Whose Devil? Which Details?*

Gordon Belot†‡

Batterman has recently argued that fundamental theories are typically explanatorily inadequate, in that there exist physical phenomena whose explanation requires that the conceptual apparatus of a fundamental theory be supplemented by that of a less fundamental theory. This paper is an extended critical commentary on that argument: situating its importance, describing its structure, and developing a line of objection to it. The objection is that in the examples Batterman considers, the mathematics of the less fundamental theory is definable in terms of the mathematics of the fundamental theory, and that only the latter need be given a physical interpretation—so we can view the desired explanation as drawing only upon resources internal to the more fundamental physical theory.

1. Introduction. Batterman's remarkably insightful recent book, *The Devil in the Details: Asymptotic Reasoning in Explanation, Reduction, and Emergence*, develops a number of salutary themes—among them, the importance of stability results in the explanation of physical phenomena and the fruitfulness of simultaneous engagement with issues in general philosophy of science and with details of a variety of physical theories (Batterman 2002).[1]

But the most provocative theme of the book is that fundamental physical theories can be expected to be explanatorily inadequate—that there exist phenomena that can be understood only when the conceptual apparatus of a fundamental theory is supplemented by concepts whose nat-

*Received June 2003; revised January 2004.

†To contact the author, please write to: Department of Philosophy, University of Pittsburgh, Pittsburgh, PA, 15260; e-mail: gbelot@pitt.edu.

‡Earlier versions of this paper were presented at Columbia University and at the 2003 Pacific APA. I would like to thank two anonymous referees, Mark Wilson, Larry Sklar, Tim Maudlin, and especially Bob Batterman for helpful comments and suggestions. This paper is based upon work supported by the National Science Foundation under Grant No. SES-0135445. The online version of the paper, available as PITT-PHIL-SCI preprint 00001515, includes an appendix, omitted from the present version, that surveys recent results on quantum chaos.

1. All citations of Batterman refer to this work.

Philosophy of Science, 72 (January 2005) pp. 128–153. 0031-8248/2005/7201-0007$10.00
Copyright 2005 by the Philosophy of Science Association. All rights reserved.

https://doi.org/10.1086/428072 Published online by Cambridge University Press

ural home is in a less fundamental theory. Batterman's route to this claim begins with a discussion of the classical philosophical accounts of explanation and reduction, leads through considerations surrounding the asymptotic analysis of the differential equations of optics and quantum mechanics, and ends with an argument that attention to such considerations should transform our view of the role of less-than-fundamental theories in our knowledge of the world.

This paper follows a similar trajectory. By way of providing a backdrop for Batterman's project, I begin in Section 2 with a discussion of explanation and reduction. I then consider, in Sections 3 and 4, the examples from optics and mechanics upon which Batterman bases his case. Finally in Section 5, I turn to Batterman's argument for his thesis that fundamental theories are typically explanatorily inadequate, and suggest a line of objection to this argument.

Perhaps it will be helpful if I provide a preliminary indication here of this line of objection. Batterman holds it to be a shortcoming of the older methodological literature on explanation and reduction that it proceeds at a high level of abstraction. Batterman is one of a number of contemporary writers on these topics who insists that adequate methodological accounts must be built upon a careful examination of the complexities of genuine examples of scientific research. This is one of the senses in which the devil is in the details for Batterman: only by working through detailed examples will we learn certain important lessons that we cannot afford to ignore.[2] In Sections 3 and 4 I am going to discuss in some detail the examples that Batterman bases his case upon—but I will in Section 5 draw a lesson different from his. To put the matter roughly and tendentiously: I am going to focus for the most part on the sort of details that figure in mathematicians' treatments of these examples, where Batterman tends to draw his details from physicists' accounts; I am going to charge Batterman with following physicists in thinking of less fundamental theories as making an ineliminable contribution to our understanding of certain phenomena when consideration of the mathematicians' approach would show that such theories make only a heuristic and eliminable contribution.[3] Of course, I will also have to fend off the counter-accusation that I mistake bloodless formalities for genuine physical understanding.

2. See especially Batterman 7. The details Batterman asks us to attend to feature *asymptotic reasoning*, which Batterman characterizes as involving "a type of abstraction—a means of throwing away various details" (3) in order to achieve deeper understanding. This is another sense in which the devil is in the detail—though in this second sense, the devil is something that we ought to ignore rather than attend to. Batterman notes the irony inherent in the double applicability of his title.

3. Similar criticisms of Batterman are advanced in Redhead 2004 and Sklar 2003.

https://doi.org/10.1086/428072 Published online by Cambridge University Press

2. Background. Let us say that one physical theory *supersedes* another if almost all of the phenomena treated adequately by the latter are afforded a superior treatment by the former (here the term 'phenomena' may cover, in addition to observable facts and regularities, unobservable ones whose status is taken to be unproblematic in the context at hand).

We can divide historical examples of superseded theories into two categories: *dead letters* and *theories emeritus*. Theories in the former class remain the subject of serious intellectual interest only among historians and antiquarians, while those falling in the latter continue to play a role in the education and research of mathematicians and physicists.[4] A plausible conjecture is that this distinction can be grounded as follows: Any superseded theory will be able to count among its lasting achievements some degree of empirical adequacy; for dead letters, the list will end there; but theories emeritus will also have lasting explanatory achievements to their credit. Thus, while the value of a Ptolemaic account of a celestial phenomenon (eclipses, say) is exhausted by the naked-eye accuracy of the predictions of the theory, the value of a geometrical optical account of some phenomenon (burning mirrors, say) may lie in the explanatory power of that account, as well as in its empirical adequacy.

We can view the classical approach to reduction as an attempt to provide an account of the logical relation between theories emeritus and their supersessors, in light of which the explanatory power—and pragmatic status—of each theory emeritus is fully vindicated.[5] According to this approach, we expect that, for good T_1 and T_0, if T_1 supersedes T_0, then it is possible to construct within T_1 a deductive-nomological explanation of the laws of T_0—that is, a derivation of the laws of T_0 from the laws of T_1, supplemented where necessary by assertions of particular fact. Now, in interesting cases, this will be impossible: the laws of T_0 involve terms alien to T_1 in a manner which renders the desired deduction impossible. To meet this difficulty, we allow ourselves to extend T_1 by adding new laws, informally understood as correlating the alien terms of T_0 with expressions native to T_1.[6] We then say that T_0 *reduces to* T_1 if it is possible

4. It is sometimes remarked that navigators and surveyors are taught algorithms rooted in spherical astronomy (see, e.g., Reichenbach 1951, 96 or Kuhn 1996, 68 and 102). But it would hardly be natural to say that learning these techniques involves the study of geocentric astronomy. So let us allow that a theory can be a dead letter even if some shadow of it lives on in techniques of calculation.

