x_0 , ce qui s'avère être le contraire lorsque l'on considère un jet.

Exemple : correspondance entre deux espaces euclidiens. Soit $f : R^n \rightarrow R^m$ une fonction opérant d'un espace euclidien vers un autre, et présentant au moins $(k + 1)$ dérivées successives. Là aussi, le théorème de Taylor nous permet d'écrire :

$$f(x) = f(x_0) + (Df(x_0)) \cdot (x - x_0) + \frac{1}{2}(D^2f(x_0)) \cdot (x - x_0)^{\otimes 2} + \dots + \frac{D^k f(x_0)}{k!} \cdot (x - x_0)^{\otimes k} + \frac{D^{k+1} f(x_0)}{(k+1)!} \cdot (x - x_0)^{\otimes (k+1)}.$$

Ainsi, le k -jet de la fonction f est défini par le polynôme :

$$(J_{x_0}^k f)(z) = f(x_0) + (Df(x_0)) \cdot z + \frac{1}{2}(D^2f(x_0)) \cdot z^{\otimes 2} + \dots + \frac{D^k f(x_0)}{k!} \cdot z^{\otimes k}$$

appartenant à $R[z]$, où $z = (z_1, \dots, z_n)$.

Définition du jet en un point d'un espace euclidien : plus formellement. Soit $C^\infty(R^n, R^m)$ l'espace vectoriel des fonctions lisses $f : R^n \rightarrow R^m$. Soit k un entier non négatif, et p un point de R^n . On peut définir la relation d'équivalence E_p^k sur cet espace en posant que deux fonctions f et g sont équivalentes jusque l'ordre k si f et g ont la même valeur en p , et que toutes leurs dérivées partielles correspondent en p jusque la dérivée k -ème incluse. Autrement dit, $f \sim g$ si et seulement si $f - g = 0$ jusqu'au k -ème ordre. Ainsi, l'espace de jet du k -ème ordre de $C^\infty(R^n, R^m)$ en p est défini comme l'ensemble des classes d'équivalence de E_p^k , et est noté $J_p^k(R^n, R^m)$. Le jet du k -ème ordre en p d'une fonction lisse $f \in C^\infty(R^n, R^m)$ est défini comme la classe d'équivalence de f dans $J_p^k(R^n, R^m)$.

Annexe C

Géomérodynamique et formalisme Hamiltonien multi-temps

C.1 Éléments de géomérodynamique

Si la question de savoir si le formalisme Lagrangien peut être totalement équivalent au formalisme Hamiltonien semble toujours faire l'objet de débats, la relation établie entre ces deux grandeurs demeure relativement peu contestée, indépendamment des hypothèses subséquentes qui peuvent être faites pour avancer cette équivalence. En effet, on voit l'Hamiltonien d'un système comme la transformée de Legendre du Lagrangien de ce même système.

En outre, si l'Hamiltonien nous est souvent précieux en raison de sa présence au coeur de plusieurs théories, le Lagrangien a pour avantage d'apparaître directement dans les équations dites d'Euler-Lagrange, lesquelles nous permettent de traduire formellement et indépendamment de la théorie retenue, le principe de moindre action.

C.1.1 Equations d'Euler-Lagrange à temps unique.

On sait que le Lagrangien L , défini comme

$$L : R \times R^{2n} \rightarrow R, (t, x, \dot{x}) \rightarrow L(t, x, \dot{x}),$$

où $x = (x^1(t), \dots, x^n(t))$ et $\dot{x} = (\dot{x}^1(t), \dots, \dot{x}^n(t))$, permet d'en dériver les équations d'Euler-Lagrange à temps unique

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{x}^i} = \frac{\partial L}{\partial x^i}, \quad i = 1, \dots, n. \quad (1)$$

De fait, nous voyons comment il est possible d'en déduire les équations de l'Hamiltonien à temps unique, via la transformation de Legendre.

C.1.2 Hamiltonien à temps unique.

Supposons que le Lagrangien L définissent le difféomorphisme suivant : $\dot{x}^i \rightarrow p_i = \frac{\partial L}{\partial \dot{x}^i}$ que l'on appelle par ailleurs la transformation de *Legendre*. La transformation de Legendre du Lagrangien L est l'Hamiltonien H défini comme :

$$H : R \times R^{2n} \rightarrow R (t, x, p) \rightarrow H(t, x, p),$$

$$H(t, x, p) = \dot{x}^i p_i - L(t, x, \dot{x}).$$

Ainsi, l'équation (1) devient

$$\frac{dx^i}{dt} = \frac{\partial H}{\partial p_i}, \quad \frac{dp_i}{dt} = -\frac{\partial H}{\partial x^i}, \quad i = 1, \dots, n \quad (2)$$

car

$$\frac{\partial H}{\partial x^i} = p_j \frac{\partial \dot{x}^j}{\partial x^i} - \frac{\partial L}{\partial x^i} - \frac{\partial L}{\partial \dot{x}^j} \frac{\partial \dot{x}^j}{\partial x^i} = -\frac{\partial L}{\partial x^i} = -\frac{d}{dt} p_i$$

et

$$\frac{\partial H}{\partial p_i} = \dot{x}^i + p_j \frac{\partial \dot{x}^j}{\partial p_i} - \frac{\partial L}{\partial \dot{x}^j} \frac{\partial \dot{x}^j}{\partial p_i} = \dot{x}^i. \quad (3)$$

De cette sorte, les équations différentielles ordinaires (1) de second ordre définies dans un espace de dimension n sont équivalentes aux équations différentielles ordinaires (2) définies dans un espace de dimension $2n$. On peut également écrire les équations d'Hamilton sous forme matricielle, en utilisant

$$\text{la structure symplectique } J = \begin{pmatrix} 0 & \delta_j^i \\ -\delta_j^i & 0 \end{pmatrix}, \quad (\delta_j^i) \in M_n(R),$$

$$J \begin{pmatrix} \dot{x}^j \\ \dot{p}_i \end{pmatrix} + \begin{pmatrix} \frac{\partial H}{\partial x^i} \\ \frac{\partial H}{\partial p_i} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

C.1.3 Equations d'Euler-Lagrange multi-temps

Soit la variable multi-temps $t = (t^1, \dots, t^p) \in R^p$, les fonctions x^i telles que :

$$x^i : R^p \rightarrow R \quad (t^1, \dots, t^p) \rightarrow x^i(t^1, \dots, t^p),$$

$i = 1, \dots, n$. On pose également $x_\alpha^i = \frac{\partial x^i}{\partial t^\alpha}$, $\alpha = 1, \dots, p$. Le Lagrangien

$$L : R^{p+n+np} \rightarrow R, \quad (t^\alpha, x^i, x_\alpha^i) \rightarrow L(t^\alpha, x^i, x_\alpha^i)$$

permet de dériver les équations d'Euler-Lagrange correspondantes

$$\frac{\partial}{\partial t^\alpha} \frac{\partial L}{\partial x_\alpha^i} = \frac{\partial L}{\partial x^i}, \quad i = 1, \dots, n, \quad \alpha = 1, \dots, p$$

C.1.4 Equations d'Hamilton multi-temps

Lors des premières approches d'écriture d'un Hamiltonien multi-temps (de Donder, 1935), celui-ci était obtenu à partir du Lagrangien en utilisant les dérivées partielles suivantes

$$p_k^\alpha = \frac{\partial L}{\partial x_\alpha^k} \tag{4}$$

L'Hamiltonien s'écrit alors : $H = p_k^\alpha x_\alpha^k - L$. Les équations (4) permettent de définir une bijection de classe C^1 $x_\alpha^i \rightarrow p_i^\alpha$, que l'on appellera transformation de Legendre dans le cas multi-temps.

Cette transformation nous permet d'écrire :

$$\frac{\partial H}{\partial p_i^\alpha} = x_\alpha^i + p_k^\beta \frac{\partial x_\beta^k}{\partial p_i^\alpha} - \frac{\partial L}{\partial x_\beta^k} \frac{\partial x_\beta^k}{\partial p_i^\alpha} = x_\alpha^i$$

$$\frac{\partial H}{\partial x^i} = p_k^\alpha \frac{\partial x_\alpha^k}{\partial x^i} - \frac{\partial L}{\partial x^i} - \frac{\partial L}{\partial x_\alpha^k} \frac{\partial x_\alpha^k}{\partial x^i} = -\frac{\partial L}{\partial x^i}.$$

Les $np + n$ équations d'Hamilton $\frac{\partial x^i}{\partial t^\alpha} = \frac{\partial H}{\partial p_i^\alpha}$, $\frac{\partial p_i^\alpha}{\partial t^\alpha} = -\frac{\partial H}{\partial x^i}$, $i = 1, \dots, n$, $\alpha = 1, \dots, p$ sont alors des équations aux dérivées partielles du premier ordre, définies sur l'espace R^{n+pn} , et équivalentes aux équations d'Euler-Lagrange définies sur R^n .

Ainsi, de la même façon que précédemment, on peut écrire les équations d'Hamilton dans le cas multi-temps

$$\delta_\beta^\alpha \delta_j^i \frac{\partial p_i^\beta}{\partial t^\alpha} + \frac{\partial H}{\partial x^j} = 0, \quad -\delta_\beta^\alpha \delta_j^i \frac{\partial x^j}{\partial t^\alpha} + \frac{\partial H}{\partial p_i^\beta} = 0$$

ce qui donne sous forme matricielle :

$$(\delta \otimes J) \begin{pmatrix} \frac{\partial x^j}{\partial t^\alpha} \\ \frac{\partial p_i^\beta}{\partial t^\alpha} \end{pmatrix} + \begin{pmatrix} \frac{\partial H}{\partial x^j} \\ \frac{\partial H}{\partial p_j^\beta} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \quad (5)$$

$j = 1, \dots, n$; $i = 1, \dots, n$; $\alpha = 1, \dots, p$; $\beta = 1, \dots, p$, où

$$\delta \otimes J = \begin{pmatrix} 0 & \delta_\beta^\alpha \delta_j^i \\ -\delta_\beta^\alpha \delta_j^i & 0 \end{pmatrix}$$

est une structure polysymplectique opérant sur R^{np+np^2} et à valeur dans R^{n+np} .

