University of São Paulo
Faculty of Philosophy, Languages and Human Sciences
Graduate Degrees in Philosophy Division

The Thermodynamic Origins and Dynamical Foundations of Quantum Discontinuity

Diana Taschetto

São Paulo, 2024

Diana Taschetto

The Thermodynamic Origins and Dynamical Foundations of Quantum Discontinuity

PhD thesis submitted to the University of São Paulo in accordance with the requirements of the degree of Doctor in Philosophy in the Philosophy Department of the Faculty of Philosophy, Languages and Human Sciences.

Thesis approved. São Paulo, February 9th, 2024:

Nick Huggett
Supervisor

Osvaldo Pessoa Jr.
Supervisor

Don Howard

Professor
Dr. Don Howard

Professor
Dr. George Smith

DocuSigned by:
Décio krause
F6221A34BFBF49D...

Professor
Dr. Décio Krause

São Paulo
2024

Catalogação na Publicação
Serviço de Biblioteca e Documentação
Faculdade de Filosofia, Letras e Ciências Humanas da Universidade de São Paulo

T197t

Taschetto, Diana
The Thermodynamic Origins and Dynamical Foundations of Quantum Discontinuity / Diana Taschetto; orientador Osvaldo Pessoa Jr.; coorientador Nick Huggett - São Paulo, 2024.
139 f.

Tese (Doutorado)- Faculdade de Filosofia, Letras e Ciências Humanas da Universidade de São Paulo. Departamento de Filosofia. Área de concentração: Filosofia.

1. Foundations of Quantum Mechanics. 2. Quantum Discontinuity. 3. Philosophy of Physics. 4. Origin of the Quantum Theory. I. Pessoa Jr., Osvaldo, orient. II. Título.

Acknowledgments

The research that resulted in this work was funded by Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP), grants nº 2018/00275-3 and 2020/02969-2. I am very grateful to FAPESP for the privilege of its financial support. I am also grateful to my supervisor in Brazil, Osvaldo Pessoa Jr., for having accepted me as his graduate student at the University of São Paulo, and for giving me support, and intellectual freedom, during all these years. I thank Décio Krause. Décio's cheerful encouragement in my intellectual endeavors since I was a master's student meant a lot to me. I thank Guido Bacciagaluppi. His criticism of the draft of the third paper was extraordinarily helpful—and I am humbled and honored by the gift of the opportunity he gave me of working, as from February 2024, under his leadership as a post-doc at the University of Utrecht. I thank Harvey Brown. Harvey had faith in me. The two months I spent with him at the University of Oxford in 2022 had the most profound, electrifying, and enduring impact on me——the quality of my work improved significantly as a direct result.

I thank Don Howard. The multifarious nature of Don's influence on me is difficult to describe—but the readers of these pages will note it. It is also difficult to quantify the debt I owe him for his mentorship, encouragement, and the way he supported my work and career. Thank you, Don. Having met you was one of the most momentous things that happened to me during my PhD.

I thank George Smith. It is quite unlikely that I would have succeeded in finding the solution to the problem of the origin of the quantum theory without the discussions I had with George. His book with R. Seth, *Brownian Motion and Molecular Reality*, encoded the fundamental fact that set me off on the right track. I am deeply grateful for his careful readings of drafts of the first paper, for his patience, his criticism, and for helping me seeing the paper to its publication. George showed me, among other things, what it takes to undertake a proper investigation—he showed me what it takes to be a great scholar.

I thank Ricardo Correa da Silva. It is a rare privilege to have a collaborator of this caliber. Our collaboration, which resulted in the third paper of this thesis, has been, so far, the most exciting intellectual adventure of my life; the interesting results we have found would have remained hidden without his immense intelligence, mathematical insight, and tenacity. He is the author of the proof of the existence of the dual action.

My gratitude to my co-supervisor, Nick Huggett, is much too deep for words. Nick taught me the ropes of philosophy of physics; he walked me through all my intellectual struggles; he helped me find the way that corresponds to the essence of my talents. I revere this most extraordinary man as the most important figure of my intellectual life.

I owe personal thanks to my parents. It is from them that I inherited the single personal element that led me to my goal: my tenacity. I thank Sueli Borrely. Her motherly support was crucial in many moments of personal difficulties. I also thank David Hilbert and his wife, Katie, and also Anne Eaton, for their hospitality when I was in Chicago.

And there is of course Thales Borrely. This unspeakably generous, strong, wise, brilliant, superhero of a man. Without the sacrifices and love coming from him, my tower of strength, this thesis would not have been written. Thank you, Thales. My goodness, I love you.

Diana Taschetto

São Paulo, January 2024

It is beautiful to seek truth,
to propagate science and, as far as one can,
to add to it.

Lorentz, Rector's Speech, 1900

Abstract

Taschetto, D. The Thermodynamic Origins and Dynamical Foundations of Quantum Discontinuity. 2024. Doctoral Thesis—Faculty of Philosophy, Languages, and Human Sciences. Philosophy Department, University of São Paulo, São Paulo, 2024.

Doubts and misconceptions about the origin and justification of a physical theory's most fundamental concept are bound to spill over into doubts and misconceptions about the origin and foundation of the theory as a whole: such is the case, in point of fact, of the concept of quantum discontinuity of orthodox quantum mechanics. This work, which comprises three chapters, each of which is an independent paper, is a comprehensive investigation of the origin and nature of the concept of quantum discontinuity. The first paper, entitled "Rewriting the Quantum 'Revolution'," gives a novel account of the origin of quantum theory. It refutes, by direct counter-example, Thomas Kuhn's well-known thesis that revolutions in science do not happen on the basis of theory and experiment alone; the quantum revolution happened exactly in this way. The second paper, "On the Lawfulness of Quanta," is a sequel to "Rewriting the Quantum 'Revolution'." It evinces that and how the existence of the energy quanta is a principled result; it refutes the traditional view that the existence of the energy quanta was originally introduced in physics as an ad hoc hypothesis in a self-contradictory argument by showing that it is instead a necessary consequence of the laws of thermodynamics of black-bodies. This result affects our understanding of what is actually involved in the classical-to-quantum paradigm-shift, of scientific method, and of the specific nature of the quantum hypothesis which was carried over into contemporary physics in the course of the quantum revolution. The third paper, "The Dual Dynamical Foundation of Orthodox Quantum Mechanics," was written in collaboration with the mathematical physicist Dr. Ricardo Correa da Silva. It is a detailed critical study of the dynamical foundations of orthodox quantum mechanics; it contains several original results of mathematical, historical, and theoretical interest. The most important of them is the discovery of a dual action functional from which both of the dynamical postulates of von Neumann's quantum mechanics, process 1 and process 2, necessarily follow. It also contains the following results that have interest on their own: (i) a derivation of the canonical commutation relations from Matrix Mechanics; (ii) an analysis of Schrödinger's first derivation of the wave equation that improves on the understanding of it that is now current in the literature; (iii) a demonstration of how the canonical commutation relations follow from Wave Mechanics; (iv) a discussion, that improves on F. A. Muller's work, about the theoretical equivalence of Matrix and Wave Mechanics; (v) a revaluation of von Neumann's statement of the measurement problem. Broadly speaking, the three papers, taken together, drive home the lesson that the concept of quantum discontinuity stands on a more solid foundation than the scholarship on the foundations of quantum mechanics has hitherto acknowledged.

Keywords: Quantum Discontinuity; Origin of the Quantum Theory; Adiabatic Invariants; Dual Action.

Resumo

Taschetto, D. As origens termodinâmicas e os fundamentos dinâmicos da descontinuidade quântica. 2024. Tese de Doutorado – Faculdade de Filosofia, Letras e Ciências Humanas. Departamento de Filosofia, Universidade de São Paulo, São Paulo, 2024.

Dúvidas e equívocos sobre a origem e justificação do conceito mais fundamental de uma teoria física estão fadados a se transformar em dúvidas e equívocos sobre a origem e o fundamento da teoria como um todo: este é o caso do conceito de descontinuidade quântica da mecânica quântica ortodoxa. Este trabalho, que compreende três capítulos, cada um dos quais é um artigo independente, é uma investigação abrangente da origem e natureza do conceito de descontinuidade quântica. O primeiro artigo, intitulado "Reescrevendo a 'Revolução' Quântica", apresenta um novo relato da origem da teoria quântica. Refuta-se, por meio de um contraexemplo direto, a conhecida tese de Thomas Kuhn de que as revoluções na ciência não acontecem apenas com base na teoria e na experiência empírica; a revolução quântica aconteceu exatamente desta forma. O segundo artigo, "Sobre o Status de Lei dos Quanta", é uma sequência de "Reescrevendo a 'Revolução' Quântica". Evidencia-se que, e como, a existência dos quanta de energia é um resultado que se segue de princípios; refuta-se a visão tradicional de que a existência dos quanta de energia foi originalmente introduzida na física como uma hipótese ad hoc em um argumento autocontraditório mostrando que é, em vez disso, uma consequência necessária das leis da termodinâmica dos corpos negros. Tal resultado afeta a nossa compreensão do que está realmente envolvido na mudança do paradigma clássico para o quântico, do método científico e da natureza específica da hipótese quântica que foi importada para a física contemporânea no decurso da revolução quântica. O terceiro artigo, "O Fundamento Dinâmico Dual da Mecânica Quântica Ortodoxa", foi escrito em colaboração com o físico matemático Dr. Ricardo Correa da Silva. É um estudo crítico detalhado dos fundamentos dinâmicos da mecânica quântica ortodoxa; contém vários resultados originais de interesse matemático, histórico e teórico. O mais importante deles é a descoberta de um funcional de ação dual do qual ambos os postulados dinâmicos da mecânica quântica de von Neumann, processo 1 e processo 2, são consequência. O estudo contém também os seguintes resultados de interesse: (i) uma derivação das relações canônicas de comutação a partir da Mecânica Matricial; (ii) uma análise da primeira derivação de Schrödinger da equação de onda que refina o entendimento da literatura contemporânea sobre o assunto; (iii) uma demonstração de como as relações canônicas de comutação se seguem da Mecânica Ondulatória; (iv) uma discussão, que complementa o trabalho de F. A. Muller, sobre a equivalência teórica entre a Mecânica Matricial e a Ondulatória; (v) uma reavaliação da afirmação de von Neumann sobre o problema de medição. Em termos gerais, os três artigos, tomados em conjunto, levam à conclusão de que o conceito de descontinuidade quântica está assentado sobre uma base mais sólida do que a literatura sobre os fundamentos da mecânica quântica reconheceu até agora.

Palavras-chave: descontinuidade quântica; origem da teoria quântica; invariantes adiabáticos; ação dual.

List of Figures

List of Tables

Contents

1 Rewriting the Quantum "Revolution"

> The soundest fact may fail or prevail in the style of its telling.
>
> Ursulla L. Le Guin, cited in (BECKER, 2018)

For over a century now there has been almost complete consensus among the international community of scholars that the quantum theory started in 1900 when Max Planck introduced in physics the concept of energy quantization to derive the correct blackbody radiation law. The following quotations are representative:

1. Max von Laue, in his Memorial Address delivered at Planck's funeral in 1947, at the Albani Church in Göttingen:

   > On December 14, 1900, [. . .] before the German Physical Society, [Planck] was able to present the theoretic deduction of the law of radiation. *This was the birthday of the quantum theory.* This achievement will perpetuate his name forever. (LAUE, 1947, 10, my italics)

2. I. Duck and E. C. G. Sudarshan, in their "100 years of Planck's quantum," published in 2000:

   > When in the year 5000 people look back three thousand years to our era, we all hope that they will find some epochal events as myth-making for them as the Trojan War for us. Many events of such lasting significance are to be found among the achievements of our twentieth century scientific revolution in physics. The very first of these revolutionary events—certainly in time and arguably in significance—is Planck's invention of the energy quantum in 1900. (DUCK; SUDARSHAN, 2000, 1)

A survey at the physics library evinces the ubiquity of this view: The statement that Max Planck inaugurated the quantum theory in 1900 is found in textbooks on quantum mechanics,[1] in historical works,[2] in philosophical-foundational works,[3]

[1] Kramers (1923, 115), Bohm (1951, 18), Tomonaga (1962, 1), Peres (1993, 3), Liboff (2003, 33), Phillips (2003, 1), Griffiths (2004, 14), Serway, Moses e Moyer (2005, 1).

[2] Rosenfeld (1936), d'Abro (1952, 438, 457-463), Whittaker (1954, 81), Klein (1961, 32), Klein (1963, 81), Klein (1985, 12, 217-240), Waerden (1967, 1), Haar (1967, preface, 3-14), Peierls (1965, 199), Jammer (1966, 21), Hermann (1971, 1, 11), Kangro (1976), Weinberg (1977, 20), Mehra e Rechenberg (1982, 79), Gamow (1985, preface, 1, 6-22), Duck e Sudarshan (2000), Nauenberg (2016, 709-711).

[3] Reichenbach (1944, v), Bub (1974, 2), Fraassen (1991, 377), Omnês (1994, 5), Sklar (1995, 159), Beller (1999, 99, 131, 214).

in mathematical-foundational treatises,[4] in scientific biographies, [5] and—even more significantly, historically speaking—in the works of founding fathers.[6]

So when Thomas Kuhn adopted the course of maintaining, in his iconoclastic *Black-Body Theory and the Quantum Discontinuity* (KUHN, 1978), that Max Planck's work in 1900 was instead *classical*, many, if not all, scholars found his thesis incredible. Step-by-step, through painstaking archival study, Kuhn tried to establish continuity between Planck's early works on blackbody radiation and the 1900 derivation that turned the tide of history, trying, thereby, to establish the claim that if quantization was not in Planck's thinking before 1900—not, in fact, until after 1908—, then it is difficult to believe that it was *only* in 1900.[7] Such a conclusion was contrary to what was universally thought to be a firmly-established and well-known and understood historical matter-of-fact, "the point of view taken by physicists and by the historians of science who have written about these events," as Martin Klein put it in his book review (KLEIN; SHIMONY; PINCH, 1979, 430-431)—that Max Planck introduced the quantum.

Kuhn's "historiographic heresy," as he described it, awoke the criticism of other historians, who dissented in different ways. The catalogue of conflicting opinions is described in (DARRIGOL, 2001; GEARGART, 2002; BADINO, 2009)—positions range from reaffirming the orthodoxy (JOST, 1995) to denying the meaningfulness of the classical-quantum distinction (NEEDEL, 1980; BADINO, 2015). At present, the experts disagree. And behind difference in conclusion, according to (DARRIGOL, 2001, §2), is difference in approach: Darrigol says that Kuhn and himself "have most closely studied Planck's intricate considerations from the beginning of his program in 1894, through the quantum papers of 1900 and 1901, to later elaborations between 1906 and 1914" (DARRIGOL, 2001, 228); that indeterminists like Needell "have focused on Planck's expressed goals" (ibid); that orthodoxy subscribers "have given more weight to historiographical tradition" and "to the formal structure of Planck's reasonings of 1900 and 1901" (ibid). Martin Klein, scholar of the orthodox story, remarked that his misgivings with Kuhn's *Black-Body Theory* stem from Kuhn's lack of general consistency:

> By insisting so strongly on this radical revision of all previous accounts, Kuhn has put himself in a position that requires him to explain away part of the available evidence and to look at the rest of it from the standpoint of his central thesis. (KLEIN; SHIMONY; PINCH, 1979, 431)

4 Neumann (1932, 5), Mackey (1963, 59), Emch (1987, 209), Moretti (2017, 290).
5 Pais (1982b, 15, 27, 368-378), Moore (1989, 3, 118), Pais (1993, 80)
6 Bohr (1913), Bohr, Kramers e Slater (1924a, 787), Bohr (1925, 846), Born (1938, 7), Born (1953, 502), Sommerfeld e Bopp (1951, 85), Schrödinger (1952, 112), Heisenberg (1958, 31), Dirac (1963, 46).
7 Kuhn contends that "Planck did not publicly acknowledge the need for discontinuity until 1909" (KUHN, 1978, 140).

In the view of such disagreement it may conceivably be protested that the statement made at the outset of this essay, about the ubiquity of the orthodox story, is outdated, but this is not the case. The controversy is restricted to a small community of historians. Firmly entrenched in the scientific literature, and useful for pedagogical purposes, the standard story naturally lingers, in the lack of a shared and complete alternative to replace it.

But there can be no denying that how, and why, and when, the quantum was introduced in physics, is a *factual* question. The quantum revolution occurred exactly one way. So what the disagreement between the historians does show is that the origin of the quantum theory *is as yet an unresolved historical problem*.

Good. The foregoing discussion has now served its purpose by preparing the reader for the following remarks. Taking stock of just how the different approaches to Planck's work were articulated, and how they have conditioned different accounts of the origin of the quantum theory, it may well be asked what happens if, *unlike* the orthodoxy subscribers (which comprise, let me repeat, most people) we do not take tradition for granted, but *like* them (and unlike the other approaches) we focus attention on Planck's seminal 1900 and 1901 works. The answer is that by working in this way the orthodox story appears so implausible it is very difficult to believe it. But if we reach this conclusion, then we are required to explain away part of the available evidence, as Klein has correctly pointed out. We are required to *undo* the connections between the facts that support the orthodox story so that they can be connected differently. This burden I have tried to meet here, by taking an approach that is different from any in the existing literature.

History, let us carefully note, is not merely a list of facts. It is a list of *connected* facts. So, if it be admitted that the quantum revolution happened exactly one way, then like a jigsaw puzzle (the analogy serves us remarkably well) the right, complete picture can only emerge when all the relevant facts are found and interlocked the right way. Logical analysis of Planck's 1900-1901 works shall enable us to break the connections supporting the orthodox story so that we can see what are the mistakes, and where are the information gaps. This shall then enable us to look for further facts, the missing puzzle pieces that fit into the spots where we know that there's something missing. And we shall be able to find the answers, and therewith assemble the pieces together—and this is one of the key features of this paper—without detailed archival studies. A careful reading of the well-known papers and books by Planck and Einstein, while considering factual data about studies in black-body radiation, and the status of the probabilistic kinetic theory of heat, as of the beginning of the twentieth century, shall suffice.

In saying this I am committing a historiographic heresy of my own so I must proceed by making its implications clear.

There are two main points. First, it will turn out that the orthodox story can be refuted *in its own terms, by general considerations, which fact renders the alternative detailed historical accounts superfluous to this task*.[8] To appreciate what is involved in the second, let us first recall that historians, with a few exceptions, concentrate on meticulous projects on tightly circumscribed topics and periods; after Kuhn's *Black-Body Theory,* "the only major work available with an explicit historical theory of scientific revolutions in the background" (BüTTNER; RENN; SCHEMMEL, 2003, 37), historians have adopted the course of criticizing, or developing further, specific results of Kuhn's book, but issues about the structure of the quantum revolution have played little, if any, role in their controversy.[9] But as a philosopher I am not interested in the historical minutiae of the isolated parts: it is the complete assembled picture, "to grasp the idea of the whole correctly, and thence see the parts in their mutual relations," as Kant would put it,[10] that matters. And it will turn out that by taking the aforementioned route—this is the second point—*an account of the quantum revolution that differs from Kuhn's in all significant aspects will naturally and straightforwardly emerge*.

Let me say a few words about why I think this result is relevant for the philosophy of science. Kuhn would probably agree when I say that the question of the origin of the quantum theory transcends historiographic concerns. It affects our understanding of theory-change, of scientific method, and of the specific nature of the quantum hypothesis which was carried over into contemporary physics in the course of the revolution. All of this may have an impact on current research. Now, with all this in the back of our thinking, let us remember that Kuhn's trailblazing *The Structure of the Scientific Revolutions* (1962) is perhaps the most influential work in philosophy in the second half of the twentieth century; Kuhn did not, it is true, use the terminology of the *Structure* in *Black-Body Theory,* but it doesn't take much effort to recognize, in the latter, the explicative pattern of the former. But the best of evidence is of course Kuhn's own words, and in (KUHN, 1984), Kuhn explicitly remarked that the arguments of *Black-Body Theory* presuppose and reinforce the historiographic/philosophical views of the *Structure*. This is no trivial fact. For, once again, just how the classical gave way to the quantum provides data about the nature of knowledge and the methods employed in its pursuit. And for Kuhn, as is well-known, a choice of paradigm "can never be unequivocally settled by logic and experiment alone" (KUHN, 1962, 95).

[8] Let there be no misunderstanding: This is not to say that the study undertaken here will not conform to the elementary standards of historical scholarship. It is to say that we shall look for no further facts, go into no more details, than such as those that are true and sufficient to a correct theoretical understanding of Planck's and Einstein's seminal works in their original historical context.

[9] Exceptions are (BüTTNER; RENN; SCHEMMEL, 2003) and (BADINO, 2009). My account is however completely different from theirs.

[10] Preface to the *Critique of Practical Reason*, Cambridge University Press, 1997, p. 8.

Yet, in this case, *it was*.

For Kuhn, "the explanation [for scientific progress] must, in the final analysis, be psychological or sociological" (KUHN, 1970, 21). And perhaps because the *Structure* was so tremendously influential, it became so customary to embrace this view, and connect the facts pertaining to the history of science by muddling together arguments from authority, psychology, and sociology, that some of Kuhn's *Black-Body Theory* critics accused *Kuhn* of "trying too hard to make the thought of Planck coherent" (see, e.g., (GALISON, 1981)). Isn't that the ultimate irony? It seems that it has been generally taken for granted that there cannot be, as a matter of *principle*, a description of the classical-to-quantum paradigm shift based on the authority of rational argument and experiment alone. But this is not right. The account of the origin of the quantum theory that I put forward here constitutes a direct counter-example.[11]

One last remark. This paper is self-contained, so that previous acquaintance with the historians' works on Planck (the literature is indeed massive) is not needed. The only demand on the reader is perhaps a tolerance for gear-shifting. The road that leads to the understanding we seek shall not be easy—we have to be prepared for many twists and turns.[12]

So much for vindicating a philosopher's right to critically analyze the structure of the quantum revolution. Let us now come to grips with the subject proper. I start with a brief discussion of the facts that have motivated this study. The relevant places at which my interpretation of the facts differs from Klein's, Kuhn's, and others', will be, in what follows, carefully indicated.

1.1 The Historical Problem Suggested by Einstein's Critique to Planck

In work done jointly in the late 1850s R. Bunsen and G. Kirchhoff laid down the basis of what they took, correctly, to be a novel and powerful method of chemical analysis: spectroscopy. From investigating solar spectra and the spectral lines of various ele-

[11] I should remark here that since the orthodox story, and Kuhn's alternative to it, make full use of psychological and sociological considerations, the latter will be, on occasion, mentioned in the text—but only insofar as they are needed to show that such considerations are actually irrelevant for a correct understanding of how and why the quantum revolution was set in motion, and, furthermore, that Kuhn and the orthodox story, by giving them undue importance, have created decisive handicaps for a correct understanding of the origin of the quantum theory which have been as misleading in the search for historical truth as pernicious for philosophical judgment.

[12] The strategy adopted here is not unlike that used by W. V. Quine, for example, in his path-breaking "Two Dogmas of Empiricism." More than once we'll take a line of thought, take it to its last consequences, reach conclusions that are manifestly implausible, and start all over, until the final, correct picture emerges.

ments Kirchhoff observed that the relative intensity of the emission and absorption of spectral lines obeys "a general law, which seems to be of importance in several aspects [...]: it expresses a property of all bodies which is connected with the emission and absorption of heat and light" (KIRCHHOFF, 1860, 664). This law states that the ratio $\frac{\epsilon_\nu}{\alpha_\nu}$ between the power of emission ϵ_ν and absorption $\alpha_\nu$is a universal—i.e., substance-independent—function of the frequency ν of the radiation and of the temperature T of the substance considered. It reads $\frac{\epsilon_\nu}{\alpha_\nu} = K_\nu(T)$ and it is interpreted as the equation describing, or governing, matter and radiation at equilibrium. The question that then remains is that of determining the function $K_\nu(T)$ theoretically. This is what Max Planck succeeded in doing.

It will be recalled that by definition, a black-body is a body which absorbs all radiation falling on it. Hence, for a black-body, $\alpha_\nu = 1$ in the above formula and $\epsilon_\nu = K_\nu(T)$. Planck's black-body radiation law is the explicit form of Kirchhoff's function $K_\nu(T)$ for $\alpha_\nu = 1$. It reads

$$\varepsilon_\nu = \frac{8\pi}{c^3} h\nu^3 \frac{1}{e^{\frac{h\nu}{kT}} - 1}. \tag{1.1}$$

In his approach, since the material in the cavity is irrelevant, Planck considered a collection of oscillators (he called them "resonators") in static equilibrium with radiation in an enclosure. For the purposes of this study it is convenient to analyze Planck's derivation of (1.1) by dividing the premises into the different categories within physics:

Premise 1. (Thermodynamics). The entropy S, the internal energy U, and the temperature T of a system in thermodynamical equilibrium are related by

$$\frac{\partial S}{\partial U} = \frac{1}{T} \tag{1.2}$$

and also, since U is an extensive quantity, the equilibrium condition enables one to write, for the total energy $E = \epsilon P$ of the system, the average energy per oscillator as

$$U = \frac{E}{N} = \varepsilon \frac{P}{N}, \tag{1.3}$$

with $\epsilon \in \mathbb{R}^+$, $N, P \in \mathbb{N}$.

Premise 2. (Statistical Mechanics). The entropy S depends on the number W of possible states the system can be in as

$$S = k \ln W. \tag{1.4}$$

Premise 3. (Classical Electrodynamics). An ion oscillating about its equilibrium position emits and absorbs, on average, an equal amount of energy if its resonating frequency ν, its average energy U, and the energy density ϵ_ν of the radiation are related by

$$U = \frac{c^3}{8\pi\nu^2}\varepsilon_v \tag{1.5}$$

and Planck's own personal twists, namely

Hypothesis 1. The formula

$$W = \frac{(N+P-1)!}{(N-1)!P!} \tag{1.6}$$

for distributing P objects of energy ϵ amongst N oscillators, and finally

Hypothesis 2. A formal postulate discretizing the energy of the individual oscillators,

$$\epsilon = h\nu. \tag{1.7}$$

I must emphasize that Planck did not explicitly pin down and distinguish premises and hypotheses in this way. The next concise proof is therefore not to be found verbatim in his work—but everything just said is there.

Proof. Upon using Stirling's approximation in (1.6) and inserting the resulting equation into (1.3), we have the asymptotic formula

$$S = \frac{k}{N}\ln\left[\frac{N+P^{N+P}}{N^N P^P}\right] \tag{1.8}$$

for the average entropy per oscillator. Using then (1.3) to substitute for P into this expression, it follows that

$$S = k\ln\left(1+\frac{U}{\varepsilon}\right)\left(1+\frac{U}{\varepsilon}\right) - \frac{U}{\varepsilon}\ln\left(\frac{U}{\varepsilon}\right), \tag{1.9}$$

whence

$$\frac{\partial S}{\partial U} = \frac{k}{\varepsilon}\ln\left(1+\frac{\varepsilon}{U}\right). \tag{1.10}$$

By putting together the latter and (1.2), we have

$$U = \varepsilon\frac{1}{e^{\frac{\varepsilon}{RT}}-1}. \tag{1.11}$$

Inserting (1.7) into (1.11) then yields for the quantity U in (1.5) the desired equation (1.1).

□

The date Planck presented his proof to the German Physical Society—namely, December 14, 1900—pins down, it is generally claimed, the birth of the quantum theory.

It follows, therefore, trivially, that

to understand whence quantum theory is to understand Planck's demonstration of the black-body radiation law.

We shall faithfully follow this statement as a heuristic principle in our study. For if Planck did, or if he did not, introduce the quantum in his 1900 derivation of (1.1), *then either conclusion must follow straightforwardly from a correct understanding of his 1900 derivation alone.*

Good. To start with, consider the following remark that Albert Einstein made on Planck's seminal work.

> It should not be forgotten that the Planck radiation formula is incompatible with the theoretical foundation from which Planck started out. (EINSTEIN, 1909b, 363)

Einstein's suggestion is that Planck's formula and its proof should be considered separately. The partition is purely logical; the blackbody radiation law is true, but not the set of premises from which it was inferred. Hence, Planck's demonstration of (1.1)—the demonstration, that is, traditionally alleged to have inaugurated the quantum theory—is mistaken; the correct derivation of the blackbody radiation law is *not* Planck's.

Now let us think this over.

If Einstein is wrong, then his contention has only historical import—it need not detain us. But if Einstein is *right*, then the very deeply entrenched tradition of unqualifiedly identifying the origin of the quantum theory with Planck's derivation of the blackbody radiation law suddenly appears incomprehensible. For to say, "Planck's proof is wrong, it must be replaced," and to say, "it pins down the birth of quantum theory," is to say that Planck's proof is an untrustworthy source of knowledge, and yet it reveals the experimental truth of quantization upon which present-day physics rests.

And this seems uncomfortably *incoherent*.

And since an incorrect proof would hardly be taken seriously by the scientific community, it would surely appear to be a riddle, wrapped in a mystery, inside an enigma that thinkers trained for centuries to reason *with* Newtonian physics convinced themselves that though Planck's derivation of the blackbody radiation law is, as Einstein remarked, wrong, the proof is, at the same time, *so on target*, that it is time-tested classical physics that ought to be abandoned in favor of the new *a priori* quantum hypothesis that the mistaken proof conveyed.

In the view of the strong stress on empiricism, characteristic of the times, it seems utterly incredible that they would.

Many of the historians who have worked on this subject have thought this to be an unjustified concern. Planck's proof, it is maintained, is "heuristic," and hence the predicates "consistent" and "inconsistent" cannot be affirmed of it. Evidently, the difficulty, thereby, conveniently disappears. But to say that Planck's proof is neither right nor wrong, to arbitrarily exempt it from the requirements of exact mathematical demonstrability that proofs are set out to meet, sounds very much like desperation. Why assume, in the search for historical truth, that this is the case? Might it not be that hitherto overlooked facts have the information that will bring coherence to the situation? For Planck clearly intended it as a serious proof; he finished it by cross-checking the values he obtained with it for the constant $k = \frac{R}{N}$ (the molar gas constant divided by Avogadro's number), with values for Avogadro's number, Loschmidt's number, and "the quantum of electricity e" that were available in the literature *to show how reliable his derivation was*. Here's Planck, concluding his proof:

> If the theory is at all correct, all these relations should be not approximately, but absolutely, valid. The accuracy of the calculated number [for e] is thus essentially the same as that of the relatively worst known, the radiation constant k, and is thus much better than all determinations up to now. To test it by more direct methods should be both an important and a necessary task for further research. (PLANCK, 1900b, 89-89)

So Planck closes his 1900 presentation of his derivation of the blackbody radiation law asking for further independent measurements of N and e to confirm his derivation further still. I don't see how the man could have possibly been more serious about his work. This notwithstanding, historians have gone as far as to say, to avoid the aforementioned difficulties, that the renowned theoretician Max Planck, in deriving the blackbody radiation law, "did not really know what he was doing" ((DARRIGOL, 2001, 223); see also (KLEIN; SHIMONY; PINCH, 1979, 431) and (GALISON, 1981, 84)). Can this really be? Is there really not a better and instructive explanation out there? But we can—and must—go still further with our critical analysis; the statement that Planck's proof is "heuristic" also involves saying that Einstein's critique of Planck is unjustified. And since new ideas are invariably judged in the light of prevailing beliefs, which fact inevitably entails that they are invariably greatly resisted at first, the statement that Planck's proof is "heuristic" also involves saying that the quantum concept was an exception to the rule, and thereby allow it to be an inexplicable miraculous fact that the scientific community was ready to overthrow classical physics, and accept Planck's suggestion that the right blackbody radiation law *can only follow* if energy comes in chunks, without Planck having, well, *proved it*.

All of this strikes me as very implausible, but I will not pursue the argument further. I cannot disprove the historians' right to call Planck's proof "heuristic" and

Planck's reasoning, "mad" (PAIS, 1982b, 372). But I remark that this approach gives up the task of bringing coherence to this historical episode and gives us no new insights. Whereas if we grant Planck his intelligence, and his proof its formal status, we discover implications about the road that paved the way to the quantum revolution that may lead us to a new understanding of this all-too-important historical episode.

So let us make the effort to take seriously what is apparently a neat paradox: it seems that we cannot both maintain the orthodox story of the origin of the quantum, and admit, with Einstein, that Planck's derivation of the blackbody law is wrong, without contradicting ourselves, without being committed to the proposition that the proof is simultaneously right and wrong and so make our statement self-nullifying. Something, we feel, is missing in this picture; something has got to give. But to establish that Planck's proof is, in this sense, both right and wrong, a historical problem of the very deepest import must be brought to the fore. I will try to state this problem very precisely; the strategy will be this:

§1.2.1 to briefly review how outstanding historians have appraised Planck's 1900 work;

§1.2.2 to contrast their accounts with how Einstein's critique of Planck's work has reverberated in the physics community, and then see what results.

It will take us only a little effort to uncover historical oversights and conceptual entanglements that will indicate—but not yet prove—that both the orthodox story of the birth of the quantum theory, and Kuhn's alternative to it, cannot be squared with the facts. In §1.2.3, this conclusion will be shown to rest on a more solid basis of historical evidence; in §1.2.4, a clear and well-explained answer to the question of what, then, Planck did do in 1900 will be given, and the orthodox story, and Kuhn's alternative to it, then conclusively refuted. In §1.3, we will put all the pieces of the origin-of-the-quantum-theory jigsaw puzzle together, and, by showing that they fall into place objectively, we will claim to have explained, on the level of historical insight and philosophical analysis, that the structure of the quantum revolution was unequivocally settled by logic and experiment, so that Kuhn was demonstrably wrong in assuming that this is impossible.

1.2 Thinking Through Planck's Derivation of the Blackbody Radiation Law, and The Factual Ground of Einstein's Critique of It

1.2.1 The review of the standard historiographies is straightforward.

Olivier Darrigol's excellent, detailed, and very perceptive essay, "On the Historians' Disagreement Over the Meaning of Planck's Quantum," gives us exactly the information we are looking for. Consider his opening remarks:

> During the past twenty years, historians have disagreed over the meaning of the quanta which Max Planck introduced in his black-body theory of 1900. The source of this confusion is the publication, in 1978, of Thomas Kuhn's iconoclastic thesis that Planck did not mean his energy quanta to express a quantum discontinuity (KUHN, 1978). The aim of the present essay is a comparison of the opinions of various historians on this issue. (DARRIGOL, 2001, 219)

The reader will now want to recall that in 1918 Max Planck won the Nobel Prize in Physics "in recognition of the services he rendered to the advancement of Physics by his discovery of energy quanta." This fact alone may lead one to be overhasty, and readily conclude that Kuhn's claim cannot be right. Otherwise we would apparently have to accept it as a mysterious inexplicable fact that Planck won the Nobel Prize for something he did not do. But, we shall see, the plot is a good bit thicker; nevertheless, it is factually true that at first, very few were ready to accept Kuhn's contention. Darrigol systematizes the opposing opinions in this chart (DARRIGOL, 2001, 222):

Table 1 – Statements on Planck's quantum

<table>
<tr><th colspan="2">Discontinuity</th><th colspan="3">Indetermination</th><th>Continuity</th></tr>
<tr><td colspan="2">Textbooks
Klein 1961-7
Hund 1967
Jost 1995</td><td colspan="3">Kangro 1970
Needell 1980, 1988
Galison 1981</td><td>Kuhn 1978</td></tr>
<tr><td></td><td colspan="2">Rosenfeld 1936
Jammer 1966</td><td></td><td colspan="2">Planck 1920, 1943, 1948
Darrigol 1988, 1992</td></tr>
</table>

which is to be read, says Darrigol, as follows:

> According to the first category [left to right], in 1900 Planck introduced the idea that some microphysical entities (his resonators) could only have

> discrete energy values. According to the second, Planck did not know himself what the precise meaning of the energy quanta was. According to the third, Planck still believed that the energy of his resonators varied continuously, and he had no intention to revolutionize the laws of dynamics. (DARRIGOL, 2001, 221-222)

A few comments: (1) Darrigol emphasizes that Martin Klein—most prominent and influential expert on Planck, Einstein and Boltzmann—is the "most powerful discontinuist," and adds that "physics historians have usually agreed with Klein's view" (223). The first category is, then, by virtue of an unbalanced combination of tradition and Klein's authority, "the most abundantly represented" and "the typical textbook story of Planck's discovery belongs to it" (ibid). These considerations make it plain that the question of the interpretation of Planck's radiation theory is dominated, if not defined, by Klein and Kuhn; this is the reason Darrigol dubs the debate the "Klein-Kuhn controversy" (224). (2) With respect to the second category, let it suffice here to note that Allan Needell argues for "leav[ing] the meaning of the energy quanta in the dark" (223) and that Peter Galison "adopts a similar stance" (ibid). (3) The bottom boxes are meant to indicate an intermediate or nuanced or (in the case of Planck's own writing) unclear position.