5. The locus classicus of this approach is Nagel [1949] 1960; 1961, Chapter 11; it can also be found in Hempel 1966, Chapter 8.

6. The status of these new laws is problematic. Are they conventional explicit definitions? Analytic truths? Statements of synthetic identities? On this issue, see Nagel [1949] 1960, 302–304; 1961, 354–358; [1970] 1979, 105–107, Hempel 1966, Section 8.2;

https://doi.org/10.1086/428072 Published online by Cambridge University Press

to explain, in the desired fashion, the laws of T_0 within the extended theory.[7]

Reductionism is sometimes associated with eliminativism—the idea being that to show that the laws of a given theory are just consequences of the laws of one of its supersessors is to show that the theory is in principle dispensable and its ontology derivative. But Nagel and Hempel themselves appear to deny that a successful reduction of, say, biology to chemistry, or of psychology to neurophysics, or of the framework of every day experience to fundamental physics, would show that reality accrued only to the objects of the more fundamental member of the pair.[8] On their view, a successful reduction shows not that the reduced theory need be discarded, but that it is safe to go on using it. For, one has on the authority of the superseding theory the trustworthiness of the predictions of the superseded theory.[9] And it will be safe to go on employing the superseded theory in generating deductive-nomological explanations—since the laws of the superseded theory are just theorems of (a pertinent extension of) the superseding theory; explanations arising in this way will count as explanations from the perspective of the superseding theory.[10] When this obtains, one has license to continue employing the concepts and laws of the superseded theory. Thus reductions of our theories emeritus to fundamental physical theories would explain why theories emeritus are kept on the books.

But here a well-known problem arises—there are, in fact very few examples of reduction in the strict classical sense.

Consider the relation between Galileo's law of free fall and Newton's theory of gravitation. Nagel offers this as an example of a successful

---

[1969] 2001, Section 6, and Sklar 1967, Section 4. NB: All citations of Hempel and Nagel page numbers refer to the reprinted versions.

7. Of course, if we are allowed to introduce supplemental laws hand-crafted to force through the requisite explanations, then any theory will reduce to any other. Nagel expects us to restrict attention to reasonable reduction schemes, in which the supplemental laws are offered in good faith as mere bridges between the respective vocabularies of the two theories ([1949] 1960, 304–305; 1961, 358–359.)

8. See Nagel [1949] 1960, 309, 311; 1961, 366; [1970] 1979, 107; and Hempel 1966, 77–79.

9. Note that this holds throughout the range of the superseded theory, and not, as in the case of Ptolemaic astronomy, merely in a part of this range.

10. Of course, nothing like this is true in the Ptolemaic example—since there the laws of the superseded theory are false according to the superseding theory, we would have no business employing them in a deductive-nomological explanation. There is no sense in which the Ptolemaic explanation of the bounded elongation of Venus and Mercury is vindicated within heliocentric astronomy.

https://doi.org/10.1086/428072 Published online by Cambridge University Press

reduction.[11] In Newtonian mechanics, Galileo's law of free-fall holds only for bodies subject to a uniform gravitational force. But Galileo's treatment is meant to hold of bodies near the earth, where the strength of the gravitational field increases as the surface of the earth is approached. Now, Nagel's account of explanation of laws (1961, 42–43), like Hempel's canonical deductive-nomological account, requires the truth of the explanans. And so, strictly speaking, there is no explanation of Galileo's law of free-fall in Newtonian mechanics. Rather, one can either: (i) derive the superseded law by supplementing the laws of the superseding theory by false empirical assumptions; or (ii) derive from true empirical assumptions together with the laws of the superseding theory a statement that, in a suitable sense, approximates the superseded law, for a range of conditions.

Similar objections apply to the other candidates for reductions proposed by Nagel—the relation between the kinetic theory of gases on the one hand and the Boyle-Charles law and the second law of thermodynamics on the other.[12] Indeed, the problem is very general: A theory reduces to one of its supersessors only if it is in fact a special case of that supersessor. Perhaps the wave theory of light can be understood as being a special case of Maxwell's theory in the relevant sense. But this phenomenon is very rare—it is far more common for a theory to be *corrected* by its supersessors, even within its range of empirical adequacy. This will be the case, for instance, whenever we say that one theory is the limit of the other—as in the relation between geometrical optics and wave optics, classical mechanics and relativistic mechanics, classical mechanics and quantum mechanics, or special relativity and general relativity. In each of these cases, reduction fails because the supersessor, rather than implying the laws of the superseded theory, offers corrections to them.[13]

It is at this point that the eliminativism of Feyerabend (1962, 1965) and Sellars (1961, 1965) gets its foot in the door. The fact that our less fundamental theories (including the 'theory' of common sense) do not reduce to our more fundamental theories shows that the former are,

11. Nagel ([1949] 1960, 291; 1961, 339).

12. See Nagel [1949] 1960, 294 ff.; 1961, 342 ff.; [1949] 1960, 289, 295; 1961, 337, 343, 360, 362 for these examples.

13. This objection to Nagelian reduction was developed in the influential papers of Kemeny and Oppenheim (1956), Popper (1957), Sellars (1961), and Feyerabend (1962). Here is Hempel's Battermanly view of the upshot: "The construal of theoretical reduction as a strictly deductive relation between the principles of two theories, based on general laws that connect the theoretical terms, is indeed an untenable oversimplification which has no strict application in science and which, moreover, conceals some highly important aspects of the relationship to be analyzed" ([1969] 2001, 206).

https://doi.org/10.1086/428072 Published online by Cambridge University Press

strictly speaking, false. At present, we keep them on the books for pragmatic reasons—they simplify calculations. But it will ultimately be desirable to strike them from the books, and educate ourselves to think in the framework of fundamental physics—only thus will we be able to see our way past the pseudo-problems and philosophical dead ends that are suggested by these false conceptualizations of the world.