C.2 Le fibré de jets $J^1(T, M)$ et le fibré dual, support de l'Hamiltonien multi-temps

Si nous avons vu formellement la définition d'un fibré de jets, nous exposons ici la construction et le formalisme usuel pour le fibré de jets du premier ordre, qui sert de support au Lagrangien multi-temps, et de manière similaire, pour le fibré dual $J^{1*}(T, M)$, servant de support à l'Hamiltonien multi-temps. Nous n'irons pas jusque les détails fins de ces constructions : nous nous limitons ici à fournir l'espace nécessaire à asseoir la structure de l'espace informationnel. Le lecteur pourra consulter les travaux qui sont essentiellement le fait de Neagu et Udriste, [322, 323, 307, 324] sur des recherches et résultats préliminaires de de Donder, Dirac, ou plus récemment Saunders. [310, 304]

Soient les variétés lisses T et M de dimensions respectives p et n , et leurs systèmes de coordonnées respectifs $(t^\alpha)_{\alpha=\overline{1,p}}$, et $(x^i)_{i=\overline{1,n}}$. Ici, les ensembles $\{1, 2, \dots, p\}$, respectivement $\{1, 2, \dots, n\}$, sont indicés par $\alpha, \beta, \gamma, \dots$, respectivement i, j, k, \dots

Soit (t_0, x_0) un point quelconque de la variété produit $T \times M$. On note $C^\infty(T, M)$ l'ensemble de toutes les cartes lisses entre T et M , et l'on définit la relation d'équivalence sur $C^\infty(T, M)$:

$$f \sim_{(t_0, x_0)} g \Leftrightarrow \begin{cases} f(t_0) = g(t_0) = x_0 \\ df_{t_0} = dg_{t_0}, \end{cases} \quad (\text{C.1})$$

Il est possible de donner une expression locale de cette relation d'équivalence $f \sim_{(t_0, x_0)} g$ pour toutes fonctions $f, g \in C^\infty(T, M)$, comme suit :

$$\begin{cases} x^i(t_0^\beta) = y^i(t_0^\beta) = x_0^i \\ \frac{\partial x^i}{\partial t^\alpha}(t_0^\beta) = \frac{\partial y^i}{\partial t^\alpha}(t_0^\beta), \end{cases} \quad (\text{C.2})$$

où $t^\beta(t_0) = t_0^\beta$, $x^i(x_0) = x_0^i$, $x^i = x^i \circ f$ et $y^i = x^i \circ g$.

On note $[f]_{(t_0, x_0)} = \{g \in C^\infty(T, M) \mid g \sim_{(t_0, x_0)} f\}$ la classe d'équivalence d'une carte lisse $f \in C^\infty(T, M)$.

Considérons le quotient $C^\infty(T, M) / \sim_{(t_0, x_0)}$ par la relation d'équivalence $\sim_{(t_0, x_0)}$, et notons-le $J_{t_0, x_0}^1(T, M)$. On peut alors construire l'espace ensemble des 1-jets (ou des jets du premier ordre) de la manière suivante :

$$J^1(T, M) = \bigcup_{(t_0, x_0) \in T \times M} J_{t_0, x_0}^1(T, M). \quad (\text{C.3})$$

Il est possible de structurer cet espace des 1-jets $J^1(T, M)$ comme un fibré vectoriel construit sur la base de l'espace $T \times M$. Pour ce faire, prenons une carte lisse $f \in C^\infty(T, M)$, $(t^1, \dots, t^p) \rightarrow (x^1(t^1, \dots, t^p), \dots, x^n(t^1, \dots, t^p))$, et procédons au développement de Taylor des cartes x^i au point $(t_0^1, \dots, t_0^p) \in R^p$.

Le développement donne :

$$x^i(t^1, \dots, t^p) = x_0^i + (t^\alpha - t_0^\alpha) \frac{\partial x^i}{\partial t^\alpha}(t_0^1, \dots, t_0^p) + \mathcal{O}(2), \quad \|(t^1 - t_0^1, \dots, t^p - t_0^p)\| < \varepsilon.$$

Si l'on considère la carte lisse $\tilde{f} \in C^\infty(T, M)$ définies par localement par l'ensemble de fonctions

$$\tilde{x}^i(t^1, \dots, t^p) = x_0^i + (t^\alpha - t_0^\alpha) \frac{\partial x^i}{\partial t^\alpha}(t_0^1, \dots, t_0^p), \quad \|(t^1 - t_0^1, \dots, t^p - t_0^p)\| < \varepsilon,$$

on montre que $\tilde{f} \sim_{(t_0, x_0)} f$, (qui est l'approximation linéaire \tilde{f} de f) fournit une façon d'écrire la classe d'équivalence $[f]_{(t_0, x_0)}$.

Soit maintenant $\pi : J^1(T, M) \rightarrow T \times M$ la projection définie par $\pi([f]_{(t_0, x_0)}) = (t_0, f(t_0))$. De manière évidente, la carte π est bien définie et surjective. Si l'on utilise cette projection pour chaque carte locale $U \times V \subset T \times M$ définie

sur la variété $T \times M$, il devient possible de définir la bijection

$$\Phi_{U \times V} : \pi^{-1}(U \times V) \rightarrow U \times V \times R^{np},$$

$$\Phi_{U \times V}([f]_{(t_0, x_0)}) = (t_0, x_0, \frac{\partial x^i}{\partial t^\alpha}(t_0^\beta)), \quad x_0 = f(t_0).$$

Autrement dit encore, l'ensemble des 1-jets $J^1(T, M)$ peut être muni d'une structure différentiable de dimension $p + n + pn$, telle que les cartes $\Phi_{U \times V}$ soient des difféomorphismes. Notamment, les coordonnées locales de $J^1(T, M)$ sont données par $(t^\alpha, x^i, x_\alpha^i)$, où

$$\begin{cases} t^\alpha([f]_{(t_0, x_0)}) = t^\alpha(t_0) \\ x^i([f]_{(t_0, x_0)}) = x^i(x_0) \\ x_\alpha^i([f]_{(t_0, x_0)}) = \frac{\partial x^i}{\partial t^\alpha}(t_0^\beta). \end{cases} \quad (\text{C.4})$$

Pour ces coordonnées de $J^1(T, M)$, la projection $\pi : J^1(T \times M) \rightarrow T \times M$ s'écrit localement $\pi(t^\alpha, x^i, x_\alpha^i) = (t^\alpha, x^i)$. De même, la différentielle π_* de la carte π est localement donnée par le Jacobien

$$\begin{pmatrix} \delta_{\alpha\beta} & 0 & 0 \\ 0 & \delta_{ij} & 0 \end{pmatrix} \in M_{p+n, p+n+pn}.$$

Il vient alors que π_* est une surjection ($\text{rang}(\pi_*) = p + n$), et donc que la projection π est une submersion. De fait, l'espace des 1-jets $J^1(T, M)$ est un fibré vectoriel sur l'espace de base $T \times M$, de fibre R^{pn} .

Les relations de transformations des coordonnées locales $(t^\alpha, x^i, x_\alpha^i) \longleftrightarrow (\tilde{t}^\alpha, \tilde{x}^i, \tilde{x}_\alpha^i)$ du fibré vectoriel $E = J^1(T, M)$ sont données par :

$$\begin{cases} \tilde{t}^\alpha = \tilde{t}^\alpha(t^\beta) \\ \tilde{x}^i = \tilde{x}^i(x^j) \\ \tilde{x}_\alpha^i = \frac{\partial \tilde{x}^i}{\partial x^j} \frac{\partial t^\beta}{\partial \tilde{t}^\alpha} x_\beta^j, \end{cases} \quad (\text{C.5})$$

où $\det(\partial\tilde{t}^\alpha/\partial t^\beta) \neq 0$ et $\det(\partial\tilde{x}^i/\partial x^j) \neq 0$. Ainsi, E est une variété toujours orientable.

Considérons la base canonique $\left\{ \frac{\partial}{\partial t^\alpha}, \frac{\partial}{\partial x^i}, \frac{\partial}{\partial x_\alpha^i} \right\}$ des champs vectoriels définis sur E et $\{dt^\alpha, dx^i, dx_\alpha^i\}$ sa base duale.

Si l'on change de coordonnées sur E , on obtient les relations de transformations suivantes :

$$\left\{ \begin{array}{l} \frac{\partial}{\partial t^\alpha} = \frac{\partial\tilde{t}^\beta}{\partial t^\alpha} \frac{\partial}{\partial\tilde{t}^\beta} + \frac{\partial\tilde{x}_\beta^j}{\partial t^\alpha} \frac{\partial}{\partial\tilde{x}_\beta^j} \\ \frac{\partial}{\partial x^i} = \frac{\partial\tilde{x}^j}{\partial x^i} \frac{\partial}{\partial\tilde{x}^j} + \frac{\partial\tilde{x}_\beta^j}{\partial x^i} \frac{\partial}{\partial\tilde{x}_\beta^j} \\ \frac{\partial}{\partial x_\alpha^i} = \frac{\partial\tilde{x}^j}{\partial x_\alpha^i} \frac{\partial t^\alpha}{\partial\tilde{t}^\beta} \frac{\partial}{\partial\tilde{x}_\beta^j}, \end{array} \right. \quad (\text{C.6})$$

$$\left\{ \begin{array}{l} dt^\alpha = \frac{\partial t^\alpha}{\partial\tilde{t}^\beta} d\tilde{t}^\beta \\ dx^i = \frac{\partial x^i}{\partial\tilde{x}^j} d\tilde{x}^j \\ dx_\alpha^i = \frac{\partial x_\alpha^i}{\partial\tilde{t}^\beta} d\tilde{t}^\beta + \frac{\partial x_\alpha^i}{\partial\tilde{x}^j} d\tilde{x}^j + \frac{\partial x^i}{\partial\tilde{x}^j} \frac{\partial\tilde{t}^\beta}{\partial t^\alpha} d\tilde{x}_\beta^j. \end{array} \right. \quad (\text{C.7})$$

D'un point de vue plus physique, l'espace T peut bien sûr être vu comme une variété de type "*temporel*" ou en l'occurrence "*multi-temps*" tandis que la variété M peut être vue comme de type "*spatial*". Le fibré vectoriel $J^1(T, M) \rightarrow T \times M$ peut lui représenter le *fibré des configurations*, au sens d'un espace de configurations comme on peut le définir usuellement en mécanique.

Il est possible d'examiner le cas particulier où T est l'ensemble des réels R , reflétant le cas classique, à temps unique. Supposons alors que l'on ait $J^1(R, M) \equiv R \times TM$, dont les éléments sont de coordonnées (t, x^i, y^i) . Le groupe de jauge du fibré

$$\pi : J^1(R, M) \rightarrow R \times M, \quad (t, x^i, y^i) \rightarrow (t, x^i), \quad (\text{C.8})$$

est alors donné par

$$\begin{cases} \tilde{t} = \tilde{t}(t) \\ \tilde{x}^i = \tilde{x}^i(x^j) \\ \tilde{y}^i = \frac{\partial \tilde{x}^i}{\partial x^j} \frac{dt}{d\tilde{t}} y^j. \end{cases} \quad (\text{C.9})$$

Dans ce cas, le groupe de jauge se distingue par la caractère relativiste de la variable temporelle t . Le fibré de jet d'ordre 1 $J^1(R, M)$ peut de cette sorte être considéré comme l'ensemble naturel des configurations de ce plusieurs auteurs appellent la mécanique Lagrangienne rhéonomique relativiste (on parle de contraintes rhéonomiques lorsque celles-ci mentionnement et dépendent explicitement d'une variable temporelle).

