From canonical transformations to transformation theory, 1926–1927: The road to Jordan's *Neue Begründung* ⋆

Anthony Duncan [a], Michel Janssen [b,∗]

[a] *Department of Physics and Astronomy, University of Pittsburgh*

[b] *Program in the History of Science, Technology, and Medicine, University of Minnesota*

Abstract

We sketch the development from matrix mechanics as formulated in the *Dreimännerarbeit* of Born, Heisenberg, and Jordan, completed in late 1925, to transformation theory developed independently by Jordan and Dirac in late 1926. Focusing on Jordan, we distinguish three strands in this development: the implementation of canonical transformations in matrix mechanics (the main focus of our paper), the clarification of the relation between the different forms of the new quantum theory (matrix mechanics, wave mechanics, q-numbers, and operator calculus), and the generalization of Born's probability interpretation of the Schrödinger wave function. These three strands come together in a two-part paper by Jordan published in 1927, "On a new foundation [*neue Begründung*] of quantum mechanics."

Key words: canonical transformations, transformation theory, Pascual Jordan, Fritz London

1 Introduction

Matrix mechanics as presented in the famous *Dreimännerarbeit* of Born, Heisenberg, and Jordan (1926) was still lacking various basic elements of quantum

⋆ This paper was written as part of a joint project in the history of quantum physics of the *Max Planck Institut für Wissenschaftsgeschichte* and the *Fritz-Haber-Institut* in Berlin.

∗ Corresponding author. Address: Tate Laboratory of Physics, 116 Church St. SE, Minneapolis, MN 55455, USA, Email: janss011@umn.edu

mechanics as we know it today. It was formulated in terms of transitions between states and did not provide a mathematical representation of the individual states between which such transitions were supposed to take place. It allowed one to calculate transition probabilities, but gave no inkling of the odd new rules for the composition of probabilities in modern quantum mechanics. The wave functions introduced by Schrödinger (1926a) provided a first representation of quantum states. Born's (1926a,b) statistical interpretation of the wave function was an important step in the formulation of the new probability rules.

Generalizations of the Schrödinger wave function and Born's statistical interpretation of it were incorporated into matrix mechanics and the related q-number theory of Dirac (1925) through what came to be known as *transformation theory.* Independently of one another, Dirac and Jordan developed this new formalism in late 1926 and published it in early 1927 (Jordan, 1927a,b; Dirac, 1927).[1] In modern notation, which follows Dirac rather than Jordan, the central quantities in transformation theory are complex probability amplitudes $\langle a|b\rangle$, which determine the probability of finding the value a for some observable A after finding the value b for some observable B and at the same time govern the transition of a basis of eigenvectors of A to a basis of eigenvectors of B.

Transformation theory grew out of investigations of how to implement canonical transformations, familiar from classical mechanics and the old quantum theory, in matrix mechanics and q-number theory. With hindsight, elements of transformation theory can already be discerned in the *Dreimännerarbeit.* Moreover, in a paper on canonical transformations published in the fall of 1926, London (1926b) had found a preliminary version of transformation theory. In particular, he had shown how to fit Schrödinger's wave functions into the new formalism. However, it was left to Dirac and Jordan, drawing on helpful suggestions from Pauli and Heisenberg, to work out the formalism's statistical interpretation. Dirac and Jordan also used the new formalism to clarify the relations between the four different forms of quantum mechanics existing side by side at the time: matrix mechanics, wave mechanics, q-number theory, and the operator calculus of Born and Wiener (1926). In April 1927, Hilbert, Von Neumann, and Nordheim (1928) submitted an exposition of Jordan's version of transformation theory based on lectures by Hilbert in the preceding winter semester to the *Mathematische Annalen.* Their paper highlighted some of the mathematical problems facing transformation theory. These prob-

[1] Mehra and Rechenberg (2000, p. 72) quote from two letters clearly showing that Dirac and Jordan initially arrived at their results independently (Heisenberg to Jordan, 24 November 1926; Dirac to Jordan, 24 December 1926). Jordan and Dirac did meet in Copenhagen before Jordan wrote the second of his pair of papers, in which he cited Dirac's paper (Jordan, 1927b, p. 1).

lems provided an important stimulus for Von Neumann (1927) to develop the Hilbert-space formalism for quantum mechanics.

These developments are, of course, covered in general overviews of the history of quantum mechanics (see, e.g., Jammer, 1966, Ch. 6; Mehra and Rechenberg, 2000, Chs. I and III). They are also the topic of two interesting papers by Lacki (2000, 2004). Still, this episode in 1926–1927 has not received nearly as much attention from historians of quantum mechanics as, say, the discoveries of matrix and wave mechanics in 1925–1926 or the formulation of the uncertainty and complementarity principles in 1927 (cf. note 37 below). Given its central importance for the development of modern quantum mechanics, the emergence of transformation theory and its impact on subsequent formulations of quantum mechanics certainly warrant further historical study. Our paper should be seen as a modest contribution to this enterprise.

We focus on Jordan, who presented his version of transformation theory in a two-part paper entitled "On a new foundation (*neue Begründung*) of quantum mechanics" (Jordan, 1927a,b).[2] In these papers he drew on an earlier pair of papers on canonical transformations (Jordan, 1926a,b). The transformations that matter in matrix mechanics are the ones that preserve the form of the fundamental commutation relations and diagonalize the energy matrix. In his papers of 1926, Jordan showed how such transformations are the implementation in matrix mechanics of canonical transformations in classical mechanics. The transformations that matter in quantum mechanics are unitary transformations that map one orthonormal basis of Hilbert space to another, preserving inner products as needed for the theory's probability interpretation. Unitarity is not a natural property of canonical transformations. It thus became a hindrance for the formulation of quantum transformation theory that it had important roots in the analysis of canonical transformations.

The structure of our paper is as follows. In Section 2, we briefly review the role of canonical transformations in classical mechanics, the old quantum theory, and matrix mechanics. In Section 3, we preview elements of quantum transformation theory by showing how they are implicitly contained already in the *Dreimännerarbeit*. In Sections 4–6, which form the heart of our paper, we discuss the 1926 papers by Jordan and London on canonical transformations, avoiding overlap with Lacki's (2004) paper on this topic as much as possible. We highlight results in these papers that show that the transformation matrix implementing a canonical transformation in matrix mechanics is generally not unitary. In Section 7, we sketch the transition from Jordan's analysis of canonical transformations to the transformation theory of his *Neue Begründung*. In Section 8, we collect some observations about the limited impact of *Neue Begründung*. Finally, in Section 9, we briefly summarize our findings.

[2] A preliminary version of *Neue Begründung* appeared in late 1926 (Jordan, 1926c).

2 Canonical transformations in classical mechanics, the old quantum theory and matrix mechanics

Canonical transformations in classical physics are transformations of the position and conjugate momentum variables (q, p) that preserve the form of Hamilton's equations,

$$\dot{q} = \frac{\partial H(p,q)}{\partial p}, \quad \dot{p} = -\frac{\partial H(p,q)}{\partial q}. \tag{1}$$

For convenience, we assume that the system is one-dimensional and that the Hamiltonian $H(p, q)$ does not explicitly depend on time. The canonical transformation to new coordinates and momenta (Q, P) is given through a *generating function*, which is a function of one of the old and one of the new variables. For a generating function of the form $F(q, P)$, for instance, we find the equations for the canonical transformation $(q, p) \rightarrow (Q, P)$ by solving the equations

$$p = \frac{\partial F(q,P)}{\partial q}, \quad Q = \frac{\partial F(q,P)}{\partial P} \tag{2}$$

for $Q(q, p)$ and $P(q, p)$. One readily verifies that this transformation preserves the form of Hamilton's equations:[3]

$$\dot{Q} = \frac{\partial \hat{H}(P,Q)}{\partial P}, \quad \dot{P} = -\frac{\partial \hat{H}(P,Q)}{\partial Q}, \tag{3}$$

where the Hamiltonians $H(p, q)$ and $\hat{H}(P, Q)$ are numerically equal to one another but given by different functions of their respective arguments. One way to solve the equations of motion is to find a canonical transformation such that in terms of the new variables the Hamiltonian depends only on momentum, $\hat{H}(P, Q) = \hat{H}(P)$. Such variables are called *action-angle variables* and the standard notation for them is (J, w). The basic quantization condition of the old quantum theory of Bohr and Sommerfeld restricts the value of a set of action variables for the system under consideration to integral multiples of Planck's constant, $J = nh$. Canonical transformations to action-angle variables thus played a central role in the old quantum theory. With the help of them, the energy spectrum of the system under consideration could be found.