It is only natural to wonder whether this isn't an over-reaction to the failure of Nagelian reduction. Mightn't there be a somewhat weaker relation that holds between our theories emeritus and their more fundamental successors, which offers us something more than pragmatic reason to keep the former on the books? Couldn't there still be *something* to the idea that theories emeritus differ from dead letters in offering genuine explanatory insight?

Under pressure from objections to their account of reduction, Hempel and Nagel suggested that we might want to introduce a notion of *approximative reduction*: roughly, we would say that a regularity could be approximatively explained by a given theory when the laws of that theory allow us to deduce a near-relative of the regularity; and that one theory approximatively reduces to another if the latter supports appropriate approximative explanations of the laws of the former.[14] This provides the outline of a program for grounding the status of theories emeritus. But this program remains largely unfulfilled. Indeed, there are daunting obstacles to be faced—in addition to the problems that bedevil the notion of explanation, one has to worry about what counts as an acceptable approximation to a regularity of a theory to be approximatively reduced. And this is likely to lead one away from the abstractions that were the stock in trade of the logical empiricists, and towards the messy mathematical details of the limiting relations between particular theories—undermining any hope for a clean and general account of the status of theories emeritus.

Batterman promises to cut through this Gordian knot of difficulties. He examines the relation between geometrical optics and wave optics and the relation between classical mechanics and quantum mechanics, and concludes that in each case the more fundamental superseding theory is explanatorily inadequate—that there are phenomena whose understanding requires that the conceptual resources of the more fundamental theory be supplemented by those of the less fundamental theory.[15] The considerations that drive Batterman's analysis turn crucially upon the complexity of the limiting relation between the pairs of theories that he con-

14. This suggestion is advanced at Hempel 1965, 344–345; 1966, Chapter 5; [1969] 2001, Section 7 and Nagel [1970] 1979, Section 1.

15. For this, see Batterman's Sections 6.5, 7.3, 7.4, 8.2.

https://doi.org/10.1086/428072 Published online by Cambridge University Press

siders (and it is plausible that the argument, if successful, can be carried over to other theories emeritus). Thus methodologists should find cause for joy rather than grief in the failure of classical reduction in these cases. For, this failure opens up a royal road towards the establishment of the explanatory credentials of certain superseded theories, bypassing entirely the muskeg of approximate laws and approximate truth and the strengths and weaknesses of the variety of explanation that they support.

Batterman offers a reason for keeping geometrical optics and classical mechanics on the books that goes far beyond a mere pragmatic interest in calculational tractability—namely, that these theories play an ineliminable role in our understanding of the physical world. How does this work? In the next two sections, I discuss the two examples that Batterman focuses upon.

3. Example: The Rainbow. Let's begin with (the successful bits of) Descartes' account of the rainbow in the Eighth Discourse of his *Meteors*. Suppose that you find yourself in the situation depicted in Descartes' diagram (see Figure 1): You stand wearing a velvet suit and a sword, looking towards a spherical droplet of water suspended in the air, which is illuminated by light source behind you. We will assume that the light is of a single wavelength. We are interested in the light that enters the raindrop from the left, undergoes a sequence of refractions and internal reflections, and exits the raindrop traveling towards the left. Let us call such a ray *primary* if it exits the drop after a single internal reflection, *secondary* if it exits after two internal reflections, and so on.

Descartes' strategy is to imagine 10,000 evenly spaced rays parallel to AB, each considered as an initial segment of a primary ray, and, for a plausible value of the refractive index of water, to calculate the angle between the initial and final segments of the ray. That is, he wants to calculate angles like the angle DEM of his diagram—that angle is, of course, the same as the angle between AB and DE. His conclusion: "I found that after one reflection and two refractions, very many more of them can be seen under the angle of 41° to 42° than under any lesser one; and that none of them can be seen under a larger angle" (Olscamp 1965, 339). Similarly, Descartes considers the secondary rays and finds that "very many more of them come toward the eye under a 51° to 52° angle, than under any larger one; and no such rays come under a lesser" (339).

If we suppose now that the observer faces a cloud of spherical water droplets, illuminated from behind by our light source, we can study the patterns of illumination which result from primary and secondary rays from each of the droplets. Describing the position of raindrops in the cloud by the angle that a line joining them makes to the continuation past the observer of a line from the sun to her eye, it follows from Des-

https://doi.org/10.1086/428072 Published online by Cambridge University Press

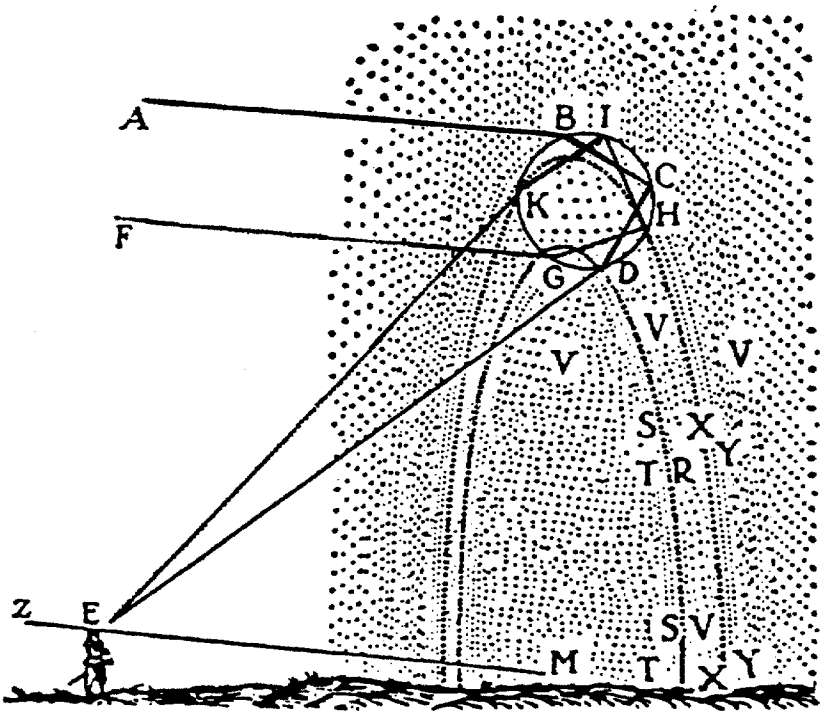


Figure 1. Descartes' diagram, from the Eighth Discourse of his *Meteors*. Here, ray ABCDE is primary, and ray FGHIKE is secondary.

cartes' analysis that many primary rays will reach the eyes from drops along the circular band corresponding to 41° to 42°. This band is the *primary bow*. Its inner side is lit by primary rays—but with the intensity falling off rapidly as the angle decreases. Similarly, the *secondary bow* consists of an illuminated band at 51° to 52°, with a lit outer side (with the intensity falling off rapidly as the angle increases). The region from 42° to 51° is unlit by primary and secondary rays.

