Unification and the Quantum Hypothesis in 1900–1913

Abstract

In this paper, I consider some of the first appearances of a hypothesis of quantized energy between the years 1900 and 1913 and provide an analysis of the nature of the unificatory power of this hypothesis in a Bayesian framework. I argue that the best way to understand the unification here is in terms of informational relevance: on the assumption of the quantum hypothesis, phenomena that were previously thought to be unrelated turned out to yield information about one another based on agreeing measurements of the numerical value of Planck’s constant.

Word Count: 4,977

1 Introduction

The idea that unification is a virtue of a scientific theory has a long history in philosophy of science, and has been presented in several guises. Accounts range over those focused on the common causal origins of various phenomena to those emphasizing a common explanatory basis. Of course, these are not mutually exclusive ideas and a combination of these elements
is common. Examples include William Whewell on the Consilience of Inductions (1989), Michel Janssen on Common Origin Inferences (2002), Philip Kitcher on explanatory unification (1989) and William Wimsatt on robustness (1981). While the details of such discussions differ, it is clear that some version of this notion has played a role in several important episodes of scientific theorizing. One period in the history of science that has perhaps been neglected in this context is the case of the years leading up to the development of the old quantum theory. While the history is well-documented (e.g. Klein (1961, 1965, 1966), ter Haar (1967), Hund (1974), Kuhn (1978), Mehra & Rechenberg (1982)) and there have been excellent discussions of the justification of particular aspects of the theory (Norton (1987, 1993)), an explicit discussion in terms of unification has not yet been provided. First, consider what I refer to as the quantum hypothesis, $QH$: this is the idea that radiation energy cannot always be treated in a continuous manner as in classical physics, but that instead, radiation of frequency $\nu$ is emitted and absorbed in packets of size $h\nu$, where $h$ refers to the universal physical constant referred to as Planck’s constant. I will consider some of the central applications of $QH$ between 1900 and 1913 and explain its unificatory role. I will argue that the best way to understand the unification present in this period is in terms of informational relevance: on the assumption of $QH$, observations performed on various diverse physical phenomena can be thought of as measuring or constraining the numerical value of $h$, and these agreeing measurements of $h$ render the physical phenomena relevant to one another by providing information about the value one is likely to obtain in the various cases. Such a feature is particularly important in early stages of scientific theory development before alternative methods of justification are available. Despite the absence of a coherent theory that incorporated $QH$, and despite its inconsistency with well-established physics, it was taken by several scientists to be a promising starting point for the development of a more adequate
theory. In this paper I will argue that they had reasonable grounds for doing so, because of empirical support for $QH$ in the form of its unificatory power.

I will first give an outline of unification in terms of informational relevance and comment briefly on how I take this concept to differ from a common cause argument. I will then give an overview of several scientific episodes invoking Planck’s constant and in doing so, I will argue for my claim that the type of unification displayed here is best understood in terms of experiments on phenomena providing information about one another by yielding agreeing measurements of the parameter $h$.

2 Unification

As mentioned previously, there are several different ways to conceive of unification as well as differences in what is being attributed unificatory power. I defend the idea that the unification should be understood in terms of informational relevance. My claim is not that this necessarily captures scientists’ actual motivations, but that we can retrospectively identify the fact that the quantum hypothesis had unificatory power in the way to be explained. Here, I have adopted the explication of unification given by Myrvold (2003).

Myrvold provides a Bayesian account of the feature of unification. He shows that on a particular understanding of what it means for a hypothesis to unify phenomena, its ability to do so contributes directly to its support by the evidence it unifies. Thus, if one accepts a Bayesian confirmational framework, the unifying hypothesis obtains support from the unifying phenomena.