And the fact that Planck himself is not put in the "discontinuity" box is, in and of itself, undoubtedly suggestive.

A couple of illustrations may now be instructive. Max Jammer writes that Planck was "not yet quite sure" whether or not his energy elements "expressed a fundamental innovation of profound physical significance" (JAMMER, 1966, 22) so his position is best catalogued between "discontinuity" and "indetermination." Darrigol places his own work between "indetermination" and "continuity," with the remark: "whereas Kuhn made Planck a persistent classical physicist, I believe with Needell that Planck left a non-classical behavior of his resonators open" (224). (4) Darrigol's analysis omits the work of Helge Kragh. Thinking this is a lack serious enough to remedy, I devote a few lines to correcting it. In 1999 (two years before Darrigol's essay came out) Kragh published an outstanding historical study, "Quantum Generations," in the preface to which he speaks of "Planck's introduction of the quantum discontinuity in 1900" (KRAGH, 2000, xi). Therefrom, one can only infer that Kragh subscribes to orthodoxy. But later it appears that this was a slip of the pen: In chapter five, §1, Kragh writes that "Planck did not really understand the energy quanta as a quantization of energy" and that he "did not recognize that the new radiation law necessitated a break with classical physics" (62). A year later Kragh published a short commemorative essay for the centenary of Planck's discovery, "Max Planck: The Reluctant Revolutionary," wherein he repeats the latter remarks from the book. So doubtless Kragh is in agreement with Kuhn. But it is noteworthy that Kragh does not make any reference to Kuhn, Klein, or any other historian, for that matter, in these works. In the very brief account that Kragh gives of Planck's discovery, the thesis that Planck did not introduce quantization in physics is *asserted*, disregarding the fact that the literature reflects considerable disagreement among prominent scholars in this respect, and leaving the question of how it was, then, that discontinuity historically came about, undecided.[13]

These considerations justify the conclusion, already stated above, that the story of the birth of the quantum theory—as the development of the theory itself—is not yet a closed file.

1.2.2 We now turn to Einstein's critique of Planck's derivation.

The first thing to be asked here is how the physics community reacted to Einstein's contention that Planck's proof is wrong. And the first indication that something is amiss is given by Einstein's 1949 remark: "Planck actually did find a derivation," he wrote in his *Autobiographical Notes*, "the imperfections of which remained at first hidden" (EINSTEIN, 1949, 39). Considering, now, the forty-year lapse between the statement expressing Einstein's critique (see above) and this later one, it may appear that Einstein thought that though he expressed his point clearly, widespread recognition came only slowly. But there are reasons to think that such an account of the facts cannot possibly be right.

Mind you, though Planck's work on blackbody radiation was met with criticism by experts in statistical mechanics such as P. Ehrenfest, J. Jeans, and H. A. Lorentz, the assertion that Planck's formula and its proof are logically incompatible was repeated, insofar as I could verify, by no other physicist. It may conceivably be protested that it appears quite incredible that such a severe critique, by an authority like Einstein, would have gone unremarked. True—but this is just the point. There is manifestly a question here, a question of how this can *be*, that suggests that a non-trivial consideration related to Planck's work was largely ignored by the scientific literature, a non-trivial consideration that seems *prima facie* incompatible with the over-100-years-old orthodox story of the origin of the quantum, and we should be very unscientific indeed not to look beneath it to make sure that we have not been hitherto overlooking important matters-of-fact in our search for historical truth.

[13] More recently, M. Badino has argued for a view that in Darrigol's schema is best characterized between "discontinuity" and "indetermination" (BADINO, 2015) Reviews of Kuhn's book are (KLEIN; SHIMONY; PINCH, 1979; NEEDEL, 1980; GALISON, 1981). Other reviews of the historians' controversy besides Darrigol's (whose treatment nevertheless rely on it) are, to repeat, (GEARGART, 2002) and (BADINO, 2009, §1).

So we raise the simple question as to what Einstein meant. The most straightforward method of demonstration will be to analyze the logical status that Planck's 1900 proof, presented in §1, can legitimately be claimed to have *vis-à-vis* another contention of Einstein's, a contention he formulated, though not in so many words himself, in a series of works from (EINSTEIN, 1905a) onward, and which we, for simplicity and future reference, indicate here in terms of a

Proposition (UVC). Planck's 1-3 premises imply another blackbody radiation law—not Planck's.

Proof. According to the equipartition theorem of statistical mechanics, an oscillating ion interacting with gas molecules must have average energy

$$U = kT. \tag{1.12}$$

Upon inserting (1.12) in (1.5), one readily concludes: mechanics and electrodynamics imply that dynamical equilibrium between matter and radiation in an enclosure holds if, and only if,

$$\varepsilon_\nu = \frac{8\pi}{c^3}\nu^2 kT. \tag{1.13}$$

□

This formula was first derived by Rayleigh in 1900. It expresses the fact Ehrenfest dubbed, in his (EHRENFEST, 1911), the "ultraviolet catastrophe." Observe: (1.13) leads to an infinite concentration of energy obtained from ϵ_ν by integration of ν over $\mathbb{R}^+$. This contradicts experience. Therefore, either premise (1.5), required by electromagnetic theory, or (1.12), required by mechanics, or both, are false.

"Objection!", the reader may now shout. "It is precisely to fix this problem that Planck introduced the quantum!" This, it is true, is the standard version of the standard story of the origin of the quantum, which has, please note, very ancient roots—here, for example, is the influential H. Poincaré, remarking in 1913:

> The old theories which seemed until recently able to account for all known phenomena have suddenly met with an unexpected check. Some modification has been seen to be necessary. A hypothesis has been suggested by M. Planck [...]. (Poincaré 1913, cited in Jeans 1914, 90)

The problem is that things cannot possibly have happened this way.

> *Was Planck aware of the radiation distribution law which Rayleigh had derived as a necessary consequence of classical physics?* Most authors answer this question in the affirmative and describe Planck's introduction of quanta as his response

> to the challenge of the "crisis" brought about by the disagreement between classical theory and experimental results and by the internal failure of classical theory as expressed in the "ultraviolet catastrophe." As a matter of fact, there was no such crisis [prior to 1900], or perhaps one should say there was no awareness of such a crisis. All of [Planck's] work on blackbody radiation prior to the summer of 1900 was done without benefit of the knowledge of just what classical physics did imply for this problem. (KLEIN, 1961, 461; Klein's italics)

We should hear warning bells about the orthodox story already.

It is noteworthy that Klein, "the most powerful discontinuist," claims that the quantum revolution *preceded* the crisis traditionally assumed to have triggered it. This nontrivial amendment to the orthodox story did not, insofar as I could verify, reach the textbooks. This is small wonder. For if Klein is right, then we no longer know how to talk about the quantum revolution.

If Klein is right, then we can no longer see *why* the quantum was introduced. The fact that the crisis in classical physics hit the physics community only later—for Planck himself explicitly stated, in his 1914 *The Theory of Heat Radiation,* that "the logical reason for the quantum theory is found in the fact that the Rayleigh-Jeans radiation formula does not agree with experiment" (PLANCK, 1914, 265, my italics)—only pushes the mystery a stage further back. And the remark that Planck, whose authority on radiation thermodynamics was acknowledged by everyone, was not cognizant of the implications of classical physics (which was, let us remember, all there was to physics) is simply crazy. Klein's explanation—that Planck did not know what he was doing—is far too easy, and it is the one that we cannot afford to accept.

So the plot thickens. The task of defining a problem is logically prior to that of finding out its solution—science, after all, *begins* with problems. Yet most, if not all, of the historians engaged in the "did Planck quantize or didn't he quantize" question (see §2.1) do not bring into focus the specific nature of the problem that the quantum was introduced to solve. The gap seems to be due to a difficulty in making sense of Klein's amendment: the historical record shows unequivocally, on the one hand, that the quantum was accepted as a way out of the ultraviolet catastrophe, and, on the other, that the ultraviolet catastrophe crisis hit the physics community only after the quantum had already been introduced. To the historian, the only solution to this predicament has seemed to be to deny that the quantum was introduced to solve the ultraviolet catastrophe; and therewith two possibilities. Either the quantum was introduced to solve a different problem than that for which it was accepted as a solution, or it was introduced to solve no problem at all.

Both options sound, of course, incredible. And since our present task is precisely to understand the implications of UVC we are now required to look beneath these facts to see if they indeed force us to adopt one of them.

They don't, as we shall see. That an ambiguous question is bound to lead to controversy, no one will deny; we have a better chance of doing something more significant rather than merely adding another story to the Klein-Kuhn controversy if we be quite clear about the work that the quantum was introduced to do. This will take us some time to set up, however. I hope the reader will bear with me through a lengthy scene-setting.

To start with, let's consider the Klein-Kuhn controversy again, but now we take a different angle from that of Darrigol. The fact that it was Kuhn, of all people, who first produced a historical account of the origin of the quantum to rival Klein's was by no means a coincidence. Kuhn published, let us remember, *The Structure of Scientific Revolutions* in the year immediately following Klein's study, 1962. The origin of the quantum is a scientific revolution *per excellence*; Klein was the most authoritative historian on Planck, Boltzmann, and Einstein, at the time that the *Structure* came out, and if Klein was right that Planck, concerned about no crisis, introduced the quantum in 1900, then Kuhn would have to admit that the fundamental thesis of the *Structure* was factually false. And so it would have been very natural for Kuhn to assume that the standard story *had* to be challenged.

And so we are driven to ask whether the facts in (KUHN, 1978) are not analyzed and organized to fit a certain preconceived pattern. It will appear, as we have remarked upon in §1, that they are. Recall that historical evidence tells us that the ultraviolet catastrophe hit the scientific community only after 1908, and that the quantum was already around at that time; given facts such as these, if one wanted to establish that scientific development follows the pattern of the *Structure*, then one would be required to show, first, that the crisis in classical physics, *and* the quantum revolution, both happened before 1908, and, consequently, that the intractable anomaly plaguing classical physics was *not* the ultraviolet catastrophe, but something else. And this is exactly what Kuhn argues for in *Black-Body Theory*.

Here's Kuhn, in his 1984 "Revisiting Planck":

> I have generally been well satisfied by the extent to which my narrative fit the developmental schema that *Structure* provides. *Black-Body Theory* is no exception. The start of the revolution that produced the old quantum theory is moved from the end of 1900 to 1906 [in *Black-Body Theory*]; most of the book deals with the period before the revolution occurred. The preceding crisis

> [i.e, the crisis that preceded the 1906 revolution], to the extent that there was one, resulted from the difficulties in reconciling Planck's derivation with the tenets of classical physics. (KUHN, 1984, 245)

Readers who know their *Structure* will immediately see a problem in the latter remark: according to the *Structure*, a single anomaly is no crisis-maker. But Kuhn's inconsistency is not the real rub here. The problem, as (BüTTNER; RENN; SCHEMMEL, 2003) correctly noted, is that there is no factual evidence whatever that Planck's derivation triggered a 1900-1905 crisis.

But then how in the world, we must now ask, can Kuhn seriously defend that it did? It is clear from Kuhn's writing that he was assuming, with Klein, that Rayleigh derived (1.13), as Klein had put it, "as a necessary a consequence of classical physics." And it is undeniable that if this is right, then Planck's 1900 derivation cannot be made fit into the Newtonian straightjacket; a tension, then, seems to have been bound to have ensued, triggering a 1900-1905 crisis.

And therefore it would seem that if all of this were indeed right, then the crisis defined by Kuhn *should* have ensued. But it is a matter of fact that it didn't. And so we have enough reason to suspect that the premise from which Kuhn derived this conclusion is not right.

Let's think about this carefully. It is noteworthy that both Klein and Kuhn, controversy notwithstanding, take for granted that (1.13) is a logical consequence of classical physics. And so it is comprehensible enough that further scholarly work in the field also took this as a given. And so, since historians could not find a logical explanation for the fact there was "no awareness" of the ultraviolet catastrophe in 1900, they explained it in cognitive (BüTTNER; RENN; SCHEMMEL, 2003; BADINO, 2015) and sociological (KUHN, 1978) terms. Indeed, if we think about it, we will quickly see that it was the assumption that (1.13) is "inevitable" that led many scholars to claim that Planck "did not know what he was doing" when he derived (1.1), that Planck "did not know the equipartition theorem," for, otherwise, "he surely would have considered its consequences" (HERMANN, 1971, 3), (KLEIN, 1961, 474). I maintain that this assumption is factually false. (1.13) was *not* a necessary consequence of classical physics. Not in 1900. But let there be no misunderstanding: when I say this, I do not mean there was no "awareness" of it in 1900. I mean it literally. And so there *is* a logical explanation for this matter, and the facts that lead to it are not at all far to seek.

First recall that between 1896 and 1900, the correct blackbody radiation formula was universally believed to be Wien's:

$$\varepsilon_{\nu} = \alpha \nu^{3} e^{-\frac{\beta v}{T}}. \tag{1.14}$$

This fact alone should enable us to dismiss Kuhn's contention that a crisis emerged in classical physics because classically only (1.13) can follow; consider how Planck opened his seminal (PLANCK, 1901):

> The recent spectral measurements of O. Lummer and E. Pringsheim and even more striking those of H. Rubens and F. Kurlbaum, both confirming more recent results obtained by H. Beckmann, would discover that the law of the energy distribution in the normal spectrum first stated by W. Wien from the molecular-kinetic consideration and later by me from the theory of electromagnetic radiation is not universally correct. (PLANCK, 1901, 553, my italics)

Wien's law carried no crisis in its train, and Wien and Planck both derived it using different models (this fact shall be important for us later on).[14] And the reader should note in this connection that it was in the course of *this* derivation (PLANCK, 1899, §23-25) that Planck introduced the constants later defined as k and h ($\alpha = 6.10 \times 10^{-56} \cong \frac{8\pi}{c^3}h, \beta = 4.866 \times 10^{11} = \frac{h}{k}$); it was, indeed, after deriving *Wien's law,* in (PLANCK, 1899, §26) with *these* constants, that Planck proposed his famous system of natural units. And since no quantum hypothesis was involved in Planck's 1899 derivation of Wien's law, it must be admitted that the appearance of the constant h, in and of itself, does not imply the quantization of anything.

Good. To continue, then, with our story: it was only when, and only because, different groups of experimentalists found, in late 1899 and early 1900, that Wien's law fails in the far infrared that its epistemic status changed from an absolute to an approximate law.[15] Motivated, then, by this fact, Lord Rayleigh—a fierce defender of Boltzmann's theory—proposed, in June 1900, his famous (1.13) in the often-quoted but otherwise little known "Remarks Upon the Law of Complete Radiation." In this paper, Rayleigh says explicitly that he is perfectly aware of "the difficulties which attend to the Boltzmann-Maxwell doctrine of the partition of energy," and that though "the doctrine fails in general" (!), it may nevertheless be possible to apply it "to the graver modes" of radiation (long wavelengths). And he hastens to add: *not theory*, but *the evidence of experiments*, suggested this possibility out to him. For H. Rubens, experimentalist at the *Physikalisch Technische Reichsanstalt*, had reported that for temperatures above 1000 degrees, there is but little increase of radiation. And so Rayleigh, "difficulties which attend to the Boltzmann-Maxwell doctrine of the partition of energy" notwithstanding, proposed (1.13) as an *effective* law *possibly* valid *only* for "the graver modes" of radiation.

[14] See (WIEN, 1986) and (PLANCK, 1899).

[15] One may ask whether this failure triggered a crisis in radiation thermodynamics; it did not. In Paul Drude's *The Theory of Optics*, the most authoritative book of optics as of 1900, the deviations are remarked upon in a footnote (DRUDE, 1900, 526).

And so it cannot be denied that it does seem to be a sort of *fact* that in no way, shape, or form, (1.13) was, as a *general law*, part of classical physics as of 1900. The specifics of Rayleigh's derivation of (1.13), and of Wien's and Planck's derivations of (1.14), both testify to this fact.

And so if (1.13) became a necessary consequence of classical physics years later, as we know that it did,[16] then it cannot *be* that physicists were making the wrong assumptions before. If, years later, (1.13) was universally accepted as *the* classical prediction, then later—and not before—the principles of classical physics forced physicists to this conclusion.

Let's see how this can be. What is behind the statement that classically only (1.13) can follow is the fact, remarked *e.g.* by Einstein in (EINSTEIN, 1909b, 186, footnote 3), that the equipartition theorem is an inevitable consequence of the kinetic theory of heat. And that it is, is surely undisputable. But it must be carefully noted and appreciated that just like today, no one will contend that it ineluctably follows from contemporary physics that we live in a 11-dimension universe, for this fact is entailed by string theory, the validity of which has not yet been established, in a similar way, as of 1900, *nobody* contended that the equipartition theorem ineluctably follows from classical physics, for, at that time, the validity of the kinetic theory of heat had not yet been established. And that is that.

Mind you, everybody knew about the equipartition theorem.[17] I find it impossible to believe, with many historians, that Planck, as of 1900, did not know about it. And everybody knew, just as well, that it "fails in general," as Rayleigh remarked. The equipartition theorem implied wrong values for the specific heat of gases and it obviously implied the wrong spectrum for blackbody radiation, just as Newton's laws implied a wrong value for the perihelion of Mercury. And nobody knew what to do about the equipartition theorem, just as nobody knew what to do about the 43 arc-seconds per century residual discrepancy in the precession of the perihelion of Mercury. But, unlike the crisis in Newtonian gravity, the crisis enforced by the equipartition theorem waited in abeyance until after 1908. Not because of human confusion or sociological forces, as historians have made it appear, but because of evidence—or lack thereof—*from experiments*;[18] for the probabilistic kinetic theory of heat earned its place

16 Recall the quote from Planck's *The Theory of Heat Radiation* given a few pages back. Consider also the remark about the undisputed validity of the Rayleigh-Jeans law as a general law made at the invitation letter for the 1911 Solvay Conference in Physics quoted in footnote 26.

17 I am indebted to Don Howard for calling my attention to this fact.

18 Kuhn, it must be remembered, defines "crisis" in psychological terms: according to the *Structure,* a crisis arises when the *confidence* of scientists is lost in the ability of the paradigm to solve worrying puzzles he calls "anomalies." But the discrepancy in the perihelion of Mercury never shook the confidence of physicists in Newtonian gravity one iota. So when I say that orbital astronomy was in a state of crisis due to the discrepancy in the perihelion of Mercury, I mean it objectively, in the

in physics only after indisputable evidence in favor of its most fundamental postulate, namely, the Gaussian distribution of displacement over time and velocity of discrete granular particles, came from Perrin's experiments after 1908. That this is so was demonstrated recently with irrefutable consistency by Smith & Seth in their masterful *Brownian Motion and Molecular Reality: A Study in Theory-Mediated Measurement* (SMITH; SETH, 2020). They make a very strong case to the effect of showing that it was the authority of evidence, stemming from Perrin's experiments, that gave the probabilistic kinetic theory of heat its scientific standing.

And therewith—I hasten to add—a crisis. Reader, please note: this is why, and when, and how, the ultraviolet catastrophe hit physics. The equipartition theorem comes along with the probabilistic kinetic theory of heat.[19] Carrying in its train all of its intractable consequences.

This is my understanding of why the crisis manifest in the ultraviolet catastrophe established itself in physics only after 1908, as historical record shows that it did. If I am right, then the psychological, sociological, and cognitive reasons heretofore adduced by historians are only of secondary importance.

And now it cannot be denied that all these considerations have a great deal of powerful cumulative suggestiveness.

For what Perrin did in 1908-1911 (or this, at least, is one way of putting it) is to confirm Einstein's 1905 predictions for the Gaussian character of Brownian motion. The Einstein who, from the same year on, kept on insisting, in a series of papers, (EINSTEIN, 1905a; EINSTEIN, 1905c; EINSTEIN, 1906; EINSTEIN, 1907; EINSTEIN, 1909b; EINSTEIN, 1909a), that current theories of radiation and matter *inexorably entail* (1.13), which fact shows (he claimed) that the foundations of physics must be considered afresh. The same Einstein who contended, as we have seen, in 1909 that Planck's 1900

sense that physicists were unable to find a robust source for the 43 arc-seconds per century residual discrepancy. When I speak of the ultraviolet catastrophe crisis I refer to the fact that theoretically (1.13) could not but be true, and yet it is evidently empirically false. I shall always use the term "crisis" in this objective sense here.

[19] I should emphasize that Perrin's granule-rotation experiment provided a theory-mediated measured for N that was meant to test the equipartition assumption *specifically*; in that case, N represented the number of granules required to have their total rotational kinetic energy matching their mean translational energy, kT, as required by the equipartition theorem (see (SMITH; SETH, 2020, §4.7), for a detailed discussion). I quote Perrin's conclusion without comment:

> The agreement is remarkable, if one thinks of the difficulties of measurement and the complete uncertainty which a priori surrounded even the order of magnitude of the rotation. The granules utilized for these measurements were about 100,000 times heavier than the granules of gamboge first studied. So, *the equipartition of energy is established throughout this great interval*. Incidentally, its verification for the rotations is an experimental confirmation of the reasoning from the kinetic theory which has enabled the ratio C/c of the specific heats of gases to be predicted. (Perrin 1910, quoted in Smith & Seth 2020, 159)

proof is *incorrect*, and who stated the content of UVC, which, at last, we seem to be in possession of the relevant facts to properly analyze.

It is easy to see that UVC trivially entails that the conclusion of Planck's 1900 derivation is "incompatible with the theoretical foundations from which Planck started out." For if premises 1-3 entail the Rayleigh-Jeans law, while the whole set of premises entail the incompatible Planck's law, then they cannot but be logically inconsistent.

But then how in the world could Planck have managed to get his minutely-confirmed-by-experiment formula out of his derivation?

Forget about the equipartition theorem for a moment. Let's face Planck's 1900 derivation head-on. Consider equation (1.5) again. Planck derived it following Maxwell's equations, assuming that the energy of the resonators varies continuously. But the assumption that energy is partitioned among the N oscillators in a way that it can only vary by jumps of (1.7) contradicts the premise of electrodynamics that radiation energy is spread throughout the whole of space, and the basic premise of mechanics according to which the states of ponderable matter vary continuously. Hence, to bring (1.5) and (1.7) together is equivalent to bringing P and $\neg P$ together. It is a flat-out contradiction. And from a contradiction, *anything* can be shown. Trivially, therefore, the right blackbody radiation law.

1.2.3 But it seems that there is something not altogether kosher about this conclusion.

Remember that what led us into this discussion of Einstein's contention that Planck's derivation of (1.1) is incorrect was its *prima facie* incompatibility with the standard story of the birth of the quantum theory. But our discussion so far has served only to confirm that the standard story leads to notoriously paradoxical consequences, without in any way removing the paradox. From showing that a matter is implausible, to proving that it is false, is a long step, and nothing in our investigation so far enables us to conclude that Planck did not introduce the quantum in 1900.

So let us see what there is to the above analysis of Planck's proof that gives it its air of sheer implausibility. If it is true indeed that Planck's 1900 derivation of (1.1) makes assertions that oppose each other, then there can be no difference in saying that the proof did, and that it did not, introduce the quantum. The question as to whether the orthodox story is right or wrong would therefore be unanswerable on the basis of Planck's proof alone. Then what are we so hard at work on right now?

There *must* be a catch here, and to find it, let's try a different strategy. Perhaps the best way to test the contention that the orthodox story is false is to examine the cases in which it seems to be confuted. One can ask: How can we say, indeed, that Planck did *not* introduce discontinuity in physics, if he won the Nobel Prize in 1918 for the discovery of energy quanta? How could Planck have arrived at formulas (1.6) and (1.7), which, according to our present conception, encode particle indistinguishability and energy quantization respectively, if he did *not* willingly break with tradition? Both Klein and Kuhn have argued that Planck did it by mistake—once again, on the grounds that he "did not know what he was doing"; we have already rejected this possibility. But the alternative would be that Planck found these formulas by consistently thinking *with* classical physics, and it is unclear how this can be. Finally, it is quite difficult to believe that Planck contradicted himself *unwittingly*, as Einstein appears to have suggested, and many historians defended. For how in the world could Planck, renowned theoretician, end his seminal 1900 presentation of his proof of (1.1) calling for more independent tests of the values of N and e *in order to confirm the absolute validity of his derivation,* not noting that his proof was, actually, so off-target, that it fails to conform to the conditions under which alone the proof could be said to be *significant*? It is vital for us to give satisfactory answers to all these questions. If we are successful in this, then shall have shown that the orthodox story is false. The missing pieces of our origin-of-the-quantum jigsaw-puzzle will be found by discussing critically the facts described next in four short steps.

1. The first thing to note is that Planck did not arrive at (1.1) by reasoning from premises to conclusion. Things were, rather, just the other way around. Planck started with Wien's (1.14) (which, remember, was shown to fail in the far infra-red in late 1899 and early 1900), while appealing to careful experimental measurements; Planck reported, years later, that H. Rubens paid him a visit in October 7, 1900 to tell him that he had determined ϵ_ν to be proportional to T for small ν (the same experimental fact that led Lord Rayleigh to propose (1.13) as an effective law in June 1900, as we have seen). Interpolating, then, between these two regimes, and following general theoretical considerations that had enabled him to derive Wien's law in 1899 (we shall come back to this point later on), Planck was able to write down an expression that rendered measurements exceptionally well,

$$\varepsilon_\nu = \frac{a\nu^3}{e^{\frac{\beta v}{T}} - 1}, \tag{1.15}$$

with α and β, of course, as in (1.14). This formula Planck reported publicly on October 19 (recall: the birth of quantum theory is said to be marked by the two-month-after presentation of the *derivation* of the formula); decades later, in his reminiscences, Planck wrote that the very next morning (October 20) he received a visit from Rubens:

> He came to tell me that after the conclusion of the meeting he had that very night checked my formula against the results of his measurements

> and found a satisfactory concordance at every point... Later measurements, too, confirmed my radiation formula again and again—the finer the methods of measurement used, the more accurate the formula was found to be. (PLANCK, 1949, 40-41)

The reader should, then, take equation (1.1) for what it was in October 1900, namely, equation (1.15). In other words, a *guess* confirmed by experiment *in need of theoretical justification*.

But wait a moment. A quantum-theoretical formula cannot be derived from classical premises. And since it is a plain fact that there was no ultraviolet catastrophe to worry about as of 1900, no reason to think that the right black-body radiation law could not be found by following the all-encompassing, time-tested classical physics, it will now seem that if Planck had reasoned *deductively,* then he could not have found (1.1). Nothing else but a prior empirical commitment, namely (1.15), could, it seems, have enabled Planck to derive the right black-body radiation law.

So just what is told, then, in textbooks about the birth of quantum theory—namely that Planck's law first appeared as a derived formula in December 1900—is simply nonsensical. Not only *was* Planck's blackbody radiation law a guess. It could only have been a guess.

2. The second point is that there is a sense in which the guess was an inductive guide for Planck. This fact is what may seem to bring out, or justify, the assumption, shared by many historians, that the problem of the origin of the quantum theory is resolved in cognitive terms, i.e., in Planck's way of thinking; from what we have already said it should be clear that this assumption is unjustified. But to make our own case we cannot afford to overlook the textual evidence that has given to such cognitive or psychological explanations their appearance of plausibility. The gist of such evidence goes as follows.

We learn from Planck's own pen that by 1900 he had been wrestling "for six years with the problem of equilibrium between radiation and matter" and that it had been "his great ambition to derive the correct form of Kirchhoff's function *from first principles*" (PAIS, 1982b, 368, my italics). These assertions spell out what we call *motivation*; consider, in this connection, the following remark by Planck:

> Anybody who has been seriously engaged in scientific work of any kind realizes that over the entrance to the gates of the temple of science are written the words: You must have faith. It is a quality which the scientists cannot dispense with.
>
> The man who handles a bulk of results obtained from an experimental

> process must have *an imaginative picture of the law that he is pursuing*. He must embody this in an imaginary hypothesis. The reasoning faculties alone will not help him forward a step, for no order can emerge from that chaos of elements *unless there is the constructive quality of mind which builds up the order by a process of elimination and choice*. Again and again, the imaginary plan on which one attempts to build up that order breaks down and then we must try another. This imaginative vision and faith in the ultimate success are indispensable. *The pure rationalist has no place here*. (PLANCK, 1933, 214-215, my italics)

In view of facts such as these we are now able to put the relevant elements of Planck's epistemic situation in their proper conceptual boxes. Recall that Planck's interests were focused on blackbody radiation *theory*, for the elaboration of which *empirical input* is required. In Planck's early works, the fact of experience which guided his train of thought, as we have seen, was Wien's law. I have already remarked upon the fact that Planck presented a derivation of it in 1899 but that soon after, deviations from Wien's distribution law were experimentally detected; the basis of Planck's work was thus thereby undercut. Planck solved the problem *himself*, then, as we have seen, within a few months' time (more specifically, in October 1900; the communication of the result, equation (1.15), had the title "An Improvement of the Wien Distribution"). The situation was then: Planck's phenomenological formula (product, as he put it (PLANCK, 1949, 41), of a "lucky intuition") was an outstanding achievement—it alone would have been sufficient to win for him enduring fame—but for the last time: his theoretician's ambition was the *derivation* of Kirchhoff's function; by October 1900, it turned out to be the case that the correct mathematical representation of the blackbody radiation law was not, after all, Wien's—it was his own. The reader may well imagine the fervent urgency with which the man felt the need to solve his problem; the weeks that followed, Planck wrote in his reminiscences, were "the most strenuous work of my life" (PLANCK, 1933). And it happens—and this is the twist—that such motivation, such urgency, may seem to hold the key—as the reader shall find next—to Planck's inconsistency.

3. Note next that Planck, as of 1900, was an outspoken critic of the probabilistic kinetic theory of heat. But unlike so many Planck did not base his criticism on general philosophical grounds; his reasons stemmed from physical theory.[20]The following passage from his *Scientific Autobiography* makes this point evident:

> I was not only indifferent but to a certain extent even hostile to the atomic theory which was the entire foundation of [Boltzmann's] research. The

[20] For a discussion about the scientific status of the probabilistic kinetic theory of heat as of the end of the nineteenth century, see (SMITH; SETH, 2020, chapter 2). See also footnote 22, below.

> reason was that at that time I regarded the principle of increase of entropy as no less immutably valid than the principle of conservation of energy itself, whereas Boltzmann treated the former merely as a law of probabilities—in other words, as a principle that could admit exceptions. (PLANCK, 1949, 32)

Note, now, that the radiation theory that Planck had been trying to develop hitherto, the theory with which he had been able to in fact derive Wien's law in 1899, was entirely based on the fundamental idea that the approach-to-thermodynamic-equilibrium of a radiation-in-a-box system could be described assuming only the laws of electrodynamics. We need not pause over details; the relevant fact here is merely qualitative, namely that one cannot, as Planck soon found out, derive his formula by this route. And all matters of fact and psychological influences and epistemological circumstances we have discussed up to this point have now served their purpose by preparing the reader for *this*: Under the sway of the emotional demand to solve his problem, Planck turned to Boltzmann's theory "in an act of desperation... I had to obtain a positive result, under any circumstance and at whatever cost" (PLANCK, 1949, 32). In an 1877 memoir of Boltzmann's, one can find a discussion of the mechanical properties of a gas of N absolutely elastic spherical molecules, each of which are allowed to take only finite values of velocity $0, \frac{1}{q}, \frac{2}{q}, \frac{3}{q}, ..., \frac{p}{q}$ and of energy $0, \epsilon, 2\epsilon, 3\epsilon, ..., p\epsilon$. The similarities between Planck's and Boltzmann's treatments are too striking to be accidental; there is no doubt that this essay was Planck's combinatorics guide. Now look again, reader, at the ingredients of the derivation of (1.1); the "non-classical" ingredients—those which Planck combined with the classical ones leading to a self-nullifying derivation, almost ubiquitously believed to have inaugurated quantum theory—those ingredients are *there*, in Boltzmann's 1877 essay, for anyone who cares to look. And mind you, Boltzmann warns the reader: the discreteness assumption "*does not correspond to any realistic mechanical model*, but it is easier to handle mathematically, and *the actual problem to be solved is re-established by letting p and q go to infinity* [hence ϵ goes to zero]" (BOLTZMANN, 1877, 1976, my italics). From here, the full picture seems to follow readily. Planck, whose mind was in the state of hunting, of searching, of trying this and that, upon using Boltzmann's formal trick did not take the limit at the end of the calculation—*contra* Boltzmann. And the motive behind the omission seems to stand out sharply; it seems to have no theoretical or logical or empirical foundation.

Planck had "a picture of the law he was pursuing": his (1.15). He was eager to derive it "at whatever cost." The motive—and this is the punchline—was, then, it seems, very simple: the law doesn't follow otherwise.

That's the route that led Planck to derive his black-body radiation law.

4. Now let's think this through with some care. Today we say that Planck's law "reduces" to (1.13) in the so-called "classical limit" (long wavelengths). But there is an alternative route: Consider, to start with, equation (1.10). Expand it in a Taylor series. Follow Boltzmann's instruction. Take the $\epsilon \to 0$ limit in order to reestablish, as Boltzmann put it, "the actual problem to be solved." (This, the reader will remember, Planck did not do.) Then we have

$$\frac{\partial S}{\partial U} = \frac{k}{U}. \tag{1.16}$$

By (1.2), this is just equation (1.12). And by (1.5), as was already shown, we are led, ineluctably, inexorably, to (1.13), the ultraviolet catastrophe, all over again.

But that can't be all there is to it. When Rayleigh derived (1.13) (June 1900) he proposed it as an *effective* law. And so the long-wavelength limit, as of 1900, was *not* the "classical" limit; Wien's law, valid for the short wavelengths, had after all been proposed in 1896. And the truth of (1.1) ("the finer the methods of measurement used, the more accurate the formula was found to be") was surely beyond doubt for Planck; the possibility of classical physics not being *right*, without qualification, for the reasons already pointed out, as I have been at pains to emphasize, could hardly be conceived. In which case it must be admitted that as far as Planck could think, as far as he could *perceive*, his formula, and the premises leading to it, were—if I may be excused to use the word anachronistically here—*classical*. Period. Mind you, Planck said it himself: his derivation of (1.1) "does not seem to contradict any of the facts established so far" (PLANCK, 1901, 555). After all,

> the fact that the chosen energy element ϵ for a given group of resonators must be proportional to the frequency ν follows immediately from the extremely important Wien displacement law. (PLANCK, 1900b, 87)

Wien's displacement law, derived by Wien in 1893, is, let us remember, a purely thermodynamic result (WIEN, 1893). It is a telling, but now almost half-forgotten, fact that it implies (see, e.g., (EHRENFEST, 1911) and (PLANCK, 1914, §3) the statement that the ratio

$$\frac{\epsilon_p}{\nu_p} \tag{1.17}$$

where ϵ_p is the energy of each of the p modes of blackbody radiation, and ν_p the associated frequency, is an adiabatic invariant. This means that relative variations of the kind

$$\frac{\epsilon_p}{\nu_p^2} \tag{1.18}$$

would lead to a conflict with the second law of thermodynamics.

So the upshot of point 3, above, misfires. It misfires because it is not as if Planck's appropriation of Boltzmann's formal maneuver were *ad hoc,* or a departure from classical theory. His words imply the very opposite.

And since thermodynamic relations are model-independent, no further justification was apparently needed. And since classical physics also requires that dynamical states of physical systems vary continuously—the laws of dynamics, after all, are *differential* equations—, if one partitions the energy of the system into $\epsilon > 0$ cells, the continuity requirement is satisfied if only the move is only, and *merely,* a formal artifice.

And it is therefore not at all surprising that this formal maneuver is exactly what Max Planck did to derive (1.1).