Aside from failing to treat chromatic phenomena, this approach gives a good rough account of some of the most conspicuous features of naturally occurring rainbows.[16]

The 19th century wave-theoretic analysis of the rainbow, due to Airy

16. One can also consider the tertiary bow, corresponding to rays undergoing three internal reflections, and so on, Such higher-order bows can be observed under laboratory conditions, but play no role in our story. Note, in particular, that they are distinct from the supernumerary bows mentioned below.

https://doi.org/10.1086/428072 Published online by Cambridge University Press

and others, predicts the following departures from the geometrical treatment.[17]

- The dark band is faintly illuminated near the outside (inside) of the primary (secondary) bow, although this tails off quickly away from the bow.
- The peak in intensity of the primary (secondary) bow occurs at an angle less than (greater than) the angle predicted by the geometrical-optical account.
- Within the illuminated region one finds, rather than the monotonic decrease in intensity predicted by the geometrical-optical account, a series of minor peaks in intensity—these are known as the *supernumerary bows*.
- In the geometrical-optical account, the positions of the primary and secondary bows are fixed by the refractive index of water (together with the assumption that the drops are perfectly spherical). In the wave-theoretic account, the position of the bows depends on the size of the water droplets.[18]

Each of these is borne out by observation—indeed, except for the first, each can be verified at the level of naked-eye observation.

So far, then, we have a familiar sort of story: A good theory giving a good account of a phenomenon of interest is superseded by an even better theory giving an even better account of that phenomenon. Batterman claims that, contrary to appearances, the wave-theoretic account of the rainbow is explanatorily inadequate, and that understanding of this phenomena requires the mobilization of resources from geometrical optics.

Why? Well, let's start by asking what we expect from a scientific explanation of a phenomenon such as the rainbow. Hempel (1966, 48) advocates the plausible view that what is characteristic of scientific explanation is that "the explanatory information adduced affords good grounds for believing that the phenomenon to be explained did, or does, indeed occur. This condition must be met if we are to be entitled to say: 'That explains it—the phenomenon in question was indeed to be expected under the circumstances!'" By way of example:

> Consider . . . the physical explanation of a rainbow. It shows that the phenomenon comes about as a result of the reflection and re-

17. See Tricker 1970, Chapter 6. For popular accounts, see Boyer 1987, Chapter 11 or Nussenzveig 1977.

18. When the light source emits white light, chromatic aspects of the bow also depend upon drop size. For drops of moderate size (of the order of a millimeter), one sees all the colors of the spectrum. But for very small droplets, one gets an almost entirely white primary bow, the so-called *fog bow*.

https://doi.org/10.1086/428072 Published online by Cambridge University Press

> fraction of the white light of the sun in spherical droplets of water such as those that occur in a cloud. By reference to the relevant optical laws, this account shows that the appearance of a rainbow is to be expected whenever a spray or mist of water droplets is illuminated by a strong white light behind the observer. (48)

Waiving questions of the truth of the laws of geometrical optics, does Descartes' account of the rainbow amount to an explanation in the desired sense? Well, Descartes shows something like: Under normal conditions, a rainbow appears when a cloud of spherical droplets is lit by a bright light behind the observer. So, it remains to show either that the drops in a mist or spray can be expected to be spherical, or that nonspherical drops lead to similar optical phenomena.[19] In fact, falling raindrops of the size of several millimeters have shapes that depart appreciably from the spherical (Pruppacher and Klett 1997, Section 10.3.2). So we are stuck trying to show something like: Any spray or mist that we come across is likely to produce a pattern of illumination very similar to that produced by a cloud of spherical drops.[20]

One way that this could be achieved would be via a structural stability result, according to which perturbations of the shape of the drops lead to relatively tame differences in the structure of appearances. Batterman (89–93) considers one such strategy, under which one shows that in the limit of high frequency, generic perturbations of the shape of the drop away from sphericity lead to a pattern of intensities closely related to that for a spherical drop. This story comes in several parts.

First Part: The Basic Integrals.[21] As a harmless simplification we study the wave equation, $[(\partial^2/\partial t^2) - \triangle]u = 0$, instead of Maxwell's equations (allowing us to focus on a scalar, rather than a vector, quantity). If we consider a point source located at $y \in \mathbb{R}^3$, emitting light of frequency k

19. Descartes adopts the first of these two strategies, arguing that water drops are normally spherical in the Fifth Discourse of his *Meteors* (Olscamp 1965, 299)—while allowing in his discussion of rainbows in the Eighth Discourse that a strong wind could deform this shape, and thus lead to aberrant bows (ibid., 343). But the argument of the Fifth Discourse is—unsatisfactory. Indeed, it is not easy to see how Descartes' eely water particles *could* form themselves into perfectly spherical drops.

20. Note that an argument of the desired type would not fit neatly into the categories of explanation considered in Hempel 1965—there, only particular facts are to be given explanations that render their occurrence merely likely, rather than certain. But in his brief discussion of the sense in which the kinetic theory affords explanations of the gas laws (1966, 68–69), Hempel implicitly liberalizes his notion of probabilistic explanations.