In matrix mechanics—as developed, after Heisenberg's (1925) ground-breaking *Umdeutung* paper, first by Born and Jordan (1925) and then by Born, Heisenberg, and Jordan (1926) in the *Dreimännerarbeit*—a canonical transformation is a transformation of the matrices (q, p) to new matrices (Q, P) that preserves the canonical commutation relations. In the one-dimensional case, this is the

[3] For elementary discussion, see, e.g., Duncan and Janssen (2007, Pt. 2, sec. 5.1).

relation:

$$[p, q] \equiv pq - qp = \frac{\hbar}{i}. \tag{4}$$

A canonical transformation is given by:

$$P = TpT^{-1}, \quad Q = TqT^{-1}, \quad \hat{H} = THT^{-1}, \tag{5}$$

where $\hat{H}$ is obtained by substituting TpT^{-1} for p and TqT^{-1} for q in the operator H given as a function p and q. One easily recognizes that this transformation preserves the form of the commutation relation (4): $[P, Q] = \hbar/i$. Solving the equations of motion in matrix mechanics boils down to finding a transformation matrix T such that the new Hamiltonian $\hat{H}$ is diagonal. The diagonal elements, $\hat{H}_{mm}$, then give the (discrete) energy spectrum.

We consider two questions raised by this procedure. First, what happens if the energy spectrum is (partly) continuous? This question is addressed in the *Dreimännerarbeit*, in a section written by Born, which anticipates elements of quantum transformation theory. We examine this issue in Section 3. Then there is the question of how the transformation matrices T in matrix mechanics are related to the generating functions F in classical mechanics and the old quantum theory. This relation was clarified in two pairs of papers, one by Jordan (1926a,b) and one by London (1926a,b). In the second of his two papers on the topic, London showed how Schrödinger's energy eigenfunctions fit into this formalism. We examine these issues in Sections 4–6.

3 The case of continuous energy spectra as treated in the *Dreimännerarbeit*

In the introduction to the *Dreimännerarbeit*, the authors briefly describe the basic strategy for solving concrete problems in matrix mechanics:

> It was found possible ... by the introduction of 'canonical transformations' to reduce the problem of integrating the equations of motion to a known mathematical formulation. From this theory of canonical transformations we were able to derive a perturbation theory (Ch. 1, sec. 4) which displays close similarity to classical perturbation theory (Van der Waerden, 1968, p. 321)

In the next sentence, they announce that they will also develop an alternative formalism, anticipating, as it turns out, elements of transformation theory:

> On the other hand we were able to trace a connection between quantum mechanics and the highly developed mathematical theory of quadratic forms of infinitely many variables (Ch. 3) (ibid.).

In an interview for the *Archive for History of Quantum Physics* (AHQP) in 1962, Kuhn asked Heisenberg why the authors of the *Dreimännerarbeit* had decided to present the theory along these parallel tracks.

> Kuhn: "Was there any discussion of the possibility of simply doing it all in Hermitian forms from the beginning? What was the reason for doing that much twice?"
>
> Heisenberg: "Well, that was more or less a historical reason . . . we felt that it might be [of] some help for other physicists to show in the one part of the paper how similar it is to those older things . . . Then we could show that there's a new technique which seems to be more powerful, which is actually identical with the other thing . . . We did realize that it was a kind of duplication of things, but still we felt that it may be useful for people—we don't know what will be the most convenient. We had the impression that there may be other still more convenient mathematical schemes behind it. One already felt that there is a great transformability in the whole thing, but we were not so far really to write down the general scheme" (AHQP, Heisenberg interview, session 7, pp. 18–19).

The alternative formalism can be found in sec. 3, "Continuous spectra," of Ch. 3, "Connection with the theory of eigenvalues of Hermitian forms," of the *Dreimännerarbeit*. In the introduction to this section, Born, who was responsible for this chapter, wrote:

> The mathematical theory of continuous spectra which occur for infinite quadratic forms has, starting from the fundamental investigations of Hilbert, explicitly been developed by Hellinger [1909] for the case of bounded quadratic forms. If we here permit ourselves to take over Hellinger's results to the unbounded forms which appear in our case, we feel ourselves to be justified by the fact that Hellinger's methods obviously conform exactly to the physical content of the problem posed (Van der Waerden, 1968, p. 358).

On the next page, Born gets down to business:

> For infinite quadratic forms, the case may arise that the form $\sum_{mn} H(mn)\, x_m x_n^*$ [4] cannot be converted into [the principal axes or diagonalized form] $\sum_n W_n y_n y_n^*$ by an orthogonal [read: unitary] transformation. We may then assume, in analogy with the results for bounded forms, that a representation with a continuous spectrum exists,
>
> $$\sum_{mn} H(mn) x_m x_n^* = \sum_n W_n y_n y_n^* + \int W(\varphi) y(\varphi) y^*(\varphi) d\varphi,$$
>
> in which the original variables are connected with new variables y_n, $y(\varphi)$ through an 'orthogonal transformation'; one only has to specify more clearly

[4] Note that the x_n's are just variables here, not components of vectors.

what is here understood by an orthogonal transformation (Van der Waerden, 1968, p. 359).

In modern terms, the 'orthogonal transformations' here are just a special case of a change of basis in Hilbert space. Put differently, this section of the *Dreimännerarbeit* can be seen as a special case of quantum transformation theory *avant la lettre*. To make this more transparent, we rephrase the argument in this section in modern language and then translate some of the results back into the language used here by Born and his co-authors.

Consider a system with an energy spectrum that is partly discrete and partly continuous. Let $\{|n\rangle, |E\rangle\}$ be a complete set of energy eigenstates that form a basis for the Hilbert space of this system. For the discrete part, we have

$$H|n\rangle = E_n|n\rangle, \text{ with } \langle n|m\rangle = \delta_{nm}, \tag{6}$$

where δ_{nm} is the Kronecker delta; for the continuous part, we have:

$$H|E\rangle = E|E\rangle, \text{ with } \langle E|E'\rangle = \mu(E)\delta(E - E'), \tag{7}$$

where $\delta(x)$ is the Dirac delta function and the normalization factor $\mu(E)$ is an arbitrary real positive function. The resolution of the identity in this basis is:[5]

$$1 = \sum_n |n\rangle\langle n| + \int \frac{dE}{\mu(E)}|E\rangle\langle E|. \tag{8}$$

Now, let $\{|k\rangle\}$ be a complete orthonormal basis (i.e., $\langle k|l\rangle = \delta_{kl}$) of eigenstates of an arbitrary Hermitian operator with a fully discrete spectrum acting on the same Hilbert space.[6] In this basis, the resolution of the identity is $1 = \sum_k |k\rangle\langle k|$. The components of $|n\rangle$ and $|E\rangle$ with respect to this fully discrete basis $\{|k\rangle\}$ are $\langle k|n\rangle$ and $\langle k|E\rangle$. The matrix elements of the Hamiltonian H relative to this basis are $\langle k|H|l\rangle$. In the notation of the *Dreimännerarbeit*, these quantities are written as x_{kn}, $x_k(E)$ and H_{kl}, respectively.

The components of $H|n\rangle$ and $H|E\rangle$ with respect to $\{|k\rangle\}$ are $\langle k|H|n\rangle$ and $\langle k|H|E\rangle$, respectively. The former can be written either as $\sum_l \langle k|H|l\rangle\langle l|n\rangle$ or as $E_n\langle k|n\rangle$. Setting these two expressions equal to one another and using the

[5] Born & Co. claim that $dE/\mu(E)$ (in our notation) gives the "a priori probability that the energy of the atom lies between [E and $E + dE$]" (Van der Waerden, 1968, p. 362). This interpretation is untenable. As we mentioned above, it is an arbitrary normalization factor.

[6] One can think, for instance, of a complete set of energy eigenstates of a one-dimensional harmonic oscillator spanning the Hilbert space, L_2, of square-integrable functions.

notation of the *Dreimännerarbeit*, we arrive at:

$$\sum_l H_{kl}\, x_{ln} = E_n\, x_{kn}, \tag{9}$$

or, in matrix notation, $\mathbf{H}\,\mathbf{x}_n = E_n\,\mathbf{x}_n$. The components $\langle k|H|E\rangle$ can likewise be written either as $\sum_l \langle k|H|l\rangle\langle l|E\rangle$ or as $E\langle k|E\rangle$ and we similarly arrive at:

$$\sum_l H_{kl}\, x_l(E) = E\, x_k(E), \tag{10}$$

or, in matrix notation, $\mathbf{H}\,\mathbf{x}(E) = E\,\mathbf{x}(E)$. Translating the relation

$$\delta_{kl} = \langle k|l\rangle = \sum_n \langle k|n\rangle\langle n|l\rangle + \int \frac{dE}{\mu(E)}\langle k|E\rangle\langle E|l\rangle, \tag{11}$$

where in the second step we used the resolution of the identity (8), into the notation of the *Dreimännerarbeit*, we arrive at:

$$\delta_{kl} = \sum_n x_{kn}\, x^*_{ln} + \int \frac{dE}{\mu(E)} x_k(E)\, x^*_k(E). \tag{12}$$

Eqs. (9), (10) and (12) can all be found in this section of the *Dreimännerarbeit* (Van der Waerden, 1968, p. 353, p. 359 (Eq. (28)), p. 361 (Eq. (35))). This confirms that the manipulations in modern language given above do capture the argument given in this section of the paper.