More specifically, he takes a common definition of the informational relevance of a proposition $p_1$ to another proposition $p_2$, conditional on background $b$,
\[ I(p_2, p_1|b) = \log_2 \frac{Pr(p_2|p_1\&b)}{Pr(p_2|b)}. \]  

He then defines the quantity \( U \) as a measure of the extent to which \( h \) unifies \( p_1 \) and \( p_2 \),

\[ U(p_1, p_2; h|b) = I(p_1, p_2|h\&b) - I(p_1, p_2|b). \]  

This generalizes straightforwardly to a set of hypotheses \( p_1 \ldots p_n \). He then shows that on two common candidates for the degree to which evidence supports a hypothesis, the quantity \( U \) contributes directly to the support of \( h \) by the evidence. I use the “degree of confirmation” in my discussion, but a similar result holds if one takes Good’s measure of the “weight of evidence.” Notice that the degree of confirmation, measured by \( \log \frac{Pr(h|e\&b)}{Pr(h|b)} \), is identical to the definition of the informational relevance of \( e \) to \( h \), so we can consider the informational relevance as a measure of evidential support. We thus obtain for the informational relevance of evidence \( e_1 \) and \( e_2 \) to hypothesis \( h \),

\[ I(h, e_1\&e_2|b) = I(h, e_1|b) + I(h, e_2|b) + U(e_1, e_2; h|b). \]  

Myrvold explains the significance as follows:

[T]he degree of support provided to \( h \) by \( e_1 \) and \( e_2 \) taken together is the sum of three terms: the degree of support of \( h \) by \( e_1 \) alone, the degree of support \( h \) by \( e_2 \) alone, and an additional term which is simply the degree of unification of the set \( \{e_1, e_2\} \) by \( h \). An analogous result holds for larger bodies of evidence. (2003, 412)

Thus, the ability of a hypothesis to unify previously unrelated phenomena contributes directly
to the likelihood of that hypothesis given the evidence. In what follows, I will provide details of how the case under consideration provides an example of this feature. Briefly, the physical phenomena to be discussed were not clearly relevant to one another before the postulation of \( QH \). However, on the assumption of such a hypothesis, numerical values of quantities obtained from observations of those phenomena could be used to calculate the numerical value of \( h \), all of which agreed to within an order of magnitude. The measured value of \( h \) via one type of phenomenon thus provided information about the measured value of \( h \) via a different phenomenon when assuming \( QH \).

The explication of unification in terms of informational relevance certainly does not preclude the existence of a common cause argument, such as the one given by Wesley Salmon in his reconstruction of Jean Perrin’s determination of Avogadro’s constant (1984). However, I would argue that characterizing the unificatory power of \( QH \) as being due to a common cause would be to overstate the strength of the information available, since at the time in question, it was not at all clear how quantization might be occurring, and no account of the underlying mechanisms was forthcoming. Despite this, it was clear that \( QH \) did possess unificatory power in the informational relevance sense, and this minimal sense is all that is required to provide the hypothesis with some confirmational force.

3 Uses of the Quantum Hypothesis

3.1 Blackbody Radiation

Planck’s constant, \( h \), was first introduced by Planck in his work on blackbody radiation (1900). He began by interpolating an expression for the equilibrium entropy based on existing laws which were only partially empirically adequate; with this formula, he was able to
determine a radiation formula that correctly described the entire emission spectrum of blackbody radiation, \( E = \frac{c_1 \lambda^{-5}}{e^{c_2/\lambda T} - 1} \). The interpretation of this formula led him to posit the quantum hypothesis, that energy of frequency \( \nu \) is absorbed and emitted in packets of size \( h \nu \).

The formula for energy distribution could then be written as

\[
E = \frac{8\pi ch}{\lambda^5} \frac{1}{e^{ch/k\lambda T} - 1}.
\] (4)

Planck was then able to use this empirically confirmed radiation formula to estimate the size of \( h \). He used his formula to calculate the amount of radiation in air, and compared this with values obtained by Ferdinand Kurlbaum in experimental work (1898). He then drew on observations made by Otto Lummer and Ernst Pringsheim, who were able to determine the wavelength of the maximum energy in air of blackbody radiation. The result was a numerical value for the parameter \( h \), \( h = 6.55 \cdot 10^{-27} \text{erg} \cdot \text{sec} \).