21. For this part, see Varadarajan 1997, Section 1.1 and Guillemin and Sternberg 1977, Chapter 1.

https://doi.org/10.1086/428072 Published online by Cambridge University Press

and amplitude $a \in \mathbb{C}$, then

$$ae^{-ikt}\frac{e^{ik|x-y|}}{|x-y|}$$

describes the radiation reaching a point $x \in \mathbb{R}^3$ ($x \neq y$) at time t. Now suppose that we have a compact (i.e., closed and bounded) surface, S, in $\mathbb{R}^3$ and that each point y on S emits light of frequency k and amplitude $a(y)$ and that these emitters are in phase with one another. Then the radiation at $x \in \mathbb{R}^3$ ($x \notin S$) at time t is given by

$$e^{-ikt}\int_S \frac{e^{ik|x-y|}}{|x-y|}a(y)dy.$$

So we have, for certain sets of emitters on S, an integral expression for a solution to the wave equation everywhere off of S. Since the only dependence on t is via the boring factor out front, we will from now on focus on the integral, calling it $I_k(x)$; $|I_k(x)|^2$ gives the intensity of light at x.

We are interested in certain qualitative aspects of the behavior of light that can be understood by studying the behavior of $I_k(x)$ in the high-frequency limit as $k \to \infty$. Notice that as k grows large, the integrand begins to oscillate wildly. But the behavior of the integral itself can be surprisingly simple—this is the purport of some 19th century results, normally grouped together under the moniker *the principle of stationary phase*.

I state some of these results for a class of integrals that includes our special $I_k(x)$.

We consider integrals of the form $\mathcal{I}_k(x) = \int_S e^{ik\varphi(x,y)}a(x, y, k)dy$, defined for $k > 0$, x in some open subset of $\mathbb{R}^3$, y in some open subset of $\mathbb{R}^2$; with φ real and a complex and vanishing for y outside of some compact set, S.[22] We call a the *amplitude* and φ the *phase*. For each x, we have the function $\varphi_x(y) := \varphi(x, y)$ defined on S. We are interested in the critical points of such $\varphi_x(y)$ (i.e., points y where the derivative along S of $\varphi_x(y)$ vanishes). In the limit $k \to \infty$ we find:

- $\mathcal{I}_k(x)$ is $O(k^{-N})$, for all $N \geq 1$ if $\varphi_x(y)$ has no critical points. Such x are said to lie in the *dark zone*.
- $\mathcal{I}_k(x)$ is $O(k^{-1})$ if $\varphi_x(y)$ has only nondegenerate critical points. Such

22. We should further assume that a has a locally uniform asymptotic expansion of the form $a(x, y, k) \sim \sum_{r=0}^{\infty} a_r(x, y)k^{\mu-r}$ as $k \to \infty$; and similarly for all of its derivatives. $a \sim \sum a_r k^{\mu-r}$ means that for each N, $a - \sum_{r=0}^{N} a_r k^{\mu-r}$ is $O(k^{N+1})$. Of course, there is nothing special about $\mathbb{R}^3$ and $\mathbb{R}^2$ for the domains of x and y—except that fixing attention on this case has allowed me to leave out some exponents from the expressions appearing in the bullet points below.

https://doi.org/10.1086/428072 Published online by Cambridge University Press

x are said to lie in the *light zone*; for k large, such points will be illuminated far more brightly than points in the dark zone.[23]

- The points x, for which $\varphi_x(y)$ has degenerate critical points, are said to form the *caustic*. Typically, at such points, $\mathcal{I}_k(x)$ is $O(k^{-d})$ for some $d < 1$—so that for k large, such points will be illuminated far more brightly than points in the light zone.

In the optical case that we are interested in, $\varphi_x(y) = |x - y|$. For fixed x, this function has a critical point at $y \in S$ iff the geodesic joining x and y is orthogonal to S—that is, iff there is a ray from y to x. And a critical point is degenerate iff x is a focal point for rays from S.[24] So we can characterize the degree of illumination at a point x in the high frequency limit in geometrical optical terms—roughly dividing points into those reached by no ray, those reached by at least one ray, and those reached by too many rays (*too many* because light burns at caustics).

Second Part: The Rainbow. Suppose that a point source emits light, which then encounters a number of mirrors and lenses. A solution to the wave equation that represents such a process may have a very complicated form. But in the limit $k \to \infty$, it can be replaced by a finite sum of our basic integrals, $\mathcal{I}_k(x) = \int_S e^{ik\varphi(x,y)} a(x, y, k)dy$, where $x \in \mathbb{R}^3$ and the integral is over auxiliary variables y; applying the principle of stationary phase will then give us a qualitative characterization of the pattern of illumination in the high-frequency limit for any such problem.[25] In the case of the rainbow, a good approximation for most purposes is given by the Airy integral, $\int_{\mathbb{R}} e^{ik(y^3+xy)} a(y)dy$.[26] Via the principle of stationary phase and numerical integration, one finds that this integral gives the pattern described above for the illumination associated with the primary bow: a relatively large peak for a value just less than 42°, quickly tailing off as one enters the dark zone, tailing off more gradually on the lit side, where

23. In fact, there is much more to be said about the contributions of nondegenerate critical points—see, e.g., Varadarajan 1997, 171. But we can ignore this.

24. Here is how this can be made precise. Define $E : S \times \mathbb{R} \to \mathbb{R}^3$ by letting $E(y, w)$ be the point reached by tracing w units along the ray through y perpendicular to S. Consider $x \in \mathbb{R}^3$ such that $x = E(y, w)$ for some y and w. Then y is a nondegenerate critical point of φ_x iff there exists a neighborhood $U \subset S$ of y on which $E(\cdot, w)$ is a diffeomorphism onto its image. This fails, for instance, with a spherical mirror: The shape of the mirror means that for certain values of w, arbitrarily nearby points on the mirror will have the same image under $E(\cdot, w)$.

25. See Duistermaat 1974, 1978 for proofs of this important result.

26. "Airy mentions that the study of this integral is at the foundation of the theory of the rainbow, and that he has been unable to integrate it in terms of the functions known at that time. We now know not too much more . . ." (Varadarajan 1997, 176). For a heuristic treatment of the rainbow in terms of the Airy integral, see Tricker 1970, Chapter 6; for a discussion of the limitations of Airy's approach, see Nussenzveig 1977.

https://doi.org/10.1086/428072 Published online by Cambridge University Press

one also finds a series of small peaks corresponding to the supernumerary bows. Here the extremal ray at 42° plays the role of a caustic (the illumination of the dark zone can be thought of as the result of diffraction around its sharp edge—see Tricker 1970, Chapter 6). As $k \to \infty$, the major peak of illumination will creep closer and closer to 42°, and the illumination of the major peak relative to the minor supernumerary peaks will grow without bound.