Looking at this argument from a modern point of view, we note that Born, Heisenberg, and Jordan—in keeping with the basic strategy of diagonalizing the Hamiltonian—restrict themselves to a basis of energy eigenstates. They did not realize yet that H can be replaced by any other Hermitian operator. So they express the position and momentum operators q and p as matrices relative to the basis of eigenstates of H. To make the connection to wave mechanics, one, conversely, needs to express H relative to a basis of eigenstates of q (or p). Given that Heisenberg had rejected the classical concept of position in his *Umdeutung* paper, it is only natural that he and his co-authors failed to consider this option. If one does use a basis of eigenstates of position, however, the equation $H|E\rangle = E|E\rangle$ turns into the time-independent Schrödinger equation:

$$H\left(q, \frac{\hbar}{i}\frac{d}{dq}\right)\langle q|E\rangle = E\langle q|E\rangle. \tag{13}$$

As Jordan (1927a, pp. 821–822) would recognize in *Neue Begründung* (cf. Eqs. (62)–(63) in Section 7 below), the quantities $\langle q|E\rangle$ are just the Schrödinger energy eigenfunctions $\varphi_E(q)$.

Discussing this section of the *Dreimännerarbeit* in the AHQP interview from which we already quoted above, Heisenberg remarked:

[I]t was all just bad luck that [Born] did not find at that time already the Schrödinger picture. He could have found the Schrödinger version especially where he treated this continuous spectrum. There it's only a short way to the Schrödinger picture (AHQP, Heisenberg interview, session 7, p. 18).

"[S]till," Heisenberg conceded, "you have to find it" (ibid.).

4 Jordan's two papers on the implementation of canonical transformations in matrix mechanics

In the abstract of the first of his two papers on canonical transformations in quantum mechanics (which, for Jordan at this point, meant matrix mechanics), Jordan (1926a) announced that he wanted to show that any canonical transformation can be implemented in quantum mechanics via a matrix T that takes the variables (p_k, q_k) to $(P_k, Q_k) = (Tp_kT^{-1}, Tq_kT^{-1})$.[7]

Asked about these two papers (Jordan, 1926a,b) in an interview for the AHQP with Kuhn in 1963, Jordan said:

> Canonical transformations in the sense of Hamilton-Jacobi were ... our daily bread in the preceding years, so to tie in the new results with those as closely as possible—that was something very natural for us to try (AHQP, Jordan interview, session 4, p. 11).

In his second paper on canonical transformations, Jordan (1926b) derived a general result that gives the relation between matrices implementing canonical transformations in matrix mechanics and generating functions implementing canonical transformations in classical mechanics. We present a streamlined proof of this result for the one-dimensional case. The generalization to n dimensions is straightforward.[8]

Problem: Given, classically, some canonical contact transformation, $(p, q) \rightarrow$

[7] Following the notation of the *Dreimännerarbeit*, Jordan still used the letter S for this matrix in this first paper; he switched to T in the second (now using S to denote the generating function). As we mentioned in Section 2, a transformation of the form $(P, Q) = (TpT^{-1}, TqT^{-1})$ preserves the commutation relations (4) and is thus by definition canonical. Jordan wanted to show that *all* canonical transformations are of this form. In other words, he wanted to prove what Lacki (2004, p. 321) calls the "converse statement."

[8] Lacki (2004, sec. 4) covers the derivation in the first paper (Jordan, 1926a) of a special case of the general result derived in the second.

(P, Q), generated by a function of the form [9]

$$F(p, Q) = \sum_n f_n(p) g_n(Q), \tag{14}$$

determine the transformation matrix $T(q, p)$ that gives the corresponding transformation of the quantum-mechanical operators, $P = TpT^{-1}$ and $Q = TqT^{-1}$.

Answer: The matrix T corresponding to the function F in Eq. (14) is given by:

$$T(q, p) = e^{\frac{i}{\hbar}\left\{(p,q) - \sum_n (f_n(p),\, g_n(q))\right\}}, \tag{15}$$

where the notation $(.\,,.)$ in the expression $(p, q) - \sum_n (f_n(p),\, g_n(q))$ in the exponential signals an ordering such that, when the exponential is expanded, all p's are put to the left of all q's in every term of the expansion. [10]

Proof: (1) We show that $TqT^{-1} = Q$. We rewrite the left-hand side as

$$TqT^{-1} = (Tq - qT + qT)T^{-1} = [T, q]T^{-1} + q. \tag{16}$$

The commutator is given by (Van der Waerden, 1968, p. 327)

$$[T, q] = \frac{\hbar}{i} \frac{\partial T}{\partial p}. \tag{17}$$

Inserting the expression for T on the right-hand side of Eq. (17), keeping track of the ordering, we find:

$$[T, q] = Tq - \sum_n f'_n(p) T g_n(q). \tag{18}$$

Substituting this result on the right-hand side of Eq. (16), we find that

$$q = \sum_n f'_n(p)\, T g_n(q) T^{-1}. \tag{19}$$

At the same time, q is given by

$$q = \frac{\partial F(p, Q)}{\partial p} = \sum_n f'_n(p) g_n(Q). \tag{20}$$

[9] In the classification of Goldstein *et al.* (2002, p. 373, table 9.1) this is generating function of type 3, $F_3(p, Q)$. The argument below can easily be adapted to generating functions that depend on (P, q), (q, Q), or (p, P), instead. In fact, Jordan considered a generating function of the form $\sum_n f_n(P) g_n(q)$ (Jordan, 1926b, p. 513, Eq. 1; cf. Lacki, 2004, p. 323, Eq. (15)).

[10] Jordan (1926b, p. 513, Eq. 3) defined this notation for the function $e^{\sum_n (x_n, y_n)}$. Introducing Jordan's result, London (1926b, p. 209) referred to this function as "the Pauli transcendent" (*die Paulische Transzendente*).

Comparing Eqs. (19) and (20), we conclude that

$$g_n(Q) = T g_n(q) T^{-1} = g_n(T q T^{-1}) \tag{21}$$

Assuming the g_n's are invertible functions, this implies $Q = TqT^{-1}$, which is the result we wanted to prove.

(2) We show that $TpT^{-1} = P$. Proceeding as before, we find

$$TpT^{-1} = [T, p]T^{-1} + p. \tag{22}$$

The commutator is given by (Van der Waerden, 1968, p. 327):

$$[T, p] = -\frac{\hbar}{i}\frac{\partial T}{\partial q} = -pT + \sum_n f_n(p) T g_n'(q). \tag{23}$$

Substituting this result on the right-hand side of Eq. (22), we find

$$TpT^{-1} = \sum_n f_n(p) T g_n'(q) T^{-1}. \tag{24}$$

Using Eq. (21) for the function g_n', we can rewrite this as

$$TpT^{-1} = \sum_n f_n(p) g_n'(Q). \tag{25}$$

Noticing that

$$\sum_n f_n(p) g_n'(Q) = \frac{\partial F(p, Q)}{\partial Q} = P, \tag{26}$$

we arrive at the result we wanted to prove.

5 London's two papers on the implementation of canonical transformations in matrix and wave mechanics

Like Jordan, Fritz London, assistant to the director of the *Technische Hochschule* in Stuttgart at the time (Lacki, 2004, p. 336, note 33), published two papers on canonical transformations in 1926 (London, 1926a,b). What strikes one immediately comparing the two is that the first has "quantum mechanics" in the title while the second has "wave mechanics" [*Undulationsmechanik*]. The first is indeed written in the context of matrix mechanics (informed by Dirac's q-number theory) and contains only a brief reference to Schrödinger toward the end (London, 1926a, p. 924). In the second, these considerations in the Göttingen tradition are connected with the formalism developed by Schrödinger. We focus on this aspect of London's contribution. For a broader discussion of these two papers we refer to Lacki (2004), who puts London rather than Jordan at the center of his analysis.

As he makes clear in the passage below, London recognized that the transformation matrices T act on Schrödinger wave functions: [11]

> Our starting point were transformations of operations ... This means the following: I have a mapping H in some domain, which takes every object x to another object y in this domain. In addition, I have another mapping T, which maps the entire domain, including its mapping H, onto a new domain: x goes to x^*, y to y^*. The "transformed mapping" H^* takes x^* to y^*. If instead of $x^* \rightarrow y^*$ one uses the detour $x^* \rightarrow x \rightarrow y \rightarrow y^*$, the well-known representation of a transformation of a transformation is recovered: $H^* = T^{-1}HT$ [this should be THT^{-1}].

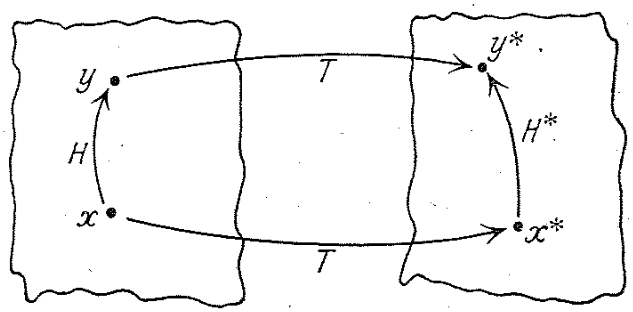

Fig. 1.