Near the beginning of his work on blackbody radiation, Planck was focused on providing observationally motivated descriptions of phenomena using a general idea of ‘resonators’ while hoping that electron theory would later be able to fill in the gaps, so to speak, on how absorption and emission of discrete energy amounts was taking place. As Gearhart has pointed out, Planck repeatedly stressed the need for a physical interpretation of the constants he introduced (2002, 200). In fact, there has been much debate on how Planck actually understood the various derivations he gives of the quantum hypothesis. For instance, Kuhn (1978) among others has argued that Planck was not literally considering quantized energy elements in his 1900, 1901 papers, but was thinking in terms of continuous energy amounts and using the mathematical apparatus of quantized energy as a calculational convenience. This is in opposition to historians such as Klein (1961), who have argued for a more robust understanding of Planck’s “energy quanta.” Gearhart has provided an overview of the history
and the various interpretive positions, and argues that it is difficult to maintain the view that
Planck himself had in mind the quantization of something like phase space as early as 1900
and 1901 (2002). It is worth noting that Planck’s own understanding of what exactly is being
quantized is not crucial to the point being made here. Regardless of how and why absorption
and emission may occur, we can still see how \( QH \) had unificatory power by examining its
application in various phenomena.

3.2 Light Quanta

Although Einstein was aware of Planck’s work on blackbody radiation, his own work in
radiation theory stemmed from a slightly different motivation and he was in fact reluctant to
fully accept Planck’s conclusions, as Einstein believed they diverged further from classical
theory than Planck himself was aware. His work led him to conclude that “monochromatic
radiation of low density . . . behaves, in a thermodynamic sense, as it if consisted of mutually
independent radiation quanta of magnitude \( R h \nu/k N_0 \)” ((Einstein, 1905, 143), translation
from (ter Haar, 1967, 102)), where I have here replaced Einstein’s constant \( \beta \) with the
equivalent \( h/k \) for ease of reference.

This paper is perhaps best known for Einstein’s treatment of the photoelectric effect in
producing “cathode rays,” or beams of electrons. One instance was the emission of such rays
from a metallic surface after the absorption of incident ultraviolet light. This was first
observed in 1887 and studied further in subsequent years, particularly by Philipp Lenard.
Einstein hypothesized that light quanta penetrating the surface layer of bodies has energy that
is transformed into electron kinetic energy within the substance; electrons then escape the
surface with a certain kinetic energy having produced some quantity of work. We can
consider the equation Einstein describes in terms of our discussion of informational relevance.
The experiments done on the photoelectric effect yield information about the size of $h$. One can derive a relation between the energy of electrons and the size of $h$ based on the kinetic energy of the electrons being emitted. Einstein reasoned that $\Pi E = Rh\nu/k - P'$, where the body under investigation is charged to positive potential $\Pi$, $E$ is the charge of a gram equivalent of an ion, and $P'$ is the potential of negative electricity. Experiments on the photoelectric effect provided observed values for the unknowns in the relation $\Pi E = Rh\nu/k - P'$. Known quantities could then be inserted into this formula: $R$ is a known constant, $E = 9.6 \cdot 10^3$, $Pt = 0$, $\nu = 1.03 \cdot 10^{15}$. ($\nu$ corresponds to frequencies of ultraviolet light, and the other values are given for an experimental setup.) The order of magnitude of $\Pi$ according to Lenard’s results = $10^7$. Einstein calculated the theoretical value of $\Pi E$ according to his theoretical assumptions, and found that his theoretical value of $\Pi$ was in good accord with the experimental results of Lenard. This provided a constraint for the value of $h$ even though at the time it could only have been given within an order of magnitude. Because Einstein’s $\beta$ was equivalent to $h/k$ and the order of magnitude of $\beta$ had $10^{-11}$, the measured value of a body’s resistance in cases of the photoelectric effect constrained $h$ to be of order of magnitude $10^{-27}$.