Dans le cas classique, c'est-à-dire, par opposition au cas relativiste, nous aurions le fibré des configurations donné par

$$\pi : R \times TM \rightarrow M, (t, x^i, y^i) \rightarrow (x^i), \quad (\text{C.10})$$

et de groupe d'invariance associé :

$$\begin{cases} \tilde{t} = t \\ \tilde{x}^i = \tilde{x}^i(x^j) \\ \tilde{y}^i = \frac{\partial \tilde{x}^i}{\partial x^j} y^j. \end{cases} \quad (\text{C.11})$$

Dans ce second cas de figure, la variable temporelle t est une variable "absolue". Par ailleurs, notons que le groupe de jauge C.11 est un sous-groupe de C.9, ce qui indique une plus grande généralité du groupe de jauge du fibré de jet. Davantage d'informations concernant les développements théoriques et techniques des mécaniques Lagrangiennes rhéonomiques classiques et relativistes peuvent être trouvées dans [307].

Les constructions relatives à un Hamiltonien multi-temps, ainsi que les objets associés à sa géométrie, sont exposées et librement accessibles dans les références suivantes [325, 309, 326, 308].

Bibliographie

- [1] Aristote, *Physique*. Paris : Flammarion, 2002.
- [2] R. Descartes, *Discours de la méthode*. Flammarion, Apr. 2000.
- [3] M. Foucault, *Naissance de la clinique*. Presses Universitaires de France - PUF, 8e édition ed., Nov. 2009.
- [4] A. Einstein and R. W. Lawson, *Relativity : the special and general theory*. 2012.
- [5] A. Einstein, *Théorie de la relativité restreinte et générale*. Paris : Éd. Payot & Rivages, 2001.
- [6] Euclide, *Les oeuvres d'Euclide*. Paris : A. Blanchard, 1993.
- [7] M. Berger, *A panoramic view of Riemannian geometry*. Berlin ; New York : Springer, 2003.
- [8] E. Durkheim, *Les règles de la méthode sociologique*. Paris : Éd. Flammarion, 2010.
- [9] E. Durkheim, *Le suicide : Étude de sociologie*. Presses Universitaires de France - PUF, Oct. 2007.
- [10] M. Weber, *Économie et société*. Paris : Pocket, 2003.
- [11] E. Durkheim, *Le socialisme sa définition, ses débuts, la doctrine saint-simonienne*. Paris : Presses universitaires de France, 2011.
- [12] C. Bernard, *Introduction à l'étude de la médecine expérimentale*. Flammarion, Aug. 2008.

- [13] B. Latour and V. A. Lépinay, *L'économie, science des intérêts passionnés : introduction à l'anthropologie économique de Gabriel Tarde*. Paris : La Découverte, 2008.
- [14] Halbwachs, *Morphologie sociale*. Armand Colin, Mar. 1999.
- [15] P. Bourdieu, *La distinction*. Les Editions de Minuit, Aug. 1979.
- [16] B. Bensaude-Vincent, *L'opinion publique et la science à chacun son ignorance*. Paris : La Découverte, 2013.
- [17] S. Shapin and S. Schaffer, *Leviathan and the air-pump Hobbes, Boyle, and the experimental life : with a new introduction by the authors*. Princeton, N.J. : Princeton University Press, 2011.
- [18] B. Latour, *Nous n'avons jamais été modernes : essai d'anthropologie symétrique*. Paris : La Découverte, 1997.
- [19] T. Keve and S. Taussig, *Trois explications du monde*. Paris : Albin Michel, 2010.
- [20] J. D. Nasio, *Introduction à la topologie de Lacan*. Paris : Payot & Rivages, 2009.
- [21] C. Shannon, "A mathematical theory of communication," *The Bell System Technical Journal*, vol. 27, pp. 379–423.
- [22] N. Wiener, *Cybernetics or Control and Communication in the Animal and the Machine*. Kessinger Publishing, Mar. 2007.
- [23] L. Brillouin, *La science et la théorie de l'information*. Sceaux, Hauts-de-Seine : Jacques Gabay, 1988.
- [24] L. v. Bertalanffy, *Théorie générale des systèmes*. Paris : Dunod, 2012.
- [25] W. Arthur, "D'Arcy thompson and the theory of transformations," *Nature Reviews Genetics*, vol. 7, pp. 401–406, Jan. 2006.
- [26] D. W. Thompson, *On growth and form*. Cambridge [Eng.] University press, 1917.

- [27] A. M. Turing, “On computable numbers, with an application to the entscheidungsproblem,” *Proceedings of the London Mathematical Society*, vol. s2-42, pp. 230–265, Jan. 1937.
- [28] A. M. Turing, “The chemical basis of morphogenesis,” *Philosophical Transactions of the Royal Society B : Biological Sciences*, vol. 237, pp. 37–72, Aug. 1952.
- [29] A. M. Turing and J.-Y. Girard, *La Machine de Turing*. Paris : Editions du Seuil, 1999.
- [30] V. Neumann, *Théorie générale et logique des automates*. Seyssel (France) : Ed. Champ Vallon, 1996.
- [31] E. Nagel, J. R. Newman, K. Gödel, *et al.*, *Le théorème de Gödel*. Paris : Éd. du Seuil, 1997.
- [32] R. Thom, *Stabilité structurelle et morphogénèse*. Dunod, 2ème ed., 1984.
- [33] R. Thom, *Stabilité structurelle et morphogénèse ; essai d’une théorie générale des modèles*. Reading, Mass. : W.A. Benjamin, 1972.
- [34] H. Dang-Vu and C. Delcarte, *Bifurcations et chaos : une introduction à la dynamique contemporaine avec des programmes en Pascal, Fortran et Mathematica*. Paris : Ellipses Ed. Marketing, 2000.
- [35] B. Mandelbrot, *The fractal geometry of nature*. San Francisco : W.H. Freeman, 1982.
- [36] B. Mandelbrot, “How long is the coast of britain? statistical self-similarity and fractional dimension,” *Science*, vol. 156, pp. 636–638, May 1967.
- [37] P. Bourguine and A. Lesne, *Morphogénèse : L’origine des formes*. Paris : Belin, 2006.

- [38] Y. Kosmann-Schwarzbach, *Groupes et symétries : groupes finis, groupes et algèbres de Lie, représentations*. Palaiseau [France] : Éd. de l'École polytechnique, 2006.
- [39] E. Noether, Y. Kosmann-Schwarzbach, and L. Meersseman, *Les théorèmes de Noether : invariance et lois de conservation au XXe siècle : avec une traduction de l'article original "Invariante Variationsprobleme"*. Palaiseau (Essonne) : Éd. de l'École polytechnique, 2004.
- [40] G. Guillard, C. Hamitouche, and C. Roux, "Using dual quaternions to describe ruled surfaces," in *Curves and Surfaces*, 2006.
- [41] J. Gil Cano, E. Stindel, G. Guillard, *et al.*, "Detection of the functional knee center using the mean helical axis : Application in computer assisted high tibial osteotomy," in *IEEE International Symposium on Biomedical Imaging : from Nano to Macro*, pp. 1557 – 1560, 2004.
- [42] G. Guillard, C. Hamitouche, J. Gil Cano, *et al.*, "Relationship between the morphology and functionality of articular surfaces. applications to 3d navigation," in *IEEE International Symposium on Biomedical Imaging : from Nano to Macro*, pp. 452 – 455, 2004.
- [43] S. Kondo and T. Miura, "Reaction-diffusion model as a framework for understanding biological pattern formation," *Science*, vol. 329, pp. 1616–1620, Sept. 2010. PMID : 20929839.
- [44] P. K. Maini, R. E. Baker, and C.-M. Chuong, "The turing model comes of molecular age," *Science*, vol. 314, pp. 1397–1398, Jan. 2006. PMID : 17138885.
- [45] G. Vogel, "Turing pattern fingered for digit formation," *Science*, vol. 338, pp. 1406–1406, Dec. 2012.
- [46] H. Laborit, *Éloge de la fuite*. Paris : Gallimard, 1985.
- [47] H. Atlan, *L'organisation biologique et la théorie de l'information*. Paris : Seuil, 2006.

- [48] K. Gödel, “über formal unentscheidbare sätze der principia mathematica und verwandter systeme, i,” *Monatshefte für Mathematik und Physik*, vol. 38, pp. 173–198, 1931.
- [49] E. C. Zeeman, *Catastrophe theory : selected papers, 1972-1977*. Reading, Mass. : Addison-Wesley Pub. Co., Advanced Book Program, 1977.
- [50] N. A. Gough, “Fractals, chaos, and fetal heart rate,” *Lancet*, vol. 339, pp. 182–183, Jan. 1992. PMID : 1346035.
- [51] E. N. Lorenz, “Deterministic nonperiodic flow,” *J. Atmos. Sci.*, vol. 20, pp. 130–141, Mar. 1963.
- [52] V. I. Arnol’d, *Mathematical methods of classical mechanics*. No. 60 in Graduate texts in mathematics, New York : Springer, 2nd ed ed., 1997.
- [53] V. I. Arnol’d and A. Avez, *Ergodic problems of classical mechanics*. Redwood City, Calif. : Addison-Wesley, 1988.
- [54] I. R. Nicola and M. Neagu, “Jet riemann-lagrange geometry and some applications in theoretical biology,” *arXiv :0802.2343 [math]*, Feb. 2008. *Journal of Dynamical Systems and Geometric Theories*, Vol. 6, No. 1 (2008), 13-25.
- [55] M. Neagu, “Jet Riemann-Lagrange geometry applied to evolution DEs systems from economy,” *arXiv :0707.3253*, July 2007. *Proceedings of Conference "Riemannian Geometry and Applications"*, Brasov, Romania, June 21-25, 2007 ; *Bulletin of Transilvania University, Brasov*, No. 14 (49), Series B (2007), 199-210.
- [56] A. L. Garner, Y. Y. Lau, D. W. Jordan, *et al.*, “Implications of a simple mathematical model to cancer cell population dynamics,” *Cell Prolif.*, vol. 39, pp. 15–28, Feb. 2006. PMID : 16426419.
- [57] C. Aslangul, *Mécanique quantique*. Bruxelles : De Boeck, 2010.