Consider the passage that has been perennially (mis)read as the introduction of quantization in physics:

> We must now give the distribution of the energy over the separate resonators of each group, first of all the distribution of the energy E over the N resonators of frequency ν. If E is considered a continuously divisible quantity, this distribution is possible in infinitely many ways. We consider, however—this is the most essential point of the whole calculation—E to be composed of a very definite number of equal parts and use thereto the constant of nature $h = 6.55 \times 10^{27}$ erg · sec. This constant multiplied by the common frequency ν of the resonators gives us the energy element ϵ in erg, and dividing E by ϵ we get the number P of energy elements which must be divided over the N resonators. *If the ratio is not an integer, we take for P an integer in the neighborhood*. (PLANCK, 1900b, 84, my italics)

If the resonator energy were actually quantized, then $\frac{E}{\epsilon}$ should be an integer. But Planck says explicitly that it need not be. The abstraction of partitioning the continuous, infinitely divisible quantity E into P energy elements was necessary to carry out the combinatorics to compute the entropy of radiation in a way that was consistent with both Wien's displacement law and the second law of thermodynamics. It was—let me repeat it expressly once more—a formal maneuver. And formal maneuvers do not have physical consequences.

An analogy with statistical mechanics will now be instructive. In statistical mechanics, the abstraction of partitioning the infinitely divisible quantities position q and momentum p into cells of definite size $\Delta q \Delta p$ to count the states is routine. No one will say that this restricts the possible values of p and q of a particle. True, in statistical mechanics, the size of the cell of phase space does not matter, whereas for radiation Planck had to fix, for the reasons given, the cell size of phase space at $h\nu$. But the point remains unscathed: the energy elements are a restriction on cell size only. It is *not* a restriction on the possible values of resonator energy, *at all.*

And so we see that just like in statistical mechanics particles move freely and continuously through and between the $\Delta q \Delta p$ cells of phase space in accordance to Newton's laws, in Planck's radiation theory the resonators move freely and continuously through and between the energy cells in accordance to Maxwell's laws, period.

And so we see that if we just look at Planck's 1900 derivation of the blackbody radiation law with unbiased eyes, if we just get out of our heads the preconception that he introduced discontinuity in physics and consider what Planck actually *said*, then the conclusion that he quantized resonator-energy will not occur to us. And so we can indeed conclude—as I mentioned at the outset—from Planck's 1900 proof *alone* that it did *not* inaugurate the quantum theory, period, end of story.

But the climax of our plot is actually *this*: By the upshot of the previous paragraphs, it must now be acknowledged that Planck *did not,* after all, make the error of which we*, following Einstein,* have accused him of committing, for in that earlier analysis we have assumed that the quantity ϵ is some *actual physical something*, *some indivisible atom of energy*, P of which are distributed amongst N oscillators. Now in this, of course, we have simply taken for granted the almost universal conviction that it is. But the present considerations show that the quantity ϵ *cannot* be given such physical interpretation. Not in the context in which it appeared originally anyway. In saying this, I am making what may be held to be an extremely controversial statement, but please note that the alternative (which is the standard account of the birth of quantum theory) admits too many irrationalities, it cannot be correct. In fact, it is surprising that outstanding scholars have gotten this wrong for so many decades when a careful reading of the above seminal passages of (PLANCK, 1900b) makes it straightforward. And that we have gotten the facts exactly right is vouched for by Planck's own words, written in a letter to his friend Robert Wood in 1931, quoted next.

> The energy elements were a purely formal assumption, and I really did not give it much thought except that, no matter what the cost, I must bring about a positive result. (quoted in (HERMANN, 1971, 23))

This passage of Planck's rules out any debate.

1.2.4 Now let's draw the significant threads together and pin down the conclusions that the above considerations seem to teach.

We began with the intuition that the orthodox story of the origin of the quantum and Einstein's contention that Planck's proof is wrong are *prima facie* incompatible.

Analyzing, then, Planck's proof, we first concluded, in §1.2.2, that it is inconsistent. But now, one step beyond, we found, by the upshot of the previous section, that, relative to the conceptual situation in which it first appeared, Planck's proof is *not* inconsistent; one can only call it "inconsistent" if a departure from classical physics is assumed, i.e., by dint of a prior appeal to discontinuity. But if Planck's proof is indeed consistent, then now it may seem that the problem we started with has somehow disappeared. What should we make, by the upshot of the previous analysis, of UVC? What was the content of Einstein's critique to Planck, then? Nothing?

Of course, that *absolutely cannot be*. For if (1.1) is, as we know it to be, a *quantum* formula, and yet Planck himself gave it a classical proof in 1900, then that proof can't be right. But how, indeed, do we get *at* this conclusion? How do we justify the assumption that (1.1) is quantum? Well, our grounds for having reevaluated Planck's 1900 proof, and concluding therewith that it is not inconsistent, was the acknowledgment that $\epsilon > 0$ was "a purely formal," and not a physical, assumption. But note that we did not ask the further question as to whether Planck was *justified* in making it. And the fact historical works have heretofore failed to make explicit, the fact that puts all the pieces of the origin-of-the-quantum jigsaw puzzle in their proper conceptual places, is that there is a factual, well-defined sense in which Planck *was*, and a factual, well-defined sense in which Planck *wasn't*, justified in making his "purely formal assumption."

These facts I shall now proceed to spell out.

First of all we ought to correct a misunderstanding that is deeply-rooted in the literature. The reader will remember from §1.2.2 that commentators have persistently alleged that (1.13) was "an inevitable consequence of classical physics." And we saw, a couple of pages back, that if one takes, in the appropriate mathematical step, the $\epsilon \to 0$ limit as *bona fide* statistical mechanics requires, then equipartition of energy, and thereby (1.13), ineluctably follows. Now the case of the orthodox story (or this, at least, is one way of putting it) was that Planck tried to *modify* the kinetic theory of heat *by* not taking the limit to correct for the ultraviolet catastrophe. So the orthodox story, on top of assuming that (1.13) was inevitable, on top of ignoring the consequences encoded in the Wien displacement law, further assumes that Planck, in deriving (1.1), was (mis)using the kinetic theory of heat. And so when (KLEIN, 1961) claimed that Planck "was not aware" of the ultraviolet catastrophe, it followed that Planck couldn't have misapplied Boltzmann's theory as a stop-gap measure. And therewith, the question: How did Planck manage to avoid the allegedly "inevitable"? Klein, *and Kuhn*, adopted the course of maintaining that Planck "misunderstood" Boltzmann's instructions ((KLEIN, 1961, 474); (KUHN, 1978, 128)); they differed only in how they interpreted (what they saw as) Planck's mistake:

> for Klein, Planck's misunderstanding allows him to go forward with quantisation without realizing the full import of what he had done. For Kuhn, Planck's misunderstanding allows him to proceed 'fully classically' without ever meaning to have quantized resonator energy. (GALISON, 1981, 79)

Thus, as the conclusion that Planck was not justified in not taking the $\epsilon \to 0$ limit in Klein's and Kuhn's accounts follows from assuming, first, that (1.13) is inevitable, and second, that Planck was following the kinetic theory of heat, and we have already shown, in §1.2.2, that the first assumption is false, all that is now required to show that Planck made no mistake in the sense meant by Klein and Kuhn is to make indubitable the fact that the latter assumption is also false. Our discussion in §2.3 already argued to this conclusion. But if any doubt on this point still remains, we may remove it with the help of the following remarks.

We saw in §1.2.2 that in his (PLANCK, 1900a) Planck said that Wien's law was first derived by "W. Wien from the molecular-kinetic consideration and later by me from the theory of electromagnetic radiation." In (PLANCK, 1900b), in the course of talking about the necessity (!) of his derivation of (1.1), Planck was careful to remark:

> [...] Apart from that, the whole deduction is based upon the theorem that the entropy of a system of resonators with given energy is proportional to the logarithm of the total number of possible complexions for the given energy. This theorem can be split into two other theorems: (1) The entropy of the system in a given state is proportional to the logarithm of the probability of that state, and (2) The probability of any state is proportional to the number of corresponding complexions, or, in other words, any definite complexion is equally probable as any other complexion. The first theorem is, *as for as radiative phenomena are concerned, just a definition of the probability of the state,* insofar as we have for energy radiation no other a priori way to define the probability that the definition of its entropy. *We have here a distinction from the corresponding situation in the kinetic theory of gases*. (PLANCK, 1900b, 87-88, my italics)

Planck's point, then, in the italicized remarks, cannot be clearer: statistical considerations are introduced differently in his theory of radiation and in the probabilistic kinetic theory of heat. Planck, after all, was applying the entropy-probability relation to *radiation*, not to atoms. Finally, still in (PLANCK, 1900b), in a footnote, after presenting the steps to derive (1.1), Planck defended his right to put forward a different microphysical model:

> When Mr. W. Wien in his Paris report about the theoretical radiation laws did not find my theory on the irreversible radiation phenomena satisfactory

> since it did not give the proof that the hypothesis of natural radiation is the only one which leads to irreversibility, he surely demanded, in my opinion, too much of this hypothesis. [...] From the same point of view *one should also declare the kinetic theory of gases to be unsatisfactory since nobody has yet proved that the atomistic hypothesis is the only one which explains irreversibility.* (PLANCK, 1900b, 88, my italics)

That Planck was *not* using the probabilistic kinetic theory of heat, but his own theory, to derive (1.1), is once again transparent. It is ironic that the contention that Planck misapplied, either wittingly or unwittingly, the kinetic theory of heat when he did not take the $\epsilon \to 0$ limit is in itself a misapplication. And by way of these remarks we see once again that explicit evidence that the orthodox story of the origin of the quantum is false is *there*, in Planck's 1900 proof of (1.1), for anybody who might have cared to look.

And we also see that the habit of leaning on assumptions about people's way of thinking to make sense of the dynamics of science may make us prone to overlook the matters-of-fact. Indeed, the foregoing mistaken assumptions, on top of the fact that Planck derived (1.1) by inferring premises from conclusion, on top of the fact that he was "desperate" to find a derivation, misled generations of scholars into unqualifiedly assuming that Planck reasoned—as physicists do, according to Kuhn's *Structure*, at times preceding a scientific revolution—without foundation. But this conclusion is wholly unjustified! In (PLANCK, 1900a), Planck said that he had been led to believe that Wien's law was "absolutely valid" not merely because it fit observations but because of the relationship between "the electromagnetic theory developed by me and the experimental data." His 1899 derivation of Wien's law, remember, had enabled him to construct his famous system of units based entirely on physical constants; when Wien's law was falsified, Planck then "looked for the *next* simpler formula" that seemed to fit the available data and his theory: this was his (1.15). Now note, reader, that this formula looks, on the surface, quite peculiar: (1.15) has two independent constants, one in the numerator, and other in the exponent of e in the denominator, while (1.1), the derived formula, has not one but both constants in the exponent of e in the denominator. Finding a way for the exponent of e to be dimensionless was no doubt a challenge in the theoretical transition from (1.15) to (1.1); to meet this challenge, Planck, "in an act of desperation," *adapted* his theory, and borrowed from Boltzmann the entropy-probability relation idea, a formal trick to erect the grid of phase space, and applied all this to *thermal radiation*. That his moves are perfectly consistent with *bona fide* classical-physics practice, I think the reader will concede without further ado. Planck had every right to expand *his* theory.

And so it seems to be a grave injustice to Planck and his dignity as a physicist to say, like so many historians have done, that he didn't know what he was doing. The contrary seems to be the truth. For Planck could and did claim that it was not, then, Wien's, but his own formula, that was "absolutely valid" not merely because it fit observations, but because of the "relationship between his theory and the experimental data." Recall the confidence with which he closed his presentation of his seminal 1900 proof of (1.1)—for the convenience of the reader, I repeat below the quote from page 33:

> If the theory is at all correct, all these relations [i.e., the values obtained for Avogadro's number N, Loschmidt's number L, and for the "quantum of electricity" e] should be not approximately, but absolutely, valid. The accuracy of the calculated number [for e] is thus essentially the same as that of the relatively worst known, the radiation constant k, and is thus much better than all determinations up to now. To test it by more direct methods should be both an important and a necessary task for further research. (PLANCK, 1900b, 89-90)

Measurements, after all, are *theory-mediated*, and the argument Planck presented at the end of his proof—namely, a cross-check between the values for the constants he obtained *via* his 1900 derivation with the values for constants obtained by other investigators *via* independent methods[21]—was the best he could possibly give in favor of the correctness of his theory.[22]

And so it is perfectly clear that Planck's 1900 proof was not inconsistent in the sense that it was self-contradictory, nor inconsistent in the sense that it misapplied the probabilistic kinetic theory of heat. On the contrary. *By the authority of the experimental evidence available at that time*, his derivation was perfectly consistent. But there's yet another objection, and this one*, in due time*, shall not miss its target. I refer to the objection that Planck should have used the probabilistic kinetic theory of heat *instead.*

Here's where Einstein comes in. In (EINSTEIN, 1905a, §1), he—the then-unknown patent officer, who sent his paper to be published at the *Annalen der Physik,* whose editor at that time was Max Planck—asks us to consider a system of gas molecules, free electrons, and a collection of resonators (he defines them as "electrons bound to points in space that emit and absorb electromagnetic waves of definite period") together in a box. The kinetic theory of heat tells us that the equipartition theorem holds for the gas, and therefore, via the dynamical equilibrium condition, for the whole system. Hence, the

21 For a discussion of the accuracy of the values Planck obtained for N, L, and e, *vis-à-vis* other values available at that time, see (SMITH; SETH, 2020, 16-18).

22 There's yet another telling fact to be detected in the above passage. Today we call the constant k "Boltzmann constant." We associate it with the probabilistic kinetic theory of heat. But Boltzmann himself never introduced it. Planck did. In 1900. To derive (1.1). Without using the atomistic hypothesis. And Planck called it—and this is the telling fact—"radiation constant" in the above quote.

energy distribution of the system is necessarily (1.13), in irreconcilable contradiction with the facts.

The critique of Planck's 1900 derivation is evidently between the lines in the argument. But Einstein's point, though implicit, to those who have followed our discussion so far, is clear enough: Planck had no right to ignore the equipartition theorem, no right in refraining from taking the $\epsilon \to 0$ limit, no right to his formal maneuver. If Planck's resonators are mechanical systems, then they obey the equipartition theorem, period. *This*, I now emphasize, is Einstein's point with UVC. We failed to see it back in §1.2.2 because hitherto we had mistaken the target of Einstein's critique: we thought it was the orthodox version of the derivation, when in fact it was the derivation argued for here. We were, back then, still under the spell of the orthodox-story myth.

But our considerations in §1.2.2 will have now served their main purpose by enabling the reader to recognize immediately that Einstein's objection—as of 1900—had little, if any, force against Planck.

For Einstein's argument patently depends on the validity of the equipartition theorem, which depends on the validity of the probabilistic kinetic theory of heat, which—as I have been at pains to emphasize—was *not* established in 1900. It is not a coincidence that in 1905 Einstein criticized Planck almost without saying it but in 1909, as we saw in §1, he was point-blank about it. The empirical authority of the theoretical ammunition he was using changed drastically between 1905 and 1909.[23] Remember Einstein's 1949 remark: "the imperfections of Planck's derivation remained at first hidden." The information that brought the imperfections of Planck's derivation to the fore—namely, conclusive empirical evidence of the validity of the probabilistic kinetic theory of heat—was not, as of 1900, literally in the field of vision.

And so we now see not only that Planck *did not* follow the probabilistic kinetic theory of heat in his 1900 derivation, notwithstanding nearly everybody's impression to the contrary. (His deduction of (1.1), to quote from the seminal 1900 paper again, "is based on the laws of electromagnetic radiation, thermodynamics, and probability calculus.") Planck, as a matter of fact, *need not to have followed* the kinetic theory of

[23] "It is not at all easy for us today," writes the kinetic theory of heat expert Marian Smoluchowski in 1914, in a very instructive paper called "Limits of Validity of the Second Law of Thermodynamics,"

> to recall the mood that was prevalent toward the end of the last century. At that time the scientific leaders in Germany and in France were convinced—with a few exceptions—that the kinetic theory of heat had ceased to be of any importance. Considering the great successes of thermodynamics, the second law had been raised to the rank of an exactly valid, absolute dogma that was true without exception. And as the kinetic theory had run into certain difficulties in the attempt to arrive at a satisfactory interpretation, especially as regards to irreversible processes, *one had condemned it together with the atomistic hypothesis* [...]. All this has changed today. (Smoluchowski 1914, cited in Feyerabend 1965, 230, my italics)

heat as of 1900, for, to borrow from the 1900 paper again, "nobody has yet proved that the atomistic hypothesis is the only one which explains irreversibility." Unlike Einstein, *Planck left the structure and motions of his resonators undetermined* (see (PLANCK, 1899, §1)); the aim was to develop a more general theory, free of special microphysical hypothesis. And if we just stop to think about the upshot of our §1.2.3 discussion, if we just remember the difficulties plaguing, at that time, the probabilistic kinetic theory of heat (see footnote 22 again), doesn't Planck's seem like a plausible—perhaps even the best—course of action?

This is the sense, to which I referred above, in which Planck *was* justified, in 1900, in not taking the $\epsilon \to 0$ limit. The sense in which he wasn't comes of course from the stringent confirmation, by Perrin, starting in 1908, of the probabilistic kinetic theory of heat predictions.[24] Thereafter, Planck's resonators *had* to be (re)interpreted as mechanical systems, in which case, Einstein's considerations apply,[25] and so we have no choice but to take the $\epsilon \to 0$ limit, and so we have equipartition, and so we have the ultraviolet catastrophe, yet again. We have the conclusion that Planck's 1900 classical derivation of (1.1) is mistaken because it ignored the dynamical equilibrium condition between matter and radiation. We have a minutely-confirmed-by-experiments

[24] In *Black-Body Theory,* and again in (KUHN, 1984), Kuhn claims that it was Lorentz's authority that led Planck to change his mind in 1908 (we know from a letter that Planck attended a talk by Lorentz on blackbody radiation that year). I think that's the wrong explanation; Lorentz had been giving talks like that for years. It is much more likely that it was the authority of *empirical evidence* that led Planck to change his mind in 1908: the general validity of (1.13), and Boltzmann's statistical interpretation of the second law of thermodynamics, had to be acknowledged after the probabilistic kinetic theory came to be part and parcel of established physics after Perrin's 1908-1911 Brownian motion experiments.

[25] To verify the correctness of the interpretation of his criticism to Planck argued for here, check (EINSTEIN, 1909b, §2-4). For the convenience of the reader, I will summarize the main points: Einstein devotes §2 to a derivation of (1.13), and starts §3 by emphasizing that "there can be no doubt" that "our current theoretical views lead inevitably" to this formula. In §4 he asks how Planck's radiation theory is related to the probabilistic kinetic theory of heat which entails (1.13). "In my view the answer to this question is made harder," he says, "by the fact Planck's own presentation of his theory suffers from a certain logical imperfection." He then goes on to argue that Planck was justified to relate entropy and probability the way he did "only if he had appended the condition" that the energy states of radiation

> had been found to be equally probable on the basis of statistical considerations. In this way he would have arrived at the formula defended by Jeans [i.e., (1.13)]. Though every physicist must rejoice that Mr. Planck disregarded these requirements in such a fortunate manner, it should not be forgotten that the Planck radiation formula is incompatible with the theoretical foundation from which Mr. Planck started out. (EINSTEIN, 1909b, 363-363)

This last statement is exactly the critique to Planck we used to motivate our study in §1. The reader will recall that probability theory is introduced in Planck's theory of natural radiation, and in Boltzmann's probabilistic kinetic theory of heat, in very different ways—Planck was, as we have seen, careful to point this out in his 1900 proof. Now, the punch line, which I cite as proof of the correctness of the account of the origin of the quantum theory that I put forward here: in a footnote in §4, Einstein adds that the fact that "the only tenable way" of distributing states statistically is the way he was arguing for, the way which coincides with Boltzmann's, the way that leads "ineluctably" to (1.13), and evinces that Planck's way of doing statistics is wrong, "follows immediately from the properties of Brownian motion" (EINSTEIN, 1909b, 361, footnote 1). See also footnote 25, below.

blackbody radiation formula with a constant h in it that doesn't appear at all in the "classical" prediction.

This is the last piece of the puzzle. It now takes only a step back to see the full picture.

1.3 The Origin of the Quantum Theory

To establish that there was a crisis in physics in the beginning of the 20th century one would have to first establish the validity of the probabilistic kinetic theory of heat. And this, we all know, is what *Einstein* tried to do in 1905 through the means of his predictions regarding Brownian motion. It is therefore little wonder that the same man, from the same year on, started repeating that there was no escape from (1.13), that Planck's 1900 derivation of (1.1) *is wrong*, that the foundations of physics must be considered afresh. That the quantum revolution was set in motion by Albert Einstein seems now to follow as a matter of logic. I shall finish by giving you evidence.

Consider (EINSTEIN, 1905a) again. In §1, Einstein argued, as we have seen, that (1.13) was inevitable. The point of his next section, though Einstein, characteristically, does not make this really explicit, is to show just what the inevitability of (1.13) implies, *given what Planck did*. Planck's (1.1) yields, Einstein remarks, $E = kT$ for high temperatures, while the kinetic theory of heat yields $E = \frac{R}{N}T$. So we can set $N = \frac{R}{k}$, and the value for N calculated with Planck's theory, and the values for N calculated with the kinetic theory of heat, are indeed in good agreement. But what this agreement shows, Einstein says, is *not*, as Planck thought, that Planck's derivation of (1.1) is "absolutely valid," but rather, on the contrary, that "Planck's determination of the elementary quanta is to some extent *independent of his theory of blackbody radiation*." Einstein leaves two points between the lines here. First, though Planck's derivation is wrong, *his constants are universal*; second, though Planck's constant k is predicated in classical physics, *his constant h is not*. And what all of this tells us, Einstein concludes, is that

> the greater the energy density and the wavelength of radiation, the more useful the theoretical principles we have been using prove to be; however, *these principles fail completely in the case of small wavelengths and small radiation densities*. (EINSTEIN, 1905a, 91, my italics)

This, of course, is the statement of a crisis. This is the statement that (1.13) is "the classical limit," as we say today. After Perrin confirmed his Brownian motion predictions, Einstein's arguments gained force, conviction, and eventually, currency.[26] And it is surely

[26] Compare footnotes 22, 24, and 26. A further evidence of the importance of Perrin's experiments for the

little wonder that the same man who noted, and found the means to empirically establish, *that* classical physics failed, was the man who came up with an explanation, *on the basis of a new hypothesis*, for *why* it failed. Just how Einstein managed to do this—just how he stepped outside classical physics to engage in this revisionary task—shows how skillfully he had schooled himself in careful epistemological thinking: "Without having a substitute for classical mechanics," he reminisced in 1949, in his *Autobiographical Notes*, "I could nevertheless see to what kind of consequences [Planck's] law of temperature-radiation leads" (EINSTEIN, 1949, 45). The question as to what Einstein's (1905a) involves, aims, and achieves, seems now to be ripe for reanalysis. I cannot undertake this detailed task here so I shall pursue it in a separate paper. But the conclusion of Einstein's (EINSTEIN, 1905a) is, anyway, known to everyone: Radiation has a granular structure. In complete contradiction with the laws of electrodynamics.

Only one must note the consequences which this fact involves.

Consider again resonators and radiation in equilibrium in a box. They are, of course, exchanging energy. If radiation is discrete, then radiation-energy can be transferred to the resonator only in discrete amounts, so that *the energy states available to the resonators are also discrete*. The process of emission and absorption of radiation, i.e., the dynamics of the resonators, is therefore discontinuous: the resonator jumps from one state to the other without passing through a continuous series of intermediate states. In complete contradiction with the laws of mechanics.

Einstein leaves all of this between the lines in (EINSTEIN, 1905a), and we can easily understand why. But a year later, in (EINSTEIN, 1906), he reiterates the point he made in (EINSTEIN, 1905a, §1): resonators in equilibrium with radiation in an enclosure should, according to classical physics, obey the Rayleigh-Jeans law, but as a matter of empirical fact they do not. Einstein then rhetorically asks how can it be that Planck arrived a different, empirically correct, law. His own answer is that "Planck's theory makes implicit use of [my] hypothesis of energy quanta" (EINSTEIN, 1906, 192).

To see this, one must only apply the foregoing conclusions to Planck's law. "To arrive at Planck's formula," Einstein remarks, "one has to postulate that, rather than assume any value whatever, the energy E of a resonator can only assume values that are integral multiples of ϵ, where $\epsilon = h\nu$" (EINSTEIN, 1906, 195). The reader must note that this yields *exactly* the self-contradictory derivation we mistakenly attributed to

whys and wherefores of the quantum revolution—for they gave to Einstein's argument in favor of the absolute inevitability of (1.13) the empirical validation that they scientifically required, and it is from such inevitability that the contention that "the logical reason for the quantum theory is found in the fact that the Rayleigh-Jeans radiation formula does not agree with experiment" (PLANCK, 1914, 265) derives its pull—is the fact that Perrin was part of the selected group of physicists that participated of the 1911 Solvay Conference. The topic of this conference, let us remember, was "Radiation Theory and the Quanta," yet Perrin was invited to talk about neither radiation nor quanta. He talked about his Brownian motion experiments. I thank George Smith for calling this latter fact to my attention.

Planck at the end of §2.2; we are finally in position to understand what *that* contradiction entails. Commentators who have described it merely as another derivation of Planck's (1.1) completely underestimate its significance. For it is a different kind of game that Einstein is playing here.

> In my opinion the above considerations do not at all disprove Planck's theory of radiation; rather, they seem to me to show that with his theory of radiation Mr. Planck introduced into physics a new hypothetical element: the hypothesis of light-quanta. (EINSTEIN, 1906, 196)

Contradictory statements cannot be either both true or both false. The truth of one necessarily implies the falsity of the other and vice-versa. Thus, if the hypothesis of light-quanta stands in contradiction with the mechanical and electrodynamical laws that, together, entail (1.13), then there can be no doubt about how to distribute truth-values between the contradictories.

Think of the fact that in 1900 Lord Rayleigh proved equipartition using only the elementary principles of mechanics. "We are here faced with a fundamental difficulty," he concluded,

> relating not to the theory of gases merely, but rather to general dynamics. [...] What would appear to be wanted is some escape from the destructive simplicity of the general conclusion relating to the partition of kinetic energy. (RAYLEIGH, 1900, 118)

Einstein said it himself: he didn't have a substitute for classical mechanics. But he could nevertheless

> see to what kind of consequences [Planck's] law of temperature-radiation leads for the photoelectric effect and for other related phenomena of the transformation of energy-radiation, as well as the specific heat of solid bodies. (EINSTEIN, 1949, 45)

The escape from the destructive simplicity of energy equipartition (which leads, let me repeat once more, to the Rayleigh-Jeans law of blackbody radiation, and to the Dulong-Petit law of specific heats) followed from the light-quanta hypothesis: the energy states of radiation, and thereby of matter, are actually discrete, and hence they cannot vary continuously. And therewith, of course—since the laws of classical physics are all differential laws that require continuity—, *on top* of the ultraviolet catastrophe crisis, *another* crisis ensued:[27]

> All my attempts [...] to adapt the theoretical foundation of physics to this [new type of] knowledge failed completely. It was as if the ground had been pulled out from under one, with no firm foundation to be seen anywhere, upon which one could have built. (ibid)

Newtonian laws would no longer do.[28] What was required was a new—quantum—mechanics.[29]

The rest, as they say, is history. Further developments only confirmed the correctness of Einstein's fundamental insight.[30] And if a final comment be desired, to understand how can it be that Max Planck won the Nobel Prize in 1918 for the discovery of energy quanta, let us note that Einstein always called the quantum hypothesis (encompassing both energy quanta and light-quanta) "Planck's theory," that everyone followed him in this regard, and that the person who suggested Planck for the Nobel Prize Committee that year was Albert Einstein.

---

[27] We can see this crisis-on-top-of-a-crisis spirit reflected in the opening paragraphs of the invitation letter to attend the 1911 Solvay Conference in Physics. I quote without comment:

> *Invitation to an 'International scientific conference to elucidate certain current questions of the kinetic theory.*
>
> It appears that we find ourselves at present in the midst of an all-encompassing reformulation of the principles on which the erstwhile kinetic theory of matter has been based.
>
> On the one hand, this theory leads to a logical formulation—which nobody contests—of a radiation formula whose validity is contradicted by all experiments; on the other, there follow from the same theory certain results on the specific heat (constancy of the specific heat of a gas with the variation of temperature, the validity of Dulong and Petit's law up to the lowest temperature), which are also completely refuted by many measurements. [...] these contradictions disappear if one places certain limits (doctrine of energy quanta) on the motion of electrons and atoms in the case of their oscillations around a position of rest. But this interpretation, in turn, is so far-removed from the equations of motion of material points employed until now, that its acceptance would incontestably lead to a far-reaching reformation of our erstwhile fundamental notions. (Solvay 1911, quoted in Mehra 1975, 6)

[28] With all of this under our belts, we can look again at Einstein's *Autobiographical Notes* and see that what he is telling us there is that it was the failure of equipartition, and the discovery of the granular structure of radiation, and therewith of energy discreteness, that led Einstein to conclude that neither mechanics nor electrodynamics could claim exact validity, whence the feeling that "the ground had been pulled out from under one." And Einstein tells us next that it was exactly *because* of this, because he understood the foundations of physics had to be rebuilt anew, that he "despaired to discover true laws by means of constructive efforts based on known facts," (Einstein 1949, 53) which path led him, first, to the special, and later to the general, theory of relativity.

[29] Einstein's dissatisfaction with how quantum mechanics developed is well-known, but I think our discussion gives us further insight into his misgivings. What Einstein was holding out for was no less than a new theory of matter and motion; a toolbox for computing probabilities of experiment outcomes would not do.

[30] To see how physicists converged to Einstein's conclusion, see e.g. (DARRIGOL, 1988).

If any doubt on this point still remains, the following excerpt, from a letter of M. Besso to Einstein, dated January 1928, will remove it.

> For my part, I was your audience in 1904 and 1905; if, when compiling your communications on the quantum problem, I robbed you of a part of your fame, in return I got you a friend in Planck.[31]

This completes my description of how and why the classical gave way to the quantum. By looking at the logic, and not at the psychology and the sociology, of research we have "rescued" Planck and Einstein from the tales told about them and we have established, *contra* Kuhn, that the structure of the quantum revolution, set in motion by Einstein in 1905, is carved at its joints by theory and experiment, and it is therefore objective throughout.

One last remark. The "clouds" obscuring 19th century physics, let us remember, to use Lord Kelvin's famous figure, were the relative motion of ether and ponderable bodies and the dynamical consequences of the equipartition theorem. The former was empirically reflected in the Michelson-Morley experiments; the latter came to a head, as we have seen, when Perrin's Brownian motion experiments did so much to establish the probabilistic kinetic theory of heat, and thereby, Boltzmann's construal of entropy. Such problems represented an internal failure of Newtonian physics (a "crisis," in an objective sense) so that the solutions, to be theoretically meaningful, should be sought outside the Newtonian framework (i.e., outside the "paradigm"). The theory of special relativity, as is well-known, proposed by Albert Einstein in 1905, solved the first problem, by a method that carefully avoids Newtonian premises (the so-called "principle-theory approach"). The same man, following the same method, proposed, in the same year, the solution to the second problem: the quantum hypothesis. And therewith it was a logical necessity to construct a new conceptual framework to accommodate it. This was what the quantum "revolution" was all about.

[31] Michelle Besso to Albert Einstein, January 17, 1928. Translation from *The Collected Papers of Albert Einstein, Vol. 16,* Document 132: "Meinerseits war ich in den Jahren 1904 und '05 Dein Publikum habe ich bei der Fassung Deiner Mitteilungen zum Quantenproblem Dich um einen Teil Deines Ruhms gebracht, Dir dafür in Planck einen Freund verschaff."

2 On the Lawfulness of Quanta

The absolute certainty of a science cannot exceed the certainty of its principles.

Isaac Newton, *letter to Oldenburg*, July 11, 1672

If the thesis I put forward in "Rewriting the Quantum 'Revolution'," henceforth RQR, is correct, then it must follow that the literature on Albert Einstein's "On a Heuristic Point of View Concerning the Production and Transformation of Light" (EINSTEIN, 1905a), henceforth LQP, though extraordinarily vast and sophisticated, has not done it justice. For even though Einstein's LQP is invariably described as revolutionary, commentators did not mean, as I did in RQR, with this qualifier that LQP started a revolution in Kuhn's technical sense. They meant only that the work is iconoclastic. A more concrete difference is that the variety of analyses available in the literature understand that Einstein's main conclusion, namely, that radiation has a granular structure, follows from—and thus, *depends on*—an analogy with the kinetic theory of heat. But if the thesis of RQR is correct, this cannot be. A revolutionary idea, conceived to change the course of physical research, and persuade a highly conservative and critical community of scientists, cannot have been based on the uncertain, precarious, and arbitrary grounds of a mere analogy.

In fact, according to the objective view of scientific revolutions put forward in RQR, the truth must be the opposite. RQR concludes with the statement that the method Einstein adopted in LQP is the same he adopted in "On the Electrodynamics of Moving Bodies" (EINSTEIN, 1905b), i.e., the so-called principle-theory approach. And it is the essence of this method to deduce conclusions from secure foundations—something one cannot do with analogies. But I did not prove such claim in RQR. I merely asserted it. The present work is therefore to be seen as Part II of RQR, for I shall now complete the task, and explain in greater detail how the quantum revolution did begin, by giving the required proof.

The claim requires further elaboration. Let us then recall, to begin with, that when Einstein famously distinguished between principle theories and constructive theories (EINSTEIN, 1919), he emphasized that the former, as opposed to the latter,

> employ not the synthetic, but the analytic method. The starting point and basis are not constituted out of hypothetical constructional elements, but out of empirically discovered, general characteristics of natural processes (principles), from which follow mathematically formulated criteria that

> the individual processes or their theoretical models have to satisfy. Thus, from the general empirical result that a perpetuum mobile is impossible, thermodynamics seeks, in an analytical manner, to determine conditions that the individual processes must satisfy. (EINSTEIN, 1919, 13)

The theory of special relativity, he says, belongs to the class of principle theories. He described the path to its discovery in his 1949 *Autobiographical Notes* as follows:

> This way of looking at the problem showed in a drastic and direct way that a kind of immediate *reality* has to be ascribed to Planck's quanta, that radiation must, therefore, possess a kind of molecular structure in energy, which of course contradicts Maxwell's theory. [...] Reflections of this type made it clear to me as long ago as shortly after 1900, i.e., shortly after Planck's trail-blazing work, that neither mechanics nor electrodynamics could (except in limiting cases) claim exact validity. Gradually I despaired of the possibility of discovering the true laws by means of constructive efforts based on known facts. The longer and more desperately I tried, the more I came to the conviction that only the discovery of a universal formal principle could lead us to assured results. The example I saw before me was thermodynamics: The laws of nature are such that it is impossible to construct a perpetuum mobile (of the first and second kind). How then could such a universal principle be found? (EINSTEIN, 1949, 51, 53, my italics)

The reader will remember that we saw in RQR that and why Max Planck indeed originally introduced the energy quanta, as Einstein's remark above clearly imply, only as a formal artifice. And we saw that and why, according to Einstein, they are nothing of the sort: "a kind of immediate reality" must be ascribed to them. Now our task is to reflect on the fact that Einstein realized this much "as long ago as shortly after 1900"—and furthermore, that it was *this*, the existence of the energy quanta, that put him in the search for a new formal universal principle from which he could deduce "assured results" that should serve as guides to find "the true laws."[1] For the question, then, cannot be avoided: Whence the certainty that gives to the existence of the energy quanta its alleged inevitability? Must we yield to its consequences because it legitimately follows from principles, from "known facts"? Einstein does not say, in plain language, that this is the case in the above passage, but it is impossible to account in any other way for the fact that he saw no escape from the truly explosive conclusion that neither mechanics nor electrodynamics could claim exact validity, that the "true" laws were yet

[1] Harvey Brown, in his influential work on the foundations of special relativity, stated: "it is impossible to understand Einstein's discovery ... of special relativity without taking on board the impact of the quantum in physics" (BROWN, 2005, 69).

to be found—an explosion that he knew would touch off a chain reaction in which the foundations of physics has been caught ever since and from which we did not yet fully recover. To win the battle of scientific persuasion, in this case, the existence of the energy quanta must be the kind of result that is difficult, almost impossible, to circumvent; it must be the kind of "certainty" that is deducible from a principle. But then we must ask: *which* principle? From Einstein's remark that the existence of the quanta put him in the search for another formal principle, there seems to be no option before us, save to conclude that the existence of the quanta follows from a principle that he already knew.