Let us now explain how this fits into the Bayesian framework by determining how the various experiments provide information about $h$. First, note that by beginning with $QH$, one can calculate the average energy of the resonators Planck was considering in order to obtain the radiation formula Equation 4. However, this equation refers only to the form of a family of equations, where the value of $h$ is not yet determined. Thus, let $e_1$ be the proposition expressing the results of Lummer and Pringsheim’s work determining the maximum wavelength of blackbody radiation in air at a given temperature, “$\lambda mT = 0.294 cm \cdot K$.” Let $e_2$ be the proposition that an experiment on the photoelectric effect would yield a result such
that $\Pi$ is of the order of magnitude $10^7$. From $e_1$, in conjunction with $QH$ as applied in
deriving Equation 4 one obtains that the value of $h = 6.55 \cdot 10^{-27}$ erg $\cdot$ sec. Similarly, the
results of $e_2$ in conjunction with $QH$ yield the result that $h$ is of the order of magnitude $10^{-27}$.
Before the suggestion of $QH$, there was no way to use $e_1$ to yield information about $e_2$. Thus,
the informational relevance of $e_1$ to $e_2$ on background $b$, given by Equation 1 was very low.
After all, there was no way that Lummer & Pringsheim’s experiments on blackbody radiation
would constrain the behaviour of cathode rays, so $Pr(e_2|e_1&b)$ should be the same as
$Pr(e_2|b)$, thus assigning $I(e_2, e_1|b)$ the value 0. Compare this with the informational
relevance value on the assumption of $QH$ along with background $b$. This is given by the
expression

$$I(e_2, e_1|QH&b) = \log_2 \frac{Pr(e_2|e_1&QH&b)}{Pr(p_2|QH&b)}.$$ 

The value of $Pr(p_2|QH&b)$ is the
probability that Lenard’s results would obtain, which does not have a particularly high value if
considered against a general background. However, once we consider $e_1$ as well, we can
calculate a value for $h$ from the blackbody spectrum, thus constraining the value we would
obtain from experiments on the photoelectric effect. This yields a very high value for
$Pr(e_2|e_1&QH&b)$, arguably a value very close to one, thus making the value of the
information relevance of $e_1$ to $e_2$ quite high.
Now recall that the unificatory power of $QH$ is given by Equation 2 which measures the
difference between the relevance of $e_1$ to $e_2$ when including $QH$ in the background
knowledge, and excluding it. This nonzero value contributes directly to the degree of
confirmation of $QH$ by $e_1$ and $e_2$ as measured by Equation 3. Thus, by positing behaviour of
radiation in terms of quanta of size $h\nu$, the form of the blackbody radiation spectrum
constrained possible values of measurements conducted on the phenomenon of the
photoelectric effect by providing information about the size of $h$.

An interesting point here is that Einstein had different ideas in mind for the understanding of
quantization than Planck; Einstein talked in terms of quantization of light, whereas Planck is somewhat noncommittal. For instance, nine years after his introduction of $h$, he writes,

> [P]revious electron theories suffer from an essential incompleteness which demands a modification, but how deeply this modification should go into the structure of the theory is a question upon which views are still widely divergent. . . . [Some physicists, including Einstein] even believe that the propagation of electromagnetic waves in a pure vacuum does not occur precisely in accordance with the Maxwellian field equations, but in definite energy quanta $h\nu$. I am of the opinion, on the other hand, that at present it is not necessary to proceed in so revolutionary a manner, and that one may come successfully through by seeking the significance of the energy quantum $h\nu$ solely in the mutual actions with which the resonators influence one another. (1915[1909], 68)

For this reason, a ‘common cause’ account of the spectrum of blackbody radiation and the various light phenomena mentioned here would be difficult to provide. QH itself does not posit any mechanisms that can be understood as causes; quantization might stem from the actions of resonators, or the constitution of light, among other possibilities. Nevertheless, we can see that the phenomena discussed above became relevant to one another on the assumption of even something as general as QH, and different interpretations of its ‘cause’ do not affect its unificatory power.