Third Part: Structural Stability of Caustics. Now, it may seem that we haven't gotten very far. The wave-theoretic account for spherical drops *corrects* the geometrical optics account. So of what interest is the observation that a consonance between the two accounts emerges in the high-frequency limit?

The point is this. We cannot, at this time, say very much about the behavior of light—even in its qualitative aspects—in problems closely related to the physical problem for perfectly spherical raindrops. We do not have any overarching theorems. It would be an analytic feat to reach an understanding of even a few relevant solutions—and whatever we were able to say about the similarities and differences between the behavior in a few such solutions and in the solution for the exactly spherical case, it would be difficult to be certain that what we were saying was generically true.

But we can do better in the high frequency limit. In particular, as we perturb the phase function of the Airy integral, we generate models of optical situations related to our original case of perfectly spherical drops ("in optics this perturbation might be produced by altering the initial surface . . . or the propagation medium" (Berry and Upstill 1980, 267)). As we do so, we will, of course, change some features of the behavior of the integral, including its asymptotic behavior in the limit $k \to \infty$. But some such features will remain unchanged. In particular, in the case that we care about, the structure of the caustic associated with the integral is unchanged for almost all perturbations, in the sense that the pre- and post-perturbation caustic sets are diffeomorphic to one another.[27] This isolates a qualitative feature of the high frequency limit that remains unchanged—albeit in a rather weak sense—as we move away from the case of perfectly spherical drops.

Batterman suggests that it is this sort analysis that allows us to say that we understand why rainbows occur in generic situations—and furthermore that this understanding must be seen as involving elements of geometric optics as well as of wave optics:

> The ray theoretic caustic, with its structural stability properties, is

27. This is a substantial result. See, e.g., Duistermaat 1978, 26.

https://doi.org/10.1086/428072 Published online by Cambridge University Press

> responsible for universal behavior in the intensity and form of various wave patterns. Thus, in the asymptotic domain characterized by the limit . . . $k \to \infty$, *elements from both theories seem to be intimately and inextricably intertwined*. (94)

While granting that there is a sense in which anything made available in an asymptotic analysis of the wave equation must be 'contained in' the wave equation we start with, Batterman (Section 6.5) insists that understanding of such an analysis requires reference to geometrical optical structures foreign to the wave theory.

4. Example: Quantum Mechanics. Quantum mechanics stands to classical mechanics as wave optics stands to geometric optics—this analogy was a guiding principle for the founders of quantum mechanics. The semiclassical domain, in which one studies the asymptotic behavior of quantum mechanical structures as $h \to 0$, corresponds to the regime in which one studies the behavior of optical phenomena in the high-frequency limit.

Batterman discusses semiclassical matters at length, and reaches a conclusion similar to that arrived at in the optics case:

> There are many aspects of the semiclassical limit of quantum mechanics that cannot be explained purely in quantum mechanical terms, though they are in some sense quantum mechanical. Just as in the case of optics . . . the fundamental theory is, in certain instances, explanatorily deficient. For instance, there are differences in the morphologies of semiclassical wave functions that depend *essentially* on the nature of the "underlying" classical motion—whether the classical motion is regular or chaotic. (109)

The explanatory deficiency of quantum mechanics consists in the fact that certain "quantum mechanical features require reference to classical properties for their full explanation" (110).

In Section 7.3, Batterman mentions, as an example of the sort of thing he has in mind, the striking role that classical structures play in the Gutzwiller trace formula, which gives semiclassical information about the energy spectra of quantizations of chaotic systems.

I'll say a bit about how this goes, working up to Gutzwiller-style results in stages.[28] The first stage will be an appetizer: a clean example of the sort of thing we are after—namely, a theorem showing how to gain information about the energy spectrum of a quantum system, as $h \to 0$, by studying facets of the geometry of the phase space of the corresponding

28. My presentation in this section omits a number of technical hypotheses, especially in the third stage below. Please consult the references for details.

https://doi.org/10.1086/428072 Published online by Cambridge University Press

classical system. The second stage consists of some remarks intended to motivate interest in trace formulae. The third stage is a description of rigorous analogs of Gutzwiller's trace formula.

We work in the following setting. We begin with a classical system whose configuration space is $\mathbb{R}^n$, with Hamiltonian $H(q, p) = \frac{1}{2}p^2 + V(q)$ (here $(q, p) = (q^1, \ldots, q^n, p_1, \ldots, p_n)$ is an element of the phase space, $T^*\mathbb{R}^n$). So the corresponding quantum theory has as Hilbert space the space $L^2(\mathbb{R}^n)$ of square-integrable wave functions on the classical configuration space, and quantum Hamiltonian $\hat{H} = -(h/2)\triangle + V$. We suppose that V is bounded away from zero, so that the spectrum of $\hat{H}$ is discrete for sufficiently small h.

First Stage: Weyl's Formula. For each $h > 0$, we write the energy spectrum as $\lambda_1(h) < \lambda_2(h) \leq \cdots$. We define $N_h : \mathbb{R} \to \mathbb{N}$ by $N_h(E) = \text{card}\{j : \lambda_j(h) \leq E\}$, and study the behavior of this function as $h \to 0$. In this limit, one finds

$$N_h(E) \sim \frac{1}{(2\pi h)^n} \text{vol}(\{(q, p) : H(q, p) \leq E\}),$$

where the volume form in question is the Liouville form, $dq^1 \wedge \cdots \wedge dq^n \wedge dp_1 \wedge \cdots \wedge dp_n$, associated with the classical phase space.[29]

Second Stage: Trace Formulae.[30] Now consider an operator, A, on our infinite-dimensional Hilbert space, $L^2(\mathbb{R}^n)$. Suppose that A has discrete spectrum, with eigenbasis $\{\phi_i\}$ and corresponding eigenvalues $\{\lambda_i\}$. Then we say that A is of *trace class* with trace $\text{Tr}(A) := \sum \lambda_i$ when this sum converges absolutely. There is a nice way of calculating the trace: Let $K_A : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{C}$ be given by $K_A(x, y) = \sum \lambda_i \phi_i(x)\overline{\phi_i(y)}$; then, so long as this sum converges uniformly and absolutely, K_A is the integral kernel of A (i.e., for $\psi \in L^2(\mathbb{R}^n)$, $A(\psi)(x) = \int_{\mathbb{R}^n} K_A(x, y)\psi(y)dy$); and in this case $\text{Tr}(A) = \int_{\mathbb{R}^n} K_A(x, x)dx$. A nice feature of this formula is that the left-hand side is algebraic and the right-hand side is analytic-geometric.