Fig. 1. Diagram from London (1926b, p. 198).

> From the beginning, this state of affairs has made the following a natural question in matrix mechanics: if the canonical transformations have the form of transformations of transformations, then on what things x does T act [*an welchen Dingen x greift dann T unmittelbar an*]? The answer is given by eq. (6c) [$T(Q, \partial/\partial Q)\Psi^*(Q) = \Psi(Q)$]: the things are Schrödinger's new state magnitudes Ψ, whose oscillating processes are described with the help of a sequence of eigenfunctions Ψ_k. The operator T maps this sequence term by term onto another sequence of eigenfunctions Ψ_k^* following (6c) (London, 1926b, pp. 197–198).

This passage is part of a section called "Canonical transformations as rotations in Hilbert space." [12] To make good on the promise in this title, London needed to show that the T's implementing the canonical transformations are

[11] This passage is also quoted and discussed by Lacki (2004, p. 334). Interestingly, London (1926b, p. 197) cites Born and Wiener (1926) and Ch. 3 of the *Dreimännerarbeit*, "Connection with the theory of eigenvalues of Hermitian forms," discussed in Section 3 above, as partial anticipations of his interpretation of the transformation matrix T.

[12] The term 'Hilbert space' as used nowadays was only introduced the following year by Von Neumann (1927) (Jammer, 1966, p. 315). London used the term loosely for the important special case of the space of square-integrable functions.

unitary. The 'proof' he offered for this claim, however, is circular.[13] As already indicated by the last sentence in the quotation above, London assumed that T maps an orthonormal basis of eigenfunctions to another orthonormal basis of eigenfunctions of the same Hilbert space. This is tantamount to assuming unitarity. That T is *not* necessarily unitary follows directly from the observation that Jordan's construction of T only determines it up to an arbitrary factor (see Section 4). If a particular T is unitary, then aT, with a an arbitrary numerical factor, clearly is not: $(aT)^\dagger = a^*T$ while $(aT)^{-1} = (1/a)T$.

In fact, one of London's own examples, that of a simple harmonic oscillator, involves transformation matrices that are not unitary (see Section 6.2). Moreover, Jordan (1926b) showed that, in general, the transformation matrix T for a point transformations is not unitary (see Section 6.1). Instead, it will in general be an isometry (i.e., a norm-preserving map) between different Hilbert spaces, L_2 and L_{2w}, with w a weight function. London touched on this problem in a footnote: "We leave it to the reader to take into account a density function [*Dichtefunktion*]" (London, 1926b, p. 198, note 1). Jordan showed how to get a unitary canonical point transformation by modifying the generating function out of which the transformation matrix T is constructed.

Yet, even though London was wrong to claim that T is always unitary, he deserves credit for recognizing the importance of unitary transformations for quantum mechanics. Moreover, in a prescient footnote added after finishing the paper, he drew attention to the relation between the rotations in Hilbert space emerging from his analysis and some "very general abstract papers ... on distributive functional operations" (London, 1926b, p. 199, note 2). As Jammer (1966) comments dramatically: "When London added this footnote ... he could hardly have been aware of its historical importance. It was the first reference to the future language of theoretical physics" (p. 298).

6 Two illustrative examples of canonical transformations in quantum mechanics

As we saw in Section 5, the transformation matrix T corresponding to the generating functions F constructed according to Jordan's general recipe (15) need not be unitary. In this section, drawing on the papers of Jordan and London discussed in Sections 3–4, we go over two examples of canonical transformations that illustrate this. The examples also illustrate how the general formalism discussed in the abstract in the preceding sections works in a few simple concrete cases.

[13] Jammer (1966, p. 297), however, accepts London's 'proof' at face value.

6.1 Point transformations

Consider a canonical transformation generated by a function of the form (Jordan 1926a, pp. 385–386; 1926b, pp. 514–515):[14]

$$F(q, P) = f(q)P. \tag{27}$$

Since (cf. Eq. (2))

$$Q = \frac{\partial F}{\partial P} = f(q), \quad p = \frac{\partial F}{\partial q} = f'(q)P, \tag{28}$$

we see that this function generates a point transformation, $Q = f(q)$, with the corresponding transformation of the conjugate momentum, $P = p/f'(q)$.

For a generating function of the general form $\sum_n f_n(q)\, g_n(P)$, the transformation matrix T that implements the corresponding canonical transformation in matrix mechanics is given by (cf. Eq. (15)):[15]

$$T(Q, P) = e^{\frac{i}{\hbar}\left\{-(Q,P)+\sum_n (f_n(Q),\, g_n(P))\right\}}, \tag{29}$$

where, as before, the notation $(.\,,.)$ signals an ordering such that, when the exponential is expanded, all Q's are put to the left of all P's in every term of the expansion.

For the specific generating function (27), this matrix becomes:

$$T(Q, P) = e^{\frac{i}{\hbar}(f(Q)-Q,\, P)} = e^{(\Delta Q,\, \partial/\partial Q)}, \tag{30}$$

where in the last step we introduced $\Delta Q \equiv f(Q) - Q$ and used that $P = (\hbar/i)\partial/\partial Q$.

Applying T to some wave function $\Psi(Q)$, expanding the exponential while keeping track of the ordering of Q's and P's, we find (London, 1926b, p. 210):

$$T\,\Psi(Q) = e^{(\Delta Q,\, d/dQ)}\Psi(Q) = \sum_n \frac{1}{n!}\Delta Q^n\, \frac{d^n\Psi(Q)}{dQ^n}. \tag{31}$$

In this last expression we recognize a Taylor series, so that we can write:

$$T\,\Psi(Q) = \Psi(Q + \Delta Q) = \Psi(f(Q)) = \hat{\Psi}(Q). \tag{32}$$

[14] So far we have been considering generating functions of the form $F(p, Q)$ (cf. Eq. (14)).

[15] In a note added in proof, London (1926b, p. 209) states Jordan's relation between generating function F and transformation matrix T in this form for the special case that the new coordinates (Q, P) are action-angle variables (w, J).

The first equality shows that T does indeed implement a point transformation.

The transformation matrix T is an isometry, a norm-preserving mapping, between two *different* Hilbert spaces:

$$T: \; \Psi(Q) \,\epsilon\, L_2(Q) \;\rightarrow\; \hat{\Psi}(Q) \,\epsilon\, L_{2w}(Q), \tag{33}$$

with the weight function $w = f'(Q)$. In other words, the norm of $\hat{\Psi}$ in L_2 is *not* equal to the norm of Ψ in L_2, but the norm of $\hat{\Psi}$ in L_{2w} *is* equal to the norm of Ψ in L_2. This can readily be verified:

$$\begin{aligned} |\hat{\Psi}|^2_{L_{2w}} &\equiv \int \hat{\Psi}^*(Q)\, \hat{\Psi}(Q)\, f'(Q)\, dQ \\ &= \int \Psi^*(f(Q))\, \Psi(f(Q))\, f'(Q)\, dQ \\ &= \int \Psi^*(\bar{Q})\, \Psi(\bar{Q})\, d\bar{Q} \equiv |\Psi|^2_{L_2}. \end{aligned} \tag{34}$$

The transformation matrix (30) is thus an example of a non-unitary transformation. As a consequence, the new momentum P is not Hermitian:

$$P = \frac{1}{f'(q)} p \;\rightarrow\; P^\dagger = p \frac{1}{f'(q)} \neq P. \tag{35}$$

In general, the new $(P, Q) = (TpT^{-1}, TqT^{-1})$ will be Hermitian if and only if the original (p, q) were Hermitian and T is unitary. One can make T unitary and P Hermitian in this case by modifying the generating function $f(q, P)$.

Jordan (1926b, p. 515) added a term to the generating function in Eq. (27):

$$F(q, P) = f(q)P + \frac{\hbar}{2i} \ln f'(q). \tag{36}$$

Instead of Eq. (28) we then get:

$$Q = \frac{\partial F}{\partial P} = f(q), \quad p = \frac{\partial F}{\partial q} = f'(q)P + \frac{\hbar}{2i} \frac{f''(q)}{f'(q)}. \tag{37}$$

We show that the last term is equal to $\frac{1}{2}[P, f'(q)]$. The new momentum is given by:

$$P = \frac{1}{f'(q)} \left(p - \frac{\hbar}{2i} \frac{f''(q)}{f'(q)} \right). \tag{38}$$

The second term commutes with $f'(q)$. The first term gives:

$$[P, f'(q)] = \frac{1}{f'(q)} \,[p, f'(q)] = \frac{\hbar}{i} \frac{f''(q)}{f'(q)}, \tag{39}$$

where we used that $[p,F] = (\hbar/i)\partial F/\partial q$ for any function $F(p,q)$ (Van der Waerden, 1968, p. 327). The expression for p in Eq. (37) can thus be rewritten as:

$$p = f'(q)P + \frac{1}{2}\left[P, f'(q)\right] = \frac{1}{2}\left(f'(q)P + Pf'(q)\right). \tag{40}$$

Since p and q are Hermitian, P must be Hermitian too.

