
*Etude de l'équivalence et de la synthèse entre
mécanique matricielle de Heisenberg &
mécanique ondulatoire de Schrödinger.*



par

Loren Coquille

Octobre 2006

*« Soudain, nous pûmes reconnaître,
entre deux couches épaisses de brouillard,
l'arête, éclairée par le soleil, d'une haute
paroi rocheuse dont nous avions déjà
soupçonné l'existence... »*

Werner Heisenberg, in *La Partie et le Tout*



Table des matières

<i>INTRODUCTION</i>	3
<i>MÉCANIQUE MATRICIELLE</i>	4
<i>CONTEXTE HISTORIQUE, MODÈLE DE BOHR – SOMMERFELD</i>	4
<i>JUIN 1925 : ARTICLE FONDATEUR DE HEISENBERG</i>	7
<i>AUTOMNE 1925 : TRAVAUX DE BORN, JORDAN ET HEISENBERG</i>	11
<i>NOVEMBRE 1925 : PREMIÈRE CONTRIBUTION DE DIRAC</i>	17
<i>JANVIER 1926 : THÉORIE OPÉRATORIELLE DE BORN ET WIENER</i>	18
<i>MÉCANIQUE ONDULATOIRE</i>	20
<i>INTRODUCTION : FORMALISME DE HAMILTON-JACOBI</i>	20
<i>HIVER 1926 : ARTICLES FONDATEURS DE SCHRÖDINGER</i>	21
<i>SYNTHÈSE OU « NOUVELLE THÉORIE QUANTIQUE »</i>	29
<i>MISE EN PARALLÈLE DES APPROCHES MATRICIELLE ET ONDULATOIRE</i>	29
<i>MARS 1926 : ARTICLE UNIFICATEUR DE SCHRÖDINGER</i>	30
<i>DÉCEMBRE 1926 : ARTICLE DE DIRAC</i>	35
<i>POINT FINAL : VON NEUMANN, FORMALISME DES ESPACES DE HILBERT</i>	38
<i>CONCLUSION</i>	39
<i>RÉFÉRENCES</i>	40

I. Introduction

Dans le présent travail, je vais étudier la période de 1925-1926, qui marque le début de la « nouvelle théorie quantique », par opposition aux modèles provisoires comme celui de Bohr-Sommerfeld. Durant un peu moins de deux ans, deux descriptions parallèles, logiquement consistantes et amenant de nouveaux cadres conceptuels pour la compréhension des phénomènes intra-atomiques ont été créés. Je présenterai tout d'abord la mécanique matricielle, dont l'article fondateur de Heisenberg a paru fin juin 1925, puis la mécanique ondulatoire, créée en 1926 par Schrödinger à la suite des travaux de De Broglie. Après m'être penchée sur les mises en formes plus précises de chacune de ces théories, j'en viendrai aux articles qui ont montré l'équivalence des deux approches, et qui ont ouvert la voie à l'interprétation moderne de la théorie quantique dans le formalisme des espaces de Hilbert.

D'un point de vue personnel, j'ai pour but d'éclaircir ma compréhension des approches matricielles et ondulatoires, et de leur équivalence. L'étude de l'émergence de nouveaux formalismes et de nouveaux concepts à partir des méthodes « archiconnues »¹ de la mécanique classique sera l'une de mes principales tâches. Dans l'objectif de mieux comprendre la signification et le statut du « principe de correspondance », je m'intéresserai à son utilisation, puisqu'il est notamment à la base de la réflexion et de l'intuition de Heisenberg dans la création de sa mécanique matricielle. Ce travail s'inscrira aussi dans une réflexion plus générale sur la place, le rôle et le sens du formalisme mathématique dans une théorie physique.

¹ Terme employé par Schrödinger dans sa « Première communication », cf [16]

II. Mécanique matricielle

a. Contexte historique, modèle de Bohr – Sommerfeld

C'est en 1900 que Planck introduit le concept de quantum d'action à la suite de ses travaux sur la thermodynamique et la mécanique statistique. L'explication de l'effet photoélectrique (par Einstein) et du comportement des chaleurs spécifiques à basse température, notamment, rendent cette nouvelle approche des interactions entre lumière et matière fort prometteuse.

En 1913, le modèle de Bohr pour les atomes constitue une contribution cruciale à la description des spectres d'émission et d'absorption des gaz. Les mesures de spectroscopie atomique de l'époque montraient clairement que ces spectres étaient discrets. Niels Bohr décrit les atomes comme formés d'un nuage électronique en orbite autour d'un noyau rigide, mais seules certaines orbites sont possibles pour les électrons et constituent les états stationnaires. Si un électron se trouve sur l'une de ces orbites, l'atome ne rayonne pas, si par contre celui-ci passe d'un état stationnaire à un autre, l'énergie émise (ou absorbée) par l'atome sous forme de lumière satisfait la relation dite de Planck – Einstein :

$$\Delta E = h\nu$$

avec ΔE la différence d'énergie entre les deux orbites, h la constante de Planck, et ν la fréquence de l'onde lumineuse émise (ou absorbée). Bohr réussit donc à lier l'énergie du rayonnement à l'énergie des orbites en utilisant l'hypothèse de Planck.

Dans le célèbre recueil d'articles de Bohr, intitulé « Physique atomique et connaissance humaine », Bohr insiste sur l'incapacité de la théorie classique à décrire les phénomènes atomiques et définit le principe qui a été à la base de ses réflexions et de son modèle, le *principe de correspondance* :

« La faillite des théories de la physique classique à rendre compte des phénomènes atomiques s'accroît encore avec le progrès de nos connaissances sur la structure des atomes. Par-dessus tout, la découverte par Rutherford du noyau atomique (1911) révéla aussitôt combien les concepts de la mécanique et de l'électromagnétisme classiques étaient impropres à exprimer la stabilité inhérente à l'atome. Là encore, la théorie quantique offrit le moyen d'éclaircir la situation : on découvrit notamment qu'il était possible de rendre compte de la stabilité atomique, aussi bien que des lois empiriques qui gouvernent le spectre des éléments, en supposant que toute réaction de l'atome conduisant à une variation de son énergie comporte une transition complète entre deux « états quantiques stationnaires ». (...) »

En fait, l'idée même d'états stationnaires est incompatible avec toute règle fixant un choix entre ces transitions et ne laisse place qu'à la notion de probabilité relative des processus individuels de transition. Pour évaluer ces probabilités, notre seul guide était le « principe de correspondance » dont l'origine avait été la recherche d'une connexion aussi étroite que possible entre la description statistique des processus atomiques et les conséquences de la théorie classique : celles-ci doivent être valables à la limite, lorsque toutes les actions mises en jeu, à tous les stades de l'analyse du phénomène, sont grand par rapport au quantum universel $[h]$. »[1]

Dès 1915, l'application des méthodes de la mécanique céleste au problème des atomes donne naissance à ce qu'on appelle la « vieille théorie quantique » ou « old quantum theory ». Cette approche traite des problèmes dits multipliement périodiques et impose des conditions de quantification à leur description classique.

Je vais ici décrire brièvement le modèle de Bohr – Sommerfeld qui consiste à quantifier les variables d'action dans la description en « variables angle – action » de l'atome de Bohr.

Le formalisme des variables angle – action s'inspire de certaines idées de la théorie de Hamilton – Jacobi développée dans la deuxième moitié du XIXe siècle. Il permet de trouver les fréquences des mouvements inhérents à un système périodique.

Pour un système périodique conservatif à un degré de liberté (problème intégrable), soit l'Hamiltonien :

$$H(q, p) = E$$

L'équation d'une orbite dans l'espace des phases est alors donnée par la solution paramétrique $p = p(q, E)$. Selon les valeurs de l'énergie, une orbite fermée ou une orbite ouverte peut être obtenue. Introduisons alors une nouvelle variable J qui remplacera E dans le rôle de la nouvelle impulsion constante. Autrement dit, on cherche à construire la transformation canonique suivante :

$$\begin{pmatrix} q \\ p \end{pmatrix} \mapsto \begin{pmatrix} Q \\ P \end{pmatrix} = \begin{pmatrix} \varphi \\ J \end{pmatrix}$$

Où J est la variable d'action définie par :

$$J = \oint p(q) dq$$

J est donc la surface d'une période complète dans l'espace des phases. Si l'orbite est ouverte, il suffit d'intégrer sur une période. Le système étant conservatif, par définition J est constante (théorème de Liouville). La transformation doit préserver les équations canoniques de Hamilton, on doit donc satisfaire :

$$\dot{\varphi} = \frac{\partial \tilde{H}}{\partial J} \quad , \quad \dot{J} = -\frac{\partial \tilde{H}}{\partial \varphi}$$

Or comme J est une constante, on a :

$$\frac{\partial \tilde{H}}{\partial \varphi} = 0 \Rightarrow \tilde{H} = \tilde{H}(J) \Rightarrow \varphi(t) = \nu t + \varphi_0 \quad \text{et} \quad J = J_0$$

$$\text{avec} \quad \nu = \nu(J) = \frac{\partial \tilde{H}(J)}{\partial J}$$

La dynamique du système dans les nouvelles variables angle – action est un mouvement périodique de fréquence ν localisé sur le périmètre d'un cercle de rayon constant J .

Pour un système conservatif intégrable à 1 degrés de liberté, on peut faire le même raisonnement que précédemment si on suppose le système périodique dans chaque variable q_i .

Les équations canoniques à satisfaire sont alors :

$$\dot{\varphi}_i = \frac{\partial \tilde{H}}{\partial J_i} \quad , \quad \dot{J}_i = -\frac{\partial \tilde{H}}{\partial \varphi_i} \quad \forall i = 1..l \quad \text{avec} \quad J_i = \oint p_i(q_i) dq_i$$

Le mouvement est donc localisé sur la surface de dimension 1 de l'hyper-tore défini par les rayons constants J_i .

Les fréquences de rotation autour de chaque rayon J_i sont données par ν_i :

$$\varphi_i(t) = \nu_i t + \varphi_{i0} \quad \text{et} \quad J_i = J_{i0} \quad \text{avec} \quad \nu_i = \nu_i(J) = \frac{\partial \tilde{H}(\vec{J})}{\partial J_i}$$

Appliquons ce formalisme au modèle de l'atome de Bohr. L'atome possède un noyau de charge Ze autour duquel tournent des électrons de charge $-e$ et de masse m . L'électron est soumis à une force centrale $\vec{F} = -Ze^2\vec{r}/r^3$ où \vec{r} est la position de l'électron relative au noyau (le centre de masse est supposé confondu, en première approximation, avec la position du noyau). Nous travaillons donc en coordonnées polaires (r, θ) .

Le modèle de Bohr proprement dit consistait à quantifier l'action correspondant à la variable radiale. Les énergies (discrètes) obtenues correspondaient à la formule empirique de Balmer pour l'atome d'Hydrogène (datant de 1885). Mais en 1891, de meilleurs dispositifs expérimentaux permettent à Michelson de voir qu'à chaque raie correspondent en fait deux raies très voisines. Pour tenter de comprendre cela, Sommerfeld réfléchit sur le sujet et quantifie également l'action liée à la coordonnée radiale. Dans la théorie de Bohr – Sommerfeld, on postule donc que les intégrales des actions sont des multiples entiers de la constante de Planck h . Les conditions de quantification sont :

$$J_r = \oint p_r dr = n_1 h$$

$$J_\theta = \oint p_\theta d\theta = n_2 h$$

où $n_1, n_2 \in \mathbb{N}$ et r, θ sont les coordonnées polaires.

Ceci revient effectivement à dire que les électrons ne peuvent emprunter que des trajectoires bien définies. En utilisant le formalisme ci-dessus, on obtient les énergies discrètes :

$$E_n = -\frac{2\pi^2 m Z^2 e^4}{n^2 h^2} \quad \text{avec} \quad n = n_1 + n_2$$

On voit que l'on obtient une dégénérescence, c'est-à-dire que plusieurs orbites (ou plusieurs mouvements orbitaux) peuvent correspondre à la même énergie. Cela n'explique pas le doublement des raies puisque l'on a toujours qu'un seul nombre quantique (les couples n_1, n_2 tels que $n = n_1 + n_2$ sont indiscernables). Cette explication viendra un peu plus tard, lorsque Sommerfeld utilisera la mécanique relativiste pour traiter le problème : il utilise la masse réduite du système noyau-électron, ce qui lui permet de prendre en compte le fait que la particule au centre n'a pas une masse infiniment plus grande que celle qui lui tourne autour (cf. « masse relativiste »). On peut donc dire que le modèle de Bohr-Sommerfeld est valable dans l'approximation non relativiste.

Insistons sur le fait que le principe même de cette description viole les lois de la mécanique et de l'électrodynamique classiques. En effet, toutes les orbites sont possibles selon la mécanique newtonienne et les électrons, subissant une accélération centripète, devraient rayonner et perdre très rapidement toute leur énergie. Ce modèle est donc considéré comme provisoire par les physiciens du début du XXe siècle.

b. Juin 1925 : article fondateur de Heisenberg

Ce modèle provisoire, cette quantification « illégale » de la physique classique qui menait à des résultats intéressants, conduit nombre de physiciens (notamment à Göttingen et à Copenhague) à développer un savoir-faire étonnamment complexe et étendu. Il s'agissait de revoir chaque problème classique, et de le résoudre à nouveau en quantifiant « les bonnes » variables de façon à obtenir un résultat en accord avec les observations. Il va sans dire que ces manipulations n'étaient pas le fruit d'une démarche systématique, il leur manquait un nouveau cadre conceptuel, et bien plus encore, pour accéder au statut de véritable théorie scientifique.

Dans son livre « La Partie et le Tout », Heisenberg raconte cette période précédant 1925, dans laquelle tout était encore très flou et où l'intuition était déterminante pour aborder et résoudre les problèmes à partir de la théorie de Bohr :

« Au cours de ces années critiques, l'évolution de la physique atomique se déroula comme Niels Bohr me l'avait prédit (...). Les difficultés et les contradictions internes qui s'opposaient à une compréhension des atomes et de leur stabilité ne purent être atténuées et encore moins éliminées. Au contraire, elles se manifestaient avec une netteté croissante. Toute tentative de s'en rendre maître à l'aide des moyens conceptuels de la physique antérieure paraissait à l'avance condamnée à l'échec.