When A has discrete spectrum, but is not of trace class, we can proceed as follows. Let $\varphi : \mathbb{R} \to \mathbb{R}$ be Schwartz (so that it and all its derivatives decrease rapidly as ∞ is approached), and suppose that $\sum \varphi(\lambda_i)$ converges absolutely. Then we switch our attention to the operator $\varphi(A)$ (i.e., that operator with eigenbasis $\{\phi_i\}$ with corresponding eigenvalues $\{\varphi(\lambda_i)\}$). This operator is of trace class, with trace given by $\sum \varphi(\lambda_i)$. We can study the integrals $\text{Tr}(\varphi(A)) = \int_{\mathbb{R}^n} K_{\varphi(A)}(x, x)dx$. This gives us information about the spectrum of A as we vary φ.

29. For this, see, e.g., Colin de Verdière 1998, Section 5.1. For a finer version, see Robert 1998, Theorem 3.11.

30. Here I follow Uribe (2000) rather closely.

https://doi.org/10.1086/428072 Published online by Cambridge University Press

Often we are interested in an operator, but we don't have even the most basic knowledge of its spectrum—let alone the sort required to calculate the trace from first principles or via the integral kernel. It turns out, though, that there are expressions which give us semiclassical information about the trace of such operators in terms of certain geometrical structures of the classical theory, without requiring us to have prior knowledge of the eigenfunctions and eigenvalues.

Third Stage: Semiclassical Trace Formulae. I am going to say just a bit about such results—just enough to give you feel for what Batterman is up to.

Here is the central result.[31] Let E be a noncritical value of the classical energy.[32] Then there exists a sequence $\{\gamma_i\}$ of distributions defined on $\mathbb{R}$ such that for any Schwartz φ with compactly supported Fourier transform $\hat{\varphi}$, we find that as $h \to \infty$

$$\sum_{\lambda_i(h) \in [E-\varepsilon, E+\varepsilon]} \varphi(h^{-1}(\lambda_i(h) - E)) = \sum \gamma_j(\hat{\varphi}) h^{-n+j} + O(h^{\infty}).$$

(Here ε is a constant that depends only upon the Hamiltonian.) This is a sort of semiclassical trace formula—it gives us information about the sum of certain ranges of eigenvalues, as $h \to \infty$. The crucial point for our purposes is that the support of each of the distributions γ_i is contained in the set of periods of the periodic classical trajectories lying on the classical energy surface $H^{-1}(E)$—in order to calculate with this formula, we need to know about the behavior of *classical* trajectories.

This trace formula is valid for a wide range of systems, regular and chaotic. Now let us assume that the periodic trajectories of the classical flow are nondegenerate—roughly, that within each energy surface they are isolated from one another rather than occurring in continuous families. This condition will obtain in chaotic classical systems. In this case, there is a semiclassical trace formula whose left-hand side is of the above form and whose right-hand side involves a number of terms with a geometrical interpretation—a term involving the Liouville measure induced on the classical energy surface, a sum over the classical periodic trajectories of a function of the classical action, and so on.[33]

5. Consequences? The picture that emerges from these examples is described thus by Batterman: "Wave theoretic or quantum mechanical as-

31. See, e.g., Helffer 1997, Theorem 4.5.1 or Robert 1998, Theorem 3.8.

32. I.e., we have that if $H(x, \xi) = E$ then $\nabla H(x, \xi) \neq 0$.

33. See Colin de Verdière 1998, Section 6.3 or Uribe 2000, Theorem 2.2. For related results, see Helffer 1997, Section 4.7 and Robert 1998, Corollary 3.9.

https://doi.org/10.1086/428072 Published online by Cambridge University Press

pects are sewn onto a skeleton of classical trajectories and their structures in the asymptotic . . . limit" (127). How is this observation supposed to carry us through to the conclusion that wave-theoretic optics and quantum mechanics are explanatorily inadequate—that there are phenomena whose understanding ineliminably involves the concepts of geometric optics and classical mechanics?

Let us explicitly separate two of Batterman's theses. The first thesis is that the successful explanation of a regularly occurring phenomenon involves two tasks: (i) finding solutions to the equations of a given theory that possess a feature corresponding to the occurrence of the phenomenon in question; and (ii) showing (something like) that the possession of this feature is stable under (generic) physically realistic perturbations of such solutions. This thesis seems entirely correct to me—and I think it one of the great virtues of Batterman's work that it calls attention to the mathematical difficulty and philosophical interest of task (ii).[34]

The second thesis, which I think we should reject, is that Batterman's examples provide cases where a given theory contains resources adequate to the first of these explanatory tasks but not to the second—but where both tasks can be accomplished when the resources of the given theory are supplemented by resources of a less fundamental theory. In the case of the rainbow, it is claimed (90-91) that although one can construct solutions to the equations of wave optics corresponding to the appearance of the rainbow, one is unable to show that this behavior is generic within the relevant range of circumstances without recourse to ray-theoretic notions such as caustics. Similarly it is claimed (109-110) that there exist (rather more abstract) quantum mechanical phenomena (involving eigenvalue statistics, etc.) that can be understood only when the resources of quantum mechanics are supplemented by knowledge of the classical mechanical trajectories in phase space. Considering the first example, Bat-

34. When Batterman says things like "asymptotic explanation does, in fact, represent a distinct form of explanation largely missed by current philosophical conceptions" (37), he intends it to be understood that his discussion of asymptotic explanation provides a sort of competitor to deductive-nomological explanation. I think of him, rather, as offering a friendly amendment to the deductive-nomological account. At one point Hempel says, speaking of the deductive-nomological model, "given this notion of explaining a particular occurrence of a solar eclipse or of a rainbow, etc., one can speak *derivatively* of a theoretical explanation of solar eclipses or rainbows in general: such an explanation is then one that accounts for any instance of an eclipse or rainbow" (1965, 423). Clearly constructing an explanation of the rainbow in general ought not to require providing a deductive-nomological explanation for each occurrence of the rainbow—one cannot consider infinitely many sets of possible initial and boundary conditions, and show that each leads to a rainbow. So charity would appear to require us to read Hempel here as (perhaps imperfectly) perceiving the necessity of something like requirement (ii) above.