For the fact that Einstein's principle-theory approach template was thermodynamics is not a coincidence. Thermodynamics was known to be "a science with secure foundations, clear definitions, and distinct boundaries" (MAXWELL, 1878, 257) long before he started his career. And the fact that the principle he took as an example of "certainty" from which to deduce "assured results" is the impossibility of constructing a perpetuum mobile of the first and second kind is not a coincidence either. For it is from this principle, I now categorically assert, that the existence of the energy quanta follows.

This is the thesis of this paper. I claim that the energy quanta, far from being, originally, an *ad hoc* hypothesis in a self-contradictory argument, instead follows from the laws of thermodynamics as "mathematically formulated criterion that the individual processes or their theoretical models have to satisfy." Both Newtonian mechanics and Maxwellian electrodynamics fail to satisfy it. And it is this principled result, this *criterion*, I submit, and the empirical confirmation of the laws of the probabilistic kinetic theory of heat, discussed in RQR, that are the truly decisive historical factors that made a scientific reality out of Einstein's insistence that the quantum revolution cannot be argued away. From the point of view of Newton's fourth law of reasoning,[2] the authorized inference is that the revolution had to happen as it did happen.[3]

The progress towards this conclusion starts with a historical investigation that traces the theoretical roots of the concept of the energy quanta back to Boltzmann's and Wien's nineteenth century works on the thermodynamics of radiation. This is §2.1. In

[2] Rules of Reasoning in Philosophy, Rule IV (Newton, Isaac, *Philosophiae Naturalis Principia Mathematica*, Bk. III, 1687):

> *In experimental philosophy we are to look upon propositions inferred by general induction from phenomena as accurately or very nearly true, notwithstanding any contrary hypotheses that may be imagined, till such time as other phenomena occur by which they may either be made more accurate or liable to exceptions.*

This rule we must follow, that the argument of induction may not be evaded by hypotheses.

[3] Don Howard's great insight that Einstein's distinction between principle theories and constructive theories was borrowed from Newton is very significant and illuminating in connection to my remarks; see (HOWARD, 2023). The fundamental philosophical conclusions I draw in this paper find their conceptual certification in Howard's meticulous and sophisticated analysis of Einstein's distinction between principle theories and constructive theories.

§2.2, these results are critically analyzed, and shown to be linked to Einstein's 1902-1904 investigations—this was the preparatory stage for the coming 1905 revolution. In §2.3, a novel analysis of Einstein's LQP is given, and the claim above is proved.

2.1 The Facts of the Case

The amount of important information that has been overlooked for over a century now, unaccounted for by scholars, and unperceived by historians, as a consequence of the handicaps created by the discontinuity myth, is truly staggering. Nothing shows this better than the following passage from Planck's seminal 1900 presentation:

> I shall now make a few short remarks about the question of the *necessity* [!] of the above given deduction. The fact that the chosen energy element ϵ for a given group of resonators must be proportional to the frequency ν follows immediately from the extremely important Wien displacement law. (PLANCK, 1900b, 87, my italics)

Planck could not have been clearer. The relation

$$\frac{\epsilon}{\nu} = constant \tag{2.1}$$

is not an arbitrary choice. It is a theoretical requirement. It "follows immediately from the extremely important Wien displacement law."

And the Wien's displacement law is a purely thermodynamic result.

The fact that Planck's justification for the energy elements was so unambiguously stated, and the paper, so widely read, shows how fantastically deep, old, and fateful are the inconsistencies that mark the discontinuity myth. In RQR we saw that it is indeed at the expense of Planck's words that it survives. Note how Martin Klein, influential historian,[4] dealt with the above remark of Planck's: "The Wien displacement law [...] played an essential part in Planck's work," Klein noted in passing, "but it was very hard to see why" (KLEIN, 1985, 248).

The task before us is therefore clear and straightforward. We must find out why. And doing so will turn out to be not so hard a task as Klein has supposed.

The first fact I want to emphasize in this respect is that it is truly almost impossible to understand the Wien displacement law, derived by Wien in 1893 (WIEN, 1893), if one ignores, as Klein apparently did, its inherent affinity and connection with Boltzmann's 1884 proof of the law for the energy density of black-body radiation,

$$u(T) = \int_0^\infty \epsilon_\nu d\nu = \sigma T^4, \tag{2.2}$$

[4] See the discussion in RQR, §2.1-2.

the so-called Stefan-Boltzmann law. Boltzmann's original derivation, like Wien's, is now half-forgotten; it has virtually disappeared from the modern literature. For Boltzmann's conclusion is now obtained directly, by calculating the integrand, the spectral energy density function

$$\epsilon_\nu(T)d\nu = \frac{8\pi\nu^3}{c^3}\frac{d\nu}{e^{\frac{h\nu}{kT}}-1}, \tag{2.3}$$

i.e., Planck's black-body radiation law. But historically the fact that the integrand of (2.2) was not known in 1884 must not be overlooked. For the one thing that jumps right at us when we consider the temporal and theoretical connection between (2.2) and (2.3) is that if Boltzmann's 1884 derivation of (2.2) was accepted as correct as of 1900, then Planck's 1900 derivation of (2.3) was logically and theoretically required to be *compatible* with Boltzmann's 1884 derivation of (2.2). And since Wien's displacement law, as remarked above, is, in actual fact, a *consequence* of Boltzmann's 1884 derivation of (2.2), then the energy quanta, which "follows from the extremely important Wien displacement law," must somehow be part and parcel of the theoretical link that connects Planck's derivation of (2.3) with Boltzmann's derivation of (2.2).

And so our problem must undergo a further reduction. We must, by the upshot of the previous paragraph, take a step further back, and look first not beneath Wien's 1893 displacement law, but beneath *Boltzmann's* 1884 proof of (2.2), to get to the roots of the energy quanta criterion in the thermodynamics of radiation.

For the necessity of quanta—to put my main point bluntly—was *there*, implicit in the thermodynamics of black bodies, sixteen years before Planck helped bring them to the surface.

So that's what we will do. We start by putting the matter in its proper context. The development that prepared the way, and allowed Boltzmann to derive (2.2) theoretically in 1884, unfolded roughly in the following stages.

1. Kirchhoff's Law.

    a) *Kirchhoff*. The reader will recall from RQR that the issue of finding the form of integrand of (2.2), i.e., the black-body radiation law, was historically connected to the issue of determining, theoretically, the form of Kirchhoff's universal function,

    $$K_\nu(T) = \frac{\epsilon_\nu}{\alpha_\nu}, \tag{2.4}$$

    ϵ_ν the radiating power per frequency ν, α_ν the absorbing power per frequency ν. Kirchhoff proposed that this function expresses a general law:

    > For rays of the same wavelength [or equivalently, of the same frequency] at the same temperature, the ratio of emissive power and

absorptive power is the same for all bodies. (KIRCHHOFF, 1859, 784)

The condition $\epsilon_\nu = \alpha_\nu$ means that a body takes up by absorption as much heat as it loses by emission, and expresses the existence of a certain temperature equilibrium for radiation; Kirchhoff designated bodies that do not emit any rays as being "completely black." The reader must note that this is a purely theoretical definition; the first black-body set-up was constructed only in 1895 (see (WIEN; LUMMER, 1895)). Kirchhoff's proof that the function $K_\nu(T)$ must exist posed, therefore, a challenge for experimentalists and theoreticians alike: Construct a body that satisfies his definition of blackness, and find the theoretical expression for the universal function $K_\nu(T)$.[5]

b) *Clausius*. R. Clausius, to whom we owe the first formulation of the second law of thermodyamics, proved in 1864 that Kirchhoff's function $K_\nu(T)$ cannot be independent of the surroundings, as claimed by Kirchhoff, without violating the second law; he proved further that no phenomena of heat radiation can produce motion that contradicts the second law (CLAUSIUS, 1864).

2. The Maxwell-Bartoli light-pressure. J. C. Maxwell proved in 1873 that if light were an electromagnetic phenomenon,[6] then pressure should result from the absorption or reflection of a beam of light of magnitude equal to its energy density (MAXWELL, 1873, 391). A. Bartoli, three years later, showed that Maxwell's theorem is actually a consequence of Clausius's: the existence of a radiation pressure numerically equal in amount to that derived by Maxwell is required by the second law of thermodynamics (BARTOLI, 1876). Neither Maxwell nor Bartoli, however, could, at that time, bring the result within the reach of methods of experimental verification by actual measurement.[7]

[5] The practical importance of the task was immediately evident. For it follows from (2.4) that for total radiation,

$$\frac{I}{A} = E, \tag{2.5}$$

I the energy of total radiation from a given body, E the same function for a black-body, A the absorbing power. And this function, if rewritten as

$$\frac{I}{E} = A, \tag{2.6}$$

yields a number that denotes the radiating power of a body in terms of its unit, the black-body. Hence the black-body plays here the role of unit of measurement. It is the concept that could make—and did make—out of the study of radiation a *bona fide* quantitative science.

[6] It is significant to remember that it is at this point in history that optical and electromagnetic phenomena were unified.

[7] Direct experimental verifications of the Maxwell-Bartoli light-pressure theorem started to appear in 1901; see (LEBEDEW, 1900) and (NICHOLS; HULL, 1903). The theoretical development that paved the way to the empirical determination of the pressure due to radiation is very well exposed in the latter work.

3. Stefan's fourth-power law. The fact J. Stefan deduced (2.2) from experiments in 1879, before the black-body became a laboratory instrument in 1895, is paradoxical only at first glance. Stefan's intuition and experience in thermal conductivity enabled him to identify a trend in the data he distilled from a host of radiation heat transfer experiments, and for a wide range of temperatures—Dulong and Petitt's (DULONG; PETIT, 1817), Provostaye and Desains' (PROVOSTAYE; DESAINS, 1846), Ericsson's (ERICSSON, 1872), and Draper's (DRAPER, 1878). It was on the basis of these results that Stefan proposed (2.2); his model flatly asserted, without qualification, that the total radiant emittance of *all* solid bodies, regardless of the color, is proportional to the fourth power of the body's absolute temperature (STEFAN, 1879). Today, of course, we know that for all bodies—real bodies—, $\frac{u(T)}{T^4} = \epsilon\sigma$, $\epsilon \leq 1$ the emissivity of the surface, which is a function of the temperature, radiation wavelength, and direction, and that in fact only for black-bodies that $\epsilon = 1$ so that (2.2) holds. We can, then, anticipate that Stefan's estimates were not, as stated, very accurate; the law was considered at first to be at best a rough approximation. It was accepted only after Boltzmann corrected its content by theoretical argument.[8]

Boltzmann's proof of the fourth-power law is far more significant than it has hitherto been acknowledged. It connected Kirchhoff's function with experience by showing, following out Clausius' result that the behavior of radiation is constrained by the second law of thermodynamics, that for black-body radiation surrounded by impenetrable walls of the temperature T, the Maxwell-Bartoli pressure implies a value for the energy density of the radiation that agrees quantitatively with the law previously discovered empirically by Stefan. And thereby, not only theory-mediated evidence for the light-pressure, and theoretical justification for (2.2), followed at one strike. Also precise mathematical relationships, which express part of the content of the equations of state of black-body radiation, *and thus must be constitutive of Kirchhoff's formula,* can be shown to deductively follow, as we shall see.

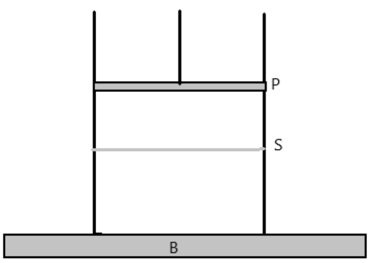

Figure 1 – Boltzmann's set-up

Here's the set-up. Boltzmann imagined a cylinder, with perfectly reflecting walls on the inside, open at one end, wherein a moving piston, the surface of which is also perfectly reflecting, encloses radiant energy (there's no gas in the cylinder) of the same temperature of the heat reservoir, a radiating black-body. If we increase the volume available to the radiation in the cylinder by moving the piston infinitely slowly, the energy density of the radiation will remain constant, for the heat radi-

[8] Compare these remarks with the historical sketch in (DOUGAL, 1979).

ation that is emitted by the body will complete the space so that it has the same energy density everywhere. This is exactly analogous to the isothermal expansion of a gas. Then, Boltzmann interposed a fixed, perfectly reflecting screen between the piston and the reservoir (see figure 1), eliminating the body emitting the radiation, and allowed the piston to move away from the screen, with a velocity that is supposed to be infinitely small in comparison with that of light. The energy density of the radiation will now decrease, first because the radiation is now occupying a greater volume, and second because work was expended by the pressure exerted by the radiation on the piston. This process is exactly analogous to the adiabatic expansion of a gas. So Boltzmann's problem, in short, is to apply Carnot's cycle to radiation, and see what follows.

Boltzmann's own presentation of his solution is very involved. To reproduce it here would serve no other purpose other than complicate our discussion. The proof given below relies on the report presented by W. Wien at the *Congrès International de Physique* in 1900 (WIEN, 1900, §II)—it leaves the logic of Boltzmann's reasoning intact and has the advantage of going to the point while avoiding the notions of "rays" and "aether" altogether.

Since the energy density of radiation u (intensity per unit volume) does not depend on its surroundings (Kirchhoff),

$$U = Vu. \tag{2.7}$$

Since the radiation in the cylinder has the same properties in all directions, the symmetry of the situation requires that just one third of the total radiation proceeds parallel to each side wall. By the Maxwell-Bartoli theorem, the pressure due to its vibrations on the surface of the piston must then coincide with one-third of the volume-density of the energy:

$$p = \frac{1}{3}u. \tag{2.8}$$

If a slow displacement of the piston changes the volume by dV, then the first law of thermodynamics,

$$dQ = dU + dW, \tag{2.9}$$

dQ if the differential for the radiation entering the cylinder,
dU the increase of internal energy,
dW the external work,

will become

$$dQ = d(Vu) + \frac{1}{3}udV = Vdu + \frac{4}{3}udV. \tag{2.10}$$

This differential equation is most straightforwardly solved by noting that an integrating factor on the right-hand side is $u^{-\frac{1}{4}}$ and, from $S = \frac{dQ}{T}$, $\frac{1}{T}$ on the left-hand side. Since the integration factors must be equal to each other,

$$T = \sqrt[4]{u},\ u = \frac{T^4}{b^4} = \sigma T^4. \tag{2.11}$$

That is the law proposed by Stefan.

Clearly the laws of thermodynamics are doing all the work here. It is the how that is not very obvious. Let's try to figure that out, then. Let's think through the thermodynamics very carefully.

Since the expansion is adiabatic, the first law requires that the process be reversible. And to make it reversible it ought to be entropy-preserving. But then, by the second law, it must now follow that the system *is always at equilibrium*—for infinitely slow adiabatic changes of density, therefore, *black-body radiation must remain black.* For infinitely slow adiabatic changes of density, the system must pass through a series of states that may be regarded, *each of them*, as a state of equilibrium.

But that of course means that at the end of the cycle we must go back to the exact same state. And that means that it must also be true that irrespective of the way in which the change of energy density proceeded—whether by adiabatic expansion, or by changes of temperature—, the spectral distribution, i.e., *the division of the energy itself into the various wavelengths per unit volume*, must also be identical for the same values of energy density. This fact, which is most important for Boltzmann's proof, was not, however, demonstrated, but was only assumed, by Boltzmann in 1884; it was Wien that proved it nine years later. Wien, by means of a simple, but extremely ingenious, argument, showed that this much is required by the second law of thermodynamics: Take, he suggested, two separate spaces, each of which filled with black-body radiation with the exact same energy density. In one, the energy density was obtained by a reversible adiabatic compression of free radiation enclosed in a space with mirror walls; in the other, the energy density in the space is the heat radiation that is emitted by a black-body of a definite temperature. If the spectral distribution were different in the two spaces, then it would be possible to construct, by means of colour filters, with the radiations in the two spaces, a perpetuum mobile of the second kind (see (WIEN, 1893, §1) to be convinced). And since that is out of question, it is, then, with a certainty that coincides with that of the second law that we can now say that the spectral composition of the black-body radiation that one gets by changing its energy density by adiabatic expansion in a space enclosed by mirror walls (please note that the radiation is free in the space–no contact with a body), *must be precisely such as that we would have obtained by radiation in equilibrium with ordinary matter of a lower temperature.*

And that is an outstanding conclusion. For it tells us, to start with, since we know that it is a consequence of the Doppler principle that radiation, incident on a wall during its motion, will have its frequency altered, that if we use the Doppler effect to calculate just how the individual wavelengths change while the radiation is compressed, adiabatically and infinitely slowly, by mirror walls, we can deduce from this the exact way that the spectrum of a black-body changes with the temperature. For we know by the second law that the states of the radiation in these situations must be exactly the same.

And that is exactly what Wien did, in §2 of his 1893 work. He determined the way that the spectrum of the radiation changes when its individual components are compressed, adiabatically and infinitely slowly, and reflected from moving mirror walls, with the help of the Doppler effect, and then found the way that the spectral composition of black-body radiation varies with the temperature. It is this: each element of the spectrum of black-body radiation varies with T in such a way that the product of T and λ always remains constant. This is the Wien displacement law.

Exactly verified later—for Wien and Lummer prepared, let us remember, the first black-body only in 1895 (WIEN; LUMMER, 1895)—by experiments (see figure 2, below).

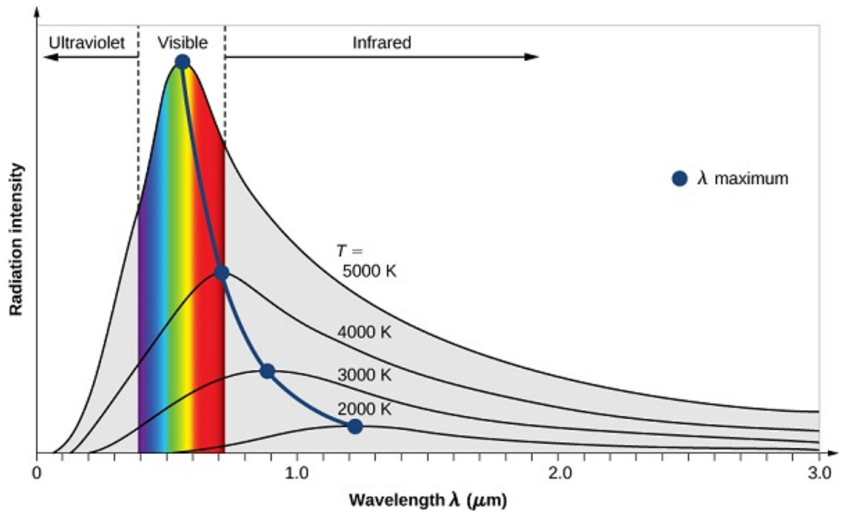

Figure 2 – The traditional spectrum curve of black-body radiation. Under a temperature alteration, every wavelength suffers a displacement such that the product of the temperature and the wavelength remains constant. Figure source: (LIBRETEXTS, 2022).

With all of this under our belts let us now reconsider the passage from Martin Klein's partially quoted above.

> Wien's displacement law was surely one of the essential properties of the equilibrium distribution—it had, for example, played an essential part in Planck's work—but it was very hard to see why. For the derivation of this law, as given in Planck's book and elsewhere, made use of the Doppler effect on the frequency of a ray reflected from a moving mirror, and other concepts

> that seemed to have very little to do with black-body radiation. (KLEIN, 1985, 248)

It is a matter of course that failure to see the significance of the Wien displacement law would fatefully lead, given that relation (2.1) follows, as Planck said in 1900 (the fact cannot be sufficiently repeated), "immediately from the extremely important Wien displacement law," to a failure to see the significance of the concept of energy quanta. But we can claim some progress now. Now we can see that and why the Doppler effect—which holds, let us bear in mind, for all kinds of vibrations—has everything to do with black-body radiation.[9]

But it is quite evident that to achieve the clarity we seek we need more than this. Theoretically, if not historically, the link between Boltzmann's, Wien's, and Planck's, laws remains vague. This whole business must be put in a more exact and straightforward form if we are going to single out the facts that establish that the energy quanta follows from the laws of thermodynamics as "a mathematical criterion that the individual processes or their theoretical models have to satisfy"—a criterion that has been made inevitable since, because not theoretically separable from, Boltzmann's 1884 derivation of (2.2). And what I now categorically assert is the following: There *is* a way of putting this whole business in more exact and straightforward a form. There *is* a way of going, from Boltzmann's derivation of the fourth-power law, to Wien displacement law, and therewith, to Planck's energy elements, directly, without lengthy calculations, and almost wholly avoiding electrodynamic considerations. For Boltzmann's proof of (2.2) admits a further conceptual reduction, suggested by the fact that any combination of the parameters that remain constant during an isentropic process defines an invariant of the system. And therewith it can be legitimately concluded that Boltzmann's 1884 derivation of (2.2) boils down to the proof of the following statement:

$$\frac{u(T)}{T^4} = \text{adiabatic invariant.} \tag{2.12}$$

From this, many important conclusions can be shown to follow.[10]

[9] It is significant to note that C. Doppler described the effect in his 1842 (DOPPLER, 1842)—decades before Maxwell unified optics and electromagnetism in the groundbreaking (MAXWELL, 1873).

[10] Etymologically the word "adiabatic" means "not passing through." The concept of adiabatic process dates from Clausius's groundbreaking *Reflections on the Motive Power of Fire,* first published in 1824; the word, however, was used first by W. J. M. Rankine in (RANKINE, 1858).

If we consider (2.8) and (2.11) in accordance to the Maxwell relation

$$\left(\frac{\partial S}{\partial V}\right)_T = \left(\frac{\partial P}{\partial T}\right)_V \tag{2.13}$$

we automatically obtain the following expression for the total entropy of black-body radiation:

$$S = \frac{4}{3} bVT^3. \tag{2.14}$$

The fact that this must, by the arguments above, be a constant, will yield

$$VT^3 = \text{adiabatic invariant}. \tag{2.15}$$

All that we now postulate is that it is possible to describe a radiation field at equilibrium in an enclosure of volume V by a set of standing waves. Then, since the thermodynamics of black-body radiation is independent of the shape of the enclosure, we can assume a shape convenient for calculation—we take

$$V = L^3. \tag{2.16}$$

Next we consider the fact that the condition for a wave to be standing is that its frequency ν must satisfy

$$\left(\frac{\nu L}{c}\right)^2 = l^2 + m^2 + n^2, \tag{2.17}$$

c is the speed of light, l, m, n integers that denote the number of nodes of the standing wave along the respective edges of the cube. They have the following invariance property (BUCKINGHAM, 1913; HAAR; WERGELAND, 1966):

If the enclosure changes its volume while retaining its shape, then the nodal planes of each wave will be displaced, but each set of numbers l, m, n stays constant.

This means that if the volume V of the enclosure is enlarged or diminished by a slow movement of the walls, the dimensions of the waves ought to change in the same linear proportion as those of the enclosure. The reader will note that this is nothing but to say that the entire system of waves and the enclosure always remain geometrically similar to itself; the number of waves present must therefore also be an invariant of the system.

Now we differentiate (2.17) at constant l, m, n. We find that the proper vibrations in the enclosure will undergo a displacement

$$\frac{\delta\nu}{\nu} = -\frac{1}{3}\frac{\delta V}{V}, \tag{2.18}$$

and so, for each independent eigenfrequency ν,

$$V\nu^3 = \text{adiabatic invariant.} \tag{2.19}$$

Our further deductions are a systematic application of the following extremely important fact: *multiplying or dividing any two known adiabatic invariants generates other forms of the adiabatic invariant.* We divide, then, (2.19) by (2.15), and get

$$\frac{\nu}{T} = \text{adiabatic invariant.} \tag{2.20}$$

Or equivalently, since $\nu = \frac{c}{\lambda}$,

$$\lambda T = \text{adiabatic invariant.} \tag{2.21}$$

The reader must note that this is nothing but the Wien's displacement law, here obtained in the most straightforward way. With this, it is now a few short steps to the fundamental result of this section.

From the equations (2.11) and (2.14),

$$S \propto \frac{U}{T} = \text{adiabatic invariant.} \tag{2.22}$$

From the principle of the independence of the elements composing the spectrum, the fact that the entropy must be the sum of the contributions of each frequency, so that its resolution,

$$S = \sum s(\nu/T), \tag{2.23}$$

contain terms that depend on the frequency only through the adiabatic invariants, and accordingly, for the resolution of the energy,

$$U = \sum \epsilon(\nu/T), \tag{2.24}$$

where ϵ is the thermal energy of each eigenvibration.

And now, *from all of this*, inserting (2.24) and (2.23) into (2.22), and dividing the latter by (2.20), the conclusion:

$$\frac{\epsilon}{\nu} = \text{adiabatic invariant.} \tag{2.25}$$

This must hold for each eigenvibration. Relationships of the kind

$$\frac{\epsilon}{\nu^2}, \frac{\epsilon}{\nu^3}, \text{ etc.} \tag{2.26}$$

are forbidden for they would violate the second law of thermodynamics.

The claim made above has therefore been proved. The fact that the line of reasoning I adopted here relies on the notion of adiabatic invariance, introduced in physics by

Paul Ehrenfest in 1916, deserves comment. For I was reassured of its essential correctness after recalling that Ehrenfest proposed his adiabatic principle as "a generalization of Planck's idea of quantization," in connection with the fact that Ehrenfest made, in 1923, in passing, the following statement, but never, insofar as I could verify, explained it:

> It was Boltzmann's radiation law and Wien's displacement law, or more precisely the mystery which was concealed behind the elegant electrodynamic and thermodynamic derivations of these laws, which baited the path that led to the Adiabatic Principle. (EHRENFEST, 1923, 543)

I hope that the above presentation has done enough to demystify the matter.

Now, to the consequences. For one cannot understand from these invariants whatever one wants.

2.2 The Verdict of Thermodynamics

The condition (2.19), which is equivalent to

$$\lambda = \frac{constant}{\sqrt[3]{V}}, \tag{2.27}$$

and equivalent, by the second law, to (2.21),[11] and that must hold true for any value of λ in the spectrum, implies that when the various wavelengths, within any interval $d\lambda$, and thence the energy density per unit volume, and thence the spectral composition, change in virtue of the movement of the walls, there is—and this is just a restatement of a previous point—no occasion for the destruction of waves, or for the formation of new ones. The number of waves is invariant (see figure 2 again) and each must remain perfectly stationary. The changes in the spectral energy distribution that occur by reflection from the moving walls, or equivalently, by changing the temperature of the radiating black-body, must therefore result only from the individual changes of the ν already present.

But let me be more explicit here, for we have arrived at the crux of the matter. Unlike gas molecules, which exchange energy (thermalize) when they interact, the energies of the waves in the cavity do not equalize if one wave superposes the other.[12]

[11] Formula (2.27) is also sometimes called Wien displacement law.

[12] It is instructive to note that in the case of gas molecules, energy is transferred from one molecule to another until the Maxwell-Boltzmann distribution sets in. Part of this whole mystery is the fact this is not true for black-body radiation; the energies of the waves, unlike those of the molecules, are independent of one another—even though we may well have a gas of molecules in equilibrium, by emission and absorption, with the radiation in the enclosure. This mystery may be given a special twist by noting that there is a striking formal similarity between the curve of chromatic distribution of black-body radiation and that of the the Maxwell-Boltzmann distribution law of velocities.

This is what I mean when I say—it is a fact of experience—that the waves are independent. And if each eigenvibration ν in the cavity, furthermore, *changes* independently, and a definite energy ϵ is associated to each ν, then it must follow, quite trivially, that each ϵ changes independently. And this—as Einstein pointed out, but did not care to explain, in (EINSTEIN, 1910)—*flatly contradicts the principle of superposition*. [13]For it is an unavoidable consequence of Maxwell's equations that the energy of an electromagnetic wave is stored in the *amplitudes* (field strengths) of the electric and magnetic fields, extending to all directions, so that the slightest, infinitesimal perturbation in the radiation field implies a change in the energy that is propagated to ever larger volumes of space, the intensity then changing proportionally to $\frac{\partial u}{\partial t}$. But this cannot be true of black-body radiation. By the above invariants, its energy states are not "interferable."

And so it will now appear that our deduction of the black-body radiation invariants has apparently enabled us to detect a perplexing and effectively hidden *inner contradiction* in the nineteenth-century method of deducing Wien's displacement law which makes it depend on the equations of the electrodynamic field. Our deduction has shown that the argument starts by assuming the validity of Maxwell's equations and, thanks to some curious thermodynamic magic, concludes a result that implies their negation. The reader will remember that when Martin Klein said that he could not understand the Wien displacement law, it is the electrodynamics that it involves that he blamed for his confusion; if one is able, however, as we have been, to strip the deduction of the Wien displacement law from (most of) its electrodynamic accouterments, and think through the thermodynamic facts that it implies, then we can understand what the law means. We can see the facts behind Planck's statement that "the energy elements [....] follow immediately from the extremely important Wien displacement law." For the Wien displacement law is nothing but a consequence of the application of the Carnot cycle to black-body radiation enclosed within movable reflecting walls, and from this, it "follows immediately" that changes in the energy density of the radiation must be such that the frequencies of the radiations in the enclosure always all alter in the same proportion,

$$\frac{\epsilon'}{\nu'} = \frac{\epsilon}{\nu}, \tag{2.28}$$

$$\nu' \sqrt[3]{V'} = \nu \sqrt[3]{V}, \tag{2.29}$$

each of the successive states the system passes through being a state of equilibrium.

A light ray can split. But the energy of black-body radiation cannot split without a change in the frequency.

[13]Einstein made the remark in connection to formula (2.45) (see (EINSTEIN, 1910, 528)), which encodes precisely the invariants we're considering.

Now what *is* all of this, if not saying, indeed, that *only mutually independent quantities*, only *packets of energy,* are capable of existing—what is all of this if not saying that radiation energy is so constituted that it "is not distributed continuously over ever-increasing spaces" (EINSTEIN, 1905a, 87) but rather "consists of a finite number of energy quanta" that "move without dividing" (ibid) and so, in the case of the reversible adiabatic compression, must "reflect undivided" by the walls and in the equivalent case of equilibrium with a body of a definite temperature, must "be absorbed or generated only as a whole" (ibid)? What is all of this if not to say, with Einstein, that "monochromatic radiation behaves thermodynamically *as if* it consisted of mutually independent energy quanta" of magnitude equal to a constant times the frequency (EINSTEIN, 1905a, 97, my italics)?

We saw in RQR that Planck's 1900 proof of (2.3) is a legitimate application of electrodynamics and thermodynamics. And we saw in §1 that if Planck's proof were to conform to Boltzmann's proof of the fourth-power law—if it were to conform, indeed, *to the second law of thermodynamics*—, then for each eigenfrequency of the radiation, one must have $\epsilon = h\nu$, period. But then, if the energy elements, on the one hand, "follow immediately from the extremely important Wien displacement law," and on the other, show that "neither mechanics nor electrodynamics can claim exact validity," no other option is before us, save to admit that it follows "immediately from the extremely important Wien displacement law" that "neither mechanics nor electrodynamics can claim exact validity." And that's that. That's the mystery Ehrenfest talked about.

And I now submit that it is against these considerations that the following remarks by Einstein in his *Autobiographical Notes* are most intelligibly interpreted.

> Planck got his radiation-formula if he chose his energy-elements ϵ of the magnitude $\epsilon = h\nu$. The decisive element in doing this lies in the fact that the result depends on taking for ϵ a definite finite value, i.e., that one does not go to the limit $\epsilon \to 0$. This form of reasoning does not make obvious the fact that it contradicts the mechanical and electrodynamic basis, upon which the derivation otherwise depends. Actually, however, the derivation presupposes implicitly that energy can be absorbed and emitted by the individual resonator only in quanta of magnitude $h\nu$, i.e., that the energy of a mechanical structure capable of oscillations as well as the energy of radiation can be transferred only in such quanta—in contradiction to the laws of mechanics and electrodynamics. [...] Reflections of this type made it clear to me shortly after 1900, i.e. shortly after Planck's trailblazing work, that neither mechanics nor electrodynamics could (except in limiting cases) claim exact validity. (EINSTEIN, 1949, 45, 51, 53)

Planck said in plain language that the energy quanta criterion is a consequence of the Wien displacement law. But the "form of reasoning" from which the latter was commonly derived (Doppler effect) "does not make obvious" why it—and by implication, the energy quanta—"is one of the essential properties of the equilibrium distribution," as Klein remarked. This form of reasoning describes a "mystery which was concealed behind the elegant electrodynamic and thermodynamic derivations" of Boltzmann's radiation law and Wien displacement law (Ehrenfest), namely: That the equilibrium condition between matter and radiation, as implied by the laws of thermodynamics, are not satisfied by, and thus, *must be independent of*, the laws of mechanics and electrodynamics.

Now I'd like to make the point that even though the adiabatic invariance of black-body radiation and other periodic phenomena were explored by Paul Ehrenfest, most notably in (EHRENFEST, 1916), Einstein's remarks in the 1911 Solvay Conference, see (SOLVAY; LANGEVIN; BROGLIE, 1912, 450), show that the man had such invariants firmly in his hand before Ehrenfest's investigations started.[14] This evinces a unique understanding and ability to scent out that which leads to the fundamentals of physical theory, and accounts for the otherwise inexplicable fact that Einstein concluded "shortly after 1900," i.e., shortly after Planck pointed out that the energy elements "follow immediately from the Wien displacement law," that "neither mechanics nor electrodynamics could claim exact validity." Which fact explains, by its turn, why Einstein directed his scientific efforts "shortly after 1900," first, *to the foundations of thermodynamics*, in the attempt to put the concepts "entropy," "temperature," and therewith, "thermal equilibrium" reached by "infinitely slow processes" in firm footing (EINSTEIN, 1902; EINSTEIN, 1903), and next, to the development of a general molecular theory of heat based on it, *in the attempt to blur the contrast between a molecule and an arbitrarily extended physical system so that the thermal equilibrium between them could be described without the kinetic hypothesis* (EINSTEIN, 1904). None of this accidental; everything was calculated for the final outcome. For it is well to remember, as remarked in the beginning of this essay, in connection to these considerations, that even though the orthodox fiction that says that the energy quanta made its first appearence as an *ad hoc* hypothesis in a self-contradictory argument was made plausible by the Kuhnian thesis of scientific revolutions, the conspicuous, if little-acknowledged, truth is that a concept of this kind would never be taken seriously by a very cautious, very conservative, and highly critical community of scientists—let alone be the catalytic agent of a revolution in the foundations of physics. No; let us not commit the fallacy of underestimating the

[14]It is worth pointing out that Boltzmann had proved as early as 1866 that, for a gas subjected to a reversible adiabatic change, $\frac{E}{\nu} = \text{constant}$; Clausius and Szily reached later the same conclusion. In 1902, Lord Rayleigh, in an impressive study that generalizes the Maxwell-Bertoli radiation pressure to general vibrations (RAYLEIGH, 1902), proved that for a simple pendulum that consists of a bob and a string, if its length is altered infinitely slowly, the energy of the pendulum increases in exact proportion to the frequency—i.e., we have, once again, $\frac{E}{\nu} = \text{constant}$.

stringency of scientific criteria. Einstein surely didn't. For if the energy quanta is, as he knew it to be, a *criterion*, a theoretical necessity required by the laws of thermodynamics, then it ought to be presented as such; it ought to be stripped from all electrodynamic and mechanical impregnation, and shown to follow, with unparalleled clarity, from unobjectionable premises, in an unobjectionable argument—purely thermodynamically. The development of Einstein's 1902-1904 works, in broad terms, is essentially and most notably the consciously and zealously forging of the theoretical weapons to mobilize and validate the 1905 revolution.[15]

Now it is natural that the literature on Einstein's 1902-1904 works did not read them in this way.[16] The realization transcends, in depth and breadth, the limitations imposed by the discontinuity myth. But now we can see further and farther. For now we know, to start with, that it follows from the second law of thermodynamics, as applied by Boltzmann in his 1884 proof of (2.2), that the states of black-body radiation in free space are equivalent to states forced upon the radiation by a ponderable body of a definite temperature. Now we know that the Wien displacement law is nothing but a direct consequence of Boltzmann's thermodynamic reasoning, and therefore, since the criterion of the energy quanta is, by its turn, a direct consequence of the Wien displacement law, we know also that the energy quanta criterion must hold true *both of free electromagnetic radiation and of the interaction between light and matter,* and so commentators who have called Einstein "extraordinarily bold" for claiming that this must the case are confusing his boldness with the validity of the second law of thermodynamics.[17] And finally, now that we know all of this, we can reflect an instant, and see that the same reasoning indicates that there has to be a certain conformity between radiation and the ponderable bodies, a certain *likeness,* which necessitates the numerical coincidence. Otherwise the consequences of the second law, and the law itself, cannot be understood. And then, once all of this sinks in, we will see that the practical task of deriving the energy quanta as a thermodynamic criterion, that is, avoiding all mechanical and electrodynamic nuisance, requires finding, first and foremost, *a purely thermodynamic link between matter and radiation*. And then, with this latter fact in the back of our minds, we can look up in Einstein's pre-1905 works, and see that the last result of (EINSTEIN, 1904) is not, as commentators have claimed it to be, a mistake on Einstein's part, a misapplication of the kinetic theory of heat. Instead it is the prelude to the quantum revolution.