[A la suite des travaux expérimentaux de Ornstein], il s'agissait de déterminer les rapports d'intensité des raies spectrales réunies dans ce que l'on appelle un multiplet. La théorie de Bohr fut utilisée pour obtenir les estimations théoriques correspondantes. Il s'avéra que, de prime abord, les formules déduites de la théorie de Bohr étaient incorrectes ; mais que, grâce à une petite modification de ces formules, on pouvait arriver à des équations nouvelles qui fournissaient apparemment, de façon précise, les résultats expérimentaux. C'est ainsi que l'on apprit à s'adapter progressivement aux difficultés qui se présentaient. On s'habitua à ce fait que les conceptions et les images que l'on avait transposées dans le domaine atomique à partir de la physique antérieure se révélaient, dans le nouveau domaine, à moitié justes et à moitié fausses en même temps ; et qu'il ne fallait donc pas fixer de critères trop stricts pour l'application de ces concepts et images. D'un autre côté, on pouvait parfois, en exploitant habilement la souplesse de la théorie, deviner par simple intuition la formulation mathématique correcte des détails. »

En 1925, un grand pas est réalisé par Heisenberg dans son article « Über quantentheoretische Umdeutung kinematischer und mechanischer Beziehungen » [9] dans lequel il présente les bases de ce qui sera appelé la « mécanique matricielle ». Un des aspects les plus importants de sa démarche est d'influence clairement positiviste. Les récentes théories d'Einstein l'ont profondément marqué, et tout comme la Relativité Restreinte commence par définir les notions d'espace, de simultanéité par le résultat d'une expérience (et non par une représentation a priori qu'en a l'être humain), Heisenberg va travailler d'emblée avec des quantités observables : la position et l'impulsion des électrons en orbite autour d'un noyau atomique ne nous sont manifestement pas accessibles, ce que l'on peut par contre observer et mesurer, ce sont les fréquences et les amplitudes optiques liées aux transitions entre deux niveaux orbitaux du modèle de Bohr. « Lorsque, au cours du semestre d'été 1925, je repris mes occupations à Göttingen – depuis juillet 1924, j'étais chargé de cours à cette université – [j'essayai] de trouver par intuition les formules correctes des intensités des raies spectrales de l'hydrogène (...). Cependant, cette tentative échoua. J'arrivai à une jungle de formules mathématiques compliquées dont je ne trouvai aucune issue. Toutefois cette tentative renforça ma conviction qu'il ne fallait pas rechercher les orbites des électrons dans l'atome, mais que la donnée des fréquences d'oscillation d'une part, des amplitudes (...) d'autre part, pouvait

remplacer pleinement la connaissance de ces orbites. En tout cas, il s'agissait là de grandeurs que l'on pouvait observer directement. » [10]

Comme nous allons le voir ci-dessous, Heisenberg fait du principe de correspondance une hypothèse de travail très clairement définie, dont la justification ne pourra se faire qu'a posteriori. C'est Dirac qui, selon moi, exprime le mieux la version « Heisenbergienne » du principe de correspondance : « [la nouvelle théorie de Heisenberg] suggère que ce ne sont pas les équations de la mécanique classique qui sont de quelque manière en faute, mais que les *opérations mathématiques* à partir desquels les résultats physiques sont déduits doivent être modifiées. » (traduit de l'anglais à partir de [6]).

Partant des considérations ci-dessus évoquées, Heisenberg commence par exprimer une quantité dépendant du temps par sa série de Fourier. Dans le cas précis du mouvement d'un électron autour d'un atome, cela permet de connaître la position $x(t)$ au moyen des amplitudes et des périodes possibles.

$$x(n, t) = \sum_{\alpha=-\infty}^{\infty} A_{\alpha}(n) e^{i\omega(n)\alpha t} = \sum_{\alpha=-\infty}^{\infty} A_{\alpha}(n) e^{2\pi i \nu(n, \alpha) t}$$

Remarque : On traite ici le cas d'un mouvement simplement périodique (on peut généraliser en faisant le remplacement suivant :

$$\nu(n, \alpha) \rightarrow \sum_{k=1}^d \nu_k(n_k, \alpha_k)$$

avec d inférieur ou égal au nombre de degrés de liberté du système, suivant s'il y a dégénérescence des périodes (plusieurs périodes identiques).

Si le mouvement n'est pas (multiplement) périodique, on travaille alors avec la transformée de Fourier correspondante :

$$x(n, t) = \int_{-\infty}^{\infty} A_{\alpha}(n) e^{i\omega(n)\alpha t} d\alpha$$

Dans cette représentation classique, il s'agit de noter que la fréquence classique $\nu(n, \alpha)$ satisfait la relation suivante :

$$\nu(n, \alpha) = \alpha \nu(n, 1) = \alpha \nu \quad (*)$$

Elle représente la fréquence mécanique du mouvement autour du noyau. Le postulat fondamental de l'article (qui permettra de travailler avec des grandeurs observables) est de faire correspondre la fréquence classique $\nu(n, \alpha)$ à la fréquence quantique $\nu(n, n-\alpha)$ qui représente la fréquence de l'onde émise (ou absorbée) lorsque l'électron passe du niveau n au niveau $n-\alpha$. De même, on identifie l'amplitude classique $A_{\alpha}(n)$ à l'amplitude quantique $A_{n, n-\alpha}$ qui, elle, est un nombre complexe car elle doit contenir les informations d'amplitude et de polarisation de l'onde émise ($A_{n, n-\alpha} = |A_{n, n-\alpha}| e^{i\varphi}$ avec $|A_{n, n-\alpha}|$ l'amplitude (réelle) de l'onde et φ l'angle de polarisation). Je qualifie cette identification formelle de postulat dans la mesure où, comme nous allons le voir, elle n'est valable que dans la limite des grands nombres quantiques (i.e. pour n grand).

J'utiliserai dorénavant les notations suivantes pour que la distinction soit claire :

- $\nu(n, \alpha)$ pour la fréquence classique
- $\nu_{n, n-\alpha}$ pour la fréquence quantique
- $A(n, \alpha)$ pour l'amplitude classique
- $A_{n, n-\alpha}$ pour l'amplitude complexe quantique

Réécrivons les développements de Fourier dans ces notations et en tenant compte de (*) :

$$\text{Classique : } x(n, t) = \sum_{\alpha=-\infty}^{\infty} A(n, \alpha) e^{i\omega(n, \alpha)t} = \sum_{\alpha=-\infty}^{\infty} A(n, \alpha) e^{2\pi i \nu(n) \alpha t}$$

$$\text{Quantique : } x(n, t) = \sum_{\alpha=-\infty}^{\infty} A_{n, n-\alpha} e^{i\omega_{n, n-\alpha} t} = \sum_{\alpha=-\infty}^{\infty} A_{n, n-\alpha} e^{2\pi i \nu_{n, n-\alpha} t}$$

Or on remarque que : $\nu(n) \cdot \alpha = \left[\frac{\partial H}{\partial J} \right]_{J=nh} \cdot \alpha \stackrel{(\#)}{=} \frac{\partial H}{\partial J} \frac{\Delta J}{h} \cong \frac{\Delta H}{h} = \nu_{\text{quantique}}(n, n-\alpha)$

L'égalité (#) est valable dans la limite des grands nombres quantiques (à α fixé, on peut toujours trouver n tel que α soit négligeable et que ce développement au premier ordre soit valable). En effet, le ΔJ est alors petit, et le J est alors grand par rapport à h (cf. principe de correspondance). Voilà l'intuition qui justifie la démarche de Heisenberg, consistant à travailler en faisant la correspondance entre fréquence classique et fréquence quantique.

Les relations fondamentales entre différentes fréquences, dans le cas classique et dans le cas quantique, sont les suivantes :

$$\text{Classique : } \nu(n, \alpha') + \nu(n, \alpha - \alpha') = \nu(n, \alpha)$$

$$\text{Quantique : } \nu_{n, n-\alpha'} + \nu_{n-\alpha', n-\alpha} = \nu_{n, n-\alpha}$$

Dans le but de pouvoir travailler avec des fonctions développées en séries de Taylor, Heisenberg se demande alors comment la fonction $x(t)^2$ doit être représentée dans les notations quantiques.

Avec les amplitudes classiques, le produit de deux séries convergentes est donné par le produit de Cauchy :

$$x^2(n, t) = \sum_{\beta=-\infty}^{\infty} B(n, \beta) e^{2\pi i \nu(n) \beta t}$$

$$\text{avec } B(n, \beta) e^{2\pi i \nu(n) \beta t} = \sum_{\alpha=-\infty}^{\infty} A(n, \alpha) A(n, \beta - \alpha) e^{2\pi i \nu(n) (\alpha + \beta - \alpha) t}$$

L'analogie quantique est le suivant :

$$x^2(n, t) = \sum_{\beta=-\infty}^{\infty} B_{n, n-\beta} e^{2\pi i v_{n, n-\beta} t}$$

$$\text{avec } B_{n, n-\beta} e^{2\pi i v_{n, n-\beta} t} = \sum_{\alpha=-\infty}^{\infty} A_{n, n-\alpha} A_{\boxed{n-\alpha, n-\beta}} e^{2\pi i v_{n, n-\beta} t}$$

En effet, pour que la correspondance entre les deux représentations fasse sens, il faut identifier les grandeurs suivantes :

$$v(n, \alpha - \alpha') \sim v_{n-\alpha', n-\alpha}$$

$$A(n, \alpha - \alpha') \sim A_{n-\alpha', n-\alpha}$$

Remarquons que, dans le modèle de Bohr, la différence d'énergie correspondant au saut $(n, n-\alpha)$ n'est pas égale à celle correspondant au saut $(m, m-\alpha)$ si $n \neq m$, c'est pourquoi :

$$v_{n, n-\alpha'} + v_{n, n-(\alpha-\alpha')} \neq v_{n, n-\alpha}$$

Si on ne s'intéresse qu'aux coefficients devant les exponentielles, i.e. les amplitudes, que je noterai cette fois $(.)$, la formule d'une grandeur au carré peut s'écrire ainsi :

$$(x^2(n))_{n, n-\beta} = \sum_{\alpha=-\infty}^{\infty} (x)_{n, n-\alpha} \cdot (x)_{n-\alpha, n-\beta}$$

De même, si on fait le produit d'une grandeur dépendant du temps, disons $a(n, t)$, par une autre $b(n, t)$, on obtient les égalités suivantes pour les amplitudes :

$$\text{Classique : } (ab(n, \beta)) = \sum_{\alpha=-\infty}^{\infty} (a(n, \alpha)) \cdot (b(n, \beta - \alpha)) \tag{*}$$

$$\text{Quantique : } (ab)_{n, n-\beta} = \sum_{\alpha=-\infty}^{\infty} (a)_{n, n-\alpha} \cdot (b)_{n-\alpha, n-\beta}$$

Alors que, pour la relation classique, le produit $a(n, t)$ par $b(n, t)$ est identique au produit de $b(n, t)$ par $a(n, t)$ (ceci se vérifie facilement en faisant le changement d'indice : $\alpha \rightarrow \beta - \alpha$), Heisenberg observe qu'en général, les grandeurs quantiques correspondantes n'ont aucune raison de commuter, c'est-à-dire :

$$(ab(n))_{n, n-\beta} \neq (ba(n))_{n, n-\beta}$$

En effet, on voit en (*) qu'on ne multiplie tout simplement pas les mêmes nombres suivant qu'on effectue l'un ou l'autre des deux produits. Ce nouveau type de multiplication se trouve donc être non commutatif. C'est Born qui, quelques mois plus tard, reconnaîtra là un produit matriciel : le coefficient (mn) d'une matrice produit est donné par le produit scalaire entre la m -ième ligne de la première matrice, et la n -ième colonne de la seconde, autrement dit,

$$(xy)_{mn} = \sum_k x_{mk} y_{kn}$$

Heisenberg continue son article en résolvant par sa nouvelle méthode les problèmes de l'oscillateur anharmonique et du rotateur (électron tournant à distance et vitesse angulaire constantes autour du noyau). Dans la vieille théorie quantique, ces problèmes étaient résolus en deux étapes :

- résoudre l'équation du mouvement $\ddot{x} + f(x) = 0$
- déterminer les constantes d'intégration des mouvements (multi-) périodiques en posant les conditions de quantification de Bohr-Sommerfeld $J = \oint pdq = \int m\dot{x}dx = \int m\dot{x}^2 dt = nh$

La démarche suivie par l'auteur consiste à utiliser la nouvelle règle de multiplication pour transformer les équations du mouvement en relations qui devront être satisfaites par les amplitudes quantiques de transitions entre états stationnaires. Comme le précise Heisenberg, les nombres $x(nm)$ qui forment le « représentant quantique » de $x(n,t)$ ne sont pas son « analogue direct », puisqu'ils ne contiennent pas d'information sur la position de l'électron en fonction du temps.

Il est important de remarquer que dans tout son article, Heisenberg n'utilise que l'opérateur (pas encore reconnu comme tel) x . Pour créer l'analogie quantique des équations du mouvement, travaille en coordonnées (x, \dot{x}) et non (x, p) . Nous verrons dans la section comment Born et Wiener sont arrivés à l'expression de l'opérateur p en fonction de q .

c. Automne 1925 : travaux de Born, Jordan et Heisenberg

Zur Quantenmechanik I

L'article de Heisenberg ouvre une nouvelle voie de recherche qui semble très prometteuse, son formalisme permet de rendre compte de nombreux résultats expérimentaux, et cette fois la méthode ne requiert pas de savoir-faire, d'intuition ou de devinette particulière relative à chaque problème. Mais il manque encore une véritable structure logique de sa théorie : nous avons ci-dessus étudié comment une nouvelle approche du sujet a pu émerger et quels étaient les éléments de base et l'intuition qui ont guidé Heisenberg. Une fois arrivé à cette ébauche de nouvelle théorie, il faut repenser les problèmes et voir quelles relations peuvent être posées comme postulats, quelle structure logique soutiendra cette nouvelle description des phénomènes atomiques.

« A Göttingen, Born et Jordan s'intéressèrent à la nouvelle perspective ainsi ouverte. Le jeune anglais Dirac, qui travaillait à Cambridge, développa des méthodes mathématiques ad hoc pour résoudre les problèmes ainsi posés ; et, au bout de quelques mois déjà, grâce au travail concentré de ces physiciens, un édifice mathématique compact et cohérent se trouvait construit, édifice dont on pouvait espérer qu'il s'adapterait vraiment aux multiples faits expérimentaux de la physique atomique. » [10]

Fin septembre 1925, un premier article de Born et Jordan est reçu dans le journal « Zeitschrift für Physik ». Il met en place le début de structure mathématique dont parle Heisenberg. Je ne vais pas étudier dans tous les détails cet article, je donnerai préférence au suivant, publié par Born, Jordan et Heisenberg quelques mois plus tard.

L'article commence par énoncer quelques définitions et théorèmes relatifs aux règles de bases du calcul matriciel (addition, multiplication, inversion, non commutativité, etc.) Puis la notion de fonction matricielle (i.e. fonction dont les arguments sont des matrices) est définie, et suivie de la règle de dérivation d'une telle fonction par rapport à l'un de ses arguments.

Le deuxième chapitre concerne la dynamique. Les auteurs travaillent explicitement avec les matrices correspondant à p et q , et non comme Heisenberg avec q et \dot{q} .