### 3.3 Spectral Phenomena

The quantum hypothesis and the quantity $h$ were crucial in early characterizations of the structure of the atom, as well as the behaviour of line spectra, specifically when heated gases
produce lines of different colours. It was observed that the radiation emitted from these heated gases were not of a continuous spectrum as classical mechanics would lead one to expect. Rather, the emitted radiation was of a number of specific frequencies, as manifested in a number of discrete lines on the spectrum. Balmer found a formula describing the emission spectrum of hydrogen gas:

$$\lambda = B \left( \frac{m^2}{m^2 - n^2} \right)$$  \hspace{1cm} (5)

where \( n = 2 \), \( m \) is an integer \( \geq 2 \), \( B \) is a constant. Written in terms of frequency and explicit values for the constant, and generalized to allow for different integers for \( n \) and \( m \), this becomes

$$\nu = \frac{2\pi^2 m e^4}{h^3} \left( \frac{1}{\tau_2^2} - \frac{1}{\tau_1^2} \right).$$  \hspace{1cm} (6)

However, this formula had no known connection with the other phenomena discussed above. Niels Bohr was able to develop a model of the atom that was able to account for the observed line spectra of different elements, which no other theory had been able to do. On his model, there were set orbits for electrons each associated with set amounts of energy. An electron making the jump from one energy level to another would emit a discrete amount of energy, \( h\nu \). Thus, the spectral lines produced by a particular gas when heated corresponded to the differences between discrete energy levels of the electrons moving from one level to another. This explained the observed discrete spectrum. There were problems with this model since it postulated the existence of stationary states, which went against certain laws of classical electrodynamics, but importantly, the preliminary model was able to account for the observed spectrum by incorporating the quantum hypothesis.

Bohr calculated relations between several observable quantities based on Planck’s radiation
theory utilizing $h$; these calculations, with observed quantities, fit with the order of magnitude of $h$. We can reinterpret this as a way to turn the observed line spectra into information about the size of $h$: we already knew that Balmer’s formula could be used to describe emission spectra. According to Bohr’s calculations,

$$\frac{2\pi^2 m e^4}{h^3} = 3.1 \cdot 10^{15}$$

(7)

The observed value was $3.290 \cdot 10^{15}$.

We can reverse the calculation in order to see how such an experiment would have constrained the value of $h$. We use the same experimental values that Bohr used for the charge of the electron $e = 4.7 \cdot 10^{-10}$ and the ratio of the charge to mass $e/m = 5.31 \cdot 10^{17}$, as well as the observed value of $3.290 \cdot 10^{15}$ and solve for $h$ in the expression above. The result is $h = 6.38 \cdot 10^{-27}$, which we see is remarkably close to Bohr’s previously calculated value. In this way, we see how Balmer’s formula carried information about the size of $h$, which was also given by the blackbody spectrum.