With the expression obtained for energy fluctuations which, according to the

[15] Including the laws of Brownian movement which, after having been experimentally verified by Perrin in 1908, as we saw in RQR, gave the revolution momentum by putting the physicists on the intolerable position of having to face the henceforth inevitable ultraviolet catastrophe.

[16] But see the excellent analyis provided by Uffink (UFFINK, 2006).

[17] (PAIS, 1982b, 377-378) is just one of the many examples that can be cited.

expression for the entropy defined in (EINSTEIN, 1903, §3, §4), and (EINSTEIN, 1904, §1), must exist,[18]

$$\overline{E^2} - \overline{E}\overline{E} = \overline{\epsilon^2} = kT^2 \frac{d\bar{E}}{dT}, \tag{2.30}$$

and with the Stefan-Boltzmann law

$$\overline{E} = bVT^4, \tag{2.31}$$

and noting that for black-body radiation enclosed in a cavity of a size of magnitude comparable to the wavelength, (2.30) reduces to

$$\overline{E^2} = \overline{\epsilon^2}, \tag{2.32}$$

one gets, by differentiating (2.31) with respect to T, inserting the result on the right-hand side of (2.30), and then taking the square of (2.31) and inserting it, following (2.32), on the left-hand side of (2.30), the formula

$$VT^3 = 4\frac{k}{b}. \tag{2.33}$$

The reader will note that this is nothing but the adiabatic invariant (2.15) *with the constant on the right-hand-side now fully determined.* Writing it as

$$\sqrt[3]{V} = 2\sqrt{2}\frac{\sqrt[3]{\frac{k}{b}}}{T}, \tag{2.34}$$

this expression must give us, by its relation with the invariant (2.21), the value of the order of magnitude of the λ that corresponds to the temperature T. Using for the constant k the value obtained from the kinetic theory of gases, and 7.06×10^{15} for the Stefan-Boltzmann constant b, Einstein obtained a result for λ and T that is in agreement, in order of magnitude, with the values observed in experiment. To his friend Habitch, he exultantly wrote: "I have now found the relationship between the magnitude of the elementary quanta of matter $[N]$ and the wavelenghts of radiation in an exceedingly simple way" (Einstein to Conrad Habitch, 15 April 1904). To his readers, he announced:

[18]The mentioned expression is formula (2.44). It is noteworthy that Einstein remarked, in the introduction of (EINSTEIN, 1904), that this formula is "completely analogous to the expression found by Boltzmann for ideal gases and assumed by Planck in his theory of radiation" (mind the use of the word "assume" in the case of Planck's). For it is very significant that Einstein, unlike Boltzmann, arrived at (2.44) in a very general way, without the use of the kinetic hypothesis, so that the formula is microphysics-independent: a true blue thermodynamic expression. It is on the basis of such derivations that Einstein, from LQP onwards, called (2.44), in analogy to the energy principle (first law), "Boltzmann's principle." And even though it is perfectly true that the statistical interpretation of the second law was not accepted in general as of 1905, the confirmation, by Perrin's 1908-1911 experiments, of the laws of Brownian motion predicted by Einstein in (EINSTEIN, 1905c), turned, as we saw in RQR, the tide drastically in favor of the statistical interpretation, giving to the Boltzmann principle the empirical warrant for its principled status. The reader can find details of the impact of Perrin's experiments in the controversy, which was still in the forefront as of 1900, concerning whether the second law holds absolutely or statistically, in the excellent investigation by Smith and Seth (SMITH; SETH, 2020). This work makes meticulous use of all source material and a very sophisticated analysis of the role of theory-mediated measurement and impact of empirical evidence in the fundamental debate over the discrete or granular structure of matter.

> Considering the broad generality of our assumptions, I believe that this agreement must not be ascribed to chance. (EINSTEIN, 1904, 362).

The fact Einstein's very next paper is LQP must not be ascribed to chance either.

2.3 What Albert Einstein Started

That the validity of the laws of thermodynamics should require a certain similarity in the structure of radiation and the various ponderable bodies is not surprising. For, without some conformity of sorts, we cannot even expect to understand what it means for radiation and ponderable bodies to be in thermal equilibrium. And it is to the condition of thermal equilibrium, and consequently, to the maximum of entropy, that black-body radiation corresponds.

And so, in view of this, and since, as Einstein pointed out at the very outset of LQP,

> There exists a profound formal difference between the theoretical conceptions physicists have formed about gases and other ponderable bodies, and Maxwell's theory of electromagnetic processes in so-called empty space.

it is not at all surprising that a consistent application of the laws of mechanics and electrodynamics entails that there should actually be *no equilibrium* between matter and radiation. There should be no black-bodies. The energy of the radiation, according to these theories, must run increasingly to the shortest wavelengths, so that we obtain, in the limit,

$$\int_0^\infty \epsilon_\nu d\nu = \frac{8\pi k}{V} \int_0^\infty \nu^2 d\nu = \infty. \tag{2.35}$$

It is not at all surprising, in other words, that the equations of mechanics and electrodynamics contradict, as we saw above, the requirements of thermodynamics.

Yet the thing, as Newton wrote in the *Scholium* of the *Principia,* is not altogether desperate. We have some arguments to guide us.

We know enough at this point about the properties of heat radiation to see the importance of the fact Einstein that devoted LQP, §3 to a review of the facts "contained in a famous study by Mr. Wien [...] presented here only for the sake of completeness." His remarks are scant in the extreme, and, in themselves, unintelligible; Einstein clearly took for granted his readers' awareness of the factual content of Boltzmann's and Wien's laws of radiation. And the fact that this section of LQP if never mentioned by the modern literature, on top of the fact that the great Paul Ehrenfest called their factual content "mysterious," on top of the oversights and misreadings we discussed in the previous sections, all indicate that it is just this awareness that we have lost. But our study in §1-2 will now have served its main purpose by enabling us to identify, and therewith, take due note of the significance of, all the fundamental facts that Einstein has left between the lines.

In view of our discussion in §1, we can summarize Einstein's whole point in LQP, §3 in a few lines. One deduces, from the principle of independence of the elements composing the spectrum, together with the invariant (2.22), that

$$\frac{\partial s_\nu}{\partial \epsilon_\nu} = \frac{1}{T} = \text{constant.} \tag{2.36}$$

The invariance entails that this quantity is the same for all ν in V. Among all conceivable distributions of energy, that of the black-body is characterized by the fact that the radiations of all frequencies have the same temperature; this alone guarantees that

> one can determine the law of black-body radiation from the function s_ν, and, vice versa, the function s_ν can be determined by integrating the former. (EINSTEIN, 1905a, 93)

If one is clever, this can be taken advantage of. Einstein considers Wien's law,

$$\epsilon_\nu = \alpha \nu^3 e^{\beta \frac{\nu}{T}}, \tag{2.37}$$

as the low-frequency limit of (2.3), i.e., with the constants α and β as determined by Planck's derivation of (2.3):

$$\alpha \approx \frac{8\pi}{c^3} h; \; \beta = \frac{h}{k}. \tag{2.38}$$

But Einstein does not, as Newton would put it, "mingle conjectures with certainties."[19] He is implicitly warning his readers that his conclusions are not tainted with the assumptions Planck used in his derivations of (2.3) when he stated, in LQP, §2, that the fact the numerical values of the constant N Planck obtained *via* his derivation of (2.3) agrees numerically with the values of N obtained *via* other methods is enough evidence that establishes that the values of the constants in the black-body radiation law are "to some extent independent of [Planck's] theory of black-body radiation" (EINSTEIN, 1905a, 90).

And so he can confidently proceed. From (2.36),

$$\frac{\partial s_\nu}{\partial \epsilon_\nu} = \frac{1}{T} = -\frac{1}{\beta \nu} ln(\frac{\epsilon_\nu}{\alpha \nu^3}) \tag{2.39}$$

[19]Compare with footnote 2. The quoted words refer to the following passage from Newton's report "New Theory of Light and Colours," communicated to the Royal Society on February 6, 1761:

> But to determine more absolutely what light is, after what manner refracted, and by what modes or actions it produces in our minds the phantasms of colors, is not so easy. And I shall not mingle conjectures with certainties. (NEWTON, 1672, 3085)

and by integration,

$$s_\nu = -\epsilon_\nu [ln(\frac{\epsilon_\nu}{\alpha\nu^3}) - 1]. \tag{2.40}$$

Now let V be the volume of the enclosure that may be diminished or enlarged by a displacement of the walls. Since the state of the radiation has the same properties in all directions, its total entropy, in that enclosure, is evidently

$$S = -E[ln(\frac{E}{V\alpha\nu^3}) - 1]. \tag{2.41}$$

Now if the enclosure is compressed to a volume V_0, the relation between the states of the radiation before and after the compression is

$$S - S_0 = \frac{E}{\beta\nu} ln(\frac{V}{V_0}), \tag{2.42}$$

or equivalently,

$$S - S_0 = \frac{R}{N} ln \left[\left(\frac{V}{V_0} \right)^{\frac{N}{R}\frac{E}{\beta\nu}} \right]. \tag{2.43}$$

To interpret this expression, we must only recall the fact proved in (EINSTEIN, 1903, §6), (EINSTEIN, 1904, §1), that the following expression for the entropy of a system holds for systems that undergo—and the reader should now recall all the facts we discussed in §1—*both thermal and adiabatic changes of state*:

$$S = \int \frac{dE}{T} = \frac{R}{N} \, lnW(E). \tag{2.44}$$

$W(E)$ denotes the probability of the instantaneous energy state of the system.[20] Comparison yields

$$W(E) = (\frac{V}{V_0})^{\frac{N}{R}\frac{E}{\beta\nu}}. \tag{2.45}$$

Now the thing we have to remember, which Einstein said explicitly in LQP, §3, and without which we can hardly understand the *grounds* of his main result, is that *compression of radiation between reflecting walls does not change its entropy*. For it is from this that it inexorably and ineluctably follows, from the above expression, that

$$\frac{N}{R}\frac{E}{\beta\nu} = \text{constant}. \tag{2.46}$$

That is: what we have here is an adiabatic invariant. And now we may also recall that it is a property *of the ideal gas* that

$$\frac{1}{N} S = C_V lnT + \frac{R}{N} lnV = \text{adiabatic invariant} \tag{2.47}$$

[20]See footnote 17.

—so after an infinitely slow expansion, the relation between the states is

$$S - S_0 = \frac{R}{N} ln \left[\left(\frac{V}{V_0} \right)^N \right], \tag{2.48}$$

and since expression for the entropy (2.44) evidently also holds for this system,

$$W(E) = (\frac{V}{V_0})^N. \tag{2.49}$$

Now we know that the adiabatic compression of the gas, just like that of the black-body radiation, does not change its entropy. And we know that the state of the black-body radiation, that of the maximum of entropy, *is not at all disturbed by the presence of a ponderable body or of an ideal gas in the enclosure.* And so it must now follow, not, as commentators have believed, from an analogy with the kinetic theory of heat, *but from the equilibrium condition*, that (2.45) must be set equal to (2.49), and so, it is with the degree of certainty that coincides with that of the second law of thermodynamics that we must now conclude that for free radiation, and for matter in equilibrium with radiation in an enclosure, the quantity

$$E = Nh\nu \tag{2.50}$$

is an invariant. We must have, by the upshot of a simple, direct, and purely thermodynamic argument, that "monochromatic radiation behaves thermodynamically as if it consisted of mutually independent energy quanta of magnitude [$h\nu$]" (EINSTEIN, 1905a, 97); we must have that the energy of radiation "is not distributed continuously over ever-increasing spaces, but consists of a finite number of energy quanta that are localized in points in space, move without dividing, and can be absorbed or generated only as a whole" (EINSTEIN, 1905a, 87). And since we know that in a reversible adiabatic compression of the walls, or equivalently, in changes of temperature, we make the black-body radiation, or equivalently, the system radiation plus gas, pass through a sequence of states that must be regarded, each of them, as a state of equilibrium, we must conclude, further, that the above invariance not only goes against the principle of superposition, and therefore, against the equations of electrodynamics, but also, against the mechanical principle that requires the system to change its state by passing through a continuous series of energy states, and therefore, against the equations of mechanics.

And so we reach, *in a deductive way,* the conclusion that whereas the equations of electrodynamics and mechanics entail that thermal equilibrium between matter and radiation does not exist (ultraviolet catastrophe), the laws that rule the thermal equilibrium between matter and radiation entail that the equations of electrodynamics and mechanics cannot claim exact validity (energy quanta). By the principle theory/constructive theory distinction, it is the principle theory that constrains the choice of possible constructive theories; it is thermodynamics that is the judge. What Einstein showed, essentially, in LQP, in the most direct, and logically impeccable, manner, is therefore, as I have remarked above, that the fundamental hypotheses of mechanics and electrodynamics do not satisfy the requirements of thermodynamics. And that is why they have to go.

3 The Dual Dynamical Foundation of Orthodox Quantum Mechanics

—written with Dr. Ricardo Correa da Silva

> Is discontinuity destined to reign over the physical universe, and will its triumph be final? Or will it finally be recognized that this discontinuity is only apparent, and a disguise for a series of continuous processes?
>
> H. Poincaré, "L'hypothèse des Quanta," 1913

3.1 Introduction

This paper is a comprehensive investigation of the dynamical foundations of the formalism of orthodox quantum mechanics. It applies the method of integrating history, mathematical analysis, and philosophy of physics to make maximal sense of foundational questions related to the dynamics of quantum mechanics. Mathematical and conceptual analyses are applied to address concerns relative to the history of quantum mechanics, and historical insights are then used to treat present-day concerns about the foundations of quantum mechanics. The most novel, most essential, of our results is the discovery of the existence of a dual action functional from which both the dynamical postulates of orthodox quantum mechanics necessarily follow. This action functional is proved to be the simple logical condition that necessitates the equivalence of Matrix and Wave Mechanics; we argue, therewith, for the following implication: The dual dynamics of von Neumann's "unified" quantum mechanics, process 1 and process 2, does not enter into the formalism of quantum mechanics mysteriously or *ad hoc*—the dual dynamics is completely determined by the dual action.

The investigation is organized as follows. We start by solving a historical problem, namely the conceptual origins of the canonical commutation relations in both Matrix Mechanics (§3.2.1) and Wave Mechanics (§3.2.2). The solution casts the problem of the equivalence of these theories, first examined by Schrödinger in 1926, in a form that has not been heretofore considered by the literature; we discuss its features in §3.2.3. We attack this problem, and find its solution, in §3.2.4: The equivalence is logically necessitated by the formal conditions of a dual action functional whose different parts have independently entailed, as a matter of historical and mathematical fact, the emergence of the canonical commutation relations in Matrix and Wave Mechanics. Therewith, consequences of theoretical interest follow, certified by the following fact: The equivalence of Matrix and Wave Mechanics is stated by von Neumann to be the logical foundation of his *The Mathematical Foundations of Quantum Mechanics*, and this work is acknowledged by the literature to have consolidated the formalism of orthodox quantum mechanics.

The subsequent discussion in §3.3.1 is devoted to a critical analysis of the arguments used by von Neumann to "unify" Matrix and Wave Mechanics in Hilbert space. This analysis leads directly to a logical understanding of von Neumann's formulation of the measurement problem that does not conform with the official understanding of the problem in terms of the so-called problem of outcomes (MAUDLIN, 1995). But then, since the problem of outcomes has given rise to much dispute in the foundations of quantum mechanics, the cogency of our results will seem to depend on our ability to explain away such discrepancy. We solve this problem in §3.3.2. The popular view that von Neumann proposed process 1 only to solve the measurement problem will be refuted. A new understanding of von Neumann's Impossibility Proofs will thereby emerge, evincing that contemporary charges against von Neumann's dignity as a mathematician, according to which his argumentation is at points "*ad hoc*," "arbitrary," and "silly" (John Bell), are unfounded.

3.2 The Dynamical Duality Between Matrix and Wave Mechanics

3.2.1 The Dynamical Origin of the Quantization Rules, Part I: Matrix Mechanics

Matrix Mechanics postulates that all the conclusions of the theory rest upon the condition

$$pq - qp = -i\hbar. \tag{3.1}$$

Today we refer to it as the "canonical commutation relation," or "Heisenberg commutation relation;" the latter attribution, however, is placed in proper historical context by (BORN, 1955), who reminisced: "I shall never forget the thrill I experienced when I succeeded in condensing Heisenberg's ideas on quantum conditions in the mysterious equation $pq - qp = -i\hbar$." The "mystery," however, seems to ensue from the fact that the premises from which the equation was derived have been omitted by the founding fathers; the "physical ideas" of (HEISENBERG, 1925) are buried in the off-beat mathematical method used in the paper[1] and Born himself did not publish a step-by-step

[1] A substantial amount of historical work has been devoted to make this paper intelligible to the modern reader; see, e.g., (MACKINNON, 1977) and, for a more recent discussion, (BLUM et al., 2017).

calculation that, as he put it, "condensed" them.[2] Indeed, (3.1) was introduced in Matrix Mechanics—and then later, by von Neumann, in orthodox quantum mechanics—as an unexplained postulate; fortunately, however, the obscurity can be eliminated by making the problem attacked in (HEISENBERG, 1925), and the premises used to solve it, absolutely explicit. Once we do this, a derivation of (3.1), and therewith a physical understanding of it, will readily follow, as the reader shall see in this section.

To understand the conceptual substratum upon which quantum mechanics was built it is indispensable to remember that older works on the quantum theory[3] prior to (HEISENBERG, 1925) uniformly show the following attitude toward the application of the quantum hypothesis: the goal was to give a mechanical account of microprocesses on the basis of two fundamental assumptions. They are:

- Assumption I. *Dynamical Discontinuity.* Regardless of classical theory, the quantum-theoretically stable orbits of a Hamiltonian system $H(p_1, ..., p_n; q_1, ..., q_n)$ are characterized by the condition imposed on the reduced action S_0 as

  $$S_0 = \oint dq_k p_k = hn_k, \tag{3.2}$$

  $k = 1, ..., n$, where h is the Planck constant and n_k are positive integers labeling the stationary states. Thus the energy of a quantum system does not run in a continuum; energy states rather transit in jumps, satisfying

  $$E(n') - E(n) = h\nu(n', n). \tag{3.3}$$

  Thus later, at their talk the 1927 Solvay Conference, M. Born and W. Heisenberg emphasized:

  > Quantum mechanics is based on the intuition that the essential difference between atomic physics and classical physics is the occurrence of discontinuities. Quantum mechanics should thus be considered a direct continuation of the quantum theory founded by Planck, Einstein, and Bohr. [....] The discontinuous nature of the atomic processes here is put into the theory from the start as empirically given. (BORN; HEISENBERG, 1927, 408, 411)

  Also John von Neumann, in opening his *The Mathematical Foundations of Quantum Mechanics*, was explicit in that the physical ground of his mathematical treatise was taken from the start as a given:

[2] The derivation in (BORN; JORDAN, 1925, 290-291) is in itself unintelligible.

[3] (EINSTEIN, 1917); (EHRENFEST, 1916); (BOHR, 1918); (BOHR; KRAMERS; SLATER, 1924b); (KRAMERS, 1924a); (KRAMERS, 1924b); (BORN, 1924), (KRAMERS; HEISENBERG, 1925).

> This is not the place to point out the great success which the quantum theory attained in the period from 1900 to 1925, a development which was dominated by the names of Planck, Einstein and Bohr. At the end of this period of development it was clear beyond doubt [!] that *all* elementary processes, i.e., all occurrences of an atomic or molecular order of magnitude, obey the *"discontinuous"* laws of quanta. [...] What was fundamentally of greater significance was that the general opinion in theoretical physics had accepted the idea that *the principle of continuity* (*"natura non facit saltus"*), prevailing in the macroscopic world, *is merely simulated by an averaging process* in a world which *in truth is discontinuous by its very nature*. (NEUMANN, 1932, 5, italics ours)

It is noteworthy that in the massive, more recent literature on the foundations of quantum mechanics, the empirical input of the theory, as von Neumann, Bohr, Heisenberg, Born, Jordan, and Dirac, saw it, seems to have been entirely forgotten; we shall return to this point in §3.

- Assumption II. *The Correspondence Principle.* Max Jammer writes (JAMMER, 1966, 118) that "there was rarely in the history of physics a comprehensive theory that owed so much to one principle as quantum mechanics owed to Bohr's correspondence principle." This principle, which is, to be sure, genetically connected with assumption I,[4] was a statement about *how* to guess, from the point of view of the framework of Bohr's model of the atom, quantum frequencies and intensities—i.e., probabilities of transitions between stationary states—from classical electrodynamics. Thus Darrigol writes:

> Bohr assumed that, even for moderately excited states, the probability of a given quantum jump was approximately given by the intensity of the 'corresponding' harmonic component of the motion in the initial stationary state. This is what Bohr called 'the correspondence principle'. (DARRIGOL, 1997, 550)

For the convenience of the reader, let us remark that

1. the "harmonic component of motion" refers to the terms in a Fourier series expansion of the kinematic variable

$$x(t) = \sum_{\alpha=-\infty}^{\infty} x_\alpha(n) e^{i\alpha\omega_n t} \tag{3.4}$$

and, furthermore,

[4] See (BOKULICH; BOKULICH, 2020) for a detailed discussion.

2. in (3.4), $\{e^{i\alpha\omega_n t}\}_{\alpha\in\mathbb{Z}}$ constitutes an orthonormal basis of the periodic functions with period $\frac{2\pi}{\omega_n}$. The coefficients of the expansion, then—i.e., the amplitudes—, can thus be readily obtained in terms of the inner product

$$x_\alpha(n) = \langle e^{i\alpha\omega_n t}, x(t)\rangle := \frac{\omega_n}{2\pi}\int_0^{\frac{2\pi}{\omega_n}} x(t)e^{-i\alpha\omega_n t}dt. \tag{3.5}$$

The quantities $|x_\alpha(n)|^2$ are the harmonic intensities.

This fact cannot be overemphasized: the chief incentive to develop, in the beginning of the 20th century, a new mathematics to describe the new mechanics was the conviction that what distinguishes classical from quantum motion is that the latter is discontinuous. Thus Poincaré wrote in 1912:

> It is hardly necessary to remark how much this concept [of quanta] differs from what we imagine to be true up to this point; physical phenomena would cease to obey the laws expressed as differential equations and this would undoubtedly be the greatest and the most profound revolution that nature has undergone since Newton. (POINCARÉ, 1912, 5)

With these preliminaries out of the way, we can now remark: (HEISENBERG, 1925) is a step-by-step attempt to develop the needed discontinuous mathematics. This is why this paper came to be considered, by general acclaim, the birth of quantum *mechanics*.

To account for the discontinuous quantum jumps mechanically, it was apparent, from assumption II, that one had to find the quantum-theoretical expressions for (3.4), and therewith, for (3.5) and $|x_\alpha(n)|^2$. How? Well, equation (3.4) is classical, and, classically speaking, the intensity of radiation $|x_\alpha(n)|^2$ is proportional to $\frac{dE}{dt}$. This suggests, if we think about it, that to find the quantum-theoretical expression for the intensity—or equivalently, the probability of a quantum jump—, the first order of business is to note that the energy of the radiation emitted in a quantum jump depends not on the time t, but on the frequency ν. For in view of this fact, given that $\nu = \frac{\partial E}{\partial S_0}$, and given that the quantum expression for the variation of the energy is evidently (3.3), all that was required, to write down the quantum-theoretical frequency $\nu(n', n)$—and therewith, a *mechanics* based fundamentally on *the observed emitted frequencies*—seems to be to write down *the difference-expression for* S_0. We now emphatically remark: *this*—ΔS_0—is precisely what the canonical commutation-relations come down to, in Matrix Mechanics, as the following remarks will demonstrate.

Having thus defined, with sufficient clarity, the problem Heisenberg dealt with in (HEISENBERG, 1925), let us now think through how he found the means to solve it.

Since (3.2) is an integral equation, and since one knows how to compute the derivative of an integral, to turn it into a *difference*-equation one could take advantage of Born's 1924 calculus of differences. For Born, in 1924, in the attempt to develop the discontinuous mathematics to describe quantum jumps, considered (see (BORN, 1924), §2) a transition from a state labeled, say, by numbers $n_1, \ldots, n_f$ to a state labeled by $n_1 + \tau_1, \ldots, n_f + \tau_f$, first, from the classical point of view. The transition takes place, of course, in this case, in a (continuous) *linear way*, as Born refers to it, going through all the intermediary states $n_k + \mu\tau_k$, where $0 \leq \mu \leq 1$. Next, Born considered the classical relation in the angle action variable $\nu_k = \frac{\partial W}{\partial S_{0k}}$ and the formula $(S_0)_k = h(n'_k + \mu\tau_k)$, and obtained

$$\tau\nu = \sum_k \tau_k \nu_k = \sum_k \tau_k \frac{\partial W}{\partial (S_0)_k} = \sum_k \frac{1}{h} \frac{d(S_0)_k}{d\mu} \frac{\partial W}{\partial (S_0)_k} = \frac{1}{h} \frac{dW}{d\mu}. \tag{3.6}$$

Born's proposal was, then, this: Whenever a classical calculated quantity yields

$$\sum_k \alpha_k \frac{\partial \Phi}{\partial (S_0)_k} = \frac{1}{h} \frac{d\Phi}{d\mu}, \tag{3.7}$$

the trick to transform it on a discontinuous (quantum), as opposed to a continuous (classical), quantity is to replace it by the difference coefficient as

$$\int_0^1 \sum_k \alpha_k \frac{\partial \Phi}{\partial (S_0)_k} d\mu = \frac{1}{h} [\Phi(n + \alpha)) - \Phi(n)]. \tag{3.8}$$

In particular, the expression of the emission of radiation in classical theory,

$$\nu(n, \alpha) = \alpha\nu(n) = \alpha \frac{\partial W}{\partial S_0} = \frac{\alpha}{h} \frac{dW}{dn} \tag{3.9}$$

corresponds, in Born's scheme, to the quantum-theoretical expression

$$\nu(n, n - \alpha) = \frac{1}{h} (W(n) - W(n - \alpha)), \tag{3.10}$$

which is indeed (3.3). Thus Born, to construct quantum mechanics, helped himself to a kind of Wittgenstenian trick. He throws away the continuity ladder after he has climbed up it.

A moment's reflection will reveal that the trouble with Born's procedure is that it was not general: one had to make the classical-to-quantum translation case-by-case. But the above considerations do exhaust "Heisenberg's ideas on quantum conditions" one needs to derive the canonical commutation relations; Born's calculus of differences is clearly built from the ground up on Assumptions I and II[5] and, from equation (3.4), we note only that it naturally implies $\dot{x}(t) = \sum_{\alpha=-\infty}^{\infty} i\alpha\omega_n x_\alpha(n) e^{i\alpha\omega_n t}$, so that we

[5] For a discussion of Heisenberg's use of the correspondence principle in (HEISENBERG, 1925) see, e.g., (MACKINNON, 1977) and (BOKULICH; BOKULICH, 2020, §4.3).

can define $p_\alpha(n) = i\alpha m\omega_n x_\alpha(n)$ to have $m\dot{x}(t) = \sum_{\alpha=-\infty}^{\infty} p_\alpha(n)e^{i\alpha\omega_n t}$. One last thing: Heisenberg also required that x be real, so that $x_\alpha(n) = \overline{x_{-\alpha}(n)}$. We can now reconstruct Heisenberg's reasoning by applying all of this in the classical expression of the reduced action S_0 to find therewith its quantum-theoretical counterpart.

$$\begin{aligned}
S_0 &= \oint p\dot{q}d\gamma \\
&= m\int_0^T \dot{x}(t)^2 dt \\
&= -m\sum_{\alpha,k=-\infty}^{\infty} \alpha\omega_n x_\alpha(n)k\omega_n x_k(n)\int_0^{\frac{2\pi}{\omega_n}} e^{i(\alpha+k)\omega_n t}dt \\
&= 2\pi m\sum_{k=-\infty}^{\infty} k^2\omega_n x_k(n)x_{-k}(n) \\
&= 2\pi i\sum_{k=-\infty}^{\infty} k\left(x_k(n)e^{ik\omega_n t}\right)\left(i(-k)m\omega_n x_{-k}(n)e^{i(-k)\omega_n t}\right) \\
&= 2\pi i\sum_{k=-\infty}^{\infty} k\left(x_k(n)e^{ik\omega_n t}\right)\left(p_{-k}(n)e^{i(-k)\omega_n t}\right),
\end{aligned}$$

where the integration, as prescribed by the old quantum theory, was performed over the closed curve corresponding to one period of the virtual oscillator, namely, $T = \frac{2\pi}{\omega}$ and, fortuitously, $\int_0^{\frac{2\pi}{\omega_n}} e^{i(\alpha+k)\omega_n t}dt = \frac{2\pi}{\omega_n}\delta_{\alpha,-k}$. Taking the derivative in the quantity S_0 yields

$$1 = 2\pi i\sum_{k=-\infty}^{\infty} k\frac{\partial}{\partial S_0}\left(\left(x_k(n)e^{ik\omega_n t}\right)\left(p_{-k}(n)e^{i(-k)\omega_n t}\right)\right).$$

So far, everything is still classical. To make it quantum-theoretical, we use Born's relation (3.8), to get

$$1 = \frac{2\pi i}{h}\sum_{k=-\infty}^{\infty}\left(x_k(n+k)e^{ik\omega_{n+k}t}\right)\left(p_{-k}(n+k)e^{i(-k)\omega_{n+k}t}\right) - \left(x_k(n)e^{ik\omega_n t}\right)\left(p_{-k}(n)e^{i(-k)\omega_n t}\right). \tag{3.11}$$

But the problem is clearly not yet solved, for the *frequencies* in (3.11) are not in agreement with (3.3). We now categorically make the statement: *it is to correct for this that Heisenberg rewrote the equations in two indices*. The harmonic frequencies can easily be expressed in two indices, but coherence, *and the correspondence principle*, demands that the amplitudes $x_k(n)$ be replaced by $x(n,k)$. The problem was how to express sums of *products* of such coefficients; the solution, Heisenberg noted, was not far to seek, for conservation of energy clearly implies, for two consecutive energy transitions in Bohr's atomic model, that

$$\nu(n,\alpha)+\nu(n,\beta)=\nu(n,\alpha+\beta)\longrightarrow\nu(n,n-\alpha)+\nu(n-\alpha,n-\alpha-\beta)=\nu(n,n-\alpha-\beta), \tag{3.12}$$

and so, analysing the expression for $x(t)^2$ for a product c of two quantities a and b, we can conclude that the rule should be indeed

$$c_\beta(n)=\sum_\alpha a_\alpha(n)b_{\alpha-\beta}(n)\longrightarrow c(n,n-\beta)=\sum_\alpha a(n,n-\alpha)b(n-\alpha,n-\beta). \tag{3.13}$$

Applying this to (3.11) readily yields

$$\begin{aligned}1=\frac{2\pi i}{h}\sum_{k=-\infty}^{\infty}&\left(x(n+k,n)e^{i\omega(n+k,n)t}\right)\left(p(n,n+k)e^{i\omega(n,n+k)t}\right)\\&-\left(x(n,n-k)e^{i\omega(n,n-k)t}\right)\left(p(n-k,n)e^{i\omega(n-k,n)t}\right).\end{aligned} \tag{3.14}$$

Recalling now that the product of two (infinite, by abuse) matrices $A=(a_{ij})_{i,j=1}^{\infty}$ and $B=(b_{ij})_{i,j=1}^{\infty}$ is given by $(AB)_{ij}=\sum_{k=1}^{\infty}a_{ik}b_{kj}$, to Born is due the credit of recognizing that the objects above behave, as far as multiplication is concerned at least, like matrices[6]—and therewith the realization that one needed not invent a new mathematics to deal with quantum phenomena after all. One could use the fully developed matrix calculus. Carrying out the identification, change of summation-index leads directly to the conclusion that the above expression reduces to

$$-i\hbar=(px-xp)_{ii},\quad\forall i\in\mathbb{N}. \tag{3.15}$$

This proves our contention that the canonical commutation relations amount to a kind of ΔS_0. The claim that it necessarily encodes discontinuity, as emphasized by (HEISENBERG, 1927), is, however, more complex than meets the eye, as the results of the next section shall unequivocally demonstrate.

3.2.2 The Dynamical Origin of the Quantization Rules, Part II: Wave Mechanics

It has been taken for granted in the history and philosophy of physics literature that the canonical commutation relations, in Wave Mechanics, have played no role whatever prior to Schrödinger's equivalence proof.[7] The standard view is that the commuta-

[6] This was how Born described his thinking years later: "I could not take my mind off Heisenberg's multiplication rule, and after a week of intense thought and trial I suddenly remembered an algebraic theory which I had learned from my teacher, Professor Rosanes, in Breslau. Such square arrays are well-known to mathematicians and, in conjunction with a specific rule for multiplication, are called matrices. I applied this rule to Heisenberg's quantum condition and found that this agreed in the diagonal terms. It was easy to guess what the remaining quantities must be, namely zero; and at once there stood before me the peculiar formula $pq-qp=-i\hbar$" (BORN, 1955, 259).

[7] (JAMMER, 1966); (MEHRA; RECHENBERG, 1987); (MULLER, 1997a); (MULLER, 1997b); (BELLER, 1999); (BITBOL, 2012); (PEROVIC, 2008).

tion relations were *posited* by Schrödinger *to obtain equivalence* with Matrix Mechanics; "Schrödinger only invented these [canonical] operators," Muller writes, for example, in his sophisticated, and widely read analysis, of the mathematical isomorphism between Matrix and Wave Mechanics, "when pursuing his equivalence proof" (MULLER, 1997a, 46). A more detailed mathematical analysis of Schrödinger's first paper on Wave Mechanics, however, will reveal that this is not the case; in fact, the canonical commutation relations are a direct consequence of Schrödinger's *first* derivation of the wave equation (but not from the second), and so have been part and parcel of Wave Mechanics since its inception. This result, to be demonstrated in this section, will lead us to ask questions about the relationship between Matrix and Wave Mechanics that Muller did not consider in his influential investigation—and therewith, instructive facts involved in the equivalence that have been hitherto hidden from sight will emerge.

The oversight mentioned in the previous paragraph is related to the fact there is a certain prejudice, in the scholarship on Wave Mechanics, to think of Schrödinger's first derivation of the wave-equation (SCHRÖDINGER, 1926b) as having only a formal significance, as it seems, on a first and second glance, to depend on seemingly physically-unmotivated steps ((WESSELS, 1979); (WESSELS, 1980); (MACKINNON, 1980); (KRAGH, 1982) (MULLER, 1997a)). We will try to remove this prejudice by proceeding here as in the previous section, and doing for Schrödinger what Schrödinger didn't do for himself: we shall be absolutely explicit about what premises we are using, and why.

Consider, by way of preamble, the opening remarks of Schrödinger's first paper on Wave Mechanics. He is warning his readers (though from a modern perspective, this may not be obvious), that he is playing a different kind of game than that of the matrix theoreticians.

> In this communication I would like first to show, in the simplest case of the (non relativistic and *unperturbed*) hydrogen atom, that the usual prescription for quantisation can be substituted by another requirement in which no word about "integer numbers" occurs anymore. Rather, the integerness emerges in the same natural way as, for example, the integerness of the number of knots of a vibrating string. The new interpretation is generalisable and touches, I believe, very deeply the true essence of the quantisation prescription. (SCHRÖDINGER, 1926b, 1, italics ours)

This quote, and the discussion in §3.2.1, tell us an exceedingly important fact, namely that while Matrix Mechanics was an attempt to improve, and generalize, the old quantum theory associated with the names of Planck, Einstein, and Bohr, *Wave Mechanics was an attempt to overcome it*. Schrödinger wanted to demonstrate, by direct counter-example, that the discontinuity hypothesis should by no means be taken as "empirically given."