Dans le but de travailler avec des équations qui sont valables dans le formalisme introduit par Heisenberg, Born et Jordan repartent du principe de Hamilton, selon lequel le mouvement effectué est celui qui extrémalise l'action, et réécrivent cette relation avec des matrices, ils en tirent des équations matricielles du mouvement qui correspondent aux équations canoniques de Hamilton.

La suite logique consiste à caractériser les matrices p et q , puis à dériver leur relation de commutation. Les auteurs montrent d'abord, à partir de la relation

$$J = \oint pdq = \int_0^{1/\nu} p\dot{q}dt$$

que les éléments diagonaux du commutateur sont les suivants (analogue quantique d'une relation connue de Thomas et Kuhn) :

$$\sum_k (p(nk)q(kn) - q(nk)p(kn)) = -i\hbar$$

Grâce à cette relation, on peut déduire que $Tr(pq) = \infty$. En effet, comme $Tr(AB) = Tr(BA)$ si la trace d'un des produits est finie, on devrait avoir $Tr(pq) - Tr(qp) = 0$. Or la relation ci-dessus livre :

$$Tr([p, q]) = Tr(pq) - Tr(qp) = \sum_n -i\hbar = \infty$$

Par conséquent, $Tr(pq) = \infty$ et les matrices p et q sont de dimension infinie. Pour montrer que les coefficients hors-diagonaux du commutateur sont nuls, il faut montrer que $\frac{d}{dt}[p, q] = 0$ et qu'une matrice de dérivée temporelle nulle est diagonale.

La première assertion n'est pas prouvée dans l'article, c'est Jordan qui en fera la preuve plus tard. La deuxième assertion est facile à prouver, en effet, on peut écrire une fonction arbitraire de q et de p de la façon suivante :

$$g = \left(g(nm) e^{2\pi i \nu(nm)t} \right)$$

On a donc : $\dot{g} = 2\pi i \left(\nu(nm) g(nm) e^{2\pi i \nu(nm)t} \right)$. Si nous acceptons que $\nu(nm) \neq 0 \quad \forall n \neq m$ (les fréquences de transition sont non nulles pour tout saut entre deux états stationnaires non confondus), alors $\dot{g} = 0$ implique que g est une matrice diagonale.

De ceci on tire :

$$pq - qp = -i\hbar 1$$

Un élément important de cet article est que la relation de commutation ci-dessus est reconnue comme ayant un statut axiomatique dans la structure logique de la mécanique matricielle.

Dans le but de faire la preuve de la conservation de l'énergie et des conditions quantiques pour les fréquences, Born et Jordan montrent que, pour un Hamiltonien pouvant s'écrire sous la forme :

$$H = H_1(p) + H_2(q)$$

On a :

$$\begin{aligned} Hq - qH &= -i\hbar \frac{\partial H}{\partial p} \\ Hp - pH &= i\hbar \frac{\partial H}{\partial q} \end{aligned} \quad (*)$$

Le raisonnement est le suivant. Connaissant la valeur de $[p,q]$, il est facile de montrer que :

$$\begin{aligned} p^n q - qp^n &= -i\hbar \cdot np^{n-1} \\ q^n p - pq^n &= i\hbar \cdot nq^{n-1} \end{aligned}$$

(utiliser la relation $[AB,C]=A[B,C]+[A,C]B$ et procéder par récurrence). Dès lors, comme l'Hamiltonien peut être décomposé en série de puissance de p et de q qui sont séparées (pas de termes mixtes), on obtient immédiatement les relations (*). Nous aurons besoin de ce résultat pour l'analyse du prochain article, dans lequel je présenterai en détails les preuves de la conservation de l'énergie et des conditions quantiques pour les fréquences.

Le troisième chapitre de « Zur Quantenmechanik I » concerne la résolution des problèmes des oscillateurs harmonique et anharmonique. Cet article est donc le premier « manuel » qui décrit comment faire de la mécanique matricielle à partir des idées fondatrices de Heisenberg.

Zur Quantenmechanik II

Reçu le 16 novembre 1925, l'article de Born, Jordan et Heisenberg intitulé « Zur Quantenmechanik II » donne suite à certaines démonstrations de l'article précédent et amène de nouveaux concepts. Bien que la théorie matricielle ait déjà eu de nombreux succès, les auteurs précisent dans l'introduction que « [leur] théorie ne peut pas encore donner une solution aux principales difficultés de la théorie quantique » (tous les analogues importants des problèmes classiques n'ont pas encore été traités et compris au moyen de cette nouvelle approche).

Le premier chapitre traite des système à un seul degré de liberté, on y trouve pour commencer deux règles de différentiation des matrices, qui se trouvent être équivalentes lorsque l'on dérive l'opérateur Hamiltonien. Soit f une fonction impliquant soit des produits soit des sommes de matrices, les deux façons de dériver f par rapport à l'un de ses arguments sont les suivantes :

$$\begin{aligned} f &= f(x_1, x_2, \dots, x_s) \\ \frac{\partial f}{\partial x_1} &= \lim_{\alpha \rightarrow 0} \frac{f(x_1 + \alpha 1, x_2, \dots, x_s) - f(x_1, x_2, \dots, x_s)}{\alpha} \\ \text{ou } \frac{\partial f}{\partial x_1}(nm) &= \frac{\partial D(f)}{\partial x_1(mn)} \quad \text{avec } D(f) = \text{Trace}(f) \end{aligned}$$

Puis, la relation fondamentale de commutation entre les opérateurs q et p est posée comme point de départ. Comme nous venons de le voir, Born et Jordan avaient démontré que satisfaire les équations canoniques de Hamilton impliquait en effet la relation de commutation ci-dessous. Nous avons donc :

$$\dot{p} = -\frac{\partial H}{\partial q} \quad \dot{q} = \frac{\partial H}{\partial p}$$

$$[p, q] = pq - qp = -i\hbar$$

Nous avons déjà montré à partir de cela que toute fonction pouvant être développée en séries de puissances de p et de q (en tout cas) satisfait :

$$[f, q] = -i\hbar \frac{\partial f}{\partial p}$$

$$[p, f] = -i\hbar \frac{\partial f}{\partial q}$$

Maintenant, la relation quantique entre les fréquences d'émission va permettre de déduire la forme de l'opérateur Energie (l'article traite évidemment en premier lieu des mouvements multipériodiques, i.e. en particulier du problème des spectres d'émission des atomes ; je parlerai des mouvements non périodiques par la suite).

En effet, $\nu(nm) + \nu(mk) = \nu(nk)$ suggère que les fréquences doivent s'écrire comme

$$\nu(nm) = \frac{W_n - W_m}{h}$$

(ceci est aussi connu du modèle de Bohr-Sommerfeld et constitue par conséquent un des points de départ de l'article). Remarquons que les W_i représentent les niveaux énergétiques entre lesquels s'effectue la transition d'un électron, ces valeurs sont donc connues (pour la plupart) expérimentalement, c'est par exemple le cas de l'atome d'Hydrogène. La mécanique matricielle va permettre de retrouver théoriquement ces valeurs, en donnant la relation entre les W_i , l'Hamiltonien et les opérateurs p et q.

On peut définir la grandeur quantique (dans l'article « quantentheoretische Grösse ») W comme une matrice diagonale telle que :

$$W(nm) = \begin{cases} W_n & \text{si } n=m \\ 0 & \text{si } n \neq m \end{cases}$$

Cela nous permet de démontrer une relation importante de la mécanique matricielle, qui lie la dérivée temporelle d'un opérateur au commutateur entre cet opérateur et W.

Puisque $a(nm) = a_{nm} e^{2\pi i \nu_{nm} t}$, l'opérateur $\frac{da}{dt} \equiv \dot{a}$ est défini comme : $\dot{a}(nm) = 2\pi i \nu_{nm} a(nm)$

$$\begin{aligned} \text{Donc } \dot{a}(nm) &= 2\pi i \nu(nm) a(nm) \\ &= 2\pi i \frac{W_n - W_m}{h} a(nm) \\ &= \frac{2\pi i}{h} (W_n a(nm) - W_m a(nm)) \\ &= \frac{i}{\hbar} [W, a](nm) \end{aligned}$$

C'est dire que :

$$\dot{a} = \frac{i}{\hbar}(Wa - aW) = \frac{i}{\hbar}[W, a]$$

Le but des trois auteurs est maintenant de prouver la conservation de l'énergie, i.e. et les relations entre fréquences de transition (ou Frequenzbedingung), i.e. :

$$\dot{H} = 0 \quad \text{et} \quad \nu(nm) = \frac{H_n - H_m}{h} ; \quad H_n = W_n + \text{const.}$$

Autrement dit, on a formé une matrice W contenant les niveaux énergétiques (encore inconnus) qui nous intéressent, mais cette matrice n'a pas encore de rapport avec l'Hamiltonien du système. Il s'agit donc de montrer qu'il existe un Hamiltonien qui commute avec W , et qui est fonction de variables P, Q canoniquement conjuguées à p et q . Je reviendrai sur la notion de transformation canonique en mécanique quantique.

Nous voulons que
$$\dot{P} = -\frac{\partial H}{\partial Q} \quad \dot{Q} = \frac{\partial H}{\partial P}$$

Donc, à l'aide des relations ci-dessus on déduit :

$$\begin{aligned} \dot{P} &= \frac{i}{\hbar}[W, P] = -\frac{\partial H}{\partial Q} = -\frac{i}{\hbar}[P, H] \\ \Rightarrow [W, P] &= [H, P] \end{aligned}$$

De même,

$$\begin{aligned} \dot{Q} &= \frac{i}{\hbar}[W, Q] = \frac{\partial H}{\partial P} = \frac{i}{\hbar}[H, Q] \\ \Rightarrow [W, Q] &= [H, Q] \\ \Rightarrow [W - H, P] &= [W - H, Q] = 0 \end{aligned}$$

$W - H$ commute avec P et Q donc avec toute fonction de P et Q , en particulier H :

$$0 = [W - H, H] = (W - H)H - H(W - H) = WH - H^2 - HW + H^2 = [W, H]$$

On obtient donc bien la conservation de l'énergie :

$$\frac{i}{\hbar}[W, H] = \boxed{\dot{H} = 0}$$

Comme un opérateur conservé dans le temps est diagonal (cf. paragraphe sur « Zur Quantenmechanik I »), on peut écrire que, dans les variables P et Q , H est de la forme :

$$H(nm) = \delta_{nm} H_n$$

et les relations entre fréquences de transition sont faciles à obtenir :

$$\begin{aligned} [W, Q] &= [H, Q] \\ \Rightarrow Q(nm)(W_n - W_m) &= Q(nm)(H_n - H_m) \\ \Rightarrow \nu(nm) &= \frac{W_n - W_m}{h} = \frac{H_n - H_m}{h} \end{aligned}$$

Suit alors un paragraphe important concernant les transformations canoniques. Effectuer une transformation canonique consiste dès lors à trouver une paire de variables P et Q qui préserve la valeur du commutateur entre p et q, i.e. :

$$[P, Q] = [p, q] = -i\hbar$$

Pour que la consistance logique de la démarche soit assurée, il s'agit de rappeler de Born et Jordan avaient trouvé la valeur du commutateur [p,q] à partir des équations canoniques de Hamilton, alors qu'ici, ayant une paire (P,Q) dont le commutateur vaut $-i\hbar$, on aimerait s'assurer que ces variables satisfont les équations canoniques. Il nous faut donc démontrer la réciproque.

Soient donc (P,Q) telles que $[P, Q] = [p, q] = -i\hbar$. Grâce aux relations ci-dessus et sachant maintenant que l'Hamiltonien en les variables (P,Q) (noté \tilde{H} pour clarifier les notations) est diagonal, on peut écrire :

$$-\left(\frac{\partial \tilde{H}}{\partial Q}\right)_{nm} = -\frac{i}{\hbar}(P\tilde{H} - \tilde{H}P)_{nm} = 2\pi i \frac{(\tilde{H}P - P\tilde{H})_{nm}}{h} = 2\pi i \frac{\tilde{H}_n P_{nm} - P_{nm} \tilde{H}_m}{h} = 2\pi i P_{nm} \underbrace{\frac{\tilde{H}_n - \tilde{H}_m}{h}}_{v(nm)} = \dot{P}_{nm}$$

De même pour la deuxième équation. La démarche est bien consistante (si l'on admet que le nouvel Hamiltonien a les mêmes valeurs propres que le premier après transformation canonique : ce sujet sera étudié et rendu rigoureux par Weyl quelques temps plus tard).

Ici intervient une hypothèse qui n'est pas des moindres et qui va faire couler beaucoup d'encre en théorie des transformations. Les auteurs « aimeraient supposer » que la transformation ci-dessous représente la transformation canonique la plus générale.

$$P = SpS^{-1}$$

$$Q = SqS^{-1}$$

Ceci n'est pas trivial à démontrer, mais l'importance des transformations canoniques formulées ainsi réside dans l'observation suivante : lorsque l'on possède une paire (P,Q) qui satisfait la relation de commutation, alors on peut résoudre le problème (trouver les niveaux d'énergie) en déterminant une matrice S, telle que si

$$P = SpS^{-1} \quad \text{et} \quad Q = SqS^{-1}$$

alors la fonction

$$H(P, Q) = H(SpS^{-1}, SqS^{-1}) = SH(p, q)S^{-1} = W$$

est une matrice diagonale.

Born, Jordan et Heisenberg remarquent que cette équation, qui peut aussi s'écrire de la façon suivante :

$$H(p, q)S^{-1} = S^{-1}W$$

est l'analogie de l'équation aux dérivées partielles de Hamilton-Jacobi en mécanique classique. Comme nous le verrons en étudiant les articles de Schrödinger, la matrice S^{-1} correspond à la fonction d'action du système considéré.

Nous voyons ici que chaque article publié donne de nombreuses pistes de réflexion supplémentaires. La théorie des transformation impliquera beaucoup de mathématiciens et

physiciens dans les années qui suivront : il s'agira de justifier rigoureusement les hypothèses employées aux débuts de la « nouvelle théorie quantique ».

Après un chapitre sur la théorie des perturbations, puis sur les systèmes dont l'Hamiltonien dépend explicitement du temps, Born, Jordan et Heisenberg attaquent la généralisation aux systèmes à plusieurs degrés de liberté.

Ceci est suivi d'un paragraphe intéressant exposant le parallèle entre mécanique matricielle et théorie des formes bilinéaires hermitiennes (NB : hermitienne et autoadjointe sont synonyme). A chaque grandeur quantique observable est associée une forme bilinéaire autoadjointe de façon à ce que $A(x, x^*)$ soit réel. (Cela équivaut à dire que les matrices de Heisenberg doivent être autoadjointes pour que l'Hamiltonien diagonalisé ne soit constitué que de coefficients réels (une énergie complexe n'aurait pas de sens physique).

Le calcul des niveaux énergétiques a donc été réduit à la recherche de valeurs propres de matrices hermitiennes en algèbre linéaire. Or comme les grandeurs quantiques ne sont en général pas de dimension finie, ni même bornées, il est nécessaire de montrer que la méthode se généralise au cas des matrices infinies. C'est ce qui est fait dans la suite de l'article. Le cas des spectres continus de valeurs propres est renvoyé aux travaux de Hilbert et Hellinger.

Le quatrième chapitre traite de l'« utilisation physique de la théorie ». La puissance de la théorie matricielle est ici démontrée : toute la théorie des moments cinétiques est exposée, et utilisée pour la description de l'effet Zeeman, et le problème des oscillateurs harmoniques couplés est également résolu.

Remarquons qu'à ce stade, la théorie matricielle était impropre à traiter des mouvements non périodiques à spectres continus. En effet, les systèmes multiples périodiques restent bornés dans l'espace de phase (même si les différentes périodes sont incommensurables), alors que tout mouvement non périodique ne l'est pas. Il faut alors travailler avec les intégrales de Fourier et c'est plutôt le formalisme des opérateurs différentiels qu'il s'agit d'utiliser. Nous verrons cela avec les travaux de Born et Wiener.

Ce travail n'ayant pas le but de décrire en détail chaque article, j'ai sélectionné les éléments importants du point de vue de leur intérêt historique et conceptuel, et du développement logique (en tout cas concernant les démonstrations) de la théorie matricielle. Il s'agit donc essentiellement des principes de base, exposés dans les premiers chapitres des articles « Zur Quantenmechanik ».

d. Novembre 1925 : première contribution de Dirac

Je dois mentionner ici l'une des premières contributions importantes de Dirac en mécanique quantique. Son premier article parut à la fin de l'année 1925. Convaincu du caractère indispensable de la mécanique Hamiltonienne pour l'étude de la physique atomique, il essaie d'adapter la théorie de Heisenberg dans le formalisme Hamiltonien. En quelques semaines il atteint son objectif et établit une des plus profondes et utiles relations entre la mécanique quantique et le formalisme classique de Hamilton-Jacobi (cf. début du chapitre suivant). Ce résultat a été obtenu en formalisant la mécanique de Heisenberg au moyen de « q-numbers », représentants formels des matrices, mais définis seulement au moyen de leurs relations de commutation. En d'autres termes, les q-numbers sont des objets algébriques qui ne commutent pas (en général) ; les représentants formels des nombres utilisés en mécanique classique sont appelés « c-numbers », ceux-ci commutent. Remarquons que Dirac n'avait pas

encore réalisé que la théorie de Heisenberg était un formalisme matriciel, bien que Born et Jordan aient déjà publié leurs travaux.

En bref, Dirac s'attelle à chercher quelle est la quantité correspondant au commutateur en mécanique classique. En utilisant le formalisme angle-action et le principe de correspondance (règles de quantification de Bohr-Sommerfeld), il montre que le coefficient nm du commutateur entre deux q-numbers x et y correspond à une constante près au crochet de Poisson entre x et y, i.e.

$$xy - yx \leftrightarrow i\hbar\{x, y\} = i\hbar\left(\frac{\partial x}{\partial \omega} \frac{\partial y}{\partial J} - \frac{\partial y}{\partial \omega} \frac{\partial x}{\partial J}\right) \quad (*)$$

La remarque fondamentale est alors la suivante : on sait que le crochet de Poisson est invariant sous une transformation canonique. Donc, comme pour des variables *classiques* q et p canoniquement conjuguées le crochet de Poisson {p,q} vaut 1, on en déduit que les q-numbers correspondants satisfont :

$$[p, q] = pq - qp = -i\hbar$$

qui est, comme nous l'avons vu, la relation de commutation canonique démontrée dans l'article de Born et Jordan.

Grâce à l'équivalence (*), Dirac peut obtenir les correspondants quantiques de toutes les relations classiques tant que celles-ci peuvent s'exprimer à l'aide de crochets de Poisson. Il obtient ainsi les conditions quantiques pour les fréquences, notamment.

Il est important de remarquer que la relation (*) a été obtenue en utilisant le modèle de Bohr-Sommerfeld, donc cette contribution de Dirac peut être vue comme une manière de quantifier la mécanique classique avec d'autres postulats que ceux de la vieille théorie quantique (typiquement, on peut postuler les relations de commutation canoniques). Le crochet de Poisson étant un outil analytique puissant et bien connu en mécanique classique, la formulation de Dirac constitue une contribution importante qui aidera au développement de la nouvelle mécanique quantique. Je renvoie à [6] pour plus de détails.

e. Janvier 1926 : théorie opératorielle de Born et Wiener

Je vais ici brièvement mentionner les travaux ultérieurs de Born et Wiener, qui ont permis une avancée considérable puisque les phénomènes non (multiplement) périodiques pouvaient dès lors être traités en mécanique matricielle.

L'idée générale consiste à partir, comme au début de l'article de Heisenberg, de développements de Fourier de fonctions « classiques ». Mais cette fois on travaille directement avec les transformées de Fourier et non avec les séries.

Born et Wiener remarquent que la grandeur qui transforme une fonction x(t) en une autre y(t) a exactement la forme de la matrice q de Heisenberg. Ils interprètent alors q comme l'*opérateur* de multiplication qui transforme x(t) en qx(t)=y(t).

$$\text{Soient } x(t) = \sum_n x_n e^{\frac{2\pi i}{h} W_n t} \quad \text{avec} \quad x_n = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T x(s) e^{-\frac{2\pi i}{h} W_n s} ds$$

$$y(t) = \sum_m y_m e^{\frac{2\pi i}{h} W_m t} \quad \text{avec} \quad y_m = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T y(s) e^{-\frac{2\pi i}{h} W_m s} ds$$

$$\text{Alors } y(t) = \sum_m \sum_n q_{nm} \underbrace{\left(\lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T x(s) e^{-\frac{2\pi i}{h} W_n s} ds \right)}_{x_n} e^{\frac{2\pi i}{h} W_m t} = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T q(t,s) x(s) ds$$

$$\text{avec } q(t,s) = \sum_m \sum_n q_{nm} e^{\frac{2\pi i}{h} (W_m t - W_n s)}$$

$$\text{Donc l'opérateur } q \text{ est tel que : } \left\{ \begin{array}{l} q : x(t) \rightarrow qx(t) \\ \left(\lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T ds q(t,s) \dots \right) x(t) = y(t) \end{array} \right\}$$

Puis ils définissent l'opérateur de différentiation par rapport au temps $D = d/dt$ agissant sur une transformée de Fourier dans le but de donner l'équivalent de la matrice \dot{q} en termes d'opérateurs. On se souvient qu'en mécanique matricielle on a :

$$(\dot{q})_{nm} = 2\pi i v_{nm} q_{nm} = 2\pi i \frac{W_n - W_m}{h} q_{nm}$$

Ils prouvent que

$$\dot{q} = Dq - qD \quad (*)$$

Dans la suite de l'article, l'accent est mis sur la linéarité des opérateurs définis ci-dessus et quelques autres propriétés sont montrées. Avant de passer à la résolution de l'oscillateur harmonique en « formalisme opératoire », les auteurs définissent l'opérateur p comme satisfaisant à la relation de commutation canonique avec q . Ils exigent que l'ensemble des opérateurs satisfaisant cette relation soient hermitiques et forment un ensemble complet.

Mais Born et Wiener étaient très proches de s'apercevoir d'une chose fondamentale : par analogie avec (*), on peut écrire l'expression suivante (le terme entre parenthèses est la matrice q différenciée par rapport à q selon les règles exposées plus haut) :

$$\frac{\partial}{\partial q} q - q \frac{\partial}{\partial q} = \left(\frac{\partial}{\partial q} q \right) = 1$$

Comme on sait que $pq - qp = -i\hbar 1$, on pourrait définir explicitement l'opérateur p comme :

$$p = -i\hbar \frac{\partial}{\partial q}$$

Et toute la mécanique ondulatoire de Schrödinger aurait été à « portée de plume » (cf. chapitre suivant). L'avis de Born à ce propos (interview de 1962, in [11]) : « Nous avons écrit la loi de commutation pour énergie et temps (...); c'était absolument la même chose pour q et p . Mais nous ne l'avons pas vu. Je ne me pardonnerai jamais, car si nous l'avions fait, nous aurions eu l'entière mécanique ondulatoire à partir de la mécanique quantique, quelques mois avant Schrödinger. »

III. Mécanique ondulatoire

a. Introduction : formalisme de Hamilton-Jacobi

Parues au tout début de l'année 1926, les deux « Premières communications » de l'article de Schrödinger intitulé « Quantification et valeurs propres » peuvent être considérées comme un exposé de la mécanique ondulatoire : il y montre comment les conditions de quantifications de la vieille théorie quantique peuvent être remplacées par un problème aux valeurs propres pour une fonction d'onde appelée ψ .

Il est important de mentionner que Schrödinger s'est beaucoup inspiré des travaux de Louis De Broglie pour créer sa théorie ondulatoire. Les références à ses travaux sont présentes dans l'article. Il serait intéressant de s'arrêter sur un des articles majeurs de De Broglie en guise d'introduction à la théorie ondulatoire, mais ceci sort malheureusement du cadre strict de mon travail.

Dans la « Première communication », l'auteur traite de l'exemple fondamental de l'atome d'Hydrogène. Il s'agit de retrouver les positions et intensités des raies spectrales qui ont été mesurées expérimentalement depuis longtemps déjà. Après une courte introduction à propos de ses postulats, Schrödinger donne les détails du calcul permettant de résoudre le problème dans le cas non-relativiste et sans perturbations, pour montrer que « les règles habituelles de quantification peuvent être remplacées par une autre condition, dans laquelle il n'est plus du tout question de « nombres entiers ». Ces nombres entiers s'introduisent de la même manière naturelle que le nombre entier des nœuds d'une corde vibrante. »

La « Deuxième communication » fournira les bases mathématiques et conceptuelles de la mécanique ondulatoire, exposées de façon claire. C'est ce deuxième article que j'analyserai plus en détails.

Comme Schrödinger utilise d'emblée le formalisme de Hamilton-Jacobi, je vais rappeler en quelques lignes le but et les points essentiels de cette théorie.

Le but est de déterminer systématiquement les intégrales premières (ou constantes du mouvement) d'un système Hamiltonien. On cherche donc à construire une transformation canonique de (q,p) vers des nouvelles variables qui sont des constantes dans le temps.

Considérons le cas général où l'Hamiltonien dépend explicitement du temps. On cherche (Q,P) tel que :

$$\begin{pmatrix} q \\ p \end{pmatrix} \mapsto \begin{pmatrix} Q = cste \\ P = cste \end{pmatrix}$$

On doit par conséquent satisfaire les équations canoniques :

$$\begin{aligned} \frac{\partial \tilde{H}}{\partial P_i} &= \dot{Q}_i = 0 \quad \forall i = 1 \dots l \\ \frac{\partial \tilde{H}}{\partial Q_i} &= -\dot{P}_i = 0 \quad \forall i = 1 \dots l \end{aligned}$$

On en conclut que \tilde{H} est égal à une constante indépendante de Q et P , que l'on peut sans perte de généralité évaluer à zéro (grandeurs intéressantes obtenues par dérivations partielles). On a donc :

$$\tilde{H}(Q, P, t) = 0$$

On choisit alors (par convention) une fonction génératrice de seconde espèce S pour décrire la transformation. Par définition (cf. théorie des fonctions génératrices):

$$Q_i = \frac{\partial S(q, P, t)}{\partial P_i} \quad i = 1 \dots l$$

$$p_i = \frac{\partial S(q, P, t)}{\partial q_i} \quad i = 1 \dots l$$

$$\tilde{H}(Q, P, t) = H(q, p, t) + \frac{\partial S(q, P, t)}{\partial t} = 0$$

En insérant les deux premières lignes dans la troisième, on obtient ladite équation de Hamilton-Jacobi :

$$H\left(q, \frac{\partial S(q, P, t)}{\partial q}, t\right) + \frac{\partial S(q, P, t)}{\partial t} = 0$$

Résoudre cette équation pour la fonction S permet d'obtenir les équations du mouvement du système.

Dans le cas indépendant du temps, l'Hamiltonien est une constante du mouvement $H=E=cste$ et on peut écrire S sous la forme :

$$S = W - Et$$

Et l'équation de Hamilton-Jacobi prend la forme :

$$H\left(q, \frac{\partial W}{\partial q}\right) = E$$

b. Hiver 1926 : articles fondateurs de Schrödinger

Première communication

Revenons au premier article historique de Schrödinger, reçu le 27 janvier 1926. Comme mentionné ci-dessus, le sujet traité ici est l'atome d'Hydrogène. Il s'agit donc de travailler sur le problème d'une particule plongée dans un potentiel central. L'équation de Hamilton-Jacobi correspondant à cette situation, dans le cas indépendant du temps, est la suivante (notations de Schrödinger) :

$$H\left(q, \frac{\partial S}{\partial q}\right) = E$$

avec S une fonction génératrice de seconde espèce. On sait que pour ce problème, on peut travailler avec une fonction S séparable, bien que l'Hamiltonien du système ne soit pas lui-même séparable [8]. Autrement dit :

$$S = \sum_i S_i(q_i)$$

Schrödinger introduit alors une nouvelle fonction,

$$\Psi = \sum_i \psi_i(q_i)$$

de telle sorte que $S = K \log(\Psi)$. L'équation de Hamilton-Jacobi devient alors :

$$H\left(q, \frac{\partial S}{\partial \psi} \frac{\partial \psi}{\partial q}\right) = H\left(q, \frac{K}{\Psi} \frac{\partial \Psi}{\partial q}\right) = E$$

Il va ensuite remplacer, comme annoncé, la recherche des conditions de quantification par un problème de calcul des variations. En toute généralité, l'expression ci-dessus peut s'écrire :

$$F(\Psi) = \text{"forme quadratique de } \Psi \text{ et de ses premières dérivées"} = 0$$

Or, il se trouve qu'il est plus simple de rendre stationnaire l'intégrale de F plutôt que de résoudre directement l'équation ci-dessus. On est amené à résoudre le problème variationnel suivant :

$$\delta J = \delta \iiint dx dy dz F(\Psi) = 0$$

En l'occurrence, pour le problème de l'atome d'Hydrogène, nous avons :

$$H = \frac{p^2}{2m} - \frac{e^2}{r} \text{ avec } p = \frac{K}{\Psi} \frac{\partial \Psi}{\partial q}$$

Ce qui nous donne pour $F(\Psi)$:

$$F(\Psi) = \left(\frac{\partial \Psi}{\partial x}\right)^2 + \left(\frac{\partial \Psi}{\partial y}\right)^2 + \left(\frac{\partial \Psi}{\partial z}\right)^2 - \frac{2m}{K^2} \left(E + \frac{e^2}{r}\right) \Psi$$

Les équations d'Euler-Lagrange correspondant au problème variationnel pour l'intégrale de F sur tout l'espace de configuration nous amène à satisfaire simultanément les deux équations suivantes :

$$\Delta \Psi + \frac{2m}{K^2} \left(E + \frac{e^2}{r}\right) \Psi = 0$$

$$\int df \delta \Psi \frac{\partial \Psi}{\partial n} = 0 \quad (*)$$

La deuxième équation assure que les conditions de bord du problème sont bien définies, l'intégrale étant ici étendue jusqu'à l'infini en chaque variable.