- [58] C. Cohen-Tannoudji, B. Diu, and F. Laloë, *Mécanique quantique. Tome I*. Paris : Hermann, 1977.
- [59] J. Hladik and M. Chrysos, *Introduction à la mécanique quantique : cours et exercices corrigés*. Paris : Dunod, 2006.
- [60] M. Le Bellac, *Physique quantique*. Paris ; Paris : EDP sciences ; CNRS, 2007.
- [61] “Mécanique quantique.” http://fr.wikipedia.org/w/index.php?title=M%C3%A9canique_quantique&oldid=100326612, Feb. 2014. [Online ; accessed 28-February-2014].
- [62] L. Nottale, *Scale relativity and fractal space-time a new approach to unifying relativity and quantum mechanics*. London ; Singapore : Imperial College Press ; World Scientific Pub. Co., 2011.
- [63] “Relativité restreinte.” http://fr.wikipedia.org/w/index.php?title=Relativit%C3%A9_restreinte&oldid=101428216, Feb. 2014. [Online ; accessed 28-February-2014].
- [64] D. McMahon, *Relativity demystified*. New York : McGraw-Hill, 2006.
- [65] M. P. Hobson, G. Efstathiou, A. N. Lasenby, *et al.*, *Relativité générale*. Bruxelles ; Paris : De Boeck, 2009.
- [66] J. Hladik, *Introduction à la relativité générale*. Paris : Ellipses, 2006.
- [67] “Transformations de lorentz.” http://fr.wikipedia.org/wiki/Transformations_de_Lorentz, Feb. 2014. [Online ; accessed 28-February-2014].
- [68] H. Poincaré, “Sur la dynamique de l’électron,” *Comptes rendus de l’Académie des Sciences*, vol. 140, pp. 1504–8, 1905.
- [69] W. Köhler, *Psychologie de la forme : introduction à de nouveaux concepts en psychologie*. Paris : Gallimard, 2000.
- [70] P. Guillaume, *La psychologie de la forme*. Paris : Flammarion, 2000.

- [71] B. Falissard, “A thought experiment reconciling neuroscience and psychoanalysis,” *J. Physiol. Paris*, vol. 105, pp. 201–206, Dec. 2011. PMID : 21864680.
- [72] H. S. Seung and D. D. Lee, “The manifold ways of perception,” *Science*, vol. 290, pp. 2268–2269, Dec. 2000.
- [73] Y. Y. Dwivedi, ed., *The Neurobiological Basis of Suicide*. Frontiers in Neuroscience, Boca Raton (FL) : CRC Press, 2012. PMID : 23035290.
- [74] B. Owens, “Predictors of suicidal behaviour found in blood,” *Nature*, Aug. 2013.
- [75] J. J. Mann, “Neurobiology of suicidal behaviour,” *Nature Reviews Neuroscience*, vol. 4, pp. 819–828, Oct. 2003.
- [76] D. Kupfer, E. Frank, and M. Phillips, “Major depressive disorder : new clinical, neurobiological, and treatment perspectives,” *Lancet*, vol. 379, pp. 1045–55, 2012.
- [77] D. J. Kupfer and D. A. Regier, “Neuroscience, clinical evidence, and the future of psychiatric classification in DSM-5,” *Am J Psychiatry*, vol. 168, pp. 672–674, July 2011.
- [78] J. E. Toettcher, C. A. Voigt, O. D. Weiner, and W. A. Lim, “The promise of optogenetics in cell biology : interrogating molecular circuits in space and time,” *Nature Methods*, vol. 8, no. 1, pp. 35–38, 2011.
- [79] K. Deisseroth, “Optogenetics,” *Nature Methods*, vol. 8, no. 1, pp. 26–29, 2011.
- [80] B. N. Cuthbert and T. R. Insel, “Toward the future of psychiatric diagnosis : the seven pillars of RDoC,” *BMC Med*, vol. 11, p. 126, 2013. PMID : 23672542.
- [81] B. N. Cuthbert and T. R. Insel, “Toward new approaches to psychotic disorders : the NIMH research domain criteria project,” *Schizophr Bull*, vol. 36, pp. 1061–1062, Nov. 2010. PMID : 20929969.

- [82] A. Zylberberg, S. Dehaene, P. R. Roelfsema, and M. Sigman, “The human turing machine : a neural framework for mental programs,” *Trends in Cognitive Sciences*, June 2011.
- [83] C. Berrou and V. Gripon, *Petite mathématique du cerveau : Une théorie de l’information mentale*. Odile Jacob, 2012.
- [84] V. Gripon and C. Berrou, “A simple and efficient way to store many messages using neural cliques,” in *Computational Intelligence, Cognitive Algorithms, Mind, and Brain (CCMB), 2011 IEEE Symposium on*, pp. 1–5, IEEE, 2011.
- [85] J. Petitot, *Neurogéométrie de la vision : modèles mathématiques et physiques des architectures fonctionnelles*. Palaiseau : les Éd. de l’École polytechnique, 2008.
- [86] J. P. Vandenbroucke, “Evidence-based medicine and "médecine d’observation",” *J Clin Epidemiol*, vol. 49, pp. 1335–1338, Dec. 1996. PMID : 8970481.
- [87] F. Godlee, “Evidence based medicine : flawed system but still the best we’ve got,” *BMJ*, vol. 348, pp. g440–g440, Jan. 2014.
- [88] D. Spence, “Evidence based medicine is broken,” *BMJ*, vol. 348, pp. g22–g22, Jan. 2014.
- [89] R. Smith and D. Rennie, “Evidence-based medicine—an oral history,” *JAMA*, vol. 311, pp. 365–367, Jan. 2014.
- [90] R. Smith and D. Rennie, “Evidence based medicine—an oral history,” *BMJ*, vol. 348, pp. g371–g371, Jan. 2014.
- [91] G. Canguilhem, *Le normal et le pathologique*. Presses Universitaires de France - PUF, 11 ed., May 2009.
- [92] R. Linder, *Les plans d’expériences : un outil indispensable à l’expérimentateur*. Paris : Presses de l’École nationale des ponts et chaussées, 2005.

- [93] B. Falissard, "Statistics in brief : when to use and when not to use a threshold p value," *Clin. Orthop. Relat. Res.*, vol. 470, pp. 315–316, Jan. 2012. PMID : 21932099 PMCID : PMC3238007.
- [94] R. Nuzzo, "Scientific method : Statistical errors," *Nature*, vol. 506, pp. 150–152, Feb. 2014.
- [95] L. Broer, C. M. Lill, M. Schuur, *et al.*, "Distinguishing true from false positives in genomic studies : p values," *Eur. J. Epidemiol.*, vol. 28, pp. 131–138, Feb. 2013. PMID : 23371043.
- [96] K. S. Button, J. P. Ioannidis, C. Mokrysz, *et al.*, "Power failure : why small sample size undermines the reliability of neuroscience," *Nat. Rev. Neurosci.*, vol. 14, pp. 365–376, May 2013. PMID : 23571845.
- [97] J. P. A. Ioannidis, "Research accomplishments that are too good to be true," *Intensive Care Med*, vol. 40, pp. 99–101, Jan. 2014. PMID : 24129497.
- [98] J. P. A. Ioannidis, "Why most published research findings are false," *PLoS Med.*, vol. 2, p. e124, Aug. 2005. PMID : 16060722.
- [99] J. P. Ioannidis, S. Greenland, M. A. Hlatky, M. J. Khoury, M. R. Macleod, D. Moher, K. F. Schulz, and R. Tibshirani, "Increasing value and reducing waste in research design, conduct, and analysis," *Lancet*, Jan. 2014.
- [100] J. P. Ioannidis, "Discussion : Why "An estimate of the science-wise false discovery rate and application to the top medical literature" is false," *Biostatistics*, Sept. 2013. PMID : 24068251.
- [101] J. P. A. Ioannidis, "Implausible results in human nutrition research," *BMJ*, vol. 347, p. f6698, 2013. PMID : 24231028.
- [102] T. Lefèvre, E. Stindel, S. Ansart, and C. Roux, "Mathematics in medicine : beyond iatromathematics," *Lancet*, vol. 383, p. 513, Feb. 2014.

- [103] M. L. Phillips and D. J. Kupfer, “Bipolar disorder diagnosis : challenges and future directions,” *Lancet*, vol. 381, pp. 1663–1671, May 2013.
- [104] G. Miller, “Beyond DSM : seeking a brain-based classification of mental illness,” *Science*, vol. 327, pp. 1437–1437, Mar. 2010.
- [105] D. A. Regier, W. E. Narrow, E. A. Kuhl, and D. J. Kupfer, “The conceptual development of DSM-V,” *Am J Psychiatry*, vol. 166, pp. 645–650, June 2009. PMID : 19487400.
- [106] A. Lin, R. L. E. P. Reniers, and S. J. Wood, “Clinical staging in severe mental disorder : evidence from neurocognition and neuroimaging,” *Br J Psychiatry*, vol. 202, pp. s11–s17, Jan. 2013.
- [107] A. Nair and R. Howard, “ENCODE and a new landscape for psychiatric genetics,” *Br J Psychiatry*, vol. 203, pp. 84–85, Aug. 2013.
- [108] J. Levine, “Risk loci with shared effects on major psychiatric disorders,” *Lancet*, vol. 382, p. 307, July 2013.
- [109] A. D. Borglum, D. Demontis, J. Grove, J. Pallesen, M. V. Hollegaard, and C. B. . Pedersen, “Genome-wide study of association and interaction with maternal cytomegalovirus infection suggests new schizophrenia loci,” *Mol Psychiatry*, Jan. 2013.
- [110] C. A. Rietveld, S. E. Medland, J. Derringer, *et al.*, “GWAS of 126,559 individuals identifies genetic variants associated with educational attainment,” *Science*, vol. 340, pp. 1467–1471, June 2013. PMID : 23722424.
- [111] L. R. Brunham and M. R. Hayden, “Whole-genome sequencing : The new standard of care ?,” *Science*, vol. 336, pp. 1112–1113, Jan. 2012.
- [112] K. Hek, A. Demirkan, J. Lahti, *et al.*, “A genome-wide association study of depressive symptoms,” *Biological Psychiatry*, vol. 73, pp. 667–678, Apr. 2013.

- [113] B. Carey, “Report on gene for depression is now faulted,” *The New York Times*, June 2009.
- [114] H. Stefansson, A. Meyer-Lindenberg, S. Steinberg, *et al.*, “CNVs conferring risk of autism or schizophrenia affect cognition in controls,” *Nature*, vol. advance online publication, Dec. 2013.
- [115] M. S. Keshavan, D. W. Morris, J. A. Sweeney, *et al.*, “A dimensional approach to the psychosis spectrum between bipolar disorder and schizophrenia : The schizo-bipolar scale,” *Schizophr Res*, vol. 133, pp. 250–254, Dec. 2011.
- [116] P. McGorry and J. van Os, “Redeeming diagnosis in psychiatry : timing versus specificity,” *Lancet*, vol. 381, pp. 343–345, Jan. 0000.
- [117] G. L. Engel, “The need for a new medical model : a challenge for biomedicine,” *Science*, vol. 196, pp. 129–136, Apr. 1977. PMID : 847460.
- [118] R. M. Nesse and D. J. Stein, “Towards a genuinely medical model for psychiatric nosology,” *BMC Medicine*, vol. 10, p. 5, Jan. 2012. PMID : 22244350.
- [119] J. E. Ferrell Jr., “Bistability, bifurcations, and waddington’s epigenetic landscape,” *Curr Biol*, vol. 22, pp. R458–R466, June 2012.
- [120] J. Wang, K. Zhang, L. Xu, and E. Wang, “Quantifying the waddington landscape and biological paths for development and differentiation,” *PNAS USA*, vol. 108, pp. 8257–8262, May 2011.
- [121] J. Baedke, “The epigenetic landscape in the course of time : Conrad hal waddington’s methodological impact on the life sciences,” *Studies in History and Philosophy of Science Part C : Studies in History and Philosophy of Biological and Biomedical Sciences*, vol. 44, pp. 756–773, Dec. 2013.
- [122] B. Horsthemke, “Waddington’s epigenetic landscape and post-darwinian biology,” *BioEssays*, vol. 34, no. 8, pp. 711–712, 2012.

- [123] A. D. Goldberg, C. D. Allis, and E. Bernstein, “Epigenetics : A landscape takes shape,” *Cell*, vol. 128, pp. 635–638, Feb. 2007.
- [124] M. Nakagawa and O. Narikiyo, “Epigenetic landscape of interacting cells : A model simulation for developmental process,” *Biosystems*, vol. 101, pp. 156–161, Sept. 2010.
- [125] R. Uher and I. C. G. Weaver, “Epigenetic traces of childhood maltreatment in peripheral blood : a new strategy to explore gene-environment interactions,” *BJP*, vol. 204, pp. 3–5, Jan. 2014. PMID : 24385457.
- [126] L. M. Lafave and R. L. Levine, “Mining the epigenetic landscape in ALL,” *Nat. Genet.*, vol. 45, pp. 1269–1270, Nov. 2013. PMID : 24165727.
- [127] M. Lock, “The lure of the epigenome,” *Lancet*, vol. 381, pp. 1896–1897, June 2013.
- [128] “Genomic and epigenomic landscapes of adult de novo acute myeloid leukemia,” *N Engl J Med*, vol. 368, no. 22, pp. 2059–2074, 2013. PMID : 23634996.
- [129] J. T. Lee, “Epigenetic regulation by long noncoding RNAs,” *Science*, vol. 338, pp. 1435–1439, Dec. 2012.
- [130] I. Maze and E. J. Nestler, “The epigenetic landscape of addiction,” *Ann. N. Y. Acad. Sci.*, vol. 1216, pp. 99–113, Jan. 2011. PMID : 21272014
PMCID : PMC3071632.
- [131] J. Baedke, “The epigenetic landscape in the course of time : Conrad hal waddington’s methodological impact on the life sciences,” *Studies in History and Philosophy of Science Part C : Studies in History and Philosophy of Biological and Biomedical Sciences*, vol. 44, pp. 756–773, Dec. 2013.

- [132] A. M. Grabiec and K. A. Reedquist, “The ascent of acetylation in the epigenetics of rheumatoid arthritis,” *Nat Rev Rheumatol*, vol. 9, pp. 311–318, May 2013.
- [133] G. Saporta, *Probabilités, analyse des données et statistique*. Paris : Éd. Technip, 2011.
- [134] S. Tufféry, *Data mining et statistique décisionnelle*. Paris : éd. Technip, 2010.
- [135] S. Theodoridis and K. Koutroumbas, *Pattern Recognition*. Academic Press Inc, 4th revised edition ed., Nov. 2008.
- [136] A. J. Izenman, *Modern multivariate statistical techniques : regression, classification, and manifold learning*. New York : Springer, 2008.
- [137] L. Kaufman and P. J. Rousseeuw, *Finding groups in data : an introduction to cluster analysis*. Hoboken, N.J. : Wiley, 2005.
- [138] S. Monti, P. Tamayo, J. Mesirov, and T. Golub, “Consensus clustering – a resampling-based method for class discovery and visualization of gene expression microarray data,” in *Machine learning, functional genomics special issue*, pp. 91–118, 2003.
- [139] E. Morin, *On complexity*. Cresskill, N.J. : Hampton Press, 2008.
- [140] T. Speed, “A correlation for the 21st century,” *Science*, vol. 334, pp. 1502–1503, Dec. 2011.
- [141] D. N. Reshef, Y. A. Reshef, H. K. Finucane, S. R. Grossman, G. McVean, P. J. Turnbaugh, E. S. Lander, M. Mitzenmacher, and P. C. Sabeti, “Detecting novel associations in large data sets,” *Science*, vol. 334, pp. 1518–1524, Dec. 2011.
- [142] J. Pearl, *Probabilistic reasoning in intelligent systems : networks of plausible inference*. San Mateo, Calif. : Morgan Kaufmann Publishers, 1991.

- [143] J. Pearl, *Causality : models, reasoning, and inference*. Cambridge : Cambridge University Press, 2010.
- [144] D. Pe'er, "Bayesian network analysis of signaling networks : A primer," *Science's STKE*, vol. 2005, pp. pl4–pl4, Apr. 2005.
- [145] G. B. West, J. H. Brown, and B. J. Enquist, "A general model for the structure and allometry of plant vascular systems," *Nature*, vol. 400, pp. 664–667, Aug. 1999.
- [146] P. K. and K. I., "Morphometrics and allometry in fishes," in *Morphometrics* (C. Wahl, ed.), InTech, Mar. 2012.
- [147] B. J. Enquist, J. H. Brown, and G. B. West, "Allometric scaling of plant energetics and population density," *Nature*, vol. 395, pp. 163–165, Sept. 1998.
- [148] M. Laguès and A. Lesne, *Invariances d'échelle : des changements d'états à la turbulence*. Paris : Belin, 2008.
- [149] D. West and B. J. West, "Physiologic time : A hypothesis," *Phys Life Rev*, Apr. 2013. PMID : 23659904.
- [150] B. J. West, M. Bologna, and P. Grigolini, *Physics of Fractal Operators*. New York, NY : Springer New York, 2003.
- [151] "Fractals and medicine," *Lancet*, vol. 338, pp. 1425–1426, Dec. 1991. PMID : 1683423.
- [152] B. J. West, "Fractal physiology and the fractional calculus : A perspective," *Front Physiol*, vol. 1, Oct. 2010. PMID : 21423355 PMCID : PMC3059975.
- [153] B. J. West, *Where Medicine Went Wrong : Rediscovering the Path to Complexity*. World Scientific Publishing Company, 1 ed., Oct. 2006.
- [154] K. Pearson, "On lines and planes of closest fit to systems of points in space," *Philosophical Magazine*, vol. 2, no. 6, pp. 559–572, 1901.

- [155] J. B. Tenenbaum, “A global geometric framework for nonlinear dimensionality reduction,” *Science*, vol. 290, pp. 2319–2323, Dec. 2000.
- [156] “A global geometric framework for nonlinear dimensionality reduction.” <http://isomap.stanford.edu/>. [Online ; accessed 28-February-2014].
- [157] T. F. Cox and M. A. A. Cox, *Multidimensional scaling*. Boca Raton : Chapman & Hall/CRC, 2001.
- [158] M. Balasubramanian and E. L. Schwartz, “The isomap algorithm and topological stability,” *Science*, vol. 295, no. 5552, p. 7, 2002.
- [159] “Matlab toolbox for dimensionality reduction.” http://homepage.tudelft.nl/19j49/Matlab_Toolbox_for_Dimensionality_Reduction.html, 2012. [Online ; accessed 28-February-2014].
- [160] T. Wittman, “MANI fold learning matlab demo.” <http://www.math.ucla.edu/~wittman/mani/>. [Online ; accessed 28-February-2014].
- [161] S. T. Roweis, “Nonlinear dimensionality reduction by locally linear embedding,” *Science*, vol. 290, pp. 2323–2326, Dec. 2000.
- [162] “Locally linear embedding.” <http://www.cs.nyu.edu/~roweis/lle/>. [Online ; accessed 28-February-2014].
- [163] D. L. Donoho, “Hessian eigenmaps : Locally linear embedding techniques for high-dimensional data,” *PNAS USA*, vol. 100, pp. 5591–5596, Apr. 2003.
- [164] Z. Zhang and H. Zha, “Principal manifolds and nonlinear dimensionality reduction via tangent space alignment,” *SIAM Journal on Scientific Computing*, vol. 26, pp. 313–338, Jan. 2004.
- [165] A. Robinson, “American cartographic association. committee on map projections. choosing a world map : attributes, distortions, classes, aspects.” (Falls Church), 1988.

- [166] F. Takens, “Detecting strange attractors in turbulence,” in *Dynamical Systems and Turbulence, Lecture Notes in Mathematics*, vol. 898, pp. 366–381, Springer-Verlag, 1981.
- [167] T. Vialar, *Dynamiques non linéaires chaotiques en finance et économie*. Paris : Economica, 2005.
- [168] L. M. Pecora, L. Moniz, J. Nichols, and T. L. Carroll, “A unified approach to attractor reconstruction,” *Chaos : An Interdisciplinary Journal of Nonlinear Science*, vol. 17, no. 1, p. 013110, 2007.
- [169] R. Bagley and L. Glass, “Counting and classifying attractors in high dimensional dynamical systems,” *J Theor Biol*, vol. 183, pp. 269–284, Dec. 1996.
- [170] J. A. Lee and M. Verleysen, *Nonlinear dimensionality reduction*. New York ; London : Springer, 2007.
- [171] R. Bellman, *Dynamic programming*. Mineola, N.Y. : Dover Publications, 2003.
- [172] F. Camastra, “Data dimensionality estimation methods : A survey,” *Pattern Recognition*, vol. 36, pp. 2945–2954, 2003.
- [173] P. Grassberger and I. Procaccia, “Measuring the strangeness of strange attractors,” *Physica D : Nonlinear Phenomena*, vol. 9, pp. 189–208, Oct. 1983.
- [174] R. Thom and E. Noël, *Prédire n’est pas expliquer : entretiens avec Émile Noël*. [Paris] : Flammarion, 1993.
- [175] Ioannidis JA, “More than a billion people taking statins? : Potential implications of the new cardiovascular guidelines,” *JAMA*, pp. –, Dec. 2013.
- [176] D. Moher, A. Liberati, J. Tetzlaff, *et al.*, “Preferred reporting items for systematic reviews and meta-analyses : the PRISMA statement,” *BMJ*, vol. 339, pp. b2535–b2535, July 2009.

- [177] A. S. Fauci and D. M. Morens, “The perpetual challenge of infectious diseases,” *N Engl J Med*, vol. 366, no. 5, pp. 454–461, 2012. PMID : 22296079.
- [178] L. Montagnier, “A history of HIV discovery,” *Science*, vol. 298, pp. 1727–1728, Nov. 2002.
- [179] F. Barre-Sinoussi, J. C. Chermann, F. Rey, *et al.*, “Isolation of a t-lymphotropic retrovirus from a patient at risk for acquired immune deficiency syndrome (AIDS),” *Science*, vol. 220, pp. 868–871, May 1983.
- [180] F. Clavel, D. Guetard, F. Brun-Vezinet, *et al.*, “Isolation of a new human retrovirus from west african patients with AIDS,” *Science*, vol. 233, pp. 343–346, July 1986.
- [181] J. L. Marx, “AIDS virus has new name—perhaps,” *Science*, vol. 232, pp. 699–700, Sept. 1986.
- [182] M. L. Gougeon and L. Montagnier, “Apoptosis in AIDS,” *Science*, vol. 260, pp. 1269–1270, May 1993.
- [183] J. Cohen, “HIV surface proteins finally caught going au naturel,” *Science*, vol. 342, pp. 546–547, Jan. 2013.
- [184] A. Cooper, M. Garcia, C. Petrovas, *et al.*, “HIV-1 causes CD4 cell death through DNA-dependent protein kinase during viral integration,” *Nature*, vol. advance online publication, June 2013.
- [185] “The durban declaration,” *Nature*, vol. 406, pp. 15–16, July 2000.
- [186] M. Stewart and G. T., “The durban declaration is not accepted by all,” *Nature*, vol. 407, pp. 286–286, Sept. 2000.
- [187] “Montagnier’s miracle cure?,” *Science*, vol. 297, pp. 1803–1803, Sept. 2002.
- [188] D. P. Culshaw, Rebecca, E. d. Harven, and H. Bialy, *La théorie VIH du Sida, incohérence scientifique ! la science a-t-elle vendu son âme au diable ?* Embourg (Belgique) : M. Pietteur, 2009.

- [189] E. d. Harven and J.-C. Roussez, *Les dix plus gros mensonges sur le sida*. Saint-Jean-de-Bray (Loiret) : Dangles, 2005.
- [190] P.-M. Girard, C. Katlama, and G. Pialoux, *VIH*. Rueil-Malmaison : Doin, 2011.
- [191] D. Lestrade and G. Pialoux, *Sida 2.0 : 1981-2011, 30 ans de regards croisés : 30 ans = 30 millions de morts du sida*. Paris : Fleuve noir, 2011.
- [192] M. D. Grmek, *Histoire du sida début et origine d'une pandémie actuelle*. Paris : Payot et Rivages, 1995.
- [193] R. T. D. Oliver, "Though HIV is the main cause of AIDS, other factors play a role," *Nature*, vol. 406, pp. 673–673, Aug. 2000.
- [194] P. Piot and T. C. Quinn, "Response to the AIDS pandemic – a global health model," *N Engl J Med*, vol. 368, no. 23, pp. 2210–2218, 2013.
- [195] A. M. Brandt, "How AIDS invented global health," *N Engl J Med*, vol. 368, no. 23, pp. 2149–2152, 2013.
- [196] P. Sendi and R. A. Proctor, "Staphylococcus aureus as an intracellular pathogen : the role of small colony variants," *Trends Microbiol.*, vol. 17, pp. 54–58, Feb. 2009. PMID : 19162480.
- [197] P. Sendi, M. Rohrbach, P. Graber, *et al.*, "Staphylococcus aureus small colony variants in prosthetic joint infection," *Clin. Infect. Dis.*, vol. 43, pp. 961–967, Oct. 2006. PMID : 16983605.
- [198] G. Sigurdsson, R. M. T. Fleming, A. Heinken, and I. Thiele, "A systems biology approach to drug targets in pseudomonas aeruginosa biofilm," *PLoS One*, vol. 7, Apr. 2012. PMID : 22523548 PMCID : PMC3327687.
- [199] J. W. Costerton, P. S. Stewart, and E. P. Greenberg, "Bacterial biofilms : a common cause of persistent infections," *Science*, vol. 284, pp. 1318–1322, May 1999. PMID : 10334980.

- [200] L. Hall-Stoodley and P. Stoodley, “Evolving concepts in biofilm infections,” *Cell. Microbiol.*, vol. 11, pp. 1034–1043, July 2009. PMID : 19374653.
- [201] R. A. Brady, J. G. Leid, J. H. Calhoun, *et al.*, “Osteomyelitis and the role of biofilms in chronic infection,” *FEMS Immunol. Med. Microbiol.*, vol. 52, pp. 13–22, Jan. 2008. PMID : 18081847.
- [202] R. M. Brucker and S. R. Bordenstein, “The hologenomic basis of speciation : Gut bacteria cause hybrid lethality in the genus *nasonia*,” *Science*, vol. 341, pp. 667–669, Sept. 2013. PMID : 23868918.
- [203] I. Cho and M. J. Blaser, “The human microbiome : at the interface of health and disease,” *Nature Reviews Genetics*, Mar. 2012.
- [204] J. F. Cryan and S. M. O’Mahony, “The microbiome-gut-brain axis : from bowel to behavior,” *Neurogastroenterology & Motility*, vol. 23, pp. 187–192, Mar. 2011.
- [205] V. O. Ezenwa, N. M. Gerardo, D. W. Inouye, *et al.*, “Animal behavior and the microbiome,” *Science*, vol. 338, pp. 198–199, Dec. 2012.
- [206] F. H. Karlsson, V. Tremaroli, and I. Nookaew, “Gut metagenome in european women with normal, impaired and diabetic glucose control,” *Nature*, vol. 498, pp. 99–103, June 2013.
- [207] S. M. Lee, G. P. Donaldson, Z. Mikulski, *et al.*, “Bacterial colonization factors control specificity and stability of the gut microbiota,” *Nature*, vol. advance online publication, Aug. 2013.
- [208] P. Vineis and C. P. Wild, “Global cancer patterns : causes and prevention,” *Lancet*, vol. 383, pp. 549–557, Feb. 2014.
- [209] L. J. Esserman, J. Thompson, Ian M, and B. Reid, “Overdiagnosis and overtreatment in cancer : an opportunity for improvement,” *JAMA*, vol. 310, pp. 797–798, Aug. 2013. PMID : 23896967.

- [210] A. Ashworth and T. J. Hudson, “Genomics : Comparisons across cancers,” *Nature*, vol. 502, pp. 306–307, Oct. 2013.
- [211] M. J. Khoury, T. K. Lam, J. P. Ioannidis, *et al.*, “Transforming epidemiology for 21st century medicine and public health,” *Cancer Epidemiol. Biomarkers Prev.*, vol. 22, pp. 508–516, Apr. 2013. PMID : 23462917.
- [212] T. Reya, S. J. Morrison, M. F. Clarke, and I. L. Weissman, “Stem cells, cancer, and cancer stem cells,” *Nature*, vol. 414, pp. 105–111, Nov. 2001.
- [213] C. Kandoth, M. D. McLellan, F. Vandin, *et al.*, “Mutational landscape and significance across 12 major cancer types,” *Nature*, vol. 502, pp. 333–339, Oct. 2013. PMID : 24132290.
- [214] D. Hanahan, “Rethinking the war on cancer,” *Lancet*, vol. 383, pp. 558–563, Feb. 2014.
- [215] M. Skipper, “Cancer genomics : A panoramic view of cancer,” *Nat Rev Genet*, vol. 14, pp. 750–750, Nov. 2013.
- [216] M. W. L. Teng and M. J. Smyth, “Can cancer trigger autoimmunity ?,” *Science*, vol. 343, pp. 147–148, Oct. 2014.
- [217] P. C. Gotzsche and K. J. Jorgensen, “Screening for breast cancer with mammography,” *Cochrane Database Syst Rev*, vol. 6, p. CD001877, 2013. PMID : 23737396.
- [218] A. Miller, C. Wall, C. Baines, *et al.*, “Twenty five year follow-up for breast cancer incidence and mortality of the canadian national breast screening study : randomised screening trial,” *BMJ*, vol. 348, pp. g366–g366, Feb. 2014.
- [219] R. A. Royce, A. Sena, J. Cates, W, and M. S. Cohen, “Sexual transmission of HIV,” *N. Engl. J. Med.*, vol. 336, pp. 1072–1078, Apr. 1997. PMID : 9091805.

- [220] E. Vynnycky and R. G. White, *An introduction to infectious disease modelling*. Oxford : Oxford University Press, 2010.
- [221] N. Bacaër, *Histoires de mathématiques et de populations*. Paris : Cassini, 2008.
- [222] F. Ader, J. Salomon, C. Perronne, and L. Bernard, “[is bone infection of endogenous or exogenous origin ? a pathophysiological approach],” *Med Mal Infect*, vol. 34, pp. 530–537, Nov. 2004. PMID : 15620057.
- [223] D. P. Lew and F. A. Waldvogel, “Osteomyelitis,” *Lancet*, vol. 364, pp. 369–379, July 2004. PMID : 15276398.
- [224] A. M. Smith, J. A. McCullers, and F. R. Adler, “Mathematical model of a three-stage innate immune response to a pneumococcal lung infection,” *J Theor Biol*, vol. 276, pp. 106–116, May 2011.
- [225] P. Pivonka, J. Zimak, D. W. Smith, *et al.*, “Model structure and control of bone remodeling : A theoretical study,” *Bone*, vol. 43, pp. 249–263, Aug. 2008.
- [226] P. Pivonka, J. Zimak, D. W. Smith, *et al.*, “Theoretical investigation of the role of the RANK–RANKL–OPG system in bone remodeling,” *J Theor Biol*, vol. 262, pp. 306–316, Jan. 2010.
- [227] M. C. Peterson and M. M. Riggs, “A physiologically based mathematical model of integrated calcium homeostasis and bone remodeling,” *Bone*, vol. 46, pp. 49–63, Jan. 2010.
- [228] M. Zumsande, D. Stiefs, S. Siegmund, and T. Gross, “General analysis of mathematical models for bone remodeling,” *Bone*, vol. 48, pp. 910–917, Apr. 2011.
- [229] S. V. Komarova, R. J. Smith, S. J. Dixon, *et al.*, “Mathematical model predicts a critical role for osteoclast autocrine regulation in the control of bone remodeling,” *Bone*, vol. 33, pp. 206–215, Aug. 2003. PMID : 14499354.

- [230] H. Khalil, R. J. Williams, G. Stenbeck, *et al.*, “Invasion of bone cells by staphylococcus epidermidis,” *Microbes Infect.*, vol. 9, pp. 460–465, Apr. 2007. PMID : 17331787.
- [231] J. K. Ellington, S. S. Reilly, W. K. Ramp, *et al.*, “Mechanisms of staphylococcus aureus invasion of cultured osteoblasts,” *Microb. Pathog.*, vol. 26, pp. 317–323, June 1999. PMID : 10343060.
- [232] T. Chambers, “The birth of the osteoclast,” *Annals of the New York Academy of Sciences*, vol. 1192, pp. 19–26, Apr. 2010.
- [233] H. C. Blair, “How the osteoclast degrades bone,” *BioEssays*, vol. 20, no. 10, pp. 837–846, 1998.
- [234] E. F. Shead, C. S. Haworth, H. Barker, *et al.*, “Osteoclast function, bone turnover and inflammatory cytokines during infective exacerbations of cystic fibrosis,” *J. Cyst. Fibros.*, vol. 9, pp. 93–98, Mar. 2010. PMID : 20006563.
- [235] A. Teti and A. Zallone, “Do osteocytes contribute to bone mineral homeostasis? osteocytic osteolysis revisited,” *Bone*, vol. 44, pp. 11–16, Jan. 2009.
- [236] S. G. Reed, “TGF-beta in infections and infectious diseases,” *Microbes Infect.*, vol. 1, pp. 1313–1325, Dec. 1999. PMID : 10611760.
- [237] D. Gong, W. Shi, S.-j. Yi, *et al.*, “TGF- β signaling plays a critical role in promoting alternative macrophage activation,” *BMC Immunol*, vol. 13, p. 31, June 2012. PMID : 22703233.
- [238] Y.-W. Lee, Y. Chung, S. K. Juhn, *et al.*, “Activation of the transforming growth factor beta pathway in bacterial otitis media,” *Ann. Otol. Rhinol. Laryngol.*, vol. 120, pp. 204–213, Mar. 2011. PMID : 21510147.
- [239] D. C. Kirouac, C. Ito, E. Csaszar, *et al.*, “Dynamic interaction networks in a hierarchically organized tissue,” *Mol Syst Biol*, vol. 6, p. 417, Oct. 2010. PMID : 20924352 PMCID : PMC2990637.

- [240] G. S. Ashcroft, “Bidirectional regulation of macrophage function by TGF-beta,” *Microbes Infect.*, vol. 1, pp. 1275–1282, Dec. 1999. PMID : 10611755.
- [241] J. T. Murphy, R. Walshe, and M. Devocelle, “A computational model of antibiotic-resistance mechanisms in methicillin-resistant staphylococcus aureus (MRSA),” *J Theor Biol*, vol. 254, pp. 284–293, Sept. 2008.
- [242] M. Chenel, A. Barbot, A. Dupuis, *et al.*, “Pharmacokinetic-pharmacodynamic modeling of the electroencephalogram effect of norfloxacin in rats,” *Antimicrob. Agents Chemother.*, vol. 47, pp. 1952–1957, June 2003. PMID : 12760873.
- [243] S. Marchand, M. Chenel, I. Lamarche, and W. Couet, “Pharmacokinetic modeling of free amoxicillin concentrations in rat muscle extracellular fluids determined by microdialysis,” *Antimicrob. Agents Chemother.*, vol. 49, pp. 3702–3706, Sept. 2005. PMID : 16127043.
- [244] T. Sasaki, H. Takane, K. Ogawa, *et al.*, “Population pharmacokinetic and pharmacodynamic analysis of linezolid and a hematologic side effect, thrombocytopenia, in japanese patients,” *Antimicrob Agents Chemother*, vol. 55, pp. 1867–1873, Feb. 2011.
- [245] R. Boudon, “Le modèle des classes latentes,” *Revue Française de Sociologie*, vol. 3, p. 259, July 1962.
- [246] T. Lefèvre and P. Chauvin, *Recherche d’une typologie de recours aux soins et des déterminants associés, à partir de la cohorte SIRS, vague 2010*. Créteil, France : Université Paris-Est Créteil, 2013.
- [247] P. Chauvin, I. Parizot, and S. Revet, *Santé et expériences de soins : De l’individu à l’environnement social*. Inserm, 2005.
- [248] J. Martin-Fernandez, F. Grillo, I. Parizot, *et al.*, “Prevalence and socioeconomic and geographical inequalities of household food insecurity

- in the paris region, france, 2010,” *BMC Public Health*, vol. 13, p. 486, May 2013. PMID : 23688296.
- [249] L. M. A. Bettencourt, “The origins of scaling in cities,” *Science*, vol. 340, pp. 1438–1441, June 2013.
- [250] M. Batty, “A theory of city size,” *Science*, vol. 340, pp. 1418–1419, June 2013.
- [251] P. James, S. Oparil, B. Carter, *et al.*, “2014 evidence-based guideline for the management of high blood pressure in adults : Report from the panel members appointed to the eighth joint national committee (jnc 8),” *JAMA*, vol. 311, no. 5, pp. 507–520, 2014.
- [252] K. Barnett, S. W. Mercer, M. Norbury, *et al.*, “Epidemiology of multimorbidity and implications for health care, research, and medical education : a cross-sectional study,” *Lancet*, vol. 380, pp. 37–43, July 2012. PMID : 22579043.
- [253] D. Bloom, E. Cafiero, E. Jané-Llopis, *et al.*, “The global economic burden of noncommunicable diseases,” tech. rep., World Economic Forum, Geneva, Switzerland, 2011.
- [254] R. Atun, S. Jaffar, S. Nishtar, *et al.*, “Improving responsiveness of health systems to non-communicable diseases,” *Lancet*, vol. 381, pp. 690–697, Feb. 2013.
- [255] C. Vogeli, A. E. Shields, T. A. Lee, *et al.*, “Multiple chronic conditions : prevalence, health consequences, and implications for quality, care management, and costs,” *J Gen Intern Med*, vol. 22 Suppl 3, pp. 391–395, Dec. 2007. PMID : 18026807.
- [256] C. Salisbury, “Multimorbidity : redesigning health care for people who use it,” *The Lancet*, vol. 380, no. 9836, pp. 7–9.
- [257] J. Y. Le Reste, P. Nabbe, B. Manceau, *et al.*, “The european general practice research network presents a comprehensive definition of

- multimorbidity in family medicine and long term care, following a systematic review of relevant literature,” *J Am Med Dir Assoc*, vol. 14, pp. 319–325, May 2013. PMID : 23411065.
- [258] M. van der Akker, F. Buntinx, and A. Knottnerus, “Comorbidity or multimorbidity what’s in a name? a review of literature,” *European Journal of General Practice*, vol. 2, no. 2, pp. 65–70, 1996.
- [259] M. van den Akker, F. Buntinx, S. Roos, and J. A. Knottnerus, “Problems in determining occurrence rates of multimorbidity,” *J Clin Epidemiol*, vol. 54, pp. 675–679, July 2001. PMID : 11438407.
- [260] M. v. d. Akker, F. Buntinx, and J. A. Knottnerus, “Comorbidity or multimorbidity,” July 2009.
- [261] H. C. Britt, C. M. Harrison, G. C. Miller, and S. A. Knox, “Prevalence and patterns of multimorbidity in australia,” *Med. J. Aust.*, vol. 189, pp. 72–77, July 2008. PMID : 18637770.
- [262] S. L. Brilleman and C. Salisbury, “Comparing measures of multimorbidity to predict outcomes in primary care : a cross sectional study,” *Fam Pract*, vol. 30, pp. 172–178, Apr. 2013. PMID : 23045354.
- [263] A. Prados-Torres, B. Poblador-Plou, A. Calderon-Larranaga, and L. A. Gimeno-Feliu, “Multimorbidity patterns in primary care : interactions among chronic diseases using factor analysis,” *PLoS ONE*, vol. 7, no. 2, p. e32190, 2012. PMID : 22393389.
- [264] A. Prados-Torres, A. Calderon-Larranaga, J. Hanco-Saavedra, *et al.*, “Multimorbidity patterns : a systematic review,” *J Clin Epidemiol*, vol. 67, pp. 254–266, Mar. 2014.
- [265] A. Frances, “The new crisis of confidence in psychiatric diagnosis,” *Ann. Intern. Med.*, vol. 159, p. 720, Nov. 2013. PMID : 24247686.
- [266] D. Lage, S. Egli, M. Riedel, *et al.*, “Combining the categorical and the dimensional perspective in a diagnostic map of psychotic disorders,”

- European Archives of Psychiatry and Clinical Neuroscience*, vol. 261, pp. 3–10, July 2010.
- [267] S. Egli, M. Riedel, H.-J. Moller, *et al.*, “Creating a map of psychiatric patients based on psychopathological symptom profiles,” *European Archives of Psychiatry and Clinical Neuroscience*, vol. 259, pp. 164–171, Jan. 2009.
- [268] T. Ketter, P. Wang, O. Becker, *et al.*, “Psychotic bipolar disorders : dimensionally similar to or categorically different from schizophrenia?,” *Journal of Psychiatric Research*, vol. 38, pp. 47–61, Jan. 2004.
- [269] T. R. Insel, “Rethinking schizophrenia,” *Nature*, vol. 468, pp. 187–193, Nov. 2010. PMID : 21068826.
- [270] T. R. Insel and P. S. Wang, “Rethinking mental illness,” *JAMA*, vol. 303, pp. 1970–1971, May 2010. PMID : 20483974.
- [271] B. Cuthbert and T. Insel, “The data of diagnosis : new approaches to psychiatric classification,” *Psychiatry*, vol. 73, no. 4, pp. 311–314, 2010. PMID : 21198381.
- [272] T. R. Insel, “Developmental psychobiology for public health : a bridge for translational research,” *Dev Psychobiol*, vol. 47, pp. 209–216, Nov. 2005. PMID : 16252289.
- [273] T. R. Insel and R. Quirion, “Psychiatry as a clinical neuroscience discipline,” *JAMA*, vol. 294, pp. 2221–2224, Nov. 2005. PMID : 16264165.
- [274] A. Chiolero, “Big data in epidemiology : too big to fail?,” *Epidemiology*, vol. 24, pp. 938–939, Nov. 2013. PMID : 24077000.
- [275] P. J. Wolfe, “Making sense of big data,” *Proc. Natl. Acad. Sci. U.S.A.*, vol. 110, pp. 18031–18032, Nov. 2013. PMID : 24145435.
- [276] S. Toh and R. Platt, “Is size the next big thing in epidemiology?,” *Epidemiology*, vol. 24, pp. 349–351, May 2013. PMID : 23549179.

- [277] V. Gainer, K. Hackett, M. Mendis, *et al.*, “Using the i2b2 hive for clinical discovery : an example,” *AMIA Annu Symp Proc*, p. 959, 2007. PMID : 18694059.
- [278] S. N. Murphy, M. E. Mendis, D. A. Berkowitz, *et al.*, “Integration of clinical and genetic data in the i2b2 architecture,” *AMIA Annu Symp Proc*, p. 1040, 2006. PMID : 17238659 PMCID : PMC1839291.
- [279] D. Segagni, V. Tibollo, A. Dagliati, *et al.*, “Clinical and research data integration : the i2b2-FSM experience,” *AMIA Summits Transl Sci Proc*, vol. 2013, pp. 239–240, 2013. PMID : 24303274 PMCID : PMC3845786.
- [280] N. Wattanasin, A. Porter, S. Ubaha, *et al.*, “Apps to display patient data, making SMART available in the i2b2 platform,” *AMIA Annu Symp Proc*, vol. 2012, pp. 960–969, 2012. PMID : 23304371 PMCID : PMC3540423.
- [281] S. Goryachev, M. Sordo, and Q. T. Zeng, “A suite of natural language processing tools developed for the I2B2 project,” *AMIA Annu Symp Proc*, p. 931, 2006. PMID : 17238550 PMCID : PMC1839726.
- [282] C. Cherry, X. Zhu, J. Martin, and B. de Bruijn, “A la recherche du temps perdu : extracting temporal relations from medical text in the 2012 i2b2 NLP challenge,” *J Am Med Inform Assoc*, vol. 20, pp. 843–848, Oct. 2013. PMID : 23523875 PMCID : PMC3756270.
- [283] E. K. Johnson, S. Broder-Fingert, P. Tanpowpong, *et al.*, “Use of the i2b2 research query tool to conduct a matched case-control clinical research study : advantages, disadvantages and methodological considerations,” *BMC Med Res Methodol*, vol. 14, no. 1, p. 16, 2014. PMID : 24479726 PMCID : PMC3909388.
- [284] W. Sun, A. Rumshisky, and O. Uzuner, “Evaluating temporal relations in clinical text : 2012 i2b2 challenge,” *J Am Med Inform Assoc*, vol. 20, pp. 806–813, Oct. 2013. PMID : 23564629 PMCID : PMC3756273.

- [285] W. G. Adams, N. Anderson, E. S. Berner, *et al.*, “Distributed health outcome monitoring and evaluation using i2b2,” *AMIA Summits Transl Sci Proc*, vol. 2013, p. 2, 2013. PMID : 24303227 PMCID : PMC3845782.
- [286] B. Roche, J.-F. Guegan, and F. Bousquet, “Multi-agent systems in epidemiology : a first step for computational biology in the study of vector-borne disease transmission,” *BMC Bioinformatics*, vol. 9, no. 1, p. 435, 2008.
- [287] C. N. Madey and Gregory, “Tools of the trade : A survey of various agent based modeling platforms,” Mar. 2009.
- [288] “NetLogo.” <http://ccl.northwestern.edu/netlogo/>. [Online; accessed 28-February-2014].
- [289] M. Laclavik, S. Dlugolinsky, M. Seleng, *et al.*, “Agent-based simulation platform evaluation in the context of human behavior modeling,” in *Advanced Agent Technology* (D. Hutchison, T. Kanade, J. Kittler, *et al.*, eds.), vol. 7068, pp. 396–410, Berlin, Heidelberg : Springer Berlin Heidelberg, 2012.
- [290] S. F. Railsback, S. L. Lytinen, and S. K. Jackson, “Agent-based simulation platforms : Review and development recommendations,” *Simulation*, vol. 82, pp. 609–623, Sept. 2006.
- [291] “FLAME website.” <http://www.flame.ac.uk/>, 2014. [Online; accessed 28-February-2014].
- [292] A. Kempf, “Spacetime could be simultaneously continuous and discrete, in the same way that information can be,” *New Journal of Physics*, vol. 12, p. 115001, Nov. 2010.
- [293] A. Kempf, “Short-distance cutoffs in curved space,” *arXiv :gr-qc/0403112*, Mar. 2004.

- [294] A. Ashoorioon, A. Kempf, and R. Mann, “Minimum length cutoff in inflation and uniqueness of the action,” *Physical Review D*, vol. 71, Jan. 2005.
- [295] A. Kempf, “Information-theoretic natural ultraviolet cutoff for space-time,” *Physical Review Letters*, vol. 103, Dec. 2009.
- [296] A. Kempf and R. Martin, “Information theory, spectral geometry, and quantum gravity,” *Physical Review Letters*, vol. 100, Jan. 2008.
- [297] M. Bojowald and A. Kempf, “Generalized uncertainty principles and localization in discrete space,” *arXiv :1112.0994 [gr-qc, physics :hep-th, physics :quant-ph]*, Dec. 2011. Phys. Rev. D 86 (2012) 085017.
- [298] C. Gignoux and B. Silvestre-Brac, *Mécanique : de la formulation Lagrangienne au chaos Hamiltonien*. Les Ulis : EDP Sciences, 2002.
- [299] G. M. Zaslavskij, *Hamiltonian chaos and fractional dynamics*. Oxford : Oxford Univ. Press, 2008.
- [300] L. D. Landau and E. M. Lifchitz, *Physique théorique*. Paris : Ellipses, 2012.
- [301] J. Jaeger, D. Irons, and N. Monk, “Regulative feedback in pattern formation : towards a general relativistic theory of positional information,” *Development*, vol. 135, pp. 3175–3183, Jan. 2008. PMID : 18776142.
- [302] G. A. Sardanashvili, *Generalized Hamiltonian formalism for field theory : constraint systems*. Singapore ; River Edge, NJ : World Scientific, 1995.
- [303] L. E. Reichl, *The transition to chaos : conservative classical systems and quantum manifestations*. New York : Springer, 2004.
- [304] P. A. M. Dirac, “Generalized hamiltonian dynamics,” *Proceedings of the Royal Society A : Mathematical, Physical and Engineering Sciences*, vol. 246, pp. 326–332, Aug. 1958.

- [305] N. Kiriushcheva and S. V. Kuzmin, “The hamiltonian formulation of general relativity : myths and reality,” *Central European Journal of Physics*, vol. 9, pp. 576–615, Sept. 2010. arXiv :0809.0097 [gr-qc, physics :hep-th].
- [306] G. Atanasiu and M. Neagu, “Distinguished torsion, curvature and deflection tensors in the multi-time hamilton geometry,” *arXiv :0807.0614*, July 2008. Electronic Journal "Differential Geometry - Dynamical Systems", Vol. 11 (2009), 20-40.
- [307] M. Neagu and C. Udriste, “Multi-time dependent sprays and harmonic maps on $j^1(t, m)$,” *arXiv :math/0009049*, Sept. 2000. M. Neagu, C. Udriste, A. Oana : "Multi-Time dependent sprays and h-traceless maps", *Balkan J. of Geom. and Appl.*, 10, 2 (2005), 76-92.
- [308] M. Neagu and C. Udriste, “The geometry of metrical multi-time lagrange spaces,” *arXiv :math/0009071*, Sept. 2000. "A Riemann-Lagrange geometrization for metrical multi-time Lagrange spaces", *Balkan J. of Geom. and Appl.*, 11, 1 (2006), 87-98.
- [309] D. Baleanu, A. Golmankhaneh, and A. Golmankhaneh, “The dual action of fractional multi time hamilton equations,” *International Journal of Theoretical Physics*, vol. 48, no. 9, pp. 2558–2569, 2009.
- [310] T. d. Donder, *Théorie invariante du calcul des variations*. Gauthier-Villars, 1935.
- [311] R. Doursat, H. Sayama, and O. Michel, “A review of morphogenetic engineering,” *Natural Computing*, vol. 12, pp. 517–535, Dec. 2013.
- [312] V. G. IVANCEVIC, *Geometrical Dynamics of Complex Systems. A Unified Modelling Approach to Physics, Control, Biomechanics, Neurodynamics and Psycho-Socio-Economical Dynamics*. Springer, 2006.
- [313] L. De Broglie, “La thermodynamique "cachée" des particules,” *Annales de l’institut Henri Poincaré (A) Physique théorique*, vol. 1, no. 1, pp. 1–

19, 1964.

- [314] A. Bejan, *Shape and structure, from engineering to nature*. New York : Cambridge University Press, 2000.
- [315] H. S. Mortveit and C. M. Reidys, *An introduction to sequential dynamical systems*. New York ; London : Springer, 2008.
- [316] J.-P. Aubin, A. M. Bayen, and P. Saint-Pierre, *Viability Theory : New Directions*. Berlin, Heidelberg : Springer-Verlag Berlin Heidelberg, 2011.
- [317] A. N. Gorban, *Principal manifolds for data visualization and dimension reduction*. Berlin : Springer, 2008.
- [318] A. N. Gorban and A. Zinovyev, "Principal manifolds and graphs in practice : from molecular biology to dynamical systems," *arXiv :1001.1122*, Jan. 2010. International Journal of Neural Systems, Vol. 20, No. 3 (2010) 219-232.
- [319] S. Weng, C. Zhang, Z. Lin, and X. Zhang, "Mining the structural knowledge of high-dimensional medical data using isomap," *Med Biol Eng Comput*, vol. 43, pp. 410–412, May 2005. PMID : 16035231.
- [320] A. Akhbardeh and M. A. Jacobs, "Comparative analysis of nonlinear dimensionality reduction techniques for breast MRI segmentation," *Med Phys*, vol. 39, pp. 2275–2289, Apr. 2012. PMID : 22482648.
- [321] A. Connes, *Noncommutative geometry*. San Diego : Academic Press, 1994.
- [322] C. Udriste and A. Udriste, "From flows and metrics to dynamics," *arXiv :math/0007059*, July 2000.
- [323] M. Neagu and C. Udriste, "Torsions and curvatures on jet fibre bundle $j^1(t, m)$," *arXiv :math/0009069*, Sept. 2000. "Torsion, curvature and deflection d-tensors on $J^1(T, M)$ ", *Balkan J. of Geom. and Appl.*, 6, 1 (2001), 29-45.

- [324] C. Ghiu and C. Udriste, "A decision functional for multitime controllability," *arXiv :1201.0259*, Dec. 2011.
- [325] C. Udriste and I. Duca, "Periodical solutions of multi-time hamilton equations," *arXiv :math/0510554*, Oct. 2005. *Analele Universitatii Bucuresti*, 55, 1 (2005), 179-188.
- [326] A. Oana and M. Neagu, "From quadratic hamiltonians of polymonta to abstract geometrical maxwell-like and einstein-like equations," *arXiv :1202.4477*, Feb. 2012.