To proceed, then, to the argument: the reader will remember that in the old quantum theory the quantum conditions, i.e., the Bohr-Sommerfeld rules, are posited over the (reduced) Hamilton-Jacobi equation

$$H\left(q, \frac{\partial S}{\partial q}\right) = E. \tag{3.16}$$

Schrödinger too will make use of the Hamilton-Jacobi formulation of mechanics. The difference is that he will not postulated over them, but rather—through the introduction of a new fundamental hypothesis he did not himself, but we will, properly discuss—, *elicit from them*, the allowed quantum-mechanical motions.

The procedure is as follows. Start by substituting the action S with a new unknown ψ[8] via the transformation

$$S = K \log \psi. \tag{3.17}$$

Inserting the latter into (3.16) yields

$$H\left(q, \frac{K}{\psi}\frac{\partial \psi}{\partial q}\right) - E = \frac{K^2}{2m}\psi^{-2}\left(\frac{\partial \psi}{\partial q}\right)^2 + V(q) - E = 0. \quad {}^{[9]} \tag{3.18}$$

Now evidently, solving (3.18) *as it stands* is merely to solve the classical equations of motion in a new variable. It doesn't change anything. And it is important to now remember that the Hamilton-Jacobi equations, in classical physics, describe *only* the motion of *localized* objects, i.e., point-particles and/or wave-packets. Schrödinger of course did not merely solve the classical equations of motion in a new variable. Rather he sought to define a *variational problem* in ψ.

Only one must note what *that* means and entails, in conceptual, and specifically in *dynamical*, terms.

But before we can get into interpretative questions let us carry through the calculation formally. We follow Schrödinger, and multiply the left-hand side of equation (3.18) by ψ^2, and then integrate it in space, thereby defining the quadratic form

$$\begin{aligned}\mathcal{J}[\psi] = \int_{\mathbb{R}^3} \frac{K^2}{2m}\left(\frac{\partial \psi}{\partial q}\right)^2 + V(q)\psi^2 - E\psi^2 dq = \lim_{R\to\mathbb{R}^3}\int_{\partial R}\frac{K^2}{2m}\psi\frac{\partial\psi}{\partial q}d\vec{S} \\ + \int_{\mathbb{R}^3} -\frac{K^2}{2m}\psi\frac{\partial^2\psi}{\partial q^2} + V\psi^2 - E\psi^2 dq. \quad {}^{[10]}\end{aligned} \tag{3.19}$$

[8] Following Schrödinger, we will consider ψ to be everywhere real, but we remark that one can complexify the space and sesqui-linearly extend the inner product to get the corresponding quadratic form (3.19).

[9] Here we are adopting the notation $\frac{\partial\psi}{\partial q} = \mathrm{grad}(\psi)$, $\left(\frac{\partial\psi}{\partial q}\right)^2 = \left(\frac{\partial\psi}{\partial q}\right)\cdot\left(\frac{\partial\psi}{\partial q}\right)$, and $\frac{\partial^2\psi}{\partial q^2} = \nabla^2\psi$.

The variational problem is then defined by requiring not that $\mathcal{J}[\psi]$ vanishes, as suggested by equation (3.18)—and from which, let us remember, the classical equations of motion for localized objects can be derived. Instead we must impose that the functional $\mathcal{J}$ be *stationary* at ψ:

$$\frac{1}{2}\delta\mathcal{J}[\psi] = \lim_{R\to\mathbb{R}^3}\int_{\partial R}\frac{K^2}{2m}\frac{\partial\psi(q)}{\partial q}\delta\psi(q)d\vec{S} + \int_{\mathbb{R}^3}\left(-\frac{K^2}{2m}\frac{\partial^2\psi(q)}{\partial q^2} + V(q)\psi(q) - E\psi(q)\right)\delta\psi(q)dq = 0. \tag{3.20}$$

A key passage of (SCHRÖDINGER, 1926b) is Schrödinger's emphasis that the quantum postulate of the old quantum theory, *i.e*, the Bohr-Sommerfeld quantization rules, should be *replaced* by this variational problem. The arbitrariness of $\delta\Psi$ does indeed imply that the terms in the integrals must vanish; the first term in (3.20) is a boundary condition, upon which Schrödinger imposes regularity conditions that imply that it too must vanish. Hence the variational problem is reduced to the second term on the right-hand side,

$$-\frac{K^2}{2m}\frac{\partial^2\psi(q)}{\partial q^2} + V(q)\psi(q) - E\psi(q) = 0, \tag{3.21}$$

which, as the reader shall immediately recognize, is the time-independent Schrödinger equation.

Now to the question of the meaning of the transition from (3.18) to (3.19). We want to bring to the open two facts that are hidden in it.

(The move to (3.20), which is traditionally described in the literature as "cryptic" and "ad hoc" ((WESSELS, 1979); (KRAGH, 1982); (JOAS; LEHNER, 2009)), necessitates more careful mathematical analysis—we shall be able to demystify it completely in §2.4.)

1. **The nature of the classical limit.** The point cannot be overemphasized: the Hamilton-Jacobi equation (3.16) describes the motion of localized objects, and replacing S by ψ in it, via transformation (3.17), does not change this fact. The physically important question is therefore what is the effect of transforming equation (3.18) into (3.19); the answer is that this transition forces the integrand to be seen as a Hamiltonian *density*, from which fact it follows that ψ must now necessarily and ineluctably acquire—*contra* Assumption I of Matrix Mechanics, and all that has been said in the previous section—the significance of a *continuous quantity distributed in space*.

   To accept Schrödinger's (SCHRÖDINGER, 1926b) derivation of the wave equation thus forces us to acknowledge that the passage from quantum to classical motion is the passage of showing just how this continuously-distributed quantity ψ behaves, in the appropriate classical limit, like localized wave-packets. The fact that this information is "hidden" in the mathematics deserves comment: Schrödinger was only explicit about the Hamiltonian analogy between Mechanics and Optics—which implies the fact just stated, namely that classical mechanics is a limiting case of a more general wave mechanics—in (SCHRÖDINGER, 1926c), his second communication on Wave Mechanics. Historians have interpreted the absence of an explicit reference to the Hamiltonian analogy in (SCHRÖDINGER, 1926b) as evidence of the fact that Schrödinger didn't even know, or, didn't fully understand, the optical-mechanical analogy when he started the quantization series (see (JOAS; LEHNER, 2009, §4)); our analysis indicates otherwise. Schrödinger ought to have known what he was doing, and the mathematics of (SCHRÖDINGER, 1926b) is unequivocal: It is an exact mathematical transcription of the optical-mechanical analogy. Schrödinger's (SCHRÖDINGER, 1926b) quantization procedure not merely allows, but rather demands, as a matter of simple self-consistency, that the classical limit be recovered in the way just described.[11]

2. **The quantum conditions**. The passage from (3.18) to (3.19) also requires that the quadratic form on the second part of the integrand on the right-hand side of (3.19) must be the result of the action of an operator[12] on ψ:

$$(\hat{H}(q,p)-E)(\psi)(q) = -\frac{K^2}{2m}\frac{\partial^2\psi(q)}{\partial q^2} + (V(q)-E)\psi(q),$$

which fact, added to the requirement that the procedure must hold for an arbitrary potential $V(q)$, implies in identifying p and q, *via* correspondence with (3.18), with operators defined by $(\hat{p}\psi)(q) = \pm iK\frac{\partial\psi}{\partial q}(q)$ and $(\hat{q}\psi)(q) = q\psi(q)$. The identification Schrödinger spelled out only later, in (SCHRÖDINGER, 1926a, §2)—namely, that the "p_l in the [classical] function is to be replaced by the operator $\frac{\partial}{\partial q_l}$" (SCHRÖDINGER, 1926a, 47)—, is therefore obtained automatically, in his first communication on Wave Mechanics, by the transition from (3.18) to (3.19). In (SCHRÖDINGER, 1926b, §2) Schrödinger remarked that experiments require that we set $K = \frac{h}{2\pi}$. It must therefore be acknowledged that the transition from (3.18) to (3.19) also inexorably and ineluctably entails that $[\hat{p},\hat{q}] = -i\hbar$.

Our assertion in the opening paragraph of this section—namely, that Schrödinger's first

[11]The matrix theoreticians used the fact that the classical limit is not recovered in this way as a strong argument against the general consistency of Schrödinger's position—and to great effect; failure to solve this and other related difficulties soon led to the downfall of Schrödinger's "interpretation." See the details in (MACKINNON, 1980, §6) and (BELLER, 1999, chapter 2).

[12]In this case the operator that defines the quadratic form in terms of the inner product is obvious, but this is in fact the general case: By standard techniques, one can construct a sesqui-linear form from which the quadratic form Q is obtained. Then, thanks to the unbounded version of the Riesz Representation theorem for the densely defined, symmetric, and closed sesqui-linear form, one obtains a unique densely defined closed unbounded operator $A : D(A) \subset \mathcal{H} \to \mathcal{H}$ such that $Q(\psi) = \langle\psi|A\psi\rangle$.

derivation of the wave equation *implies* the canonical commutation relations—has thus been proved.

3.2.3 The Dynamical Origin of the Quantization Rules, Part III: The Problem of the "Inner Connection" Between Matrix and Wave Mechanics

Equipped with the results of the two previous sections we now wish to consider the opening lines of (SCHRÖDINGER, 1926a). This is Schrödinger's so-called "equivalence paper."

> Considering the extraordinary differences between the starting-points and the concepts of Heisenberg's quantum mechanics and of the theory which has been designated "undulatory" or "physical" mechanics, and has lately been described here, it is very strange that these new two theories *agree with one another with regard to the known facts*, where they differ from the old quantum theory. I refer, in particular, to the peculiar "half-integralness" which arises in connection with the oscillator and the rotator. That is really very remarkable because starting-points, presentations, methods and in fact the whole mathematical apparatus, seem fundamentally different. [...] Above all, however, the departure from classical mechanics in the two theories seems to occur in diametrically opposed directions. In Heisenberg's work the classical continuous variables are replaced by systems of *discrete* numerical quantities (matrices), which depend on a pair of integral indices, and are defined by *algebraic* equations. The authors themselves describe the theory as a "true theory of a discontinuum". On the other hand, wave mechanics shows just the reverse tendency; it is a step from classical point-mechanics towards a *continuum*-theory. In place of a process described in terms of a finite number of dependent variables occurring in a finite number of differential equations, we have a continuous *field-like* process in configuration space, which is governed by a single partial differential equation, derived from a Principle of [Least] Action.

The fact to be explained, as Schrödinger sees it, is therefore this: "the two theories agree [numerically] with one another with respect to the known facts." There are four points to be made in this connection.

The first is that since Muller's persuasive (MULLER, 1997a; MULLER, 1997b), it has been generally taken for granted by the literature that Schrödinger didn't prove that the canonical matrix- and wave-operator algebras are isomorphic. We will proceed to question the mathematical correctness of Muller's disproof of Schrödinger's isomorphism proof in a separate paper; here, we need not, and so will not, get into it. For with or without a rigorous proof of algebra isomorphism—and this is our second point—, the further fact remains:

> [The] relation of matrices to functions is *general*; it takes no account of the *special* system considered, but is the same for all mechanical systems. (In other words: the particular Hamiltonian function does not enter the connecting law.) (SCHRÖDINGER, 1926a, 46)

In other words, kinematical equivalence does not entail dynamical equivalence. This is a fact that does not emerge with clarity in Muller's investigation, as he focuses in the isomorphism: Dynamical equivalence between Matrix and Wave Mechanics must be proved *independently* from the isomorphism between matrices and wave-operators (this fact shall be important for us in §3.3.1). This is why Schrödinger sets out to prove, in (SCHRÖDINGER, 1926a, §4), that to solve his minimization problem (3.21) is equivalent to solving the Matrix Mechanics problem of diagonalizing the matrix H.

Now, let us reflect. It will quickly appear that this argument alone cannot be taken as the ground of dynamical equivalence. For the diagonalization of the matrix H provides no indication of how the quantum system evolves in time; the dynamical posits of each theory may well be different. Indeed, if by the time Schrödinger wrote the equivalence paper (March 1926), Matrix and Wave Mechanics did not postulate the exact same dynamical laws, then it will trivially follow that the theories were not equivalent. And since Schrödinger's treatment of the transitions (intensity lines) appeared only after the equivalence paper, in (SCHRÖDINGER, 1926d), wherein he explicitly says that he *imported the rule*, *via* the mapping between matrix-coefficients and wave-functions described in the equivalence proof (see (SCHRÖDINGER, 1926d, §4)), *from* Matrix Mechanics (this fact shall also be important for us in §3.3.1.), we find, in this fact alone, conclusive evidence that establishes that by the time Schrödinger wrote the equivalence proof, they indeed were not.

The third point we want to make is that the way Schrödinger related his theory to Heisenberg, Born and Jordan's falls short of explaining their connection. For Schrödinger explicitly remarked, let us repeat it expressly once more, that matrix- and wave-operators isomorphism does not entail theoretical equivalence; furthermore, by the argument we have just given, the theories were indeed, as of March 1926, neither mathematically nor empirically equivalent. But the puzzle of how it can then *be* that they "agree [numerically] with one another with respect to the known facts" apparently remains. It is against this fact that we now ask the reader to note, and this is our fourth and final point, that by the lights of our discussion the numerical agreement of Matrix

Mechanics and Wave Mechanics for simple toy systems that depend only on p and q such as those Schrödinger mentions is *not* strange. It is an unavoidable mathematical consequence of the fact that *both theories carry the canonical commutation relations in themselves.*

For to say that Matrix and Wave Mechanics agree in that $[\hat{p}, \hat{q}] = -i\hbar$ because there is a mathematical isomorphism between them is, please note, to invert the arrow of explanation. For it is $[\hat{p}, \hat{q}] = -i\hbar$ that characterizes the algebra. It is, then, the fact that $[\hat{p}, \hat{q}] = -i\hbar$ holds for both theories that will allow an isomorphism between them to be established—not the other way around.

And so it is *this*, and not the numerical agreement itself, as Schrödinger thought, that is the fact to be explained. The fact that theories that imply and are implied by mutually contradictory hypotheses both *entail* $[\hat{p}, \hat{q}] = -i\hbar$ is a truly astonishing fact that must itself have a deeper significance. It suggests that there is something here, a deeper *physical* fact, a common underlying structure, that supports or necessitates the link between Matrix and Wave Mechanics—the link upon which von Neumann erected the mathematical structure of orthodox quantum mechanics (we shall discuss this point in detail in §3.3). The purpose of the next section is to bring out this deeper, and effectively forgotten, matter-of-fact finally to the surface.

3.2.4 The Dynamical Origin of the Quantization Rules, Part IV: The Dual Action

If it is, as we have remarked above, a common error to regard Schrödinger's first derivation of the wave equation as "cryptic" and "*ad hoc*," this is because, let us note, Schrödinger encouraged the mistake himself: In the opening lines of his second paper of Wave Mechanics, (SCHRÖDINGER, 1926c), wherein a different, "more intuitive," derivation of the wave equation is given, Schrödinger, in referring to the first paper, remarked that the transition from *the equating to zero* of a certain expression to the postulation that the *space integral* of the said expression shall be *stationary*—i.e., the passage from (3.18) to (3.20), the very construction of his variation problem—is in itself incompreensible (SCHRÖDINGER, 1926c, 13). But let us not isolate the word "incomprehensible" from the qualifier "in itself." A principled justification for Schrödinger's variational problem must exist, even though he did not state it. And it will be our task to find it.

For we cannot afford to bypass this historical/mathematical/conceptual difficulty in our study. The relation $[\hat{p}, \hat{q}] = -i\hbar$ follows, as a matter of mathematical necessity, in Wave Mechanics only from the apparently "cryptic," "in itself incomprehensible," derivation of the wave equation given in (SCHRÖDINGER, 1926b)—it doesn't follow from the "intuitive" derivation given in (SCHRÖDINGER, 1926c).[13] Nothing in the latter paper implies identifying classical functions with operators. And since we can expect that the fact that necessitates the emergence of the commutation relations in Wave Mechanics is mathematically and conceptually tied, by the upshot of the result of the previous section, to the fact that necessitates the emergence of the commutation relations in *Matrix Mechanics*, it appears that we must only find the justification for Schrödinger's "incomprehensible" transition to find the key to the solution of the problem we have posed at the end of the previous section. And this, as the reader shall soon see, is exactly the case.

We start by thinking through the assumptions involved in the derivation discussed in §3.2.2 in a bit more detail. Schrödinger's first ingredient, let us remember, is the reduced Hamilton-Jacobi equation. A word about the Hamilton-Jacobi Theory: In the Hamiltonian formalism, the Hamilton-Jacobi equation is obtained by a canonical transformation that makes the transformed Hamiltonian vanish (the transformed action must still satisfy Hamilton's principle, of course). The *reduced* Hamilton-Jacobi equation is a particular case of this, which appears when (a) the Hamiltonian does not depend on time, or (b) one proceeds by additive separation of variables. A few pages back we saw that the passage from (3.18) into (3.19) involves replacing functions by functionals. The question comes naturally, then, as to whether the form of the Hamilton-Jacobi equation should not be adapted accordingly. So let us verify what follows from doing precisely this: imposing Hamilton's principle and separation of variables upon replacing functions by functionals in the new action function (3.17).

Let's start slow. The quantity action, by definition,

$$S = \int L dt = \int p\dot{q} dt - \int H dt = S_0 - \int H dt, \tag{3.22}$$

where $S_0 := \int p\dot{q}dt$ is the reduced action and H is the Hamiltonian, encodes, let us remember, the change of a physical system over time. The equations of motion of a physical system are then derivable by the postulation that the action be stationary. If one is in the particular case, however, as Schrödinger indeed was, wherein one has to proceed by separation of variables (for Schrödinger was assuming conservation of energy), then the quantity S reads instead

$$S = \int L dt = S_0 - Et, \tag{3.23}$$

where E is a constant. In the standard formulation of Hamiltonian mechanics, equations (3.22) and (3.23) lead to $H(q(t), p(t)) = E$; here, we must adapt it. We must follow out

[13]Note indeed that in the passage from the equivalence paper quoted above Schrödinger refers to the first derivation of the wave-equation, and not the second, as "the" derivation of the wave equation. This is in good agreement with the fact that the first derivation—and not the second, let us reiterate—implies the canonical commutation relations.

the fact that the transition from (3.18) to (3.19) forces us to reinterpret such equations in terms of functionals. We then have

$$S = S_0 - \int H dt \leftrightarrow \mathcal{S} = \mathcal{S}_0 - \int \tilde{\mathcal{J}} dt, \quad \text{and} \quad S = S_0 - Et \leftrightarrow \mathcal{S} = \mathcal{S}_0 - \int \tilde{\mathcal{E}} dt, \tag{3.24}$$

where

$$\tilde{\mathcal{J}}[\psi] = \int_{\mathbb{R}^3} -\frac{K^2}{2m}\psi(q)\frac{\partial^2\psi(q)}{\partial q^2} + V(q)\psi(q)^2 dq, \quad \text{and} \quad \tilde{\mathcal{E}}[\psi] = \int_{\mathbb{R}^3} E\psi(q)^2 dq. \tag{3.25}$$

This makes it clear that it is an immediate consequence of equations (3.24), and Hamilton's principle *alone*, that

$$\begin{aligned} \delta\mathcal{S}_0 - \delta\left(\int \tilde{\mathcal{J}} dt\right) = \delta S = 0 = \delta\mathcal{S}_0 - \delta\left(\int \tilde{\mathcal{E}} dt\right) \Rightarrow \delta\left(\int (\tilde{\mathcal{J}} - \tilde{\mathcal{E}}) dt\right) = 0 \Rightarrow \\ \delta\mathcal{J} = \delta(\tilde{\mathcal{J}} - \tilde{\mathcal{E}}) = 0.^{[14]} \end{aligned} \tag{3.26}$$

This is exactly Schrödinger's variational problem. Though Schrödinger called it "incomprehensible," once the steps are explained, nothing can seem more justified. The stationary condition (3.20), or equivalently, the Schrödinger equation, *is a logical consequence of applying Hamilton's principle to the action written in terms of the continuously spatially-distributed quantity ψ.*

The reader may wonder if this procedure does not amount to finding an action for a field. We categorically assert: it does not. For there is no direct generalization of the theory of canonical transformations for fields, hence no counterpart of the Hamilton-Jacobi equation in field theory.[15] Recall, indeed, from §3.2.2: it is a peculiarity of *this* procedure of Schrödinger's that it involves "transforming" the quantity ψ from a discrete to a continuous object (optical-mechanical analogy). The elegance of it is undeniable; one must agree with Born: "it would have been beautiful if you [Schrödinger] were right" (Born to Schrödinger, 6 November 1926, quoted in (BELLER, 1999, 36)). For it is precisely this transformation, the transformation that does the work of the optical-mechanical analogy, that "quantizes" the system.

This completes our analysis of how the canonical commutation relations are implied by Schrödinger's derivation of the wave equation from a principle of Least Action. If we were right in expecting that the facts that lead to the emergence of the canonical commutation relations in Matrix Mechanics, and the facts that lead to the emergence of the canonical commutation relations in Wave Mechanics, are theoretically

[15]The Hamiltonian formulation for classical fields requires an infinite-dimensional manifold M and the introduction of functional derivatives. This is why it is much harder to establish a symplectic manifold structure, and therefore, a theory of canonical transformations, in the cotangent bundle T^*M. For more details, we refer to (ABRAHAM; MARSDEN, 2008), in particular, it is worth checking remarks (3) and (4) on p. 383.

linked, then we should now be able to identify the link exactly. That is exactly the case—consider equations (3.24) and (3.26) again.

Schrödinger focused on the *second* term of the action; he promoted H to $\tilde{\mathcal{J}}$ and, as we just saw, the condition $\delta\tilde{\mathcal{J}}[\psi] = 0$ is equivalent to Schrödinger's equation. Consistency, however, suggests—perhaps demands—that the first term of the action also be considered. And all the detailed historical and technical analyses we have carried out up to now in this paper have finally served their purpose by enabling the reader to recall at once that the matrix theorists derived the canonical commutation relations starting from $\int p\dot{q}dt$, which is just S_0. That is precisely the first term of the action.

We thus now categorically conjecture the following: A *duality* between Matrix and Wave Mechanics, *grounded on the action*, must hold. To prove it we shall extend Schrödinger's treatment and apply it to the first term of the action, defining thereby the functional $S_0 \leftrightarrow \mathcal{S}_0$. The canonical commutation relations will be shown to follow provided $\mathcal{S}_0$ satisfies a remarkably important physical condition.

We begin by noting that if one decides to follow Schrödinger's (SCHRÖDINGER, 1926b) procedure naively,[16] then, as we saw in §3.2.2, the replacement rule $p \to \hat{p} := -iK\frac{\partial\psi}{\partial q}$ and $q \to \hat{q} := q\psi$ must hold. But there is an ambiguity here regarding $\dot{q} = \frac{1}{m}p$; we shall assume that the correct replacement rule is $\dot{q} = \frac{\partial\hat{q}}{\partial t}$. Carrying out, then, the substitutions, and integrating by parts, we have

$$\begin{aligned}
\mathcal{S}_0[\psi] &= \int_0^{\frac{2\pi}{\omega_n}} \int \hat{p}\frac{\partial}{\partial t}\hat{q}dqdt \\
&= \int_0^{\frac{2\pi}{\omega_n}} \int -iK\frac{\partial\psi}{\partial q}\frac{\partial}{\partial t}(q\psi)dqdt \\
&= \int_0^{\frac{2\pi}{\omega_n}} \int -iK\frac{\partial\psi_n}{\partial q}q(-i\alpha\omega_n)\psi_n dqdt \\
&= -2\pi K\alpha \int \frac{\partial\psi_n}{\partial q}q\psi_n dq \\
&= -2\pi K\alpha \lim_{V\to\mathbb{R}^3}\int_{\partial V}\psi^2 q d\vec{S} + 2\pi K\alpha\int\psi_n^2 dq + 2\pi K\alpha\int\psi_n q\frac{\partial\psi_n}{\partial q}dq \\
&= 2\pi K\alpha - \mathcal{S}_0[\psi],
\end{aligned} \tag{3.27}$$

where we have used, as Schrödinger did in (3.20), that the boundary integral vanishes, namely, $\lim_{V\to\mathbb{R}^3}\int_{\partial V}\psi^2 q d\vec{S} = 0$, and the normalization of the wave function. Hence,

$$2\mathcal{S}_0[\psi] = 2\pi K\alpha = \alpha h. \tag{3.28}$$

[16]We are still following Schrödinger's assumption that the wave-function must be real. The case of complex wave functions will become clear a few steps later.

In modern notation, (3.27) reads

$$2\mathcal{S}_0[\psi] = \int_0^{\frac{2\pi}{\omega_n}} \left\langle \psi \middle| \hat{p}\frac{\partial}{\partial t}\hat{q}\psi \right\rangle + \overline{\left\langle \psi \middle| \hat{p}\frac{\partial}{\partial t}\hat{q}\psi \right\rangle} dt \tag{3.29}$$

$$= \int_0^{\frac{2\pi}{\omega_n}} \left\langle \psi \middle| \hat{p}\frac{\partial}{\partial t}\hat{q}\psi \right\rangle + \left\langle \psi \middle| \left(\hat{p}\frac{\partial}{\partial t}\hat{q}\right)^* \psi \right\rangle dt \tag{3.30}$$

$$= \int_0^{\frac{2\pi}{\omega_n}} \left\langle \psi \middle| (\hat{p}\hat{q} - \hat{q}\hat{p}) \frac{\partial}{\partial t}\psi \right\rangle dt. \tag{3.31}$$

Now if we rewrite the foregoing expression for $\psi = \psi(x,t) = \psi_n(x)e^{-i\alpha\omega_n t}$,[17], we have

$$2\mathcal{S}_0[\psi] = \langle\psi|[\hat{p},\hat{q}]\psi\rangle \int_0^{\frac{2\pi}{\omega_n}} (i\omega_n\alpha)dt \tag{3.32}$$

$$= \langle\psi|[\hat{p},\hat{q}]\psi\rangle\,(i2\pi\alpha), \tag{3.33}$$

from which it follows that

$$2\mathcal{S}_0[\psi] = \alpha h \Leftrightarrow [p,q] = -i\hbar. \tag{3.34}$$

This proves that Schödinger's procedure, applied to the reduced action $\mathcal{S}_0$, *recovers the Bohr-Sommerfeld rule with the condition that the commutation relations are satisfied*. Since Matrix Mechanics deduced the commutation relations precisely from $\mathcal{S}_0$ by assuming the Bohr-Sommerfeld rule (see §3.2.1)[18], we can state the conclusion:

the reason the canonical commutation relations are a common feature of Matrix and Wave Mechanics is that these theories are dual from the point of view of the action.

This is the main result of this section. We summarize it in the following

Scholium. *Let the action $\mathcal{S}[\psi] = \mathcal{S}_0[\psi] - \mathcal{J}[\psi]$ be as above. The following two conditions are equivalent:*

1. $(\delta\mathcal{S}[\psi])_{\delta\mathcal{S}_0[\psi]=0} = 0.$
2. *For every stationary state ψ, i.e. $\delta\mathcal{S}[\psi] = 0$, there exists $n \in \mathbb{N}$ such that $2\mathcal{S}_0 = nh$.*

Wave Mechanics (item (1)) and Matrix Mechanics (item (2)) are therefore "dual" theories because they were constructed upon equivalent conditions.

We now want to use this result as basis for a critical analysis of the mathematical foundations of orthodox quantum mechanics.

[17] Using (3.17), one sees that the decomposition of ψ in such a product corresponds to the additive separation of variables $\tilde{S}(x,t) = K\log\psi(x,t) = K\log\psi_n(x) - iK\alpha\omega_n t = S(x) - iE_n t$, exactly as assumed by Schrödinger.

[18] This is manifest in the fact that Born's substitution rule (3.8) for $\Phi = J$ implies $\alpha h = J(n+\alpha) - J(n)$.

3.3 The Dynamical Duality of "Unified" Quantum Mechanics

John von Neumann's *The Mathematical Foundations of Quantum Mechanics* is arguably quantum-mechanics-as-we-know-it. We ask the reader to consider, in this connection, the following facts.

1) von Neumann presented his mathematical treatise as a "unification" of the "equivalent" Matrix and Wave Mechanics;

2) It is with von Neumann's treatise that the most intensely investigated problem in the foundations of quantum mechanics, namely the "measurement problem," or "problem of outcomes" (MAUDLIN, 1995), is traditionally associated in the literature.

The main purpose of this section is to investigate, first, the link between these two facts, and then discuss the connection with our results.

3.3.1 The Equivalence of the Two Theories: Hilbert Space?

We want to begin by making explicit the specific tactics we have adopted. First, we shall indicate just what von Neumann's mathematical task, as he sees it, is—and for this, a couple of quotations from (NEUMANN, 1932, §I) will have to suffice; second, we shall unpack, since von Neumann did not, the mathematical burden of the task. Our own task will then be straightforward: Verify critically, in logical and mathematical terms, just how the burden was met in the body of the book. The logical outcome of this will be a reassessment of the mathematical foundations of quantum mechanics.

Consider the following statement from §I.1, entitled "The Origin of the Transformation Theory":

> A procedure initiated by Heisenberg [i.e., Matrix Mechanics] was developed by Born, Heisenberg, Jordan, and a little later by Dirac, into a new system of quantum theory, the first complete system of quantum theory which physics has possessed. A little later Schrödinger developed the "wave mechanics" from an entirely different starting point. This accomplished the same ends, and soon proved to be equivalent to the Heisenberg, Born, Jordan and Dirac system (at least in a mathematical sense, cf. 3 & 4 below [here a footnote refers the reader to Schrödinger's equivalence paper]). On the basis of the Born statistical interpretation of the quantum theoretical description of nature, it was possible for Dirac and Jordan to join the two theories into one, the "transformation theory," in which they make possible a grasp of physical problems which is especially simple mathematically. (NEUMANN, 1932, 6)

Now if one is interested, as we are, in pinning down the facts involved in engineering a "unified" theory of quantum processes based on the equivalence of Matrix and Wave Mechanics, one has to look beneath von Neumann's remark that it was Born's statistical interpretation that allowed Transformation Theory "to join the two theories into one." For Born introduced, as is well-known, the statistical interpretation of the wave-function in his papers on collisions and on a paper about the adiabatic theorem, all of which were published in June-October 1926 (see (BACCIAGALUPPI, 2022) for the interesting details); Schrödinger, let us recall, published his equivalence paper in March 1926—*before* Born's papers on collisions. Now consider, with this in mind, the fact that therein Schrödinger showed that the isomorphism $f \mapsto (f_z)_{z \in Z}$ defined by $f(x) = \sum_{z \in Z} f_z \psi_z(x)$ for all $x \in \Omega$, where ψ_z are an orthonormal basis of solutions to Schrödinger's equation, links the matrix-elements of Matrix Mechanics—which, with respect to their indices, "taken singly, specify the *states*, and in pairs, specify the *transitions*" (HEISENBERG; BORN; JORDAN, 1926, 339)—, to the Fourier components of the wave functions of Wave Mechanics. We already remarked in §3.2.3 upon the fact that Schrödinger made explicit use of this isomorphism to calculate the intensities of the Stark effect patterns in (SCHRÖDINGER, 1926d, §4)—a paper also published before Born's first paper on collisions, in May 1926. In "The Exchange of Energy According to Wave Mechanics," Schrödinger's last paper on Wave Mechanics, published in June 1927, Schrödinger stated the obvious: His rule to compute the intensities—which, let us repeat expressly once more, he had learned from Matrix Mechanics—, and Born's statistical interpretation of the wave function, are mathematically equivalent.

Formally speaking, therefore, the Born rule was actually anticipated by Schrödinger, and is nothing but an automatic consequence of his equivalence proof.[19]

This is the first main conclusion of this section. It shows that when von Neumann remarked that it was Born's statistical interpretation that allowed "joining the two theories into one," he was turning the tables. The mathematical fact of the matter is that once Schrödinger "joined the two theories into one," the Born rule was (formally speaking, we repeat) unavoidable.

Good. To proceed, then, with von Neumann's argument: in §I.3, called "The

[19]This conclusion is in disagreement with some of the details of (BACCIAGALUPPI, 2022). The facts described as "odd" by Abraham Pais in the Postscript to his "Max Born's Statistical Interpretation of Quantum Mechanics," however, are in full harmony with it:

> Jorgen Kalckar from Copenhagen wrote to me about his recollections of discussions with Bohr on this issue. "Bohr said that as soon as Schrödinger had demonstrated the equivalence between his wave mechanics and Heisenberg's matrix mechanics, the 'interpretation' of the wave function was obvious.... For this reason, Born's paper was received without surprise in Copenhagen. 'We had never dreamt that it could be otherwise,' Bohr said." (PAIS, 1982a, 1198)

.

Equivalence of the Two Theories: The Transformation Theory," von Neumann remarks that against Dirac and Jordan's Transformation Theory stands the determinative requirement of mathematical rigor:

> We do not desire to follow any further here this train of thought which was shaped by Dirac and Jordan into a unified theory of the quantum processes. The "improper" functions (such as $\delta(x)$, $\delta'(x)$) play a decisive role in this development——they lie beyond the scope of mathematical methods generally used, and we desire to describe quantum mechanics with the help of these latter methods. *We therefore pass over to the other* (*Schrödinger*) *method of unification of the two theories.* (NEUMANN, 1932, 20, italics ours)

In §I.4, a section entitled "The Equivalence of the Two Theories: Hilbert Space," von Neumann then argues that the reason the Dirac-Jordan Transformation Theory runs into trouble is that it tries to relate the "discrete" space $Z = \{1, 2, 3, \ldots\}$ of index values, the arena of Matrix Mechanics, with the continuous state space $\Omega \subset \mathbb{R}^d$, the arena of Wave Mechanics, but this is impossible. The spaces are completely unrelated. *But*, he says, the *functions* in these spaces—the space F_Z of complex sequences $(x_z)_{z \in Z}$ and the space F_Ω of wave-functions $\Phi : \Omega \to \mathbb{C}$—are isomorphic.[20] It is the fact that Schrödinger related matrices to functions in his equivalence paper that is behind the italicized sentence in the passage quoted above: in (NEUMANN, 1932, 22, footnote 35) von Neumann remarks that Schrödinger's equivalence proof corresponds to the part of the theorem proved by Fischer and Riesz on the isomorphism of F_Z and F_Ω which Hilbert proved in 1906. "But Schrödinger didn't really establish the equivalence of Matrix and Wave Mechanic," the reader may now object. True—but the glitz of this mathematical fact should not blind us to the historical fact of the matter: that Schrödinger's contention that "from the formal mathematical standpoint, one may speak of the *identity* of the two theories" (SCHRÖDINGER, 1926a, 46) was *accepted*. For it is Schrödinger's contention, and the full Riesz–Fischer theorem, that are von Neumann's starting points:

> Since the systems F_Z and F_Ω are isomorphic, *and since the theories of quantum mechanics constructed on them are mathematically equivalent*, it is to be expected that a unified theory, independent of the accidents of the formal framework selected at the time, and exhibiting only the really essential elements of quantum mechanics, will then be achieved if we do this: Investigate the

[20]The space-state F_Z in Matrix Mechanics corresponds to set of square summable sequences $\ell^2(Z, \mathbb{C})$; meanwhile, in Wave Mechanics, the space-state corresponds to the set of complex square-integrable functions $L^2(\Omega, \mathbb{C}, dx)$ on the measurable subset $\Omega \subset \mathbb{R}^d$. It turns out that $L^2(\Omega, \mathbb{C}, dx)$ is a separable Hilbert space with the usual inner product $\langle f|g\rangle = \int_\Omega \overline{f(x)} g(x) dx$ and all separable Hilbert spaces are unitarily isomorphic by considering two orthonormal basis $\{e_n\}_{n \in \mathbb{N}} \subset \mathcal{H}_1$ and $\{f_n\}_{n \in \mathbb{N}} \subset \mathcal{H}_2$, and the linear map such that $e_n \mapsto f_n$ for all $n \in \mathbb{N}$.

> intrinsic properties (common to F_Z and F_Ω) of these systems of functions, and choose these properties as a starting point. (NEUMANN, 1932, 24)

von Neumann is therefore clearly taking the equivalence between Matrix and Wave Mechanics *as a given*. The isomorphism between F_Z and F_Ω is also given; only one must note, and think through the fact, that such givens are independent. For kinematical equivalence—let us repeat it expressly once more—does not imply dynamical equivalence.