https://doi.org/10.1086/428072 Published online by Cambridge University Press

terman asks whether the features of the rainbow that we are interested in—the form of the bow, the interference fringes on its lit side, etc.—can be explained within wave-theoretic optics. His answer:

> Now, if by "explanation" one has in mind one of the dominant contemporary conceptions—the D-N model or some causal-mechanical model—then I think that the answer is a clear "no." . . . These emergent phenomena are not derivable in any straightforward sense from the underlying wave theory. They are not, as it were, from-first-principle solutions to the wave equation . . . . They are deeply encoded in that equation but are apparent only through its asymptotic analysis.
>
> On the other hand, the "theory" of that asymptotic domain—catastrophe optics—does provide satisfactory accounts of the phenomena of interest. So the phenomena are not inexplicable or brute tout court. One has, in a well-defined sense, an explanation that is "grounded in" the fundamental wave theory; but this kind of *asymptotic* explanation is . . . distinct from the usual types of explanation talked about in the philosophical literature. The phenomena are not explainable through *derivation*—that is, through straightforward solutions to the differential equation—from the fundamental wave theory alone. (118–119)

Batterman thinks that a delicate balance must be struck in considering each of his examples. For the phenomena he is interested in, Batterman locates some of the resources crucial to their explanation in the limiting regime between a more, and a less, fundamental theory. He thinks that while there is some sense in which these resources are contained in the more fundamental theory, there is also a sense in which they are unpredictable from it alone—since "the understanding of those mathematical representations requires reference to structures foreign to the fundamental theory" (96; cf. 109 ff.).

Against this, I claim that an examination of the sense in which rays and caustics are implicit in the apparatus of the wave theory, and of the sense in which classical phase space trajectories are implicit in the mathematical structure of quantum mechanics, saps all intuitive force from Batterman's claim that we here have cases where the understanding of phenomena requires that the resources of the more fundamental theories be supplemented by resources of the less fundamental theories. In brief outline, the argument is as follows. In Batterman's cases, the physics of the more fundamental theory is described by a certain partial differential equation whose independent variables parameterize a manifold X. Suppose that we ask some mathematicians to investigate the asymptotic behavior of approximate solutions to this equation (or the behavior of the

https://doi.org/10.1086/428072 Published online by Cambridge University Press

spectrum of the corresponding operator) as some parameter appearing in the equation grows large or small. One of the most powerful techniques for attacking such problems involves the construction and investigation of an auxiliary problem definable from knowledge of the original problem. One first constructs a new manifold T^*X, the cotangent bundle associated with X (a point of T^*X is a pair (x, ξ) consisting of a point $x \in X$ and a covector ξ at x). One also constructs a function, the *principal symbol*, on T^*X associated with the differential operator of the original problem. T^*X carries a natural geometrical structure which in concert with the principal symbol determines a set of curves in T^*X. The auxiliary problem consists in the study of these curves. Now, formally speaking, our auxiliary problem is a problem in the theory of Hamiltonian systems: T^*X can be thought of as the space of possible positions and momenta for a particle living on X, and our curves are the dynamical trajectories for such a particle if the energy of the particle in each state (x, ξ) is just given by the value of the principal symbol at that state. But in order to construct this auxiliary problem, our mathematicians need not have heard of any physical theory (such as classical mechanics), whose mathematics is given by the theory of Hamiltonian systems—the auxiliary problem can be constructed uniquely as soon as the original problem is given, the structures needed to set up the auxiliary problem being definable in terms of those used to set up the original one. And once they have the auxiliary problem at hand, our mathematicians can see that the asymptotic behavior features of the original problem will turn upon the features of the auxiliary problem—in the optics case, in the high frequency limit, approximate solutions to the wave equation blow up near points of X where the projection of certain dynamical trajectories from T^*X to X becomes singular (these are the caustic points) while in the quantum mechanics case, periodic dynamical trajectories in the auxiliary problem will play a special role in the analysis of the spectrum of the quantum Hamiltonian in the $h \to 0$ limit. So there is, as far as the mathematics goes, nothing *foreign* to the more fundamental theory about the ingredients needed to prove the asymptotic results that secure explanations of phenomena such as the rainbow—these ingredients are implicit in the mathematics of the more fundamental theory.

Two questions remain. (1) How does all of this work? (2) When one tries to understand all of this as physics rather than as mere mathematics, does one have to draw upon the conceptual resources of the less fundamental theory after all? I address these two questions, in turn, in the following subsections. (Readers uninterested in (1) should skip to Section 5.2.)

5.1. More Details. I describe in slightly more detail the construction

https://doi.org/10.1086/428072 Published online by Cambridge University Press

of the auxiliary problem associated with a given differential equation and its role in the asymptotic investigation of that equation.[35] Suppose that we are interested in the study of a linear partial differential equation whose independent variables parameterize a manifold X of dimension n. We assume that this differential equation is of order m and depends upon a parameter τ. We suppose that in coordinates we can write the equation in the form $P(x, \partial/\partial x, \tau)u(x) = 0$, with $u \in C^{\infty}(X)$ and P of the form $P = \sum_{\alpha=0}^{m} a_{\alpha}(x, \tau)(\tau(\partial/\partial x))^{\alpha}$ with each a_{α} of the form $\sum_{j=0}^{\infty} b_{\alpha,j}(x)\tau^{j}$.[36] Recall that T^*X is the cotangent bundle of X—the points (x, ξ) of this space can be thought of as the possible positions and momenta of a particle moving in X. The manifold T^*X possesses a natural symplectic structure ω (in virtue of being a cotangent bundle). For present purposes, we can think of ω as being, more or less, an antisymmetric metric on T^*X.[37] We introduce the *principal symbol* of the operator P: This is the function $P_0 : T^*X \to \mathbb{R}$, defined as $P_0(x, \xi) := \sum b_{\alpha,0}\xi^{\alpha}$.[38]

Now, in classical mechanics, if the possible configurations of a physical system are parameterized by a