Finally, we verify that the additional term in the generating function (36) ensures that the corresponding transformation matrix T is unitary. Inserting the modified generating function (36) into the general formula (29), we find:

$$T(Q,P) = e^{\frac{i}{\hbar}\left(\frac{\hbar}{2i}\ln f'(Q) - (Q,P) + (f(Q),P)\right)}. \tag{41}$$

So we need to add a factor $\sqrt{f'(Q)}$ to the expression for T in Eq. (30):

$$T(Q,P) = \sqrt{f'(Q)}e^{(\Delta Q,\,\partial/\partial Q)}. \tag{42}$$

The same factor needs to be added to Eq. (32) for the action of T on a wave function:

$$T\,\Psi(Q) = \sqrt{f'(Q)}\,\Psi(f(Q)) = \hat{\Psi}(Q). \tag{43}$$

The extra factor ensures that Ψ and $\hat{\Psi}$ have the same norm in the same Hilbert space L_2. There is no need anymore for the weight function w:

$$\begin{aligned}
|\hat{\Psi}|^2_{L_2} &\equiv \int \hat{\Psi}^*(Q)\,\hat{\Psi}(Q)\,dQ \\
&= \int \Psi^*(f(Q))\,\Psi(f(Q))\,f'(Q)\,dQ \\
&= \int \Psi^*(\bar{Q})\,\Psi(\bar{Q})\,d\bar{Q} \equiv |\Psi|^2_{L_2}.
\end{aligned} \tag{44}$$

This shows that the modified T of Eq. (42) corresponding to the modified generating function of Eq. (36) is indeed unitary.

6.2 Simple harmonic oscillator

The Hamiltonian for a simple harmonic oscillator of unit mass is:

$$H = \frac{p^2}{2} + \frac{1}{2}\omega^2 q^2. \tag{45}$$

Following London (1926b, pp. 204–205), we consider two successive canonical transformations, from (p,q) to (P,Q) and then from (P,Q) to action-angle variables (J, w).[16] The generating function for the first transformation is:[17]

$$F(q, P) = i\left(\frac{\omega}{2}q^2 + \sqrt{2\omega}qP + \frac{1}{2}P^2\right). \tag{46}$$

Solving

$$p = \frac{\partial F}{\partial q} = i(\omega q + \sqrt{2\omega}P), \quad Q = \frac{\partial F}{\partial P} = i(\sqrt{2\omega}q + P) \tag{47}$$

for $Q(q, p)$ and $P(q, p)$, we find

$$Q = \frac{1}{\sqrt{2\omega}}(p + i\omega q), \quad P = \frac{-i}{\sqrt{2\omega}}(p - i\omega q). \tag{48}$$

The quantum versions of these new variables are essentially the familiar raising and lowering operators $a = (i/\sqrt{2\hbar\omega})\,(p - i\omega q)$ and $a^\dagger = -(i/\sqrt{2\hbar\omega})\,(p + i\omega q)$:

$$P = -\sqrt{\hbar}\,a, \quad Q = i\sqrt{\hbar}\,a^\dagger. \tag{49}$$

To the best of our knowledge, this is the first time anybody introduced these operators, which would come to play such an important role in quantum mechanics.

The operators P and Q are clearly not Hermitian. Since p and q were Hermitian, it follows that the transformation matrix T that turns them into $P = TpT^{-1}$ and $Q = TqT^{-1}$ is not unitary (cf. Section 6.1).

The Hamiltonian (45) in the new coordinates is:

$$H = i\omega QP + \frac{1}{2}\hbar\omega. \tag{50}$$

With the help of Eq. (49) the quantum version can be rewritten in the familiar form $H = \hbar\omega\left(a^\dagger a + \frac{1}{2}\right)$. Substituting $(\hbar/i)\,d/dQ$ for P, we find

$$H = \hbar\omega\left(Q\frac{d}{dQ} + \frac{1}{2}\right). \tag{51}$$

The eigenfunctions of the time-independent Schrödinger equation in these coordinates,

$$H\,\hat{\Psi}(Q) = E\,\hat{\Psi}(Q), \tag{52}$$

are of the form

$$\hat{\Psi}_n(Q) \propto Q^n, \tag{53}$$

[16] Unfortunately, the angle variable w is hard to distinguish from the angular frequency ω.

[17] London (1926b, p. 204) uses the notation S_1 and S_2 for the two generating functions.

with eigenvalues $E_n = \hbar\omega\left(n + \frac{1}{2}\right)$. This shows that Q and $P \propto d/dQ$ are indeed raising and lowering operators, turning the n^{th} eigenfunction, Q^n, into the $(n+1)^{\text{th}}$ and $(n-1)^{\text{th}}$ eigenfunctions, Q^{n+1} and Q^{n-1}, respectively (modulo immaterial constants). To recover the eigenfunctions in the original coordinates (i.e., the familiar Hermite functions), one has to apply the inverse of the transformation matrix T constructed out of the generating function (46) according to Eq. (29) to the eigenfunctions, $\hat{\Psi}(Q)$, in the new coordinates.

Consider a further canonical transformation $(Q, P) \to (w, J)$ generated by the function

$$G(Q, J) = -i \ln(Q) J. \tag{54}$$

This gives

$$w = \frac{\partial G}{\partial J} = -i \ln Q, \quad P = \frac{\partial G}{\partial Q} = -i \frac{1}{Q} J. \tag{55}$$

Using the second of these equations, we can rewrite the Hamiltonian (50) as:

$$H = \omega J + \frac{1}{2}\hbar\omega. \tag{56}$$

Note that H only depends on J and not on w. As the notation suggests, these new coordinates are action-angle variables. Substituting $(\hbar/i)\, d/dw$ for J, we find

$$H = \hbar\omega \left(-i\frac{d}{dw} + \frac{1}{2}\right). \tag{57}$$

The eigenfunctions of the time-independent Schrödinger equation in these coordinates,

$$H\, \bar{\Psi}(w) = E\, \bar{\Psi}(w), \tag{58}$$

are of the form

$$\bar{\Psi}_n(w) \propto e^{inw}, \tag{59}$$

with, again, eigenvalues $E_n = \hbar\omega\left(n + \frac{1}{2}\right)$.

7 From Jordan's papers on canonical transformations to his papers on transformation theory

In his AHQP interview, Jordan recalled how he arrived at the *Neue Begründung* papers in which he presented his version of transformation theory (Jordan, 1926c, 1927a,b).[18] He, Born, and Heisenberg had initially thought that the

[18] The *Neue Begründung* papers are discussed extensively in session 3 of the interview (pp. 15–23). Kuhn returned to *Neue Begründung* in session 4 (p. 1, pp. 11–12) because he saw it as "one of the very very important stages of development" in which "so many things . . . come together" (p. 1). Talking to Oppenheimer about five months later, Kuhn called it "perhaps the most exciting of all the things in

only observable quantities in matrix mechanics were energies and transition probabilities. After Born's statistical interpretation of the wave function they realized position could also be measured. This raised the question:

> [W]hat kind of quantities in general can be regarded as somehow observable? There the idea was that in all generality some function of coordinates and momenta conceived of as q-numbers would be an observable quantity[[19]] As in the case of the Schrödinger eigenfunctions one would then also have to formulate the corresponding statistical relations. That, I believe, was roughly the train of thought that led to [*Neue Begründung*] (AHQP, Jordan interview, session 4, p. 12). [20]

As one would expect and as Kuhn confirms in the interview (session 3, p. 16), Jordan drew on his earlier work on canonical transformations to analyze the transformations $(q, p \to f(q, p), g(q, p))$ involved in the attempt to generalize the Schrödinger eigenfunctions and their statistical interpretation. [21] This is how Jordan put it later in 1927 in a semi-popular article in *Die Naturwissenschaften* on recent breakthroughs in quantum theory:

> [T]he position variables play a preferred role in Schrödinger's theory, in very unsatisfactory contrast to the great analytical generality achieved in q-number theory: with q-numbers one had the option to calculate with point coordinates as well as with arbitrary other canonical coordinates introduced through some "contact transformation" (Jordan, 1927d, p. 646). [22]

This same consideration, in more technical language, is given in the introduction of the first *Neue Begründung* paper. Jordan considers the time-independent Schrödinger equation for a particular choice of coordinates and asks how its solutions transform under canonical transformations to other coordinates. [23] The investigation of this question, Jordan continues, "led to the formulation of a very general formal network of connections among the quantum-mechanical

[the year 1926–1927 that Oppenheimer spent in Göttingen]" (AHQP, Oppenheimer interview, session 2, p. 15).

[19] Note that there is no restriction to functions producing Hermitian operators at this point.

[20] Jordan's remarks here are quoted more extensively by Mehra and Rechenberg (2000, p. 68), who devote a section (sec. I.4, pp. 55–72) to *Neue Begründung.*

[21] The transformation matrix constructed in Jordan's (1926b) paper on canonical transformations (see Section 4) returns in *Neue Begründung* (Jordan, 1927a, p. 830).