The isomorphism between F_Z and F_Ω is a relation between *static* structures; the Hilbert space itself contains no information about dynamics ("the dependence on t is not to be considered in forming the Hilbert space" (NEUMANN, 1932, 128)). Since it must thus be acknowledged that the equivalence of Matrix and Wave Mechanics is not encoded in the abstract structure of the Hilbert space, to substantiate, then, the "expectation" that a "unified theory [...] exhibiting only the really essential elements of quantum mechanics" should be erected over Hilbert space, von Neumann must present *independent arguments* to prove that the equivalence between Matrix and Wave Mechanics—their complete dynamical parallelism from the mathematical viewpoint—does follow of necessity from the Hilbert space.

We will follow this conclusion as a heuristic principle in our analysis of von Neumann's work. It is in connection with it that we now want to recall that von Neumann presented a proof of the Stone-von Neumann theorem in §II.9 of his treatise; this theorem, which is too well-known to warrant description, is referred to in the literature as "the mathematical reason behind the equivalence of the Heisenberg quantum mechanics and the Schrödinger wave mechanics" (EMCH, 1987, 333). Let us assess this statement.

1. The Stone-von Neumann theorem clearly doesn't shed any light whatsoever on how it can be that Heisenberg's and Schrödinger's apparently orthogonal initial assumptions turn out to be different aspects of the same thing—unlike our (3.2.4);

2. The Stone-von Neumann theorem is a result about the unitary equivalence of irreducible representations of the canonical commutation relations, which entails unitarily equivalent *algebras*, which entails a unitary map linking the Hamiltonian $H(p, q)$ of Wave Mechanics to Matrix Mechanics and conversely. But consider the facts described in §3.2.1-3—it will immediately appear that unitarily equivalent Hamiltonians does not establish dynamical equivalence between Matrix and Wave Mechanics. For Matrix Mechanics posited, besides a unitary evolution for the Hamiltonian matrix H, that individual transition processes happen in discontinuous jumps to stationary states. The Stone-von Neumann theorem, therefore, is not—it cannot be— "the mathematical reason behind the equivalence of the

Heisenberg quantum mechanics and the Schrödinger wave mechanics." There is more to the dynamics of quantum mechanics than the Hamiltonian.

The Stone-von Neumann theorem alone therefore is not sufficient to establish the equivalence of Matrix and Wave Mechanics. This is the second main conclusion of this section. It is with it in the back of our thinking that we now with to consider critically the following facts.

In (NEUMANN, 1932, §III) von Neumann proposed that the evolution of a quantum system occurs in two stages which he called "process 1" and "process 2": a non-unitary, stochastic "jump" induced by measurements, and a unitary evolution described by Schrödinger's time-dependent equation of motion between measurements. This duality—neither encoded in the static Hilbert space structure, nor implied by the Stone-von Neumann theorem—must then be justified. We will now consider how von Neumann framed the problem, and stated his answer, in chapters V and VI respectively.

We read in chapter V:

> We must now analyze in greater detail these two types of [processes]—–their nature, and their relation to one another. First of all, it is noteworthy that [process] 2 admits of the possibility that H is time-dependent, *so that one might expect that 2 would suffice to describe interventions caused by measurement*: indeed, a physical intervention can be nothing other than the temporary insertion of a certain energy coupling into the observed system; i.e., the introduction into H of a certain time dependency (prescribed by the observer). *Why then have we need—for measurements—of the special process 1?* The reason is this: In a measurement we cannot observe the system S by itself, but must rather investigate the system S + M in order to obtain (numerically) its interaction with the measuring apparatus M. The theory of measurement is a statement concerning S + M, and should describe how the state of S is related to certain properties of the state of M (namely, the positions of a certain pointer, since the observer reads these). Moreover, it is rather arbitrary whether one includes the observer in M, and replaces the relation between the S state and the pointer positions in M by relations between this state and chemical changes in his eye or even in his brain (i.e., to that which he has "seen" or "perceived"). We shall investigate this more precisely in VI.1. In any case, the application of 2 is of importance only for S + M. Of course, we must show that this gives the same result for S as does the direct application of 1 to S. If this proves successful, then we will have achieved a unified way of looking at the physical world on a quantum mechanical basis. (NEUMANN, 1932, 230, italics ours)

The measurement problem, as von Neumann sees it, is therefore this:

MP Formulation α. Why does process 2 not suffice, in practice, to describe measurements, given that it could in principle (if H is time-dependent)? Why do we need process 1 for measurements?

Since the completeness of the wave-function, and the eigenvector-eigenvalue link, are implicitly presupposed, we find in MP Formulation α (henceforth, MP α) the historical origin of the so-called problem of outcomes (MAUDLIN, 1995) and of the view von Neumann introduced process 1 to solve it. Now consider against this fact the further fact that when von Neumann returns to the question in VI.1, "The Measurement Process: Formulation of the Problem," he puts the matter in different, more general, terms:

> In the discussion so far we have treated the relation of quantum mechanics to the various causal and statistical methods of describing nature. In the course of this, we have found a peculiar dual nature of the quantum mechanical procedure which could not be satisfactorily explained. (NEUMANN, 1932, 269)

Here, then, the measurement problem reads:

MP Formulation β. What is the explanation for the dual nature of quantum dynamics?

MP α thus makes explicit a fundamental assumption that MP β does not, namely that it is theoretically conceivable that process 2 should hold good for all processes. We are therefore logically justified to distinguish between such formulations. Furthermore, the fact that MP β is more general than MP α makes it plain that it is not the measurement problem itself, but the "solution" von Neumann gives to it, that is investigated "more precisely" in chapter VI than in chapter V. For the answer in chapter VI reads (excuse the long quote):

> First, it is inherently correct that measurement or the related process of subjective perception is a new entity relative to the physical environment, and is not reducible to the latter. Indeed, subjective perception leads us into the intellectual inner life of the individual, which is extra-observational by its very nature, since it must be taken for granted by any conceivable observation or experiment. Nevertheless, it is a fundamental requirement of the scientific viewpoint——the so-called *principle of psycho-physical parallelism*——that it must be possible so to describe the extra-physical process of subjective perception as if it were in the reality of the physical world; i.e.,

to assign to its parts equivalent physical processes in the objective environment, in ordinary space. [...] In a simple example, these concepts might be applied as follows: We wish to measure the temperature. If we want, we can proceed numerically by looking to the mercury column in a thermometer, and then say: "This is the temperature as measured by the thermometer." But we can carry the process further, and from the properties of mercury (which can be explained in kinetic and molecular terms) we can calculate its heating, expansion, and the resultant length of the mercury column, and then say: "This length is seen by the observer." Going still further, and taking the light source into consideration, we could find out the reflection of the light quanta on the opaque mercury column, and the path taken by the reflected light quanta into the eye of the observer, their refraction in the eye lens, and the formation of an image on the retina, and then we would say: "This image is registered by the retina of the observer." And were our physiological knowledge greater than it is today, we could go still further, tracing the chemical reactions which produce the impression of this image on the retina, and in the optic nerve and in the brain, and then in the end say: "These chemical changes of his brain cells are perceived by the observer." But in any case, no matter how far we proceed——from the thermometer scale, to the mercury, to the retina, or into the brain——at some point we must say: "And this is perceived by the observer." That is, we are obliged always to divide the world into two parts, the one being the observed system, the other the observer. In the former, we can follow all physical processes (in principle at least) arbitrarily precisely. In the latter, this is meaningless. The boundary between the two is arbitrary to a very large extent. In particular, we saw in the four different possibilities considered in the preceding example that the "observer"——in this sense——need not be identified with the body of the actual observer: in one instance we included even the thermometer in it, while in another instance even the eyes and optic nerve were not included. That this boundary can be pushed arbitrarily far into the interior of the body of the actual observer is the content of the principle of psycho-physical parallelism. But this does not change the fact that in every account the boundary must be put somewhere if the principle is not to be rendered vacuous; i.e., if a comparison with experience is to be possible. (NEUMANN, 1932, 272-273, italics in the original)

To this he adds:

Quantum mechanics describes events which occur in observed portions of

> the world, so long as they do not interact with the observing portion, and does so by means of process 2 (V.1). But as soon such an interaction does occur—i.e., a measurement is made—the theory requires application of process 1. This duality is therefore fundamental to the theory. [He adds in a footnote: "Bohr was the first to point out that the duality which is necessitated by quantum formalism, by the quantum mechanical description of nature, is fully justified by the physical nature of things, and that it may be connected with the principle of psycho-physical parallelism."] Danger, however, lies in the fact that the principle of psycho-physical parallelism is violated so long as it is not shown that the boundary between the observed system and the observer can be displaced arbitrarily, in the sense given above. (NEUMANN, 1932, 273)

The problem of showing that the boundary between the system and the outside can be shifted arbitrarily, called by J. Bub "the consistency problem" (BUB, 2001, §2), is solved by von Neumann in VI.2. But the solution to the consistency problem does not imply that the principle of psycho-physical parallelism is the correct answer to the measurement problem (as defined in MP α). And the consensus in the international community of scholars today is that indeed *it is not*.

The reason is evident. To say that "the duality which is necessitated by the quantum formalism [...] is fully justified by the physical nature of things, and that it may be connected with the principle of psycho-physical parallelism" is to provide a metaphysical answer to a dynamical question. And to give to a dynamical question anything else but a dynamical answer is a mistake of a particularly infelicitous kind. It is a category mistake.

However, the sophisticated discussions of scholars about von Neumann's answer to the measurement problem need not detain us.[21] What *must* detain us is that which scholars have universally taken for granted, namely von Neumann's formulation of the question. For even though MP α has caused a flood of literature, the assumption that process 2 could, in principle, suffice for measurements if H is time-dependent has always passed over as innocent. This is surely a massive oversight, for we find reason to suspect that it may not be in von Neumann's own words; a fundamental claim von Neumann made earlier in the book is in flat contradiction with it. And two contradictories cannot, by the laws of classical logic, both be true simultaneously.

In §3.2.1 we quoted a passage from chapter I of von Neumann's book wherein he emphasized that "all occurrences of an atomic molecular order obey the discontinuous law of quanta." He reminded the reader at the outset of the fact that "the general opinion

[21]But see Abner Shimony's masterful (SHIMONY, 1963).

in theoretical physics had accepted the idea that the classical principle of continuity is merely simulated in Nature;" he underscored that we live in a world "which *in truth* is *discontinuous* by its very nature." Now this implies as a matter of course that it is the certain effect of "the insertion of an energy coupling into the observed [quantum] system"—i.e., a finite disturbance of the system from the outside, *or a measurement*—that the system will undergo a jumpwise transition. The contradiction mentioned above is now entailed by the simple fact that the Schrodinger equation, which describes the unitary dynamics, is a differential equation. It is based on the property of continuity which von Neumann asserts Nature itself lacks. How can it really be *thinkable* that the *continuous* process 2 could in principle suffice to describe the *discontinuous* transitions, as von Neumann says in MP α? Must we not necessarily maintain, "so that one might expect that 2 would suffice to describe interventions caused by measurement," that 2, *and the interventions caused by measurement*, are continuous?

We are logically obliged to say that we must. But let us press the matter further; von Neumann's inconsistency on this point can be unproblematical if, and only if, he is implicitly admitting the possibility that "the insertion of an energy coupling into the observed system" can happen continuously. Let us assume that this is the case; does MP α make sense then? We now categorically assert: *it does not*. For the insertion of a time-dependency into H does *not* change the fact that the Schrödinger equation *cannot* describe, *even in principle*, "the insertion of an energy coupling into the observed system," *even if the world is continuous all the way down*, as the following remarks will readily demonstrate.

It is at this point that we bring the results of §3.2 to the table. Condition (1), which is equivalent to Schrödinger's equation, holds only for $\delta\mathcal{S}_0[\Psi] = 0$, which means: *no change in quantum number*. We saw in §3.2.2 that Schrödinger derived the time-independent wave equation in (1926b) by starting from the time-independent Hamilton-Jacobi equation; a straightforward calculation shows, however, that it suffices to start from the time-*de*pendent Hamilton-Jacobi equation, and then follow further the same procedure, to obtain the time-*de*pendent Schrödinger equation for a Hamiltonian that depends explicitly on t. In both cases the term $\mathcal{S}_0$, which accounts for the transitions, is bypassed; what happens if H depends on t in process 2 is only that besides the wave-function $\psi(x,t)$, the energy associated to the quantum numbers n_t will then depend on t. But the condition $\delta\mathcal{S}_0[\Psi] = 0$—the point cannot be overemphasized—*still holds*. If one inserts a time-dependency into H, *it is still the case* that the wave equation *cannot* describe the transitions between any two given quantum numbers n_t and m_t at any given time.[22]Even though it is, indeed, true that process 2 admits of the possibility

[22]This is not to deny that it is in principle possible, given a initial state ψ, a final state ϕ, and a time instant t_1, to tailor a time-dependent Hamiltonian $H_{\phi,\psi,t_1}(t)$ such that the solutions of the Schrödinger

that H is time-dependent, the mathematical fact that one *cannot* expect that 2 would suffice to describe interventions caused by measurement is not changed one iota; von Neumann's statement to the contrary is plausible only so long one ignores the premises from which the time-dependent Schrödinger equation is deduced.

To protect ourselves against any charge of historical inaccuracy, we must point out to the fact that Schrödinger himself did not find the time-dependent wave equation in this way. In Schrödinger's own derivation, however, it is only Ψ, not H, that depends on t; Schrödinger, to make the dependence of Ψ on the time t explicit, merely eliminated the E parameter in the time-independent equation, upon using the trivial fact that for $\Psi(x,t) = \Psi(x)e^{\pm\frac{2\pi iEt}{h}}$, $\frac{\partial\Psi}{\partial t} = \pm\frac{2\pi i}{h}E\Psi$. That everything that was said in the above paragraph holds for his derivation too is evident without explication. But this—mind you—Schrödinger explicitly said it himself; an important, yet half-forgotten, fact is that even though Schrödinger begins (SCHRÖDINGER, 1926e), the paper wherein he proposed the time-dependent equation derived in the way just mentioned, by saying that

> there arises an urgent need for the extension of the theory for *non-conservative* systems, because *it is only in that way that we can study the behavior of the system under the influence of prescribed external forces*, e.g. a light wave, or a strange atom flying past. (SCHRÖDINGER, 1926e, 103, italics ours),

he ends the paper with the remark: "I have not succeeded in forming [the wave equation] for the non-conservative case" (SCHRÖDINGER, 1926e, 123).

Our contention has thus been proved. It is not right, it is not legitimate, to expect that 2 would suffice to describe the interventions caused by measurements even if one omits the demand, which von Neumann said is imposed on us by Nature itself, that external perturbations give rise to discontinuous transitions.

This is the third main conclusion of this section. It follows from it that MP α is, in a word, spurious. For if process 2 can't describe the interventions caused by measurement *even in principle*, then it must be recognized that it makes no sense to ask why it doesn't do it in practice.[23]

equation satisfies $\psi(x,t_1) = \phi(x)$. We deny only that this constitutes a counter-example that can be cited as proof against the correctness of our argumentation. For not only to construct such a H_{ϕ,ψ,t_1} requires knowing the time t_1 and the final state ϕ in advance. It is also transparent that the described evolution from ψ to ϕ is not a *bona fide* transition.

[23]But there is more to this than meets the eye—we shall consider the facts that reveal the kind of game von Neumann was playing in stating MP α in §3.2.

So logic will now demand that we get rid of MP α and focus on MP β. It is the exact same question Muller named "the measurement explanation problem" in his (MULLER, 2023):

> The two postulates of standard QM that mutually exclude and jointly exhaust the change of state over time (Dynamics and Projection Postulate) evoke the question: why *two*, and why *these two?* (MULLER, 2023, 25)

Let remind ourselves why von Neumann was forced to ask MP β to begin with. The reason lies in the heuristic principle defined a few pages back: The Hilbert space formalism of quantum mechanics was designed to be a self-standing structure independent of the "accidents" of the formalisms of Matrix and Wave Mechanics; the Hilbert space structure, however, being itself static, does not carry within itself the dynamical content of Matrix and Wave Mechanics. The dynamics of unified quantum mechanics had thus to be introduced independently from the outside. But unification requires proving the equivalence of Matrix and Wave Mechanics *in Hilbert space*; the Stone-von Neumann theorem alone, however, does not suffice, as we have seen, to attain equivalence. Thus von Neumann could not adequately justify, on the basis of his framework alone, why quantum mechanics should have the *dual* dynamics that it does.

But note, reader, that our Scholium 3.2.4 can.

Mind you: If historically, Heisenberg, Born, and Jordan learned how to describe the evolution of the state of the isolated system only after Schrödinger presented his wave equation, and Schrödinger, by his turn, learned how to calculate the transitions only after he noted that the Fourier coefficients of the wave equation are isomorphic to Heisenberg's matrix coefficients, so that the intensities are given by computing their square (Born rule), this was, it appears, no contingent development. By the constraints of the Scholium 3.2.4 such a course of events could hardly have been different as a matter of pure mathematics. Each rival theory started independently from the other, as we saw in §3.2.1-2, *by considering a different component of the action*; each theory then *learned* from the other *the missing piece of dynamical information required for a fuller dynamical description.* (1), the condition upon which Wave Mechanics was based, tells us that the unitary evolution holds only in so far there is no change of quantum number; (2), the condition upon which Matrix Mechanics was based, prescribes the eigenvector-eigenvalue link (the quantum condition on the dynamically allowed states) and the collapse postulate (the "empirically given" jumpwise transitions). The first condition is mathematically and conceptually tied to the second by the bounds of the canonical commutation relations, *via* $\mathcal{J}$ and $\mathcal{S}_0$; the mathematical meaning of the relationship is clear: *both conditions must be satisfied simultaneously if they hold individually.*

This answers MP β unequivocally. This tells us that the dynamics of a quantum mechanics based on the equivalence of Matrix and Wave Mechanics must have indeed *two* dynamical postulates, and it tells us that it must be *these two*.

MP β is a general question about dynamics, and an adequate, accurate answer should itself be grounded on dynamics. This is precisely what the Scholium 3.2.4 yields. It shows that and how the duality necessitated by the quantum formalism is itself necessitated by the dual nature of the quantum action.

This is the central conclusion of this section. We should like to remark in connection to it that in other branches of physics it is generally taken for granted that the dynamics of a physical theory should follow from the action functional. What we have disclosed here is the hitherto ignored fact that orthodox quantum mechanics is no exception. If orthodox quantum mechanics, based as it is on the equivalence of Matrix and Wave Mechanics, has this "peculiar dual dynamics," this is so because one has been assuming a peculiar dual action.[24]

3.3.2 The Equivalence of the Two Theories: Hilbert Space

When Muller remarked, in the passage quoted above, that the two dynamical postulates of orthodox quantum mechanics exclude one another, he was expressing a very popular view. It is notewhorthy that this view is closely related to the also very popular view that von Neumann posited process 1 in an arbitrary, *ad hoc* manner, only to solve the measurement problem. This affinity deserves our attention. For, since the dual action implies that the first view is mistaken, insofar as according to it the two dynamical postulates do not exclude, but in actual fact require, one another, it would be perplexing if the second view were true. We want to expose the fact that is is not. Regardless of the antipathy that we may have for von Neumann's solution to the measurement problem, it is illegitimate to criticize it on the grounds that it was arbitrary. It wasn't. And that should be made very clear.

The history of science evinces it. The reader familiar with the historical development of quantum mechanics will recall that the discontinuous, temporally undetermined transitions are the very idea that set the quantum revolution in motion; "the fundamental laws [of the new mechanics]," wrote James Jeans in 1914, "must be based on discontinuities, and not on the ideas of continuity involved in the classical mechanics" (JEANS, 1914, 7). The reader will also recall from §3.2.1 that Matrix Mechanics, which appeared, let us not forget, in 1925, was by design "a true theory of the discontin-

[24]It is perfectly true that the action pertains to the observable H alone, whereas the form of the collapse depends on the particular choice of measurement. But our point remains unscathed, for any observable must evolve, irrespective of the choice of measurement, according to the temporal evolution of the formalism, and that is determined in orthodox quantum mechanics by the dual action.

uum;" Schrödinger derived the continuous wave equation, which *ex hypothesi* could not describe the transitions, in 1926. In 1927, Niels Bohr, in the famous Como Lecture, remarked that

> notwithstanding the difficulties which hence are involved in the formulation of the quantum theory, it seems, as we shall see, that its *essence* may be expressed in the so-called quantum postulate, *which attributes to any atomic process an essential discontinuity* [...]. (BOHR, 1927, 88, italics ours)

and to this he added, further in the text, the important point about unifying Matrix and Wave Mechanics:

> Schrödinger has expressed the hope that the *development* of the wave theory will *eventually remove the irrational element expressed by the quantum postulate*[25] *and open the way for a complete description of atomic phenomena along the line of the classical theories*.[26] In support of this view, Schrödinger, in a recent paper (SCHRÖDINGER, 1927), emphasises the fact that *the discontinuous exchange of energy between atoms required by the quantum postulate,* from the point of view of the wave theory, is replaced by a simple resonance phenomenon.[27] In particular, the idea of individual stationary states [eigenvector-eigenvalue-link] would be an illusion and its applicability only an illustration of the resonance mentioned. It must be kept in mind, however, that just in the resonance problem mentioned we are concerned with a closed system which, according to the view presented here, is not accessible to observation. In fact, wave mechanics just as the matrix theory on this view represents a symbolic transcription of the problem of motion of classical mechanics adapted to the requirements of quantum theory[28] and [so wave mechanics is] *only to be interpreted by an explicit use of the quantum postulate*. (BOHR, 1927, 110, italics ours)

Now, reader, please note: What else can it *mean* to make explicit use of the quantum postulate, "which attributes to any atomic process an essential discontinuity," in

[25]The essence of Schrödinger's argument is that since the discrete states can be derived from his continuous wave equation, we should expect that accordingly, any transition between also happen continuously.

[26]It is noteworthy that Bohr while calls discontinuity the "essence" of quantum mechanics, he at the same time qualifies it as the "irrational element" of the theory. This seems to be the fundamental reason he preached quantum processes cannot, in principle, be understood. It is also noteworthy that Bohr is saying here that by removing discontinuity, as Schrödinger hoped, one would go back to classical physics.

[27]Discontinuity is in this view "apparent;" compare with Einstein's quote below.

[28]Note then that to Bohr the problem of quantum theory is to describe motion given that its central aspect— discontinuity—violates the mechanical principle.

Wave Mechanics, other than posit, as von Neumann did, that when there's energy exchange between two partial systems, as, say, in a measurement, Schrödinger's continuous wave-function must undergo discontinuous collapses?

Mind you, von Neumann did say it explicitly, after introducing the collapse postulate (NEUMANN, 1932, §III): "these jumps are related to the 'quantum jumps' concept of the older Bohr theory" (NEUMANN, 1932, 161, footnote 125).

This, then, seems to be the simple connection that people do not make when they say that von Neumann postulated process 1 only to eliminate the incompatibility between process 2 and the results of measurements: that von Neumann's collapses are a generalization of Niels Bohr's quantum jumps. And the idea of these jumps did not originate with Bohr. It was the very idea that set the quantum revolution in motion.

Consider briefly the fact that it became customary to take at face value the scandalous idea that Niels Bohr somehow succeeded in "brainwashing a generation of scientists" and explain the apparent hegemony of the so-called "Copenhagen Interpretation" in the pre-Bell period by appealing to sociology (BELLER, 1999). What many scholars, preoccupied with the sociological phenomenon, failed to realize is that the empirical hypothesis, which Bohr defended with great forcefulness, had its roots deep indeed in the development of the quantum theory: "The keynote of the old mechanics was continuity, *natura non facit saltus*. The keynote of the new mechanics is discontinuity" (JEANS, 1914, 89). As a consequence of this, they have also overlooked the plain fact that the protests of Schrödinger and Einstein against the completeness of quantum mechanics involved a denial of the discontinuity assumption; "it is worthwhile to hold on to this," Einstein wrote to Schrodinger as late as 1950, "the continuum, as long as one has no really sound arguments against it" (PRZIBRAM et al., 1967, 40). The passage quoted above from Bohr's Como Lecture evinces that Schrödinger was sure that *further developments*—hence not the results that he had already achieved—would eliminate the "irrational element" of quantum mechanics (Bohr), "these damned quantum jumps" (Schrödinger); Einstein, by general acclaim the great champion against the quantum orthodoxy, is nevertheless curiously ignored by contemporary investigators when he tells us, in the *Autobiographical Notes*, why it is that he, like Schrödinger, could not agree with the *status quo*:

> This is the point at which my expectation departs most widely from that of contemporary physicists. They are convinced that it is impossible to account for the *essential* aspects of quantum phenomena (*apparently discontinuous and temporally not determined changes of the situation of a system*, and at the same time corpuscular and undulatory qualities of the elementary bodies of energy) by means of a theory which describes the real state of things [objects]

> by continuous functions of space for which differential equations are valid. They are also of the opinion that in this way one can not understand the atomic structure of matter and of radiation.[29] [...] Above everything else, however, they believe that *the apparently discontinuous character of elementary events can be described only by means of an essentially statistical theory*, in which the discontinuous changes of the systems are taken into account by way of the continuous changes of the probabilities of the possible states. (EINSTEIN, 1949, 87, italics ours)

Thus in this, Einstein agrees with Bohr: The discontinuous transitions *are* a key feature of quantum phenomena. The disagreement is that while to Bohr (and virtually everyone else) discontinuity was "essential," to Einstein (and Schrödinger) discontinuity was "apparent," and thus should be derived from a more fundamental description. The following remarks by Arnold Sommerfeld, in his essay following Einstein's *Autobiographical Notes* in Schilpp's volume, are in full agreement with our conclusions:

> In the old question "continuum versus discontinuity," [Einstein] has taken his position most decisively on the side of the continuum. Everything of the nature of quanta—to which, in the final analysis, the material atoms and the elementary particles belong also—he would like to derive from a continuum-physics [...]. His unceasing efforts, since he resides in America, have been directed towards this end. Until now, however, they have led to no tangible success. [...] By far the most of today's physicists consider Einstein's aim as unachievable. (SOMMERFELD, 1949, 105)

There can then be no denying that Einstein's and Schrödinger's battle for rationality in physics was essentially a battle against the "irrational" quantum postulate (Bohr), and all the irrationalities that it implies; their warning voices confronted "the general opinion in theoretical physics [which] had accepted the idea that the principle of continuity ('*natura non facit saltus*'), prevailing in the macroscopic world, is merely simulated by an averaging process in a world which in truth is discontinuous by its very nature" (von Neumann). The nature of their disagreement could therefore hardly have been more fundamental, for if discontinuity was indeed, as Einstein and Schrödinger believed, only "apparent," then one is led to the conclusion that orthodox quantum mechanics has been erected from the ground up on the basis of a mistaken dynamical hypothesis.

But the question of completeness is not our concern. The point we are leading up by way of all this is only that it must be conceded that as of 1950, orthodox quantum mechanics was essentially the same theory it is today, i.e., the theory described in von

[29] See footnote 24.

Neumann's *The Mathematical Foundations of Quantum Mechanics*. And so, if now it became standard to view the discontinuous, temporally-not-determined quantum jumps as an *ad hoc* postulate von Neumann introduced "in an act of desperation to solve the measurement problem," a feature of quantum mechanics almost no one believes, today, that's true, whereas before, the jumpwise transitions were demonstrably seen by all but a few as the "an essential aspect of quantum phenomena," as the indestructible "essence" of quantum mechanics, then we must conclude from the disparity that the foundations of quantum mechanics has somehow lost touch with its past. We must acknowledge that it can't really *be* that the collapse postulate was introduced by von Neumann to solve the measurement problem, for the collapse has a prehistory that antedates the discovery of the Schrödinger equation. We must admit that the popular view that von Neumann engineered process 1 merely as an *ad hoc* solution to escape from the contradiction posed by the problem of outcomes is historically untenable, and therefore false.

In this way we reach coherence with the facts described in §3.3.1. For there we proved that it is mathematically illegitimate to expect that process 2 can in principle describe measurements, and this implies admitting, *contra* popular belief, that MP α is no genuine question. And the fact MP α is no genuine question is in good agreement with our conclusion that the collapse postulate was not, *contra* popular belief, manufactured to solve it.

But then we must understand what was von Neumann's *point*, then, in misleadingly formulating the apparently fallacious, apparently spurious, MP α and acting as though the collapse postulate is its solution. For von Neumann *had* a point. It would be a grave injustice to him and his dignity as a mathematician to hastily infer that he must have been oblivious to the fact that the condition $\delta\mathcal{S}_0[\Psi] = 0$ holds for the time-dependent Schrödinger equation, that he, in formulating MP α, unwittingly made a mistake.

So let us think things through. We know that Hilbert space isomorphism is static. And we saw, in the previous section, that von Neumann's mathematical burden was thus to present arguments that establish the equivalence between Matrix and Wave Mechanics in Hilbert space; we saw that to that aim the Stone-von Neumann theorem alone will not do. A linear vector space, the Riesz–Fischer theorem, the Stone-von Neumann theorem, *plus* the collapse postulate—only then one can attain equivalence between Matrix and Wave Mechanics. von Neumann's task as a mathematician was therefore to prove the mathematical necessity of the collapse postulate in Hilbert space. And to do so he availed himself of one of a mathematician's favorite tactics, the *reductio.* Having designed the linear structure to accommodate the discontinuous jumps,[30] he

[30]That the discontinuous jumps, and the statistics they imply, require linearity was emphasized e.g. by

rhetorically asked what would happen if the jumps were eliminated. He then derived contradictions.

The Impossibility Proof is designed to establish the claim: "quantum mechanics is in compelling logical contradiction with causality" (NEUMANN, 1932, 213). Since the quantum jumps are acausal, establishing acausality as an ineliminable feature of quantum mechanics automatically establishes that the discontinuous jumps are an ineliminable part of quantum mechanics. But look beneath this claim. The Impossibility Proof states this: It follows from the general validity of the expression $Exp(R) = \mathbf{Tr}(UR)$ for the expected value of an observable R, given in terms of a density matrix U, that there are no completely dispersion-free states ensembles for quantum quantities, and since hidden variables are supposed to lead to dispersion-free ensembles, it follows that they are impossible. But note that von Neumann naturally assumed that the states associated to the hidden variables behave linearly—they would not fit into Hilbert space otherwise. And note that this amounts to assuming that hidden variables are irrelevant for quantum mechanics to begin with, as pointed out by Grete Herrmann as early as 1933. And it will then follow—and this was von Neumann's whole point—that *they would also undergo discontinuous transitions upon measurements*. The argument to make the acausal jumps appear inevitable is clearly circular, but note, reader, just how von Neumann camouflaged the circle. He camouflaged it in the way he stated the *problem* the Impossibility Proof was designed to address. He contrasted the nature of the statistics of quantum mechanics with that of the kinetic theory of heat. "Although we believe that after having specified we know the state of the system completely," he emphasized, "nevertheless only statistical statements can be made concerning the values of the physical quantities involved" (NEUMANN, 1932, 134).

Heisenberg in the seminal uncertainty paper:

> [When] we carry out a determination of the stationary state, say, by use of an inhomogeneous magnetic field, then we will find that the atom has jumped from the nth state to the mth state with a probability $c_{nm}\overline{c_{nm}}$. When we find experimentally that an atom has indeed jumped to the mth state, then we have to ascribe to it in all calculations thereafter, not the function $\sum c_{nm}S_m$, but simply the function S_m with an undetermined phase. Through the experimental determination, "mth state," we select out of the multitude of different possibilities (c_{nm}) a definite one, m. [...] We can therefore deduce from one experiment the possible results of another by definite statistical rules. The other experiment itself selects out of the plenitude of all possibilities a quite definite one, and thereby limits the possibilities for all later experiments. *Such an interpretation of the equation for the transformation matrix S or the Schrödinger wave equation is only possible because the sum of solutions is again a solution. In this circumstance we see the deep significance of the linearity of Schrodinger wave equations.* On that account they can be understood only as equations for waves in phase space; and on that account *we may regard as hopeless every attempt to replace these equations by nonlinear equations*. (HEISENBERG, 1927, 71-72, italics ours)

The reader will note that if the transitions could be described by non-linear equations, then we'd have to yield to Einstein's and Schrödinger's claim that they happen continuously. It is this that Heisenberg says that is impossible.

A little reflection will now show that there's more to this comparison with the kinetic theory of heat than meets the eye. For the statistics and the dynamics are independent postulates in the kinetic theory of heat, and it goes without saying that determinism is a property not of the equations of state, but of the dynamics. And so we will see, if we stop to think about it, that von Neumann's suggestion that it should be *a priori* possible to implement causality in quantum mechanics depends logically, and therefore necessarily, on the *a priori* possibility that the unitary time-dependent Schrödinger equation *alone* exhausts the dynamics; the possibility of retaining determinism in his framework depends, in other words, on the *a priori* possibility that process 2 can describe the transitions induced by measurement. The exact same assumption von Neumann made, although in that case explicitly, in the statement of the measurement problem.

Consider the claim that von Neumann's Insolubility Proof was designed to establish: "the non-causal nature of the process 1 cannot be attributed to the incomplete knowledge of the observer" (NEUMANN, 1932, 284). The above considerations about the kinetic theory of heat apply again; von Neumann's conclusion should follow from the dynamics of the theory. But that is precisely von Neumann's point. If it is impossible to start, as von Neumann showed that is, with a ignorance-interpretation mixture in the beginning of the measurement, and finish with a ignorance-interpretation mixture in the end,[31] then it must be acknowledged that process 1 is dynamical; it must be acknowledged that the statistics is a consequence of the dynamics. von Neumann's Insolubility and Impossibility Proofs were therefore both designed to prove the mathematical necessity of process 1 as a dynamical postulate; the problem of causality, and the measurement problem, were the vehicles that gave to these proofs a logical foundation. For the reader knows how proofs by contradiction work. One first negates what one wants to establish, and then shows that the consequences of this are not possible. The suggestion that process 2 can in principle describe measurements, or equivalently that process 1 can be dispensed with for S+M mixtures, is made only to show that the opposite is the case; it is calculated to make the general self-consistency point that "the duality is therefore fundamental to the theory" (NEUMANN, 1932, 273).

And so, if one defines the measurement problem as the problem of outcomes, as in (MAUDLIN, 1995), which is nothing but an elaboration of MP α, then it will now seem that the collapse postulate was not, as a matter of historical and mathematical fact, designed by von Neumann to solve the measurement problem. *Instead the measurement problem was invented to justify the collapse postulate.*

In this way we can understand the otherwise mysterious fact that not until the late

[31]See Harvey Brown's (BROWN, 1986). See also the novel analysis given by Guido Bacciagaluppi in (BACCIAGALUPPI, 2013).

1950s—not, in point of fact, insofar as we could verify, until Hugh Everett's "The Theory of the Universal Wave Function" (1957)—one can find no reference in the literature to the contradiction defined by the problem of outcomes.[32] The shift in the foundational status of the temporally undetermined jumps, and the shift in the credibility of the so-called "Copenhagen interpretation,"[33] are, indeed, historically connected. We can identify two components to this shift.

The first is that today's scholars, unlike those from the pre-Bell period, have not been fed discontinuity with their mothers' milk; to these thinkers the "Copenhagen interpretation," which follows out the logic of this "irrational" concept (Bohr), appears then pseudomystical, lawless nonsense. The other component is the "formalism *vs.* interpretation" myth Heisenberg invented in his 1926-1927 confrontations with Schrödinger. For the myth had the effect of giving to von Neumann's edifice a façade of "formalism"—as if physical theories could be invented without making, to start with, physical hypotheses; and since von Neumann pretended, for the reasons just explained, that it was *a priori* possible that process 2 could describe the transitions induced by measurements, he thereby opened the possibility of inferring that it should be *a priori* possible to get rid of the collapse postulate.

In this way we can understand how it can be that the problem of outcomes seems, on a first and second glance, and out of historical context, a serious and legitimate formal difficulty, but in the context of von Neumann's construction it was a set up for *reductio* proofs that together with the Stone-von Neumann theorem, and the Riesz–Fischer theorem, assure that Matrix and Wave Mechanics have indeed been, as they must be *ex hypothesi*, properly unified in Hilbert Space.