In order to make the informational relevance explicit, let us take $e_1$, as above, to be the statement of Lummer & Pringsheim’s results on the maximum wavelength of blackbody radiation in air, $\lambda_m T = 0.294 cm \cdot K$. Let $e_2$ here be that the constant in Equation [7] in front of the brackets, takes on a value around $3.290 \cdot 10^{15}$. As before, a value of this constant without the assumption of $QH$ could a priori have taken on an infinite range of values, and the result of measurements on blackbody radiation would not be expected to be informative about this. Thus, the informational relevance of $e_1$ to $e_2$ was low, if not zero. However, by assuming $QH$, the blackbody spectrum provides information about the size of $h$, thus constraining the possible values that the constant could take. This makes it much more likely that the value of the constant should be the one found (on the reasonable assumption that values close to the
After Bohr’s success with the hydrogen spectrum, other phenomena related to atomic spectra were used as explicit tests for the value of $h$. James Franck and Gustav Ludwig Hertz performed experiments on the energy of electrons colliding with molecules of an inert gas or metal vapour (1914–1967). In particular, their experiments with mercury vapour were able to help determine value of $h$. Here, electrons of a certain kinetic energy were introduced into mercury vapour. It was known that at relatively high energies, the mercury gas became ionised. However, below this level but at certain energy thresholds, the electrons lost their kinetic energy; this was attributed to inelastic collisions between the free electrons and those bound to mercury atoms. The fact that these only occurred at discrete levels of energy of the introduced electrons was evidence for the idea that the mercury gas atoms could only absorb energy in those discrete quantities. These energy levels corresponded to the observed spectrum lines emitted by mercury gas.

Since the experiment involved only quantities that were pre-determined or measurable such as the energy of the introduced electrons, the voltage drop corresponding to the loss of the electrons’ kinetic energy, and the frequency of emitted energy in the spectrum, these results were used to calculate a value for Planck’s constant. Franck and Hertz calculated that $h$ had the value $6.59 \cdot 10^{-27}$. An analysis of the informational relevance of this experiment is analogous to the one given above.
3.4 Summary of Informational Relevance

I have presented several phenomena that were unified by the quantum hypothesis, namely, the frequency spectrum of blackbody radiation, light phenomena, atomic spectral phenomena, and the specific heat of diamond. One important feature that I have emphasized is the ability of several of these phenomena to help constrain and measure the numerical value of Planck’s constant which was an integral feature of the quantum hypothesis. Below is a table summarizing the values obtained from each of the phenomena discussed above.

<table>
<thead>
<tr>
<th>Phenomenon</th>
<th>Value of $h$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Blackbody radiation</td>
<td>$6.55 \cdot 10^{-27}$</td>
</tr>
<tr>
<td>Light quanta</td>
<td>Order of $10^{-27}$</td>
</tr>
<tr>
<td>Hydrogen emission spectrum</td>
<td>$6.38 \cdot 10^{-27}$</td>
</tr>
<tr>
<td>Mercury gas resonance radiation</td>
<td>$6.59 \cdot 10^{-27}$</td>
</tr>
</tbody>
</table>

These measurements are significant because they demonstrate the idea that various observations, understood in terms of constraining information about a parameter, were able to render previously unrelated phenomena relevant to one another by yielding information implicitly contained in those observations. By increasing the informational relevance of each phenomenon to the other, the unificatory power of $QH$ is raised. My previous discussion considered only pairwise informational relevance relations, but the generalization to several phenomena yields the following, taking each of the $e$’s below to represent the results of experiments from the four phenomena listed in the table.

$$U(e_1, e_2, e_3, e_4; QH|b) = I(e_1, e_2, e_3, e_4|QH&b) - I(e_1, e_2, e_3, e_4|b)$$  \hspace{1cm} (8)
Thus, the Bayesian notion of unificatory power of the quantum hypothesis is nonzero, and the degree of confirmation of $QH$ receives support not only from the individual phenomena, but from the fact that $QH$ makes those phenomena relevant to one another.

4 Conclusion

In this paper, I have argued that the type of unification displayed by the old quantum theory can be understood in terms of informational relevance, which yields the result that in a Bayesian confirmational framework, this unificatory power contributed to the confirmation of a quantum hypothesis over and above the evidence taken individually. I have argued that in many of these cases, an account of the mechanisms that would explain the observed behaviour were not available, which makes a causal story for the unification more difficult to provide. While not denying that causal explanations have their place in theoretical justification, I hope to have shown that there is at least one case where even when such unification is not available to us, there is an alternative sense that has epistemic force. Thus, despite the lack of a fully acceptable quantum theory, it was epistemically justified for scientists of the time to pursue the quantum hypothesis.
References