[22] Jordan cited both Dirac's work and his own and also acknowledged London's contribution: "the formal connections between Schrödinger's theory and the theory of canonical transformations had already been revealed in part by London [1926b]" (Jordan, 1927d, pp. 646–647).

[23] See Section 6.2 for an example of such a transformation in the special case of a simple harmonic oscillator.

laws, which contain the formal facts captured by the earlier formulations as special cases" (Jordan, 1927a, p. 810). Jordan's clarification of the "connection between the different representations of the theory" (ibid.) went well beyond Schrödinger's (1926b) famous equivalence proof of wave and matrix mechanics:[24]

> As is well known, quantum mechanics was developed in four different and autonomous forms: in addition to the original matrix theory, we have Born and Wiener's theory, wave mechanics, and q-number theory. How the last three formulations relate to matrix mechanics is known; every formulation gives the same end results as matrix mechanics in the area covered by the latter. However, there was essentially no intrinsic connection between the three later formulations; even the general proof was lacking that these formulations also lead to equivalent results in those areas where they go beyond matrix mechanics (Jordan, 1927a, p. 810).[25]

Transformation theory resolved these issues. To conclude their discussion of the subject, Mehra and Rechenberg (2000, p. 89) appropriately quote Oskar Klein, who told Kuhn in his AHQP interview that the papers on transformation theory by Jordan and Dirac "were regarded as the end of the fight between matrix and wave mechanics, because they covered the whole thing and showed that they were just different points of view" (AHQP, Klein interview, session 6, p. 2).[26]

To generalize the Schrödinger eigenfunctions Jordan also had to generalize their statistical interpretation. There he drew on ideas from Pauli—to what extent is not entirely clear.

> Following considerations by Born [1926a,b], Pauli proposed the following interpretation of the Schrödinger eigenfunctions. If $\varphi_n(q)$ is normalized, then $|\varphi_n(q)|^2 dq$ gives the probability that, if the system is in the state n, the coordinate q has a value between q and $q + dq$ (Jordan, 1927a, p. 811).

In a footnote, Jordan refers to a forthcoming paper by Pauli on gas degeneracy. The relevant passage occurs in a footnote in that paper (Pauli, 1927, p. 83), which was submitted 16 December 1926, two days before Jordan submitted his.

[24] For discussion of (the cogency of) Schrödinger's proof, see Muller (1997–1999) and Perovic (2008).
[25] Jordan took this passage almost verbatim from his preliminary note in the *Göttinger Nachrichten* (Jordan, 1926c, p. 161) and recycled it once more for his article in *Die Naturwissenschaften* (Jordan, 1927d, p. 646).
[26] Jammer (1966, p. 307) also emphasizes this element of unification, though in his view it was only fully achieved by Von Neumann (ibid., p. 316).

Jordan continues, repeating almost verbatim what he wrote in his preliminary note in the *Göttinger Nachrichten* (Jordan, 1926c, pp. 161–162; cf. note 25):

> Pauli considers the following generalization: Let q, β be two Hermitian quantum-mechanical quantities, which for convenience we assume to be continuous. Then there is always a function $\varphi(q, \beta)$, such that $|\varphi(q_0, \beta_0)|^2 dq$ measures the (conditional) probability that, for a given value β_0 of β, the quantity q has a value in the interval q_0, $q_0 + dq$. Pauli calls this function the probability amplitude (Jordan, 1927a, p. 811).

In the preliminary note (Jordan, 1926c, p. 162), though not in *Neue Begründung* itself, Jordan also attributes the related notion of the "interference of probabilities" (see below) to Pauli. [27] Commentators have looked in vain through Pauli's papers and correspondence for these central elements of Jordan's statistical interpretation of his formalism (Jammer, 1966, p. 305; Mehra and Rechenberg, 2000, p. 66; Lacki, 2004, p. 336, note 34). The closest thing they could find is a letter from Pauli to Heisenberg of 19 October 1926, which contains the special case of the generalization mentioned in the quotation above where q is the momentum and β is the energy (Pauli, 1979, Doc. 143, p. 348). [28]

In the AHQP interview with Jordan Kuhn also asked about Pauli's input in *Neue Begründung*. Jordan did not remember exactly what he got from Pauli but pointed out that they had had many conversations during this period. They went on vacation together on Neuwerk, an island on the German North Sea coast, and their paths crossed regularly in Göttingen, Hamburg, and Copenhagen (AHQP, Jordan interview, session 3, p. 15). [29] Maybe the ideas Jordan attributed to Pauli did all come out of those conversations; maybe Jordan was overly generous to Pauli. Presenting his transformation theory in *Die Naturwissenschaften* a few months later, Jordan (1927d, p. 647), after introducing the basic probabilistic ideas that went into it, only says: "as was suspected by Pauli" [*wie von Pauli vermutet wurde*]. But maybe Jordan simply felt that he did not have to be more specific in this case since this was a popular article.

Jordan presented the statistical interpretation of his formalism in the form of two postulates that have to be satisfied by the probability amplitude $\varphi(q, \beta)$.

[27] In the second part of *Neue Begründung*, Jordan (1927b, p. 19, note 1) once again stresses that Pauli deserves much of the credit for the statistical interpretation of his formalism.
[28] Jammer (1966, p. 305) quotes a paraphrase of the relevant passage in later reminiscences by Heisenberg (1960, p. 44).
[29] Quoted by Mehra and Rechenberg (2000, p. 66)

Postulate I: The function $\varphi(q, \beta)$ is independent of the mechanical nature (the Hamiltonian) of the system and is determined only by the kinematical relation between q and β.[30]

Postulate II: If $\psi(Q_0, q_0)$ is the probability amplitude for finding the value Q_0 for Q given $q = q_0$, then the amplitude $\Phi(Q_0, \beta_0)$ for a certain Q_0 given β_0 is:

$$\Phi(Q_0, \beta_0) = \int \psi(Q_0, q)\varphi(q, \beta)dq,$$

where the integration runs over all possible values of q.[31]

From a modern point of view, this second postulate becomes immediately obvious once we translate it into modern Dirac notation. Writing $\langle x|y\rangle$ for Jordan's probability amplitudes $\varphi(x, y)$, we can rewrite the expression for $\Phi(Q_0, \beta_0)$ as:

$$\langle Q_0|\beta_0\rangle = \int \langle Q_0|q\rangle\langle q|\beta_0\rangle dq. \tag{60}$$

One has to be careful, however, not to read too much into this notation. Neither Jordan nor Dirac—who used the notation (./.) at this point (Dirac, 1927, p. 631)[32]—thought of their probability amplitudes/transformation functions as inner products of vectors in a Hilbert space. That further step is due to Von Neumann (1927).

Jordan's second postulate captures an important new feature: "The circumstance that it is thus not the probabilities themselves but their amplitudes that follow the usual composition rule of the probability calculus can appropriately be called the *interference of probabilities*" (Jordan, 1927a, pp. 811–812).

In sec. 4 of the paper, Jordan (1927a, p. 821–822) gives Schrödinger-type equations for his probability amplitudes (see also Hilbert, Von Neumann, and Nordheim, 1928, sec. 10). Once again, it is easy to see how one arrives at these equations when one translates Jordan's equations into modern Dirac notation. Consider some canonical transformation $(q, p) \to (\alpha, \beta)$. The corresponding quantum operators are related via:

$$\hat{\alpha} = T\hat{q}T^{-1} = f(\hat{p}, \hat{q}), \quad \hat{\beta} = T\hat{p}T^{-1} = g(\hat{p}, \hat{q}). \tag{61}$$

[30] In the AHQP interview with Jordan, Kuhn emphasized the importance of this idea: "That seems to me such a big step. The terribly important step here is throwing the particular Hamiltonian function away and saying that the relationship is only in the kinematics" (session 3, p. 15). It is in this context that he raised the question about Pauli's involvement.

[31] There is a typo in the last equation: $\varphi(q, \beta)$ should be $\varphi(q, \beta_0)$.

[32] For an analysis of Dirac's path to his bra-ket notation, see Borrelli (2009).