This closes our analysis of von Neumann's mathematical foundations of quantum mechanics. We want only to emphasize that our discussion must in no way be confounded with a defense of any version of the "Copenhagen interpretation." Our aim here only to set the facts straight, for we cannot overcome the *fait accompli* of the Copenhagen heritage at their expense. And it is by all means a *fact* that von Neumann's choice of dynamical postulates was not arbitrary. The dual dynamics is necessitated by the equivalence of Matrix and Wave Mechanics, which he took as a given from the outset. The further, deeper question is *why* Matrix and Wave Mechanics are equivalent;

[32] Einstein and Schrödinger always explicitly acknowledged the logical consistency of the formalism of quantum mechanics; so did David Bohm when presenting his hidden-variables interpretation (see the opening remarks of his seminal (BOHM, 1952)). Indeed, the first to rephrase MP α in terms of the now-familiar problem of outcomes, thence starting the trend of understanding "interpretations" of quantum mechanics as a way out of such contradiction, seems to have been Hugh Everett.

[33] The scare quotes are meant to represent the known fact that even though the Copenhagen interpretation is a *fait accompli* as far as textbooks and philosophical discussions are concerned, it is no unified point of view; see the details in (HOWARD, 2022). We must however emphasize that we disagree with Howard in an important point: he claims that Niels Bohr did not subscribe to von Neumann's collapse postulate.

why, in other words, the dual dynamics is fundamental to unified quantum mechanics. It is to this question that we claim to have found an answer. For whereas von Neumann only assumed that Matrix and Wave Mechanics are mathematically equivalent theories, and constructed unified quantum mechanics accordingly, we have been able to trace the matter a step further back, and found the dynamical duality entailed by the equivalence between Matrix and Wave Mechanics to have its origins in the dual nature of a certain action functional.

Bibliography

ABRAHAM, R.; MARSDEN, J. *Foundations of Mechanics*. Providence, R.I.: AMS Chelsea Pub./American Mathematical Society, 2008. Cited on page 105.

BACCIAGALUPPI, G. Insolubility Theorems and EPR Argument. *European Journal for the Philosophy of Science*, p. 87–100, 2013. Cited on page 125.

BACCIAGALUPPI, G. The Statistical Interpretation: Born, Heisenberg, and von Neumann, 1926-27. In: OLIVAL FREIRE JR. *The Oxford Handbook of the History of Quantum Interpretations*. Oxford: Oxford University Press, 2022. p. 203–231. Cited on page 109.

BADINO, M. The Odd Couple: Boltzmann, Planck and the application of statistics to physics. *Ann. Phys.*, v. 18, p. 81–101, 2009. Cited 3 times on pages 26, 28, and 37.

BADINO, M. *The Bumpy Road: Max Planck from Radiation Theory to the Quantum (1896–1906)*. New York: Springer, 2015. Cited 3 times on pages 26, 37, and 41.

BARTOLI, A. *Sopra i movimenti prodotti dalla luce e dal calore*. [S.l.]: Le Monnier, 1876. Cited on page 70.

BECKER, A. *What is Real? The Unfinished Quest for the Meaning of Quantum Physics*. Basic Books: Basic Books, 2018. Cited on page 25.

BELLER, M. *Quantum Dialogue: The Making of a Revolution*. Chicago: University of Chicago Press, 1999. Cited 5 times on pages 25, 96, 100, 105, and 121.

BITBOL, M. *Schrödinger's Philosophy of Quantum Mechanics*. The Netherlands: Springer, 2012. Cited on page 96.

BLUM, A. et al. Translation as heuristics: Heisenberg's turn to matrix mechanics. *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics*, v. 60, p. 3–22, 2017. Cited on page 90.

BOHM, D. *Quantum Theory*. New York: Dover Publications, 1951. Cited on page 25.

BOHM, D. A suggested interpretation of the quantum theory in terms of "hidden" variables. i. *Phys. Rev.*, American Physical Society, v. 85, p. 166–179, 1952. Cited on page 126.

BOHR, N. On the Constitution of Atoms and Molecules. *Philosophical Magazine*, v. 26, p. 1–25, 476–502, 857–75, 1913. Cited on page 26.

BOHR, N. On the quantum theory of line-spectra, part I. *Kongelige Danske Videnskabernes Selskabs Skrifter Naturoidenskabelig og mathematisk afdeling*, n. 8, p. 1–118, 1918. Cited on page 91.

BOHR, N. Atomic Theory and Mechanics. *Nature*, n. 116, p. 845–852, 1925. Cited on page 26.

BOHR, N. The Quantum Postulate and the Recent Development of Quantum Theory. In: J. A. WHEELER AND W. H. ZUREK. *Quantum Theory and Measurement*. Princeton: Princeton University Press, 1927. (Princeton series in physics). Cited on page 120.

BOHR, N.; KRAMERS, H. A.; SLATER, J. C. The Quantum Theory of Radiation. *Philosophical Magazine*, v. 47, p. 785–802, 1924. Cited on page 26.

BOHR, N.; KRAMERS, H. A.; SLATER, J. C. The Quantum Theory of Radiation. *Phil. Mag.*, n. 47, p. 785–802, 1924. Cited on page 91.

BOKULICH, A.; BOKULICH, P. Bohr's correspondence principle. In: ZALTA, E. N. (Ed.). *The Stanford Encyclopedia of Philosophy*. Fall 2020. [S.l.]: Metaphysics Research Lab, Stanford University, 2020. Cited 2 times on pages 92 and 94.

BOLTZMANN, L. Sitzungberichte der Kaiserlichen Akademie der 1517 Wissenschaften. *Mathematisch. Naturwissen Classe*, n. 76, p. 373–435, 1877. Page numbers and quotations taken from the English translation by K. Sharp and F. Matschinsky as: "On the Relationship between the Second Fundamental Theorem of the Mechanical Theory of Heat and Probability Calculations Regarding the Conditions for Thermal Equilibrium." *Entropy* 17 (2015), 1971-2009. Cited on page 49.

BORN, M. über Quantenmechanik. *Z. Phys.*, n. 26, p. 379–395, 1924. Cited 2 times on pages 91 and 94.

BORN, M. Some Philosophical Aspects of Modern Physics. *Proceedings of the Royal Society of Eddinburgh*, v. 57, p. 1–18, 1938. Cited on page 26.

BORN, M. The Conceptual Situation in Physics and the Prospects of its Future Development. *Proceedings of the Physical Society. Section A*, v. 66, p. 501–513, 1953. Cited on page 26.

BORN, M. Statistical Interpretation of Quantum Mechanics. *Science*, American Association for the Advancement of Science, v. 122, n. 3172, p. 675–679, 1955. Nobel laureate lecture. Cited 2 times on pages 90 and 96.

BORN, M.; HEISENBERG, W. On Quantum Mechanics. In: BACCIAGALUPPI, G. AND VALENTINI, A. *Quantum Theory at the Crossroads: Reconsidering the 1927 Solvay Conference*. Cambridge: Cambridge University Press, 1927. p. 408–447. Cited on page 91.

BORN, M.; JORDAN, P. Zur Quantenmechanik. *Z. Phys.*, n. 34, p. 858–888, 1925. Page numbers and quotations taken from the English translation by B. L. van der Waerden as: "On Quantum Mechanics." In: van der Waerden, B. L. (1967), *Sources of Quantum Mechanics*. New York: Dover. Cited on page 91.

BROWN, H. The Insolubility Proof of the Quantum Measurement Problem. *Foundations of Physics*, v. 16, p. 857–870, 1986. Cited on page 125.

BROWN, H. *Physical Relativity: Space-time Structure from a Dynamical Perspective*. Oxford: Oxford University Press, 2005. Cited on page 66.

BUB, J. *The Interpretation of Quantum Mechanics*. Holland: Reidel, 1974. Cited on page 25.

BUB, J. Von Neumann's Theory of Quantum Measurement. In: MIKLÓS RÉDEI AND MICHAEL STÖLTZNER. *John von Neumann and the Foundations of Quantum Physics*. Dordrecht: Springer Netherlands, 2001. p. 63–74. Cited on page 115.

BUCKINGHAM, E. *On the Deduction of Wien's Displacement Law*. [S.l.]: U.S. Department of Commerce and Labor, Bureau of Standards, 1913. Cited on page 76.

BüTTNER, J.; RENN, J.; SCHEMMEL, M. Exploring the limits of classical physics: Planck, Einstein, and the structure of a scientific revolution. *Studies in History and Philosophy of Modern Physics*, v. 34, p. 37–59, 2003. Cited 2 times on pages 28 and 41.

CLAUSIUS, R. Ueber die Concentration von Wärme- und Lichtstrahlen und die gränzen ihrer Wirkung. *Annalen der Physik und Chemie. Hrsg. von J.C. Poggendorff*, v. 121, p. 1–44, 1864. Cited on page 70.

D'ABRO, A. *The Rise of the New Physics Volume 2: Quantum Physics*. New York: Dover, 1952. Cited on page 25.

DARRIGOL, O. Statistics and Combinations in Early Quantum Theory. *Historical Studies in the Physical and Biological Sciences*, p. 17–80, 1988. Cited on page 62.

DARRIGOL, O. Classical concepts in Bohr's atomic theory (1913–1925). *Physis: Riv. Internaz. di Storia della Scienza*, n. 34, p. 545–567, 1997. Cited on page 92.

DARRIGOL, O. On the Historians' Disagreement over the Meaning of Planck's Quanta. *Centaurus*, p. 219–239, 2001. Cited 4 times on pages 26, 33, 35, and 36.

DIRAC, P. A. M. The Evolution of the Physicist's Picture of Nature. *Scientific American*, v. 208, p. 45–53, 1963. Cited on page 26.

DOPPLER, C. *Ueber das farbige Licht der Doppelsterne und einiger anderer Gestirne des Himmels: Versuch einer das Bradley'sche Aberrations-Theorem als integrirenden Theil in sich schliessenden allgemeineren Theorie*. [S.l.]: In Commission bei Borrosch & André, 1842. Cited on page 75.

DOUGAL, R. C. The centenary of the fourth-power law. *Physics Education*, v. 14, n. 4, p. 234, 1979. Cited on page 71.

DRAPER, J. *Scientific Memoirs: Being Experimental Contributions to a Knowledge of Radiant Energy*. London: Sampson Low, Marston, Searle & Rivington, 1878. Cited on page 71.

DRUDE, P. *Lehrbuch der Optik*. Leipzig: Hirzel, 1900. Page numbers and quotations taken from the English translation by C. R. Mann and Robert Millikan as: *The Theory of Optics*. (1959) New York: Dover. Cited on page 42.

DUCK, I.; SUDARSHAN, E. C. G. *100 Years of Planck's Quantum*. New Jersey: World Scientific Publishing Company, 2000. Cited on page 25.

DULONG, P.; PETIT, A. T. Des Recherches sur la Mesure des Températures et sur les Lois de la communication de la chaleur. *Annales de Chimie et de Physique*, v. 7, p. 225–264, 1817. Cited on page 71.

EHRENFEST, P. Welche Züge der Lichtquantenhypothese spielen in der Theorie der Wärmestrahlung eine wesentliche Rolle? *Annalen der Physik*, v. 341, n. 11, p. 91–118, 1911. Cited 2 times on pages 38 and 50.

EHRENFEST, P. Adiabatic Invariants and the Theory of Quanta. *Phil. Mag.*, n. 33, p. 500–513, 1916. Cited 2 times on pages 81 and 91.

EHRENFEST, P. Adiabatische Transformationen in der Quantentheorie und ihre Behandlung durch Niels Bohr. *Naturwissenschaften*, v. 11, p. 543–550, 1923. Cited on page 78.

EINSTEIN, A. Kinetic Theory of Kinetic Equilibrium and of the Second Law of Thermodynamics. *Annalen der Physik*, v. 9, p. 417–433, 1902. Cited on page 81.

EINSTEIN, A. A Theory of the Foundations of Thermodynamics. *Annalen der Physik*, v. 11, p. 107–187, 1903. Cited 3 times on pages 81, 83, and 86.

EINSTEIN, A. On the General Molecular Theory of Heat. *Annalen der Physik*, v. 14, p. 354–362, 1904. Page numbers and quotations taken from the English translation by Anna Beck at *The Collected Papers by Albert Einstein, Vol. 2, The Swiss Years: Writings, 1900-1909*. <https://einsteinpapers.press.princeton.edu/vol2-trans/>. Cited 5 times on pages 81, 82, 83, 84, and 86.

EINSTEIN, A. On a Heuristic Point of View Concerning the Production and Transformation of Light. *Annalen der Physik*, p. 132–148, 1905. Page numbers and quotations taken from the English translation by Anna Beck at *The Collected Papers by Albert Einstein, Vol. 2, The Swiss Years: Writings, 1900-1909*. <https://einsteinpapers.press.princeton.edu/vol2-trans/>. Cited 9 times on pages 38, 44, 56, 59, 60, 65, 80, 85, and 87.

EINSTEIN, A. On the Electrodynamics of Moving Bodies. *Annalen der Physik*, p. 891–921, 1905. Cited on page 65.

EINSTEIN, A. On the Movement of Small Particles Suspended in Stationary Liquids Required by the Molecular-Kinetic Theory of Heat. *Annalen der Physik*, v. 17, p. 549–560, 1905. Page numbers and quotations taken from the English translation by Anna Beck at *The Collected Papers by Albert Einstein, Vol. 2, The Swiss Years: Writings, 1900-1909*. <https://einsteinpapers.press.princeton.edu/vol2-trans/>. Cited 2 times on pages 44 and 83.

EINSTEIN, A. On the Theory of Light Production and Light Absorption. *Annalen der Physik*, v. 20, p. 199–206, 1906. Page numbers and quotations taken from the English translation by Anna Beck at *The Collected Papers by Albert Einstein, Vol. 2, The Swiss Years: Writings, 1900-1909*. <https://einsteinpapers.press.princeton.edu/vol2-trans/>. Cited 3 times on pages 44, 60, and 61.

EINSTEIN, A. Planck's Theory of Radiation and the Theory of Specific Heat. *Annalen der Physik*, v. 22, p. 180–190, 1907. Cited on page 44.

EINSTEIN, A. On the Development of Our Views Concerning the Nature and Constitution of Radiation. *Physikalische Zeitsschrift*, v. 10, p. 817–826, 1909. Cited on page 44.

EINSTEIN, A. On the Present Status of the Radiation Problem. *Physikalische Zeitsschrift*, v. 10, p. 185–193, 1909. Cited 4 times on pages 32, 43, 44, and 58.

EINSTEIN, A. On the Theory of Light Quanta and the Question of Localization of Energy. *Archives des sciences physiques et naturelles*, v. 29, p. 525–528, 1910. Cited on page 79.

EINSTEIN, A. Zur Quantentheorie der Strahlung. *Z. Phys.*, n. 18, p. 121–128, 1917. Cited on page 91.

EINSTEIN, A. Space, Time, and Gravitation. *Times* (*London*), p. 13–14, 1919. Cited 2 times on pages 65 and 66.

EINSTEIN, A. Autobiographical Notes. In: PAUL ARTHUR SCHILPP. *Albert Einstein: Philosopher-Scientist*. [S.l.]: Library of Living Philosophers, MJF Books, 1949. p. 1–96. Cited 6 times on pages 37, 60, 61, 66, 80, and 122.

EMCH, G. G. *Mathematical and Conceptual Foundations of 20th-century Physics*. Amsterdam: North-Holland Publishing Company, 1987. Cited 2 times on pages 26 and 111.

ERICSSON, J. The temperature of the surface of the sun. *Nature*, p. 505–507, 1872. Cited on page 71.

FRAASSEN, B. van. *Quantum Mechanics: An Empiricist View*. New York: Clarendon Press, 1991. Cited on page 25.

GALISON, P. Kuhn and the Quantum Controversy. *British Journal for the Philosophy of Science*, v. 32, p. 71–84, 1981. Cited 4 times on pages 29, 33, 37, and 54.

GAMOW, G. *Thirty Years that Shook Physics: The Story of Quantum Theory*. New York: Dover, 1985. Cited on page 25.

GEARGART, C. A. Planck, the Quantum, and the Historians. *Phys. Perspect.*, v. 4, p. 170–215, 2002. Cited 2 times on pages 26 and 37.

GRIFFITHS, D. *Introduction to Elementary Particles*. New Jersey: Wiley, 2004. Cited on page 25.

HAAR, D. *The Old Quantum Theory*. Great Britain: Pergamon Press, 1967. Cited on page 25.

HAAR, D.; WERGELAND, H. *Elements of Thermodynamics*. Boston: Addison-Wesley Publishing Company, 1966. (A-W series in advanced physics). Cited on page 76.

HEISENBERG, W. über quantentheoretische Umdeutung kinematischer und mechanischer Beziehungen. *Z. Phys.*, n. 33, p. 694–705, 1925. Page numbers and quotations taken from the English translation by B. L. van der Waerden as: "Quantum-theoretical Reinterpretation of Kinematic and Mechanical Relations." In: van der Waerden, B. L. (1967), *Sources of Quantum Mechanics*. New York: Dover. Cited 4 times on pages 90, 91, 93, and 94.

HEISENBERG, W. The physical content of quantum kinematics and mechanics. In: J. A. WHEELER AND W. H. ZUREK. *Quantum Theory and Measurement*. Princeton: Princeton University Press, 1927. (Princeton series in physics), p. 62–83. Cited 2 times on pages 96 and 124.

HEISENBERG, W. *Physics and Philosophy: The Revolution in Modern Science*. New York: Harper, 1958. Cited on page 26.

HEISENBERG, W.; BORN, M.; JORDAN, P. Zur Quantenmechanik II. *Z. Phys.*, n. 35, p. 557–615, 1926. Page numbers and quotations taken from the English translation by B. L. van der Waerden as: "On Quantum Mechanics." In: van der Waerden, B. L. (1967), *Sources of Quantum Mechanics*. New York: Dover. Cited on page 109.

HERMANN, A. *The Genesis of quantum theory (1899-1913)*. Massachusetts: MIT Press, 1971. Cited 3 times on pages 25, 41, and 52.

HOWARD, D. The Copenhagen Interpretation. In: OLIVAL FREIRE JR. *The Oxford Handbook of the History of Quantum Interpretations*. Oxford: Oxford University Press, 2022. p. 521–542. Cited on page 126.

HOWARD, D. "And I Shall not Mingle Conjectures with Certainties": On the Interpretation of and Intellectual Background to Einstein's Distinction between Principle Theories and Constructive Theories. Unpublished Manuscript. 2023. Cited on page 67.

JAMMER, M. *The Conceptual Development of Quantum Mechanics*. New York: McGraw Hill, 1966. Cited 4 times on pages 25, 36, 92, and 96.

JEANS, J. *Report on Radiation and the Quantum-theory*. London: "The Electrician" Printing & Publishing Company, Limited, 1914. Cited 2 times on pages 119 and 121.

JOAS, C.; LEHNER, C. The classical roots of wave mechanics: Schrödinger's transformations of the optical-mechanical analogy. *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics*, v. 40, n. 4, p. 338–351, 2009. Cited 2 times on pages 99 and 100.

JOST, R. Planck-Kritik des T. Kuhn. *Das Märchen vom Elfenbeinernen Turm. Reden und Aufsätze, Berlin*, p. 67–78, 1995. Cited on page 26.

KANGRO, H. *Early History of Planck's Radiation Law*. UK: Taylor & Francis, 1976. Cited on page 25.

KIRCHHOFF, G. Ueber den Zusammenhang von Emission und Absorption von Licht und Wärme. *Monatsberichte der Akademie der Wissenschaften zu Berlin*, p. 783–787, 1859. Cited on page 70.

KIRCHHOFF, G. Ueber die Fraunhoferschen Linien. *Monatsberichte, Akademie 1615 der Wissenschaften*, p. 662–665, 1860. Cited on page 30.

KLEIN, M. *Paul Ehrenfest: The making of a theoretical physicist*. Amsterdam: North-Holland, 1985. Cited 3 times on pages 25, 68, and 75.

KLEIN, M. J. Max Planck and The Beginnings of Quantum Theory. *Arch. Hist. Exact Sci.*, v. 1, p. 459–479, 1961. Cited 4 times on pages 25, 39, 41, and 53.

KLEIN, M. J. Einstein's First Paper on Light Quanta. *The Natural Philosopher*, v. 2, p. 57–86, 1963. Cited on page 25.

KLEIN, M. J.; SHIMONY, A.; PINCH, T. J. Paradigm Lost? a Review Symposium. *Isis*, v. 70, p. 429–440, 1979. Cited 3 times on pages 26, 33, and 37.

KRAGH, H. Erwin Schrödinger and the Wave Equation: The Crucial Phase. *Centaurus*, v. 26, n. 2, p. 154–197, 1982. Cited 2 times on pages 97 and 99.

KRAGH, H. Max Planck: The Reluctant Revolutionary. *The Physics World*, v. 13, p. 31–36, 2000. Cited on page 36.

KRAMERS, A. H. *The Atom and the Bohr Theory of Its Structure: An Elementary Presentation*. New York: Knopf, 1923. Cited on page 25.

KRAMERS, H. A. The law of dispersion and Bohr's theory of spectra. *Nature*, n. 133, p. 673–674, 1924. Cited on page 91.

KRAMERS, H. A. The quantum theory of dispersion. *Nature*, n. 114, p. 310–311, 1924. Cited on page 91.

KRAMERS, H. A.; HEISENBERG, W. Über die Streuung von Strahlen durch Atome. *Z. Phys.*, n. 31, p. 681–708, 1925. Cited on page 91.

KUHN, T. *The Structure of Scientific Revolutions*. Chicago: University of Chicago Press, 1962. Cited on page 28.

KUHN, T. Logic of discovery or psychology of research? In: LAKATOS, I. AND MUSGRAVE, A. *Criticism and the Growth of Knowledge*. [S.l.]: University Press, 1970. p. 1–24. Cited on page 29.

KUHN, T. *Black-Body Theory and the Quantum Discontinuity, 1894-1912*. Oxford: Clarendon Press, 1978. Cited 5 times on pages 26, 35, 40, 41, and 53.

KUHN, T. Revisiting Planck. *Historical Studies in the Physical Sciences*, v. 14, p. 231–252, 1984. Cited 3 times on pages 28, 41, and 58.

LAUE, M. von. Max Planck. *Physikalische Blätter*, v. 3, p. 249–252, 1947. Page numbers and quotations taken from the English translation by Frank Gaynor as "Memorial Address by Max von Laue"' in: *Scientific Autobiography and Other Papers by Max Planck*. (1950) Williams & Norgate. Cited on page 25.

LEBEDEW, P. Les forces de Maxwell-Bartoli dues a la pression de la lumière. In: POINCARÉ, L.A. AND GUILLAUME, C.É. AND SOCIÉTÉ FRANÇAISE DE PHYSIQUE. *Rapports présentés au Congrès international de physique: Optique ; Électricité ; Magnétisme*. Paris: Gauthier-Villars, 1900. p. 133–140. Cited on page 70.

LIBOFF, R. L. *Introductory Quantum Mechanics*. New York: Pearson, 2003. Cited on page 25.

LIBRETEXTS. *Blackbody Radiation*. 2022. <https://phys.libretexts.org/@go/page/4520>. Cited 2 times on pages 11 and 74.

MACKEY, G. W. *The Mathematical Foundations of Quantum Mechanics*. Chicago: Benjamin, 1963. Cited on page 26.

MACKINNON, E. Heisenberg, Models, and the Rise of Matrix Mechanics. *Historical Studies in the Physical Sciences*, v. 8, p. 137–188, 1977. Cited 2 times on pages 90 and 94.

MACKINNON, E. The rise and fall of the Schrödinger interpretation. In: PATRICK SUPPES. *Studies in the Foundations of Quantum Mechanics*. Cincinnati: Philosophy of Science Association, 1980. p. 1–58. Cited 2 times on pages 97 and 100.

MAUDLIN, T. Three Measurement Problems. *Topoi*, v. 14, n. 1, p. 7–15, 1995. Cited 4 times on pages 90, 108, 113, and 125.

MAXWELL, J. *A Treatise on Electricity and Magnetism*. Oxford: Clarendon Press, 1873. Cited 2 times on pages 70 and 75.

MAXWELL, J. Tait's Thermodynamics. *Nature*, v. 17, p. 257–259, 1878. Cited on page 67.

MEHRA, J.; RECHENBERG, H. *The historical development of quantum theory, part I: The quantum theory of Planck, Einstein, Bohr, and Sommerfeld: Its foundations and the rise of its difficulties 1900-1925, Vol I*. New York: Springer, 1982. Cited on page 25.

MEHRA, J.; RECHENBERG, H. *Erwin Schrödinger and the Rise of Wave Mechanics*. [S.l.]: Springer-Verlag, 1987. Cited on page 96.

MOORE, W. *Schrödinger: Life and Thought*. Cambridge: Cambridge University Press, 1989. Cited on page 26.

MORETTI, V. *Spectral Theory and Quantum Mechanics: Mathematical Foundations of Quantum Theories, Symmetries, and Introduction to the Algebraic Formulation*. New York: Springer, 2017. Cited on page 26.

MULLER, F. A. The equivalence myth of quantum mechanics –part I. *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics*, Elsevier, v. 28, n. 1, p. 35–61, 1997. Cited 3 times on pages 96, 97, and 101.

MULLER, F. A. The equivalence myth of quantum mechanics—part II. *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics*, v. 28, n. 2, p. 219–247, 1997. Cited 2 times on pages 96 and 101.

MULLER, F. A. *Six Measurement Problems of Quantum Mechanics*. 2023. ArXiv: 2305.10206. Cited on page 118.

NAUENBERG, M. Max Planck and the birth of the quantum hypothesis. *American Journal of Physics*, v. 84, p. 709–720, 2016. Cited on page 25.

NEEDEL, A. Irreversibility and the Failure of Classical Dynamics: Max Planck's Work on the Quantum Theory, 1900–1915. Npublished PhD dissertation, Yale University. 1980. Cited 2 times on pages 26 and 37.

NEUMANN, J. von. *Mathematische Grundlagen der Quantenmechanik*. Heidelberg: Julius Springer, 1932. Page numbers and quotations taken from the third edition of the English translation by Robert T. Beyer as: *Mathematical Foundations of Quantum Mechanics.* (2018) Princeton: Princeton University Press. Cited 12 times on pages 26, 92, 108, 110, 111, 112, 113, 114, 115, 121, 124, and 125.

NEWTON, I. A Letter of Mr. Isaac Newton, Mathematick Professor in the University of Cambridge; Containing his New Theory about Light and Colors. *Philosophical Transactions of the Royal Society*, v. 80, p. 3075–3087, 1672. Cited on page 85.

NICHOLS, E. F.; HULL, G. F. The Pressure Due to Radiation. *Proceedings of the American Academy of Arts and Sciences*, v. 38, n. 20, p. 559–599, 1903. Cited on page 70.

OMNêS, R. *The Interpretation of Quantum Mechanics*. Princeton: Princeton University Press, 1994. Cited on page 25.

PAIS, A. Max Born's statistical interpretation of quantum mechanics. *Science*, v. 218, n. 4578, p. 1193–1198, 1982. Cited on page 109.

PAIS, A. *Subtle is the Lord : The Science and the Life of Albert Einstein: The Science and the Life of Albert Einstein*. Oxford: Oxford University Press, 1982. Cited 4 times on pages 26, 34, 47, and 82.

PAIS, A. *Niels Bohr Times. In Physics, Philosophy, and Politics*. Oxford: Oxford University Press, 1993. Cited on page 26.

PEIERLS, R. E. The development of quantum theory. *Contemporary Physics*, v. 6, p. 192–205, 1965. Cited on page 25.

PERES, A. *Quantum Theory: Concepts and Methods*. Philadelphia: Kluwer, 1993. Cited on page 25.

PEROVIC, S. Why were Matrix Mechanics and Wave Mechanics considered equivalent? *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics*, v. 39, n. 2, p. 444–461, 2008. Cited on page 96.

PHILLIPS, A. C. *Introduction to Quantum Mechanics*. New Jersey: John Wiley & Sons, 2003. Cited on page 25.

PLANCK, M. Über irreversible Strahlungvsvorgänge. *Annalen der Physik*, v. 5, p. 440–480, 1899. Cited 2 times on pages 42 and 58.

PLANCK, M. On an Improvement of Wien's Equation for the Spectrum. In: HAAR, D. *The Old Quantum Theory*. Great Britain: Pergamon Press, 1900. p. 79–81. Cited 2 times on pages 54 and 55.

PLANCK, M. On the Theory of the Energy Distribution Law of the Normal Spectrum. In: HAAR, D. *The Old Quantum Theory*. Great Britain: Pergamon Press, 1900. p. 82–90. Cited 8 times on pages 33, 50, 51, 52, 54, 55, 56, and 68.

PLANCK, M. On the Law of the Energy Distribution in the Normal Spectrum. *Annalen der Physik*, v. 4, p. 553–562, 1901. Cited 2 times on pages 42 and 50.

PLANCK, M. *The Theory of Heat Radiation*. Philadelphia: Blakiston, 1914. Cited 3 times on pages 39, 50, and 60.

PLANCK, M. *Where is Science Going?* Connecticut: Ox Bow Press, 1933. Cited on page 48.

PLANCK, M. *Scientific Autobiography and Other Essays*. Edinburgh: Williams Norgate, 1949. Cited 3 times on pages 47, 48, and 49.

POINCARÉ, H. Sur la théorie des quanta. *J. Phys. Theor. Appl.*, v. 2, n. 1, p. 5–34, 1912. Cited on page 93.

PROVOSTAYE, F. de la; DESAINS, P. Mémoire sur le rayonnement de la chaleur. *Annales de Chimie et de Physique*, v. 16, p. 337–425, 1846. Cited on page 71.

PRZIBRAM, K. et al. *Letters on Wave Mechanics: Schrödinger, Planck, Einstein, Lorentz*. New York: Philosophical Library, 1967. Cited on page 121.

RANKINE, W. J. M. *A manual of applied mechanics*. London and Glasgow: Pergamon Press, 1858. Cited on page 75.

RAYLEIGH, L. The Law of Partition of Kinetic Energy. *Philosophical Magazine*, v. 49, p. 98–118, 1900. Cited on page 61.

RAYLEIGH, L. On the Pressure of Vibrations. *Philosophical Maganize*, v. 6, p. 338–346, 1902. Cited on page 81.

REICHENBACH, H. *Philosophical Foundations of Quantum Mechanics*. New York: Dover, 1944. Cited on page 25.

ROSENFELD, L. La premiere phase de revolution de la theorie des quanta. *Osiris*, v. 2, p. 149–196, 1936. Cited on page 25.

SCHRÖDINGER, E. On the Relation Between the Quantum Mechanics of Heisenberg, Born, and Jordan, and that of Schrödinger. In: ERWIN SCHRÖDINGER. *Collected Papers on Wave Mechanics*. London and Glasgow: Blackie & Son limited, 1926. p. 45–61. Cited 4 times on pages 100, 101, 102, and 110.

SCHRÖDINGER, E. Quantisation as a Problem of Proper Values, Part I. In: ERWIN SCHRÖDINGER. *Collected Papers on Wave Mechanics*. London and Glasgow: Blackie & Son limited, 1926. p. 1–12. Cited 5 times on pages 97, 99, 100, 103, and 106.

SCHRÖDINGER, E. Quantisation as a Problem of Proper Values, Part II. In: ERWIN SCHRÖDINGER. *Collected Papers on Wave Mechanics*. London and Glasgow: Blackie & Son limited, 1926. p. 13–40. Cited 3 times on pages 100, 103, and 104.

SCHRÖDINGER, E. Quantisation as a Problem of Proper Values, Part III. In: ERWIN SCHRÖDINGER. *Collected Papers on Wave Mechanics*. London and Glasgow: Blackie & Son limited, 1926. p. 62–101. Cited 2 times on pages 102 and 109.

SCHRÖDINGER, E. Quantisation as a Problem of Proper Values, Part IV. In: ERWIN SCHRÖDINGER. *Collected Papers on Wave Mechanics*. London and Glasgow: Blackie & Son limited, 1926. p. 102–123. Cited on page 117.

SCHRÖDINGER, E. The Exchange of Energy according to Wave Mechanics. In: ERWIN SCHRÖDINGER. *Collected Papers on Wave Mechanics*. London and Glasgow: Blackie & Son limited, 1927. p. 137–146. Cited on page 120.

SCHRÖDINGER, E. Are there quantum jumps? part I. *he British Journal for the Philosophy of Science*, v. 3, p. 109–123, 1952. Cited on page 26.

SERWAY, R.; MOSES, C.; MOYER, C. *Modern Physics*. New York: Belmont Books, 2005. Cited on page 25.

SHIMONY, A. Role of the Observer in Quantum Theory. *American Journal of Physics*, v. 31, n. 10, p. 755–773, 1963. Cited on page 115.

SKLAR, L. *Philosophy of Physics*. Oxford: Oxford University Press, 1995. Cited on page 25.

SMITH, G.; SETH, R. *Brownian Motion and Molecular Reality*. Oxford: Oxford University Press, 2020. (Oxford studies in philosophy of science). Cited 4 times on pages 44, 48, 56, and 83.

SOLVAY, E.; LANGEVIN, P.; BROGLIE, M. de. *La théorie du rayonnement et les quanta. Rapports et discussions de la réunion tenue à Bruxelles, du 30 octobre au 3 novembre 1911*. Paris: Gauthier-Villars, 1912. Cited on page 81.

SOMMERFELD, A. To Albert Einstein's seventieth birthday. In: PAUL ARTHUR SCHILPP. *Albert Einstein: Philosopher-Scientist*. New York: Library of Living Philosophers, MJF Books, 1949. p. 97–106. Cited on page 122.

SOMMERFELD, A.; BOPP, F. Fifty Years of Quantum Theory. *Science*, v. 113, p. 85–92, 1951. Cited on page 26.

STEFAN, J. über die Beziehung zwischen der Wärmestrahlung und der Temperatur. *Sitzungsber. Kaiserl. Akad. Wiss. Math. Naturwiss. Cl. II. Abth.*, v. 79, n. 3, p. 391–428, 1879. Cited on page 71.

TOMONAGA, S. *Quantum Mechanics, Vol. I*. Amsterdam: North Holland, 1962. Cited on page 25.

UFFINK, J. Insuperable difficulties: Einstein's statistical road to molecular physics. *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics*, v. 37, n. 1, p. 36–70, 2006. Cited on page 82.

WAERDEN, B. *Sources of Quantum Mechanics*. New York, NY: Dover Books, 1967. Cited on page 25.

WEINBERG, S. The Search for Unity: Notes for a History of Quantum Field Theory. *Dedalus*, v. 106, p. 17–35, 1977. Cited on page 25.

WESSELS, L. Schrödinger's route to wave mechanics. *Studies in History and Philosophy of Science Part A*, v. 10, n. 4, p. 311–340, 1979. Cited 2 times on pages 97 and 99.

WESSELS, L. The intellectual sources of Schrödinger's interpretations. In: PATRICK SUPPES. *Studies in the Foundations of Quantum Mechanics*. Cincinnati: Philosophy of Science Association, 1980. p. 59–76. Cited on page 97.

WHITTAKER, E. T. *A History of the Theories of Aether and Electricity: The Modern Theories, 1900-1926*. New York: Philosophical Library, 1954. Cited on page 25.

WIEN, W. Eine neue Beziehung der Strahlung schwarzer Körper zum zweiten Hauptsatz der Wärmetheorie. *Sitzungsber. Berl. Akad.*, v. 6, p. 55–62, 1893. Cited 3 times on pages 50, 68, and 73.

WIEN, W. Les lois théoriques du rayonnement. In: POINCARÉ, L.A. AND GUILLAUME, C.É. AND SOCIÉTÉ FRANÇAISE DE PHYSIQUE. *Rapports présentés au Congrès international de physique: Optique ; Électricité ; Magnétisme*. Paris: Gauthier-Villars, 1900. p. 23–40. Cited on page 72.

WIEN, W. Ueber die Energievertheilung im Emissionsspectrum eines schwarzen Körpers. *Annalen der Physik*, v. 58, p. 662–669, 1986. Page numbers and quotations taken from the English translation by Mr. J. Burke as: "On the division of energy in the emission-spectrum of a black body," *Philosophical Magazine Series 5,* 43(262), pp. 214-220, 1987. Cited on page 42.

WIEN, W.; LUMMER, O. Methode zur Prüfung des Strahlungsgesetzes absolut schwarzer Körper. *Annalen der Physik*, v. 292, n. 11, p. 451–456, 1895. Cited 2 times on pages 70 and 74.