La suite de l'article est plus technique que conceptuel, le point intéressant est que Schrödinger recherche des fonctions ψ réelles, finies, à détermination unique et deux fois dérivables dans tout l'espace de configuration. Il montre que l'on peut satisfaire à ces conditions ainsi qu'à (*) pour toutes les valeurs positives de E (électron non lié au noyau, trajectoires hyperboliques), mais dans le cas où $E < 0$ (états liés), il existe seulement des valeurs discrètes de E qui permettent de résoudre le problème. « Autrement dit, le problème de calcul des variations

énoncé a un spectre discret et en même temps un spectre continu de valeurs propres. (...) Pour obtenir une concordance numérique effective [avec les données expérimentales], il faut donner à la constante K la valeur $h/2\pi$ ».

L'équation d'onde d'écrit donc :

$$\Delta\psi + \frac{8\pi^2 m}{h^2} \left(E - \frac{e^2}{r} \right) \psi = 0$$

A la fin de l'article, Schrödinger donne une interprétation partielle et provisoire de cette fonction d'onde ψ . « On est évidemment très tenté, dit-il, de rattacher la fonction y à un phénomène de vibrations inter-atomiques, ayant un caractère de réalité beaucoup plus prononcé que celui, si souvent mis en doute actuellement, des trajectoires électroniques. » Il justifie le choix de se baser avant tout sur une nouvelle forme mathématique de résolution du problème, plutôt que sur des postulats liés à une quelconque intuition physique à propos de ψ , en disant que ce procédé « a l'avantage de mettre plus clairement en évidence l'essentiel de la question. ». On remarquera que l'introduction de ψ , qui est une fonction continue, mène à une interprétation physique « classicisante » : si « le réel » est contenu dans la fonction d'onde, alors ce « réel » est continu, sans discontinuités essentielles. Les valeurs discrètes apparaissent seulement au moment du passage d'un état stationnaire de ψ à un autre, en totale analogie avec les fréquences propres d'une corde vibrante (problème classique).

Deuxième communication

La deuxième communication de Schrödinger, qui fait toujours partie de son article « Quantification et valeurs propres », a été reçue le 23 février 1926. Elle donne une justification et un contenu physiques à l'équation d'onde qui provient de l'équation aux dérivées partielles de Hamilton. Schrödinger montre avec grands détails que le principe variationnel de Hamilton peut être considéré comme un principe de Fermat (principe de temps minimal) pour une propagation d'ondes dans l'espace de configuration (espace des q_i) et que l'équation aux dérivées partielles de Hamilton-Jacobi exprime le principe de Huygens pour cette propagation. Ensuite, le passage de la description des systèmes macroscopiques à celle des systèmes microscopiques (phénomènes intra-atomiques) est comparé à la transition entre optique géométrique et optique ondulatoire. Remarquons que les travaux de Hamilton ont eu leur origine dans ses tentatives d'unification entre l'optique ondulatoire et la mécanique, [11], ce n'est donc pas un hasard si ce formalisme met en évidence des phénomènes ondulatoires. Le mérite de Schrödinger réside donc surtout dans l'interprétation et l'utilisation à bon escient de ce formalisme. Les intuitions, ébauches d'interprétations, espoirs et craintes de l'auteur quant à l'avenir de cette nouvelle description des phénomènes quantiques sont développées sur de nombreux paragraphes. Je vais ici donner l'idée générale du développement de Schrödinger.

Nous avons vu, dans le paragraphe sur le formalisme de Hamilton-Jacobi, que l'équation aux dérivées partielles qui nous intéresse est la suivante (cas où $H = T + V$ avec $T=T(p,q)$ l'énergie cinétique $V=V(q)$ l'énergie potentielle, et $\partial H / \partial t = 0$).

$$\frac{\partial S}{\partial t} + T \left(q_i, \frac{\partial S}{\partial q_i} \right) + V(q_i) = 0$$

avec $S = W - Et$

[Remarque : je garde ici une notation cohérente par rapport aux paragraphes précédents. S et W sont inversés par rapport à l'article de Schrödinger.]

Notons tout de suite que la fonction S représente l'action, i.e. l'intégrale par rapport au temps du Lagrangien $L = T - V$. En effet, puisque $S=S(q,P,t)$ avec $\dot{P} = 0$, on a :

$$\begin{aligned}\frac{dS}{dt} &= \sum_i \frac{\partial S}{\partial q_i} \dot{q}_i + \frac{\partial S}{\partial t} \\ &= \sum_i p_i \dot{q}_i - H \\ &\equiv L\end{aligned}$$

Nous avons donc $S = \int L dt + cste$ (et nous avons montré que : $\frac{\partial S}{\partial t} = -H$; $\frac{dS}{dt} = L$).

On peut donc écrire :

$$\begin{aligned}\frac{\partial S}{\partial t} + T\left(q_i, \frac{\partial S}{\partial q_i}\right) + V(q_i) &= -E + T\left(q_i, \frac{\partial S}{\partial q_i}\right) + V(q_i) \\ \Rightarrow 2T\left(q_i, \frac{\partial S}{\partial q_i}\right) &= 2(E - V(q_i))\end{aligned}$$

Schrödinger montre que cette expression devient « particulièrement simple et intuitive » si l'on définit dans l'espace de configuration une métrique non-euclidienne à l'aide de l'énergie cinétique T du système. Pour ce faire, il faut exprimer T en fonction des vitesses et non plus en fonction des moments :

$$T(p, q) \rightarrow \bar{T}(q, \dot{q})$$

Et on travaille avec l'élément linéaire

$$ds^2 = 2\bar{T}(q, \dot{q}) dt.$$

Il s'agit donc de travailler avec les coordonnées (q, \dot{q}) comme avec des vecteurs de \mathbb{R}^3 en remplaçant bien sûr l'élément linéaire euclidien $(dr^2 = dx^2 + dy^2 + dz^2)$ par ds .

Comme les $\partial S / \partial q_i$ sont les composantes du vecteur $grad(S)$ on peut écrire (5) comme

$$\begin{aligned}(grad(S))^2 &= 2(E - V) \\ \text{ou } |grad(S)| &= \sqrt{2(E - V)}\end{aligned}$$

[Remarque : pour être tout à fait juste, il manque la constante $1/m$ devant $(grad(S))^2$, puisque $T = p^2 / 2m$ (cf (8)). Nous pourrions la rajouter dans l'équation d'onde qui se trouve plus bas, mais pour l'instant je garde les notations de Schrödinger.]

Maintenant, on peut tracer mentalement dans l'espace q l'ensemble des surfaces $S = cste$ en indexant chacune d'elles la valeur de S qui lui correspond. Ainsi, grâce à la relation ci-dessus, si on a une condition initiale S_0 et si on connaît la surface lui correspondant, il est possible de construire toutes les autres surfaces $S_0 + dS$ à un temps t fixé, puisque la valeur du gradient de $S(q, t)$ est connue en chaque point de l'espace de configuration. Il suffit de prendre la normale en chaque point de la surface initiale (après avoir choisit un « côté positif ») et d'y construire un vecteur de longueur

$$d\vec{n} = dS_0 \frac{grad(S)}{\sqrt{2(E - V)}}$$

$$\text{En norme : } dn = \frac{dS_0}{\sqrt{2(E-V)}}$$

On a ici normalisé le vecteur gradient pour obtenir un vecteur unitaire pointant dans la direction normale. Comme S dépend des q , le dS_0 varie en chaque point.

La variation temporelle de S étant simple ($S = W - Et$), les surfaces $S = \text{cste}$ obtenues avec la méthode ci-dessus vont simplement se déplacer en fonction du temps « pour prendre l'aspect et la place de celles qui suivent », il suffit d'écrire :

$$S(t_0 + t) = W - E(t_0 + t) = S_0 - Et$$

Comme S est l'action (produit d'une énergie par un temps), chacun des points de la surface S_0 se déplace durant l'intervalle de temps dt sur la normale positive d'une distance infinitésimale égale à :

$$dn = \frac{Edt}{\sqrt{2(E-V)}}$$

C'est dire que le mouvement des surfaces se fait à la vitesse :

$$u = \frac{dn}{dt} = \frac{E}{\sqrt{2(E-V)}}$$

Schrödinger fait les comparaisons suivantes : « notre ensemble de surfaces $S = \text{cste}$ peut être considéré comme un ensemble de surfaces d'ondes progressives dans l'espace des q , qui restent les mêmes au cours du temps, et dont la valeur de la vitesse de phase en chacun des points est donnée par $[u]$. En effet, la construction indiquée qui emploie les normales de l'onde peut évidemment être remplacée par la construction des ondes élémentaires de Huygens et de leur enveloppe. L'indice de réfraction est proportionnel à l'inverse $[de u]$; il dépend du point considéré mais il est indépendant de la direction. L'espace des q est donc optiquement isotrope mais non homogène. »

Il est important de remarquer que la fonction S , dans cette théorie, est assimilée à la phase du système d'ondes, et il n'est pas question d'amplitudes. Cela joue un rôle important dans l'interprétation qui suivra.

On peut maintenant parler du « mouvement mécanique » dont le formalisme de Hamilton-Jacobi a pour but de simplifier la recherche (i.e. l'évolution des q en fonction du temps). Il se trouve que le point représentatif du système mécanique ne se déplace pas le long des normales aux surfaces avec une vitesse de propagation égale à u . Montrons que sa vitesse est proportionnelle à $1/u$:

$$\begin{aligned} ds^2 &= 2\bar{T}(q, \dot{q}) dt \\ \Rightarrow v &= \frac{ds}{dt} = \sqrt{2\bar{T}} = \sqrt{2(E-V)} \end{aligned}$$

On voit donc qu'il y a une différence fondamentale entre la vitesse de propagation des ondes et celle du point représentatif du système : comme $p_i = m\dot{q}_i = \partial S / \partial q_i$, le mouvement mécanique a une vitesse grande où le gradient de l'action est grand, i.e. où la variation de S est grande ; par contre, comme on l'a vu, c'est là où la vitesse u des ondes est petite (petite

distance entre chaque surface). Cette différence explique pourquoi l'analogie mécanique ne fait pas intervenir de grandeur comme la longueur d'onde, la fréquence, ou alors l'amplitude : ces grandeurs n'ont en effet pas d'analogue mécanique. Schrödinger dit que la signification de cette « phase » des ondes que représente S n'est pas encore très claire. On sait que l'optique géométrique est une approximation de l'optique ondulatoire lorsque les longueurs d'onde sont grandes par rapport aux obstacles que la lumière traverse. Dans le cas où la longueur d'onde est de l'ordre de grandeur de l'obstacle, la notion de rayon lumineux ne peut plus être utilisée avec succès : les phénomènes observés (diffraction, interférences) sont en contradiction avec une telle approche. Schrödinger voit dans sa nouvelle théorie une analogie profonde avec le problème de l'optique. Pour lui, les notions utilisées pour décrire un mouvement mécanique macroscopique sont obsolètes dans le cas des phénomènes intra-atomiques. Il faut alors utiliser non plus les équations aux dérivées partielles de Hamilton-Jacobi, mais directement une équation d'onde, qui peut être dérivée de l'analyse du phénomène ondulatoire venant d'être mis en évidence.

Le point de départ d'une nouvelle mécanique ondulatoire est donné dans le deuxième paragraphe de l'article. Il faut tout d'abord imposer une forme aux ondes définies plus haut pour pouvoir aboutir à une équation d'onde. Schrödinger les choisit sinusoïdales, pour simplifier le problème. (« Cette hypothèse est la plus simple et la première qui se présente à l'esprit ; cependant, à cause de son importance fondamentale, il faut souligner expressément l'arbitraire qu'elle contient. ») Schrödinger fait donc les hypothèses suivantes :

- la fonction d'onde ne dépend du temps que par un facteur du type $\sin(\dots)$
- l'argument du sinus est une fonction linéaire de l'action S .
- Le coefficient de S , qui doit avoir les dimensions de l'inverse d'une action, est choisit égal à $2\pi/h$

Le facteur contenant la dépendance temporelle peut être écrit de la manière suivante :

$$\sin\left(\frac{2\pi S}{h} + cste\right) = \sin\left(-\frac{2\pi Et}{h} + \frac{2\pi W}{h} + cste\right)$$

La fréquence des ondes est donc égale à

$$\nu = \frac{E}{h}$$

Cette fréquence est proportionnelle à l'énergie du système et correspond à la relation de Planck-Einstein. Tout l'enjeu sera de donner une interprétation probante et cohérente correspondant au nouveau formalisme.

On peut maintenant trouver la dépendance de la vitesse u de propagation des ondes en fonction de la fréquence, autrement dit trouver la relation de dispersion pour ces ondes.

$$u = \frac{E}{\sqrt{2(E-V)}} = \frac{h\nu}{\sqrt{2(h\nu-V)}} \quad \left(\Rightarrow \nu = \frac{h}{2} \left(\frac{\nu}{u} \right)^2 + V \right)$$

$$\text{Or } \nu = \sqrt{2(E-V)} = \sqrt{2(h\nu-V)} \Rightarrow \nu = \frac{h\nu}{u} = \frac{d\nu}{d(\nu/u)}$$

On voit donc que « la vitesse du point représentatif du système est égale à la vitesse d'un groupe d'ondes, réparties dans un domaine de fréquences peu étendu. » Schrödinger fait alors le lien avec la théorie de De Broglie, où un résultat similaire faisant appel à la Relativité est démontré pour les « ondes de phase » d'un électron.

Ces résultats permettent à Schrödinger de faire une comparaison plus précise entre la validité de sa théorie et celle de la mécanique classique : si on construit un paquet d'onde (en superposant des ondes sinusoïdales) qui a des dimensions spatiales assez restreintes (i.e. la superposition est telle que l'interférence est destructive partout sauf dans une région très limitée de l'espace), de façon à ce que cela se rapproche d'un point matériel, alors on a besoin d'un grand spectre de fréquences. Or la vitesse de groupe d'un tel paquet d'onde n'est plus égale à $d\omega/dk = dv/d(v/u)$, cela est une approximation au premier ordre pour un paquet d'onde à spectre peu étendu en fréquences. Donc la notion de point représentatif comme défini ci-dessus n'a plus de sens physique quand le paquet d'ondes est trop restreint dans l'espace : il faut alors utiliser exclusivement la mécanique ondulatoire dans ce cas.

La suite de l'article montre comment passer de la représentation ondulatoire aux trajectoires classiques lorsque l'on a affaire à un paquet d'onde comme ci-dessus :

« A une constante universelle près ($1/h$), S représente la phase de la fonction d'onde. Supposons qu'au lieu d'un seul système d'onde, il y en ait tout un ensemble continu, défini par la variation continue d'un paramètre α_i quelconque. Dans ce cas, les conditions $\partial S/\partial \alpha_i = cste$ expriment le fait que tous les éléments infiniment voisins de cet ensemble ont même phase. Ces équations déterminent par conséquent le lieu géométrique des points de concordance de phase. Ce lieu géométrique se réduit à un point si les équations sont en nombre suffisant ; elles déterminent alors en fonction du temps le point unique de concordance de phase. »

Schrödinger termine la première partie de son article par des réflexions sur ses hypothèses et le développement potentiel de sa « mécanique ondulatoire ». Les bases conceptuelles d'une telle approche doivent être nouvelles, au même titre que celles de l'optique ondulatoire par rapport à l'optique géométrique.

Je cite ici une phrase qui me paraît importante :

« L'hypothèse à laquelle je crois pouvoir attacher un haut degré de certitude est la suivante : La manière correcte de concevoir ou de représenter les phénomènes mécaniques consiste à les rattacher à une propagation d'ondes dans l'espace des q et non à un mouvement de points représentatifs dans le même espace. »

Comme déjà mentionné, Schrödinger insiste sur le fait que les équations de Hamilton-Jacobi devront être remplacées par l'équation d'onde, c'est-à-dire l'équation que devra satisfaire la fonction d'onde ci-dessus introduite. A l'instar de Heisenberg, Schrödinger considère la notion de trajectoire comme ne faisant pas sens dans la description des phénomènes intra-atomiques, mais on remarquera que sa démarche ne la rejetait pas a priori : c'est le fait qu'une fonction d'onde puisse décrire les choses de manière fructueuse qui mène au rejet de ce concept.

De façon à développer cette conception ondulatoire de la mécanique, il faut pouvoir travailler avec une équation d'onde. Ce que l'on a à disposition est la vitesse de propagation des ondes en fonction de l'énergie ou de la fréquence, mais ces informations ne suffisent pas pour fixer l'équation d'onde. Schrödinger choisit arbitrairement l'équation bien connue du second ordre (a priori il n'y a pas de raison pour que cette équation soit du second ordre) :

$$\Delta \psi - \frac{1}{u^2} \ddot{\psi} = 0$$

Comme on le sait, cette équation est valable uniquement pour les phénomènes dont la dépendance temporelle est de la forme $e^{2\pi i \nu t}$.

Connaissant l'expression de u , et en se rappelant qu'il manque un facteur m depuis la première remarque faite au début du paragraphe, on peut écrire :

$$\Delta\psi + \frac{8\pi^2 m}{h^2} (h\nu - V)\psi = 0 \quad (*)$$

ou $\Delta\psi + \frac{8\pi^2 m}{h^2} (E - V)\psi = 0$

Cette équation correspond à la fonction d'onde obtenue dans la première communication pour le problème de l'atome d'Hydrogène. L'argument est cette fois plus profond que le principe variationnel précédemment utilisé.

Après avoir proposé une généralisation possible de l'équation et s'être inquiété du trop grand nombre de solution qu'elle pourrait permettre, Schrödinger remarque que les conditions de quantification de Bohr-Sommerfeld sont déjà incluses dans l'équation ci-dessus :

« Le calcul des niveaux quantiques n'a plus lieu en deux étapes distinctes :

- 1) calcul de toutes les trajectoires dynamiquement possibles
- 2) élimination de la plupart des solutions obtenues dans 1, dont on ne retient qu'un petit nombre, satisfaisant à des conditions particulières, données a priori.

Les niveaux quantiques sont déterminés tous à la fois, comme valeurs propres de l'équation (*) qui contient en elle-même ses propres conditions aux limites. »

Avant de passer à divers exemples d'application de sa mécanique ondulatoire (Oscillateur de Planck, Rotateur à axe fixe, Rotateur rigide dont l'axe est libre, Rotateur déformable (molécule biatomique)), comme l'avait fait Heisenberg dans le premier article que j'ai cité et résumé [9], Schrödinger parle des récentes publications sur la mécanique matricielle. Voici quel est sa vision de la situation :

« Je ne saurais passer sous silence ici les recherches en cours, entreprises par Heisenberg, Born, Jordan et d'autres chercheurs éminents en vue d'aplanir les nombreuses difficultés de la théorie des quanta ; ces recherches ont déjà à leur actif un certain nombre de succès si remarquables, qu'il est difficile de ne pas croire qu'elles contiennent au moins une partie de la vérité. Les recherches de Heisenberg se rapprochent singulièrement des miennes, en ce qui concerne le but à atteindre (...). Mais la méthode qu'il emploie est à tel point différente de la mienne qu'il ne m'a pas été possible de trouver jusqu'à présent l'élément qui permettrait de les réunir. »

IV. Synthèse ou « nouvelle théorie quantique »

a. Mise en parallèle des approches matricielle et ondulatoire

Comme nous l'avons vu, il aura suffi de moins de huit mois pour que deux nouvelles approches des phénomènes atomiques naissent et commencent à faire leurs preuves de façon sérieuse (physiquement et mathématiquement) au point d'être considérées comme durables, c'est-à-dire comme devant faire partie intégrante de la théorie quantique que l'on « sent » venir. Outre les sujets qui n'avaient pas encore pu être cernés par l'une des deux théories, le lien entre mécanique matricielle et mécanique ondulatoire avait commencé à faire couler de l'encre depuis quelque temps déjà. Lanczos, Pauli, et Schrödinger sont parmi les premiers à avoir vu que les deux approches se trouvaient être équivalentes, mais l'article de Schrödinger analysé au paragraphe suivant est l'un des plus connus.

Il est très important de rappeler que les physiciens de l'époque avaient un bagage mathématique relativement faible en algèbre linéaire, mais un grand savoir-faire dans le domaine des équations différentielles. L'analyse fonctionnelle (équations intégrales, etc) leur était également moins familière. Cela explique que les structures algébriques sous-tendant chaque théorie n'aient pas été reconnues immédiatement.

Comparons tout d'abord les démarches respectives, que j'ai détaillées dans les deux premiers chapitres de ce travail.

<i>Mécanique matricielle</i>	<i>Mécanique ondulatoire</i>
Modèle de Bohr-Sommerfeld, formalisme angle-action	Formalisme de Hamilton-Jacobi
Développement en série de Fourier des fonctions classiques. Identification : $v_{classique}(n, \alpha) \sim v_{quantique}(n, n - \alpha)$	La fonction génératrice S, qui représente l'action, peut être considérée comme la phase d'une « fonction d'onde » ψ
Via cette identification, les coefficients de Fourier des fonctions q, p, etc. peuvent être représentés par des matrices.	On peut remplacer les équations du mouvement classiques par une équation d'onde (i.e. au lieu de chercher S(q,P,t) on cherche ψ dont S est la phase).
Dans le formalisme matriciel, les équations du mouvement classiques sont remplacées par la diagonalisation de l'Hamiltonien H : il faut résoudre une équation matricielle du type : $H\psi = E\psi$.	Modulo une hypothèse simplificatrice sur la « forme » de ψ , l'équation d'onde à résoudre est une équation aux valeurs propres dont les valeurs propres donnent les fréquences du mouvement.

Le début de la synthèse consistera, du côté de la mécanique matricielle, à reconnaître que les opérateurs peuvent être considérés comme agissant sur « quelque chose ayant un sens physique ». Du côté de la mécanique ondulatoire, reconnaître la structure linéaire des opérateurs différentiels qui agissent sur ψ mènera à faire un parallèle direct avec les équations de type matriciel (i.e. écrire $A\Delta\psi + B\psi = 0$ comme $L\psi = 0$ permet de reconnaître $(\lambda I - \alpha)\psi = 0 \rightarrow \det(\lambda I - \alpha) = \det(L) = 0$). Nous verrons plus bas que c'est le formalisme des espaces de Hilbert qui permet de faire l'identification formelle entre les deux théories.

D'un point de vue physique, la fonction d'onde ψ devra acquérir un sens clair dans cette « nouvelle théorie quantique ». Schrödinger et Dirac font quelques tentatives dans les articles

analysés ci-dessous, mais c'est Born qui, un peu plus tard, proposera l'interprétation probabiliste bien connue : « le carré du module de ψ représente la probabilité de présence de la particule (modulo normalisation) ».

b. Mars 1926 : article unificateur de Schrödinger

Le 18 mars 1926 est reçu l'article de Schrödinger intitulé « Sur les rapports qui existent entre la mécanique quantique de Heisenberg-Born-Jordan et la mienne » [18]. C'est un des articles fondamentaux qui a mené à l'unification des deux théories exposées plus haut en une seule « théorie quantique ». J'utilise ici le terme « unifier » dans le sens de « voir les deux théories comme une seule ». Je vais résumer les points importants de son développement.

Dans son introduction, Schrödinger commence par remarquer que mécanique matricielle et mécanique ondulatoire « conduisent aux mêmes résultats, au moins pour les cas particuliers connus jusqu'à présent, même lorsque ces résultats diffèrent de ceux que donne l'ancienne théorie des quanta », et ceci bien que « tout, point de départ, conception, méthode, appareil mathématique utilisé, parai[sse] radicalement différent. » Pour approfondir la différence de « conception », il rappelle que la mécanique ondulatoire est appelée « la vraie théorie du discontinu » : elle met en effet au premier plan les amplitudes et fréquences de transition entre niveaux quantiques. C'est une réalité considérée comme discrète qui est à décrire. Par contre, comme nous l'avons déjà vu, la mécanique ondulatoire décrit ces sauts discrets au moyen d'une fonction d'onde, et manipule des objets représentant une réalité continue.

Pourtant, Schrödinger a compris quel est le lien entre les deux théories, et a pour but dans cet article de montrer que « formellement, d'un point de vue purement mathématique, ces deux théories sont identiques ». Il va montrer comment obtenir une fonction d'onde (dépendant des coordonnées et des moments conjugués) à partir d'une matrice de Heisenberg et vice versa, en passant par une écriture des équations fondamentales de la mécanique matricielle (commutateur entre p et q , dérivée temporelle, règle de multiplication, etc.) dans le formalisme de opérateurs agissant sur des fonctions d'onde. Voyons ceci plus en détails :

Le premier pas consiste à démontrer que les règles de calcul applicables aux matrices de Heisenberg (qui dépendent des $2l$ coordonnées q et p) sont identiques aux règles de calcul de l'analyse ordinaire, applicables à des opérateurs différentiels linéaires qui dépendent des l variables q seulement. Pour assurer cette correspondance, il faut associer à la matrice q_k de Heisenberg l'opérateur de multiplication par q_k , et à la matrice p_k l'opérateur de différentiation par q_k .

$$\begin{aligned} \{q_k\} &\rightarrow (\psi \mapsto q_k \psi) \\ \{p_k\} &\rightarrow \left(\psi \mapsto \frac{\partial}{\partial q_k} \psi \right) \end{aligned}$$

Ces opérateurs agissant sur les fonctions d'onde, (qui sont continues, plusieurs fois continûment différentiables, etc.), on voit facilement que l'opérateur $\partial/\partial q_k$ commute avec $\partial/\partial q_m$ et ceci $\forall m \neq k$. Dans le cas $m = k$, on a :

$$\left(K \frac{\partial}{\partial q_m} q_m - q_m K \frac{\partial}{\partial q_m} \right) \psi = K \frac{\partial q_m}{\partial q_m} \psi + q_m K \frac{\partial \psi}{\partial q_m} - q_m K \frac{\partial \psi}{\partial q_m} = K \psi$$

$$\Rightarrow K \frac{\partial}{\partial q_m} q_m - q_m K \frac{\partial}{\partial q_m} = K \cdot Id$$

En remplaçant K par $-i\hbar$, on obtient l'équivalent en termes d'opérateurs de la relation de commutation entre les matrices q et p de Heisenberg.

Comme les opérateurs q et p ne commutent pas lorsqu'ils ont le même indice, il est indispensable de noter les « fonctions de p et q » de façon convenable, c'est-à-dire en respectant l'ordre des opérations que l'on veut effectuer. Une telle fonction est appelée par Schrödinger fonction bien ordonnée. (Des règles de symétrisation sont connues à l'époque et permettent de rendre les opérateurs hermitiques s'ils ne le sont pas déjà.)

Prenons un exemple :

"Fonction des p et q bien ordonnée" $F(p, q) = p_k^2 + q_k^2 - q_m$

"Opérateur correspondant" $\tilde{F} : \psi \mapsto -\hbar^2 \frac{\partial^2}{\partial q_k^2} \psi + (q_k^2 - q_m) \psi$ (*)

Il montre alors que l'on peut faire correspondre une matrice à toute fonction bien ordonnée comme décrite ci-dessus, ceci par l'intermédiaire de l'opérateur qui lui correspond (i.e. grâce à la correspondance (*)) et par un système complet de fonctions orthogonales arbitraires définies dans l'espace de configuration.

Si l'on désigne les fonctions du système complet orthogonal par $\{u_k(q_1 \dots q_l)\} = \{u_k(x)\}$, on doit avoir :

$$\int \rho(x) u_k(x) u_m(x) dx = \begin{cases} 0 & \text{si } k \neq m \\ 1 & \text{si } k = m \end{cases}$$

avec $\rho(x)$ une fonction de poids pour le produit scalaire, ou une fonction de normalisation dont la racine devrait être mise en facteur de chaque fonction $u_k(x)$. On suppose de plus que les fonctions $u_k(x)$ « s'annulent d'une façon suffisamment rapide sur la frontière naturelle de l'espace q ». Cela servira à éliminer des intégrales gênantes qui apparaissent notamment en faisant des intégrations par parties comme nous le verrons par la suite. Remarquons que cette hypothèse supplémentaire n'est pas anodine puisqu'elle signifie que Schrödinger travaille dans un espace de Hilbert (cf. point c).

L'idée est donc de faire correspondre à un opérateur comme dans l'exemple (7) une matrice dont les coefficients sont les suivants :

$$F^{km} = \int \rho(x) u_k(x) \tilde{F}(u_m(x)) dx$$

c'est-à-dire, comme on peut le lire dans une note de bas de page importante : « F^{km} est le k -ième coefficient du développement en série de fonctions orthogonales de la fonction qu'on obtient en appliquant l'opérateur $[\tilde{F}]$ à la fonction u_m ».

Schrödinger montre ensuite que l'addition et la multiplication des fonctions bien ordonnées ou de leurs opérateurs correspondent à l'addition et au produit matriciel. Je renvoie à l'article pour le détail du calcul.

Le pas important était dans un premier temps de voir que les « fonctions de p et q » agissent sur les fonctions d'onde de manière linéaire et peuvent avoir un équivalent matriciel via la correspondance (*) et les hypothèses décrites plus haut concernant le système de fonctions formant la base dans laquelle on travaille. L'équivalent opératoire des relations matricielles de Heisenberg est ensuite assez facile à prouver.

En effet, les relations suivantes :

$$[p, f] = -i\hbar \frac{\partial f}{\partial q}$$

$$[f, q] = -i\hbar \frac{\partial f}{\partial p}$$

se démontrent ainsi dans le formalisme des opérateurs :

- 1) Puisque l'opérateur p consiste à dériver par rapport à q et multiplier par $-i\hbar$, on pourrait écrire le membre de droite de la première relation pf ; mais comme la fonction f dépend des p et des q , on va en plus dériver par q les composantes de f . C'est pourquoi il faut retrancher fp au résultat.
- 2) Pour prouver la seconde relation, on regarde ce que donne le résultat de fq , et sachant que dans les « fonctions de p et q » les p_k n'apparaissent que sous forme de produits de puissances (sinon on développe en série de puissances), lorsque se présente un terme en $p_k^n q_k$ l'idée est de ramener pas à pas l'opérateur q_k vers la gauche en rajoutant la valeur du commutateur. On va forcément trouver qf et les autres termes seront le commutateur de f avec q , qui sera égal à $\partial f / \partial p$.

Exemple : Soit $f(p, q) = p_k^3$

Alors (en oubliant le coefficient $-i\hbar$) :

$$fq_k(u) = p_k p_k p_k q_k u = p_k p_k \left(q_k p_k u + \underset{=1}{u} \right) = p_k (q_k p_k (p_k u) + p_k u) + p_k p_k u$$

$$= q_k p_k (p_k p_k u) + p_k p_k u + p_k p_k u + p_k p_k u = q_k f(u) + 3p_k^2 u$$

$$\Rightarrow (fq_k - q_k f)u = 3p_k^2 u = \frac{\partial f}{\partial p_k}$$

Schrödinger montre aussi dans une note de bas de page la réciproque suivante : à une matrice donnée correspond un et un seul opérateur différentiel linéaire si le système orthogonal de fonctions de base et la fonction de poids du produit scalaire sont donnés.

Il s'attaque ensuite à la résolution des équations du mouvement de Heisenberg. Pour transcrire un Hamiltonien en termes d'opérateurs, il faut veiller à obtenir une « fonction bien ordonnée ». Schrödinger fait remarquer que les procédés de symétrisation employés jusqu'à présent en mécanique matricielle (qui servent à obtenir une matrice autoadjointe pour H , sachant que p et q le sont) ne sont pas univoques, et que son processus de variation employé en mécanique ondulatoire, dont il donne l'équivalence avec la diagonalisation d'un Hamiltonien, fournit automatiquement une symétrisation unique et bien déterminée de l'opérateur correspondant à H .

Précisons la méthode annoncée. Il s'agit de résoudre les équations du mouvement de Hamilton :

$$\left\{ \begin{array}{l} \left(\frac{dq_k}{dt} \right)^{nm} = \left(\frac{\partial H}{\partial p_k} \right)^{nm} \\ \left(\frac{dp_k}{dt} \right)^{nm} = \left(-\frac{\partial H}{\partial q_k} \right)^{nm} \end{array} \right\} k = 1 \dots l ; n, m = 1, 2, 3, \dots \text{ad inf}$$

L'interprétation de la dérivation temporelle en mécanique matricielle est la suivante : « il existe une suite de nombres $\{v_k\}$ telle que les équations [ci-dessus] soient satisfaites si

l'opérateur d/dt signifie simplement : multiplication de l'élément de matrice (n,m) par $2\pi i(v_n - v_m)$ » comme nous l'avons vu dans la première partie de ce travail. Il s'agit donc de résoudre les équations suivantes avec comme inconnues les fréquences quantiques $\{v_k\}$:

$$\left\{ \begin{array}{l} (v_n - v_m) q_k^{nm} = \frac{1}{h} (Hq_k - q_k H)^{nm} \\ (v_n - v_m) p_k^{nm} = \frac{1}{h} (Hp_k - p_k H)^{nm} \end{array} \right\} k = 1 \dots l ; n, m = 1, 2, 3, \dots \text{ad inf}$$

Et le point important est ici de voir que ce sera le choix convenable du système de fonctions orthogonales qui permettra de résoudre ces équations. C'est-à-dire que l'on peut satisfaire les équations ci-dessus en se plaçant dans la base des fonctions propres du « problème aux limites normal » de l'équation aux dérivées partielles suivante :

$$H(\psi) = E\psi \quad \text{ou} \quad -H(\psi) + E\psi = 0$$

Cette équation étant « identique à l'équation des ondes qui forme la base de la mécanique ondulatoire ». Schrödinger montre ensuite explicitement que « l'opérateur des ondes de la mécanique ondulatoire est identique à l'opérateur de la fonction d'Hamilton, changée de signe et convenablement symétrisée ».

En partant de l'Hamiltonien général suivant,

$$H = T(p_k, q_k) + V(q_k)$$

nous avons vu, dans la première communication de « Quantification et valeurs propres », que l'équation d'onde correspondant au problème de l'atome d'Hydrogène était obtenue à partir du calcul variationnel suivant

$$\delta J_1 = \delta \int \left\{ T \left(-i\hbar \frac{\partial \psi}{\partial q_k}, q_k \right) + \psi^2 V(q_k) \right\} d\tau = 0$$

avec la condition (*)

$$J_2 = \int \psi^2 d\tau = 1$$

Le $d\tau$ est la mesure du volume de l'espace de configuration, en tenant compte du changement de variables $(p, q) \rightarrow (-i\hbar \partial / \partial q, q)$.

Je rappelle que l'équation d'onde obtenue dans la première communication par un simple principe variationnel est la même que celle (générale) obtenue dans la deuxième communication en raisonnant sur la fonction d'action comme phase d'une fonction d'onde.

De l'équation (*) on tire :

$$0 = \frac{1}{2} (\delta J_1 - E \delta J_2) = \int \left\{ -\frac{h^2}{8\pi^2} \sum_k \frac{\partial}{\partial q_k} \left(|d\tau| \frac{\partial T}{\partial p_k} \left(\frac{\partial}{\partial q_k}, q_k \right) \right) + (V(q_k) - E) |d\tau| \psi \right\} \delta \psi dq_1 \dots dq_l$$

On construit l'équivalent matriciel (qui va constituer une version symétrisée univoque de H) en utilisant le théorème d'Euler sur les fonctions homogènes (puisque T est une forme quadratique homogène en les p_k), que je rappelle :

$$T(p_k, q_k) = \frac{1}{2} \sum_k p_k \frac{\partial T}{\partial p_k}(p_k, q_k)$$

On obtient donc :

$$\underbrace{\frac{h^2}{8\pi^2} \frac{1}{|d\tau|} \sum_k \frac{\partial}{\partial q_k} \left\{ |d\tau| \frac{\partial T}{\partial p_k} \left(\frac{\partial}{\partial q_k}, q_k \right) \right\}}_H - V(q_k) \psi + E\psi = 0$$

Schrödinger explicite ce qu'il entend par « problème aux limites normal » en disant qu'en général, on constate que la « frontière naturelle » de l'espace de configuration, située à l'infini, constitue une singularité pour notre équation différentielle et que la seule condition aux limites admissible est que la fonction d'onde cherchée doit rester finie sur cette frontière. Ceci signifie que l'on travaille dans l'espace $L^2(\mathbb{C})$ des fonctions de carré sommable, mais Schrödinger n'en est pas encore là.

A la fin de l'article, Schrödinger parle de l'interprétation physique des deux théories et expose son interprétation de l'intensité et de la polarisation du rayonnement émis. Il va plaider en la faveur de son interprétation ondulatoire qu'il estime fructueuse d'un point de vue physique, notamment pour ce qui concerne le couplage entre ondes électromagnétiques et fonctions d'onde. Après avoir montré l'équivalence mathématique des deux théories, il s'agit de savoir laquelle des deux on va effectivement utiliser pour résoudre les problèmes physiques et interpréter les résultats. Selon le courant positiviste influent au début du XXe siècle, dont le représentant principal est Mach, « le rôle des théories physiques consiste uniquement à fournir une description mathématique des relations expérimentales entre grandeurs observables ». Dans ce cadre, la meilleure théorie est celle qui ne contient pas d'éléments inobservables et qui n'utilise pas de concepts propres à faire porter un faux caractère intuitif à certaines grandeurs. La mécanique matricielle serait alors le candidat désigné, mais Schrödinger ne pense pas qu'une équivalence mathématique soit identique à une équivalence du point de vue physique, il donne l'exemple des deux expressions mathématiques donnant l'énergie électrostatique d'un système de conducteurs chargés, et montre qu'elle sont absolument équivalentes si l'on reste dans le cadre de l'électrostatique. Par contre, lorsque l'on passe à l'électrodynamique, une de ces expressions reste valable tandis que l'autre ne l'est plus. Il défend donc l'idée intéressante selon laquelle la pertinence d'une représentation mathématique d'une grandeur physique est liée à sa validité lors d'une généralisation de la théorie. Cet argument lui permet de justifier sa préférence pour l'interprétation ondulatoire, car il voit déjà une généralisation possible en ces termes à la question du couplage entre les phénomènes intra-atomiques et le champ électromagnétique « ou ce qui doit le remplacer dans la nouvelle théorie ». Il montre ensuite d'une façon plus précise (pour le cas du problème à un seul électron) qu'il « devrait être possible (...) de trouver une interprétation réellement intuitive de l'intensité et de la polarisation de la lumière émise ». Nous ne nous étendrons pas sur son développement mathématique étant donné que c'est la démarche « philosophique » qui nous intéresse en premier lieu.

Pour finir, il est fort intéressant de se pencher sur une note de bas de page vers la fin de l'article de Schrödinger, qui concerne les travaux de Lanczos [14]. En bref, ceux-ci montrent que la théorie matricielle de Heisenberg-Born-Jordan peut être formulée non seulement dans langage matriciel mais aussi, de manière équivalente, dans celui des équations intégrales. Le

premier livre de Jammer [11] mentionne qu'en fait, la reformulation par Lanczos de la mécanique de Heisenberg en termes d'équations intégrales précède la première communication de Schrödinger d'environ quatre semaines et fut donc, strictement parlant, le premier formalisme « théorique-continu » de la mécanique quantique. Mais les physiciens n'étant pas très familiers avec les équations intégrales et les communications de Schrödinger ayant paru quelques semaines plus tard, les travaux de Lanczos n'ont pas eu beaucoup de « succès ». Schrödinger a pris connaissance de la récente publication du mathématicien. Il en parle en ces termes :

« Dans une note très intéressante récemment parue (...) K. Lanczos exprime des opinions analogues et y ajoute en même temps l'observation précieuse que la dynamique atomique de Heisenberg est également susceptible d'une interprétation continue. Cependant, pour le reste, les points de contact immédiats entre le travail de Lanczos et celui-ci sont beaucoup moins nombreux qu'on ne le penserait au premier abord.

En ce qui concerne le calcul effectif du système des formules de Lanczos, que celui-ci a laissé provisoirement indéterminé, il ne faut pas chercher à identifier le noyau symétrique $K(s, \sigma)$ de Lanczos avec la fonction de Green de notre équation des ondes. En effet, cette fonction de Green, lorsqu'elle existe, a comme valeurs propres les niveaux quantiques eux-mêmes, tandis que le noyau de Lanczos doit avoir comme valeurs propres les inverses de ces mêmes niveaux. »

On peut faire le parallèle avec une lettre de Pauli à Jordan [15], écrite le 12 avril 1926, juste après la publication de la première communication de Schrödinger mais juste avant l'article d'équivalence que je viens de commenter. Elle établit la connexion entre mécanique matricielle et ondulatoire, en prenant un chemin logique différent et indépendant de celui de Schrödinger. Pauli ne publia jamais le contenu de cette lettre; sa femme, après sa mort, permit à un physicien de la faire publier. Il est également brièvement question des travaux de Lanczos dans cette lettre. Pauli écrit que son approche n'a pas beaucoup de valeur, essentiellement pour les mêmes raisons que Schrödinger...

c. Décembre 1926 : article de Dirac

A la fin de l'année 1926, Dirac publie un article concernant plus précisément la théorie des transformations. Après les démonstrations de l'équivalence entre mécanique matricielle et ondulatoire, Dirac considère son papier comme une généralisation des travaux de Lanczos sur les équations intégrales. On peut dire qu'il « donna à sa théorie un fondement logique nouveau et indépendant en introduisant l'un des outils mathématiques les plus utiles en théorie quantique : la fameuse fonction delta. » [11]

L'introduction de son article est assez difficile à suivre dans le sens où il semble parfois se répéter, ou alors décrire comme non trivial quelque chose qui nous semble aller de soi. Le contexte historique ne doit pas être oublié : le formalisme utilisé était nouveau pour la plupart des physiciens, c'est précisément la période où tout était en train d'être mis en place de façon rigoureuse et logique.

Dirac a pour but de répondre à la question suivante dans son article : « Comment peut-on obtenir à partir de la théorie un c-number susceptible d'être comparé avec des valeurs expérimentales ? » En d'autres termes, quelles sont les prédictions théoriques que peut fournir la mécanique quantique ? Il commence par dire qu'en mécanique matricielle, les éléments de la matrice diagonale qui représente l'énergie du système sont les niveaux d'énergie possibles, et la matrice de polarisation totale donne les fréquences de transition et les intensités des raies spectrales. En mécanique ondulatoire, le carré de la fonction d'onde peut « dans certains cas »

être interprété comme une probabilité. On a donc accès aux probabilités de transition. L'article va proposer une interprétation qui englobe et généralise ces deux « points de vue ».

Pour commencer, Dirac annonce qu'il va travailler avec des matrices dont les indices varient de façon continue. Il définit alors sa fonction delta :

$$\begin{aligned}\delta(x) &= 0 \quad \forall x \neq 0 \\ \int \delta(x) dx &= 1\end{aligned}$$

Notons que Dirac reconnaissait que sa fonction delta était seulement un artifice mathématique pratique : « Strictement parlant, bien sûr, $\delta(x)$ n'est pas une fonction propre de x , mais ne peut être regardée que comme la limite d'une certaine suite de fonctions. Mais tout de même, on peut utiliser $\delta(x)$ comme si c'était une fonction propre pour pratiquement tous les buts de la mécanique quantique sans obtenir de faux résultats. »

Il montre ensuite que : $\int_{-\infty}^{\infty} f(x) \delta^n(a-x) dx = f^n(a)$.

Et introduit la notation suivante : les grandeurs non primées sont des q-numbers, les grandeurs primées sont des c-numbers qui indiquent de façon continue les lignes et colonnes des matrices. Par exemple, $\alpha(\alpha' \alpha'')$ est le coefficient de la matrice α dont l'indice de la ligne est α' et celui de la colonne est α'' .

La matrice identité est la suivante : $1(\alpha' \alpha'') = \delta(\alpha' - \alpha'')$

Et une matrice quelconque x s'écrit : $x(\alpha' \alpha'') = x(\alpha') \delta(\alpha' - \alpha'')$

Les transformations canoniques sont écrites de la manière suivante :

$$\begin{aligned}G &= b g b^{-1} \\ G(\xi' \xi'') &= \iint b(\xi' \alpha') d\alpha' g(\alpha' \alpha'') d\alpha'' b^{-1}(\alpha'' \xi'') \\ &\equiv \iint (\xi' / \alpha') d\alpha' g(\alpha' \alpha'') d\alpha'' (\alpha'' / \xi'')\end{aligned}$$

On peut donc écrire les relations équivalentes :

$$\begin{aligned}Gb &= b g \\ \int g(\xi' \xi'') d\xi'' (\xi'' / \alpha') &= \int (\xi' / \alpha'') d\alpha'' g(\alpha'' \alpha') \equiv g(\xi' \alpha') \\ \text{De meme } b^{-1}G &= g b^{-1} \\ \int (\alpha' / \xi'') d\xi'' g(\xi'' \xi') &= \int g(\alpha' \alpha'') d\alpha'' (\alpha'' / \xi') \equiv g(\alpha' \xi')\end{aligned}$$

Où l'on a utilisé la convention de changement de lettre pour les indices (mais pas pour les q-numbers) lorsqu'on change de base. Les deux expressions ci-dessus peuvent être regardées comme représentant la variable g par rapport à deux nouveaux points de vue différents, i.e. où les lignes et les colonnes de la matrice se réfèrent à deux points de vue différents.

Notant une matrice diagonale comme ceci :

$$\xi(\xi' \xi'') = \xi' \delta(\xi' - \xi'')$$

Il veut trouver quelle est la matrice représentant la variable canoniquement conjuguée à ξ , nommée η . Grâce aux définitions ci-dessus et à des intégrations par parties, il trouve :

$$\eta(\xi' \alpha') = -i\hbar \frac{\partial(\xi' \alpha')}{\partial \xi'}$$

De plus, une fonction arbitraire de ξ et η peut s'écrire :

$$f(\xi, \eta)(\xi' \alpha') = f\left(\xi', -i\hbar \frac{\partial}{\partial \xi'}\right)(\xi' \alpha')$$

Le résultat important et très général est donc le suivant : pour passer dans une base où la matrice f est diagonale, il faut résoudre les équations différentielles suivantes :

$$f\left(\xi', -i\hbar \frac{\partial}{\partial \xi'}\right)(\xi' \alpha') = f(\xi, \eta)(\xi' \alpha') = \int (\xi' \alpha'') d\alpha'' f(\alpha'', \alpha') = f(\alpha')(\xi' \alpha')$$

On peut les résoudre pour ξ' , en considérant α' comme un paramètre.

L'interprétation de cette dernière relation est la suivante, c'est la généralisation de l'équation de Schrödinger pour n'importe quel système de fonctions orthogonales (représentées ici comme des matrices de changement de base) et pour n'importe quelle fonction f . On obtient l'équation de Schrödinger en prenant $f = H$, $f(\alpha') = E$ et en travaillant avec les variables position et impulsion. On voit donc en partant d'une représentation matricielle continue et en imposant la relation de commutation comment on peut arriver à une équation aux valeurs propres qui a pour but de rendre une fonction arbitraire f diagonale. Dirac voit que les fonctions propres ψ de Schrödinger sont les fonctions qui constituent la base dans laquelle le Hamiltonien est diagonal. Résoudre l'équation ci-dessus pour x' revient donc bien à chercher les fonctions propres de H .

Ayant montré que la matrice de changement de base $(\xi' \alpha')$ est la généralisation des fonctions d'onde de Schrödinger et que l'équation ci-dessus est la généralisation de son équation d'onde, le but est maintenant de répondre à la question de départ : quelles prédictions peut-on faire théoriquement et qui puissent être vérifiées par l'expérience. L'interprétation probabiliste de Born consiste à décomposer une fonction d'onde quelconque y sur les états stationnaires (fonctions propres de H) et à voir le carré du module des coefficients comme la probabilité de transition vers l'état stationnaire correspondant :

$$\psi(q, t) = \sum_n c_n(t) \psi_n(q) \Rightarrow \text{Pr ob}(\text{transition vers } \psi_n) = |c_n(t)|^2$$

La généralisation de Dirac consiste à décomposer une fonction y sur n'importe quel ensemble de fonctions correspondant à $(\xi' \alpha')$ et à interpréter le carré du module des coefficients comme la probabilité de transition vers l'état correspondant. Cela donne un sens physique à tout ensemble de fonctions de base et pas seulement aux fonctions propres de l'Hamiltonien.

d. Point final : Von Neumann, formalisme des espaces de Hilbert

Quelques années plus tard (après 1930), John Von Neumann montre que la théorie quantique peut être vue comme du calcul d'opérateurs hermitiens (auto-adjoints) dans un espace de Hilbert, et montre que la théorie matricielle de Heisenberg comme la mécanique ondulatoire de Schrödinger sont des représentations particulières de ce calcul. En effet, Heisenberg travaillait dans l'espace $\ell^2(\mathbb{C})$ des suites infinies de nombres complexes dont la somme des carrés des modules est finie,

$$\ell^2(\mathbb{C}) = \left\{ (a_i)_{i=-\infty}^{\infty} \mid \sum_{i=-\infty}^{\infty} |a_i|^2 < \infty \right\}$$

Schrödinger, lui, travaillait dans l'espace $L^2(\mathbb{C})$ des fonctions de carré sommable et à valeurs complexes

$$L^2(\mathbb{C}) = \left\{ f : \mathbb{R} \rightarrow \mathbb{C} \mid \int_{-\infty}^{\infty} |f(x)|^2 dx < \infty \right\}$$

Mais comme ces deux espaces sont deux réalisations de dimension infinie du même espace de Hilbert, ils sont en particulier isomorphes et il existe une bijection entre les fonctions d'ondes de $L^2(\mathbb{C})$ et les suites infinies de $\ell^2(\mathbb{C})$, ainsi qu'entre les opérateurs différentiels autoadjoints et les matrices hermitiennes. C'est donc dire que résoudre un problème aux valeurs propres dans $L^2(\mathbb{C})$ est équivalent à diagonaliser une matrice autoadjointe dont les coefficients sont dans $\ell^2(\mathbb{C})$. L'analyse détaillée des « Fondements mathématiques de la mécanique quantique » de Von Neumann sort du cadre de ce travail puisque cet écrit a été publié bien après la période critique 1925-26. Cependant, cette lecture est évidemment d'un grand intérêt pour la compréhension de la théorie quantique enseignée de nos jours. Je renvoie à [19].

Comme je l'ai déjà mentionné, l'interprétation probabiliste de Born de la fonction ψ joue un rôle crucial pour donner une assise physique à la « nouvelle » théorie quantique.

V. Conclusion

On peut conclure ce travail en discutant le rôle du formalisme mathématique au sein d'une théorie physique. L'étude des mécaniques matricielle et ondulatoire fournit un exemple d'un grand intérêt. J'aimerais montrer que le formalisme fait partie intégrante d'une théorie physique, dans le sens où celui-ci peut conditionner une réflexion, lui fixer un cadre, mais aussi et surtout donner une ouverture nouvelle ou être initiateur d'un développement conceptuel.

Nous avons vu que mécanique matricielle et ondulatoire s'opposaient en bien des points sur le plan des idées. Dès le moment où l'identification entre fréquences classiques et fréquences quantiques a mené à la reconnaissance par Born d'une structure matricielle des coefficients de Fourier des grandeurs classiques, le développement de la mécanique de Heisenberg s'est fait dans la direction du calcul algébrique de quantités non commutatives. Les grandeurs que l'on cherchait étaient les fréquences de transition entre états stationnaires, les amplitudes et la polarisation des ondes émises, les intensités des raies spectrales. Tous les résultats expérimentaux demandaient à ce que l'on justifie un comportement discret, quantifié de la nature. Dès lors, cette approche algébrique du problème a mis l'accent sur la discontinuité et en a fait un élément fondateur. Le formalisme utilisé ne donnait pas l'occasion de créer des images ou des interprétations intuitives des phénomènes. Bien que les matrices de Heisenberg aient été plus tard manipulées par Born et Wiener en tant qu'opérateurs, dans le sens très général du terme (i.e. sans nécessairement demander de représentation particulière de ces opérateurs), la notion d'*état* d'un système physique, ou de « ce sur quoi les opérateurs agissent » n'était pas nécessaire à la compréhension des phénomènes à ce stade. Ce sont les coefficients des opérateurs qui varient avec le temps, fournissant ainsi une dynamique.

En mécanique ondulatoire, au contraire, Schrödinger travaille avec le formalisme de Hamilton-Jacobi, dans lequel il interprète la fonction d'action comme la phase d'une fonction d'onde. L'appareil mathématique consiste en la résolution d'équations différentielles linéaires, donc impliquant un principe de superposition. Rien ne mène à penser, dans cette approche, que les discontinuités observées expérimentalement sont intrinsèques à la nature. On travaille avec une grandeur (ψ) continue, finie, plusieurs fois continûment différentiable, à détermination unique, et qui satisfait une équation d'onde au caractère classicisant : « [les sauts discrets apparaissent] de la même manière naturelle que le nombre entier des nœuds d'une corde vibrante ».

Les deux approches sont donc conceptuellement radicalement différentes. On peut dire que, dans un premier temps, la présence de deux formalismes si différents est un obstacle à la pleine compréhension des phénomènes. Mais les similitudes des deux théories au niveau des résultats obtenus et de la structure des équations à résoudre (équations aux valeurs propres) mènent à des tentatives d'unification (sans le sens : essayer de détecter ce qu'il y a de commun). Le fruit de ces recherches oriente les travaux des physiciens vers des outils mathématiques qui leur sont peu familiers, et qui, par là même, permettront une ouverture vers de nouvelles idées. Donner une interprétation physique claire à la fonction d'onde ψ , tout en créant une théorie des transformations qui soit une « théorie des changements de bases dans un espace fonctionnel », est dès lors inévitable, pour ne citer qu'un exemple.

La reformulation des fondements mathématiques de la physique quantique par Von Neumann, que j'ai très brièvement abordée à la fin de ce travail, est un réel facteur de développement des idées. Les théories dérivées exclusivement de l'approche de De Broglie – Schrödinger, je pense aux « ondes pilotes » de Bohm notamment, se révéleront trop faibles du point de vue de leur « pouvoir explicatif » (concept qui serait à définir...) et laisseront place à des approches modernes plus algébriques.

Nous voyons donc que le formalisme mathématique, indéniablement nécessaire à la physique moderne, loin de n'être qu'un outil, est inhérent à la synthèse physique d'une nouvelle théorie, et fait partie intégrante de ce que l'on appelle « comprendre ».

VI. Références

- [1] BOHR Niels, *Physique atomique et connaissance humaine*, Traduit de l'allemand par Edmond Bauer et Roland Omnes, Editions Gonthier, Paris, 1961.
- [2] BORN Max, JORDAN Pascual, *Zur Quantenmechanik*, Zeitschrift für Physik, 34, 858-888, article reçu le 27 septembre 1926.
- [3] BORN Max, HEISENBERG Werner, JORDAN Pascual, *Zur Quantenmechanik II*, Zeitschrift für Physik, 35, 557-615, article reçu le 16 novembre 1926.
- [4] BORN Max, WIENER Norbert, *Eine neue Formulierung der Quantengesetze für periodische und nicht periodische Vorgänge*, Zeitschrift für Physik, 36, 174-187, article reçu le 5 janvier 1926.
- [5] COHEN-TANNOUJDI Claude, DIU Bernard, LALOE Franck, *Mécanique Quantique I*, Collection Enseignement des sciences, Hermann, Paris, 1973 (Nouveau tirage 1998).
- [6] DIRAC P.A.M., *The fundamental Equations of Quantum Mechanics*, Exhibition Senior Research Student, St-John's College, Cambridge, article reçu le 7 novembre 1925.
- [7] DIRAC P.A.M., *The Physical Interpretation of the Quantum Dynamics*, Proceedings of the Royal Society of London (A), 113, 621-641 (1926), article reçu le 2 décembre 1926.
- [8] DROZ Michel, *Cours de Mécanique II*, Université de Genève, donné au semestre d'hiver de l'année 2005-2006.
- [9] HEISENBERG Werner, *Über quantentheoretische Umdeutung kinematischer und mechanischer Beziehungen*, Zeitschrift für Physik, 33, 879-893, article reçu le 29 juin 1925.
- [10] HEISENBERG Werner, *La Partie et le Tout, Le monde de la physique atomique (Souvenirs, 1920-1965)*, traduit de l'allemand par Paul Kessler, Collection Les Savants et le Monde, Editions Albin-Michel, Paris, 1972.
- [11] JAMMER Max, *The conceptual development of Quantum Mechanics*, McGraw-Hill Book Company, USA, 1966.
- [12] JAMMER Max, *The Philosophy of Quantum Mechanics, The Interpretations of Quantum Mechanics in Historical Perspective*, John Wiley & Sons, USA, 1974.
- [13] LACKI Jan, *The puzzle of canonical transformations in early quantum mechanics*, History and Philosophy of Science, University of Geneva.
- [14] LANCZOS Kornel, *On a Field Theoretical Representation of the New Quantum Mechanics*, Frankfurt, article reçu le 22 décembre 1925.
- [15] PAULI Wolfgang, *Lettre à Pascual JORDAN du 12 avril 1926*, traduction anglaise.

[16] SCHRODINGER Erwin, *Quantification et valeurs propres, Première communication*, Annalen der Physik (4) Bd.79. 1926, article reçu le 27 janvier 1926. (²)

[17] SCHRODINGER Erwin, *Quantification et valeurs propres, Deuxième communication*, Annalen der Physik (4) Bd.79. 1926, article reçu le 23 février 1926. (*)

[18] SCHRODINGER Erwin, *Sur les rapports qui existent entre la mécanique quantique de Heisenberg-Born-Jordan et la mienne*, Annalen der Physik (4) Bd.79. 1926, article reçu le 18 mars 1926. (*)

[19] VON NEUMANN John, *Les fondements mathématiques de la mécanique quantique*, Traduit par A. Proca, Librairie Alcan, Paris, 1946.

² Articles traduits de l'allemand par Al. Proca dans le livre intitulé Mémoires sur la Mécanique Ondulatoire publié par la librairie Alcan en 1933, réimpression aux Editions Jacques Gabay, Sceaux, 1988.