The matrix element $\langle q|\hat{\beta}|\beta\rangle$ can be written either as $\beta\langle q|\beta\rangle$ or as $\langle q|g(\hat{p},\hat{q})|\beta\rangle$. Substituting $(\hbar/i)d/dq$ for $\hat{p}$ and q for $\hat{q}$, we arrive at the equation:

$$g\left(\frac{\hbar}{i}\frac{d}{dq},q\right)\langle q|\beta\rangle = \beta\langle q|\beta\rangle. \tag{62}$$

For the special case that $\hat{\beta}$ is the Hamiltonian $\hat{H}$, this is just the (time-independent) Schrödinger equation (cf. eq. (13)). The transformation functions/probability amplitudes $\langle q|E\rangle$—or $\varphi(q,E)$ in Jordan's notation—are just the Schrödinger energy eigenfunctions $\varphi_E(q)$:

$$H\left(\frac{\hbar}{i}\frac{d}{dq},q\right)\varphi_E(q) = E\varphi_E(q). \tag{63}$$

To conclude this section, we turn to sec. 2 of Jordan's paper, which we deliberately skipped above and in which Jordan presented the probability interpretation of his formalism in more detail. It is at this point that we see how the analysis of canonical transformation that had helped him arrive at the general formalism of transformation theory became something of an obstacle to its detailed elaboration. Rather than defining probabilities as the square of the norm of the amplitudes, Jordan assigned what he called a "supplementary amplitude" [*Ergänzungsamplitude*] $\psi(x,y)$ to the amplitude $\varphi(x,y)$ and defined the probability that the quantity q would have a value between x and $x+dx$ given the value y for the quantity β as

$$\varphi(x,y)\,\psi^*(x,y)\,dx \tag{64}$$

where "[t]he star * means that one has to take the function that is the complex conjugate of $\psi(x,y)$ for unchanged (i.e., not complex conjugated) arguments x, y" (Jordan, 1927a, p. 813). If the quantities q and β are Hermitian operators, the two amplitudes coincide: $\psi(x,y) = \varphi(x,y)$. Given the general framework of canonical transformations, Jordan, however, wanted to allow transformations to quantities that are *not* Hermitian operators (see Section 6, Eqs. (35) and (49) for examples of such quantities). The peculiar notion of an *Ergänzungsamplitude* was dropped in Part II of *Neue Begründung* (Jordan, 1927b, pp. 5–6). Jordan must have quickly realized that his attempt at a more general formalism does not work. Hilbert, Von Neumann, and Nordheim (1928) also do not mention the *Ergänzungsamplitude* in their paper on transformation although they closely follow Part I of *Neue Begründung* otherwise (Part II had not appeared yet when they wrote their paper).[33]

[33] Jordan's more general formalism only works for so-called *normal operators*, which are not necessarily Hermitian operators that commute with their adjoints, $[O,O^\dagger] = 0$. This follows from two results that we state here without proof. We are planning a more detailed paper on Jordan's *Neue Begründung* in which we will provide these

8 The impact of Jordan's version of transformation theory

As in the case of his pioneering work on field quantization, Jordan's version of transformation theory was largely overshadowed by Dirac's.[34] Dirac's vastly superior notation was undoubtedly an important factor in this. Despite his enthusiasm for *Neue Begründung* (see note 18), even Kuhn took Jordan to task for his notation:

> Kuhn: "In the first of the two papers on transformation theory, you use, if you will excuse me, a dreadful notation."
>
> Jordan: "This was just clumsiness [*nur eine Ungeschicktheit*]. I had rummaged [*herumgewühlt*] through these considerations extensively before I achieved clarity; the formulae were complicated and there was a large mass of formulae ... The notation found by Dirac was very beautiful and transparent" (AHQP, Jordan interview, session 3, p. 17).

A little later Jordan told Kuhn that he was unhappy with the presentation in the first *Neue Begründung* paper and that he largely wrote the second to give a "prettier and clearer" [*schöner und übersichtlicher*] exposition of the same material (session 3, p. 22).[35] It must be said, however, that the second paper is not much better on this score than the first.

Another reason for the limited appeal of *Neue Begründung*, at least to physicists, was its axiomatic structure. As recorded in the transcript of the AHQP interview, Jordan chuckled as he recalled Ehrenfest's reaction to *Neue Begründung*: "Well, since you wrote the paper axiomatically, that only means that one has to read it back to front" (session 3, p. 17, see also p. 19).[36] Ehrenfest was not alone. After emphasizing the importance of transforma-

---

proofs. Consider a canonical transformation $(\hat{p}, \hat{q}) \rightarrow (\hat{\alpha}, \hat{\beta})$. (1) If α and β are normal operators, then their spectra will in general be complex. However, in this very special case, $\psi^* = \bar{\varphi}$ (where the bar stands for ordinary complex conjugation), so $\varphi\,\psi^*$ gives the right probability density. (2) If α and β are *not* normal operators (e.g., raising and lowering operators for a simple harmonic oscillator), then $\varphi\,\psi^*$ gives the wrong probability density (e.g., for the coherent states which are eigenstates of the lowering operators).

[34] For the lack of recognition of Jordan's work on field quantization, see Duncan and Janssen (2008, pp. 642–643).

[35] Compare this statement to the following characterization of the relation between the two parts of *Neue Begründung*: "Part I contained the physical motivation and the outline of the theory, and in Part II, Jordan developed the mathematical details and methods" (Mehra and Rechenberg, 2000, p. 67). In the abstract of Part II, Jordan (1927b, p. 1) announces a "simplified and generalized presentation" of the theory presented in Part I.

[36] Quoted by Mehra and Rechenberg (2000, p. 69).

tion theory for the clarification of the relation between the different forms of quantum theory, Heisenberg told Kuhn in his AHQP interview:

> Heisenberg: "Jordan used this transformation theory for deriving what he called the axiomatics of quantum theory ... This I disliked intensely ... [It] was leading us a bit away from the physical content of quantum theory. I could not object to it, because after all it was correct physics, but I felt a bit uneasy about it."
>
> Kuhn: "Now does this mean that you felt happier with the Dirac paper? The results in the Dirac paper are almost identical, but the whole spirit of the paper is very different."
>
> Heisenberg: "Yes ... Dirac kept within the spirit of quantum theory while Jordan, together with Born, went into the spirit of the mathematicians" (AHQP, Heisenberg interview, session 11, pp. 7–8).[37]

Jordan's axiomatic approach was indeed more congenial to mathematicians. As emphasized by Lacki (2000, p. 295, p. 298) in his paper on early axiomatizations of quantum mechanics, Hilbert, Von Neumann, and Nordheim (1928) closely followed *Neue Begründung* in their paper on transformation theory.[38] Nordheim told Heilbron in his AHQP interview that he prepared Hilbert's lectures on quantum theory in the winter semester of 1926–1927 and that he was the main author of the paper resulting from them.[39] Despite his reliance on *Neue Begründung*, even Nordheim mentioned Dirac's paper first and called Jordan's formulation of transformation theory "very tortuous" (AHQP, Nordheim interview, p. 13).

In an annotated list of his publications prepared years later, Nordheim wrote about his paper with Hilbert and Von Neumann: "While ... not mathematically rigorous, this paper stimulated von Neumann to his later fundamental development of this topic" (AHQP, Nordheim folder). In a footnote at the end of their paper, Hilbert, Von Neumann, and Nordheim (1928, p. 30) already refer to a forthcoming paper by Von Neumann addressing some of the unresolved mathematical problems of transformation theory. In this paper, which

[37] Despite Heisenberg's aversion against Jordan's axiomatics, the latter's views—especially as expressed in his *Habilitationsvortrag* (Jordan, 1927c)—strongly influenced Heisenberg in the process of articulating the uncertainty principle (Beller, 1985; 1999, 91–95). Jammer (1966) writes that the uncertainty principle "had its origin in the Dirac-Jordan transformation theory" (p. 326) and, even more strongly, that "Heisenberg derived his principle from the Dirac-Jordan transformation theory" (p. 345).

[38] The paper was submitted in April 1927 and Jordan refers to it in the second part of *Neue Begründung* submitted in June 1927.

[39] For some quotations from the interview with Nordheim and discussion of the paper by Hilbert, Von Neumann, and Nordheim (1928), see Mehra and Rechenberg (2000, pp. 404–411).

ended up appearing in print before his paper with Hilbert and Nordheim, Von Neumann gives Jordan credit for turning the ideas of Born and Pauli about the probability interpretation of quantum mechanics into a "closed system" (Von Neumann, 1927, p. 2). In the same sentence, however, he goes on to say that this system is "facing serious mathematical objections" [*schwere mathematischen Bedenken ausgesetzt*] (ibid., pp. 2–3). Von Neumann's paper nonetheless shows that, for all its notational, organizational and mathematical shortcomings, Jordan's *Neue Begründung* played an important role in the development of quantum mechanics.

In the process of providing sound mathematical underpinnings of Jordan's transformation theory, Von Neumann introduced the idea of representing quantum-mechanical states by vectors or rays in Hilbert space. Vectors or rays in Hilbert space thus replaced Jordan's probability amplitudes—which, from the new point of view are inner products of such vectors—as the fundamental elements of the theory. Jordan continued to prefer thinking in terms of probability amplitudes.[40] In the second part of *Neue Begründung*, Jordan (1927b, p. 2, pp. 20–21) already distanced himself from Von Neumann's approach (cf. Lacki, 2000, p. 292, note 33). A decade later, in the preface to his texbook on quantum mechanics, Jordan accordingly described his and Dirac's formulation of transformation theory "as the pinnacle of the development of quantum mechanics" [*in deren Aufstellung die . . . Entwicklung der Quantummechanik gipfelte*](Jordan, 1936, p. VI). In the section devoted to transformation theory, he calls it "the most comprehensive and profound version of the quantum laws" [*die umfassendste und tiefste Fassung der Quantumgesetze*] (ibid., p. 171).

9 Conclusion

Three strands in the early development of quantum mechanics come together in the transformation theory of Jordan's *Neue Begründung* (Jordan, 1927a,b): the implementation of canonical transformations, the generalization of Born's probability interpretation, and the clarification of the relation between the four different forms of quantum theory co-existing in 1925–1926. In this paper we largely focused on the first of these strands. Our main conclusion is based on the observation that canonical transformations are not naturally restricted to unitary transformations of Hermitian operators (a restriction necessary in view of the probability interpretation of the quantum formalism). This ex-

[40] As Darrigol (1992) admonishes: "Perhaps modern-day interpreters of quantum mechanics should . . . remember that there exists a formulation of quantum mechanics without state vectors, and with transition amplitudes (transformations) only" (p. 344). Bub and Pitowsky (2007) seem to be heeding this admonition.

plains both why canonical transformations gave way to the picture of rotations in Hilbert space and why Jordan, still wedded to canonical transformations, initially but unsuccessfully tried to set up his transformation theory in a way that could accommodate observables represented by non-Hermitian operators.

Acknowledgments

This paper is based on our presentation at the second International Conference on the History of Quantum Physics (HQ2), Utrecht, July 14–17, 2008. The paper benefited from discussion in the quantum reading group at the *Max-Planck-Institut für Wissenschaftsgeschichte* (MPIWG) in Berlin, especially from comments by Arianna Borrelli and Christoph Lehner. We thank Michael Jordan and Jochen Heisenberg for permission to quote from the transcripts of the unpublished interviews with their fathers in the *Archive for History of Quantum Physics* (AHQP). We thank the staff at Walter Library at the University of Minnesota, where we consulted a copy of the AHQP. Michel Janssen gratefully acknowledges support of the MPIWG. The research of Anthony Duncan is supported in part by the National Science Foundation under grant PHY-0554660.

References

Beller, M. (1985). Pascual Jordan's influence on the discovery of Heisenberg's indeterminacy principle. *Archive for History of Exact Sciences* 33: 337–349.

Beller, M. (1999). *Quantum dialogue. The making of a revolution.* Chicago: University of Chicago Press.

Born, M. (1926a). Zur Quantenmechanik der Stoßvorgänge. *Zeitschrift für Physik* 37: 863–867.

Born, M. (1926b). Quantenmechanik der Stoßvorgänge. *Zeitschrift für Physik* 38: 803–827.

Born, M., Heisenberg, W., and Jordan, P. (1926). Zur Quantenmechanik II. *Zeitschrift für Physik* 35: 557–615. English translation in Van der Waerden (1968, pp. 321–385).

Born, M., and Jordan, P. (1925). Zur Quantenmechanik. *Zeitschrift für Physik* 34: 858–888. English translation of chs. 1–3 in Van der Waerden (1968, pp. 277–306).

Born, M., and Wiener, N. (1926), Eine neue Formulierung der Quantengesetze für periodische und nicht periodische Vorgänge. *Zeitschrift für Physik* 36: 174–187.

Borrelli, A. (2009). *Co-evolution of physics and mathematics: the case of Dirac's bra-ket notation.* Unpublished manuscript.

Bub, J., and Pitowsky, I. (2007). Two dogmas about quantum mechanics. In S. Saunders, J. Barrett, A. Kent, and D. Wallace (Eds.), *The Everett interpretation at the crossroads.* Oxford: Oxford University Press, to appear.

Darrigol, O. (1992). *From c-numbers to q-numbers: the classical analogy in the history of quantum theory.* Berkeley: University of California Press.

Dirac, P. A. M. (1925). The fundamental equations of quantum mechanics. *Proceedings of Royal Society of London. Series A* 109, 642–653. Reprinted in Van der Waerden (1968, pp. 307–320)

Dirac, P. A. M. (1927). The physical interpretation of the quantum dynamics. *Proceedings of the Royal Society of London. Series A* 113: 621–641.

Duncan, A., and Janssen, M. (2007). On the verge of *Umdeutung* in Minnesota: Van Vleck and the correspondence principle. 2 Pts. *Archive for History of Exact Sciences* 61: 553–624, 625–671.

Duncan, A., and Janssen, M. (2008). Pascual Jordan's resolution of the conundrum of the wave-particle duality of light. *Studies in History and Philosophy of Modern Physics* 39: 634–666.

Goldstein, H., Poole, C., and Safko, J. (2002). *Classical mechanics.* 3rd ed. San Francisco: Addison Wesley.

Heisenberg, W. (1925). Über die quantentheoretische Umdeutung kinematischer und mechanischer Beziehungen. *Zeitschrift für Physik* 33: 879–893. English translation in Van der Waerden (1968, pp. 261–276).

Heisenberg, W. (1960). Erinnerungen an die Zeit der Entwicklung der Quantenmechanik. Pp. 40–47 in: M. Fierz and V. F. Weisskopf (eds.), *Theoretical physics in the twentieth century. A memorial volume to Wolfgang Pauli.* New York: Interscience Publishers.

Hellinger, E. (1909). "Neue Begründung der Theorie quadratischer Formen von unendlich vielen Veränderlichen." *Journal für die reine und angewandte Mathematik* 136: 210–271.

Hilbert, D., Von Neumann, J., and Nordheim, L. (1928). Über die Grundlagen der Quantenmechanik. *Mathematische Annalen* 98: 1–30.

Jammer, M. (1966). *The conceptual development of quantum mechanics.* New York: McGraw-Hill.

Jordan, P. (1926a). Über kanonische Transformationen in der Quantenmechanik. *Zeitschrift für Physik* 37: 383–386.

Jordan, P. (1926b). Über kanonische Transformationen in der Quantenmechanik. II. *Zeitschrift für Physik* 38: 513–517.

Jordan, P. (1926c). Über eine neue Begründung der Quantenmechanik. *Königliche Gesellschaft der Wissenschaften zu Göttingen. Mathematisch-physikalische Klasse. Nachrichten* 161–169.

Jordan, P. (1927a). Über eine neue Begründung der Quantenmechanik. *Zeitschrift für Physik* 40: 809–838.

Jordan, P. (1927b). Über eine neue Begründung der Quantenmechanik II. *Zeitschrift für Physik* 44: 1–25.

Jordan, P. (1927c). Kausalität und Statistik in der modernen Physik. *Die Naturwissenschaften* 15: 105–110. Translated (by J. R. Oppenheimer) as:

Philosophical foundations of quantum theory. *Nature* 119: 566–569.
Jordan, P. (1927d). Die Entwicklung der neuen Quantenmechanik. 2 Pts. *Die Naturwissenschaften* 15: 614–623, 636–649.
Jordan, P. (1936). *Anschauliche Quantentheorie. Eine Einführung in die moderne Auffassung der Quantenerscheinungen.* Berlin: Julius Springer.
Lacki, J. (2000). The early axiomatizations of quantum mechanics: Jordan, von Neumann and the continuation of Hilbert's program. *Archive for History of Exact Sciences* 54: 279–318.
Lacki, J. (2004). The puzzle of canonical transformations in early quantum mechanics. *Studies in History and Philosophy of Modern Physics* 35: 317–344.
London, F. (1926a). Über die Jacobischen Transformationen der Quantenmechanik. *Zeitschrift für Physik* 37: 915–925.
London, F. (1926b). Winkelvariable und kanonische Transformationen in der Undulationsmechanik. *Zeitschrift für Physik* 40: 193–210.
Mehra, J., and Rechenberg, H. (2000). *The historical development of quantum theory.* Vol. 6. *The completion of quantum mechanics 1926–1941.* Part I. *The probability interpretation and the statistical transformation theory, the physical interpretation, and the empirical and mathematical foundations of quantum mechanics 1926–1932.* New York, Berlin: Springer.
Muller, F. A. (1997–1999). The equivalence myth of quantum mechanics. 2 Parts plus addendum. *Studies in History and Philosophy of Modern Physics* 28 (1997) 35–61, 219–247; 30 (1999) 543–545.
Pauli, W. (1927a). Über Gasentartung und Paramagnetismus. *Zeitschrift für Physik* 41: 81–102.
Pauli, W. (1979). *Scientific correspondence with Bohr, Einstein, Heisenberg a.o.* Vol. I: *1919–1929.* Edited by A. Hermann, K. von Meyenn, and V. F. Weisskopf. Berlin: Springer.
Perovic, S. (2008). Why were matrix mechanics and wave mechanics considered equivalent? *Studies in History and Philosophy of Modern Physics* 39: 444–461.
Schrödinger, E. (1926a). Quantisierung als Eigenwertproblem. 4 Pts. *Annalen der Physik* 79: 1–16, 17–55; 80: 437–490; 81: 100–139.
Schrödinger, E. (1926b). Über das Verhältnis der Heisenberg-Born-Jordanschen Quantenmechanik zu der meinen. *Annalen der Physik* 79: 734–756.
Van der Waerden, B. L., ed. (1968). *Sources of quantum mechanics.* New York: Dover.
Von Neumann, J. (1927). Mathematische Begründung der Quantenmechanik. *Königliche Gesellschaft der Wissenschaften zu Göttingen. Mathematisch-physikalische Klasse. Nachrichten* 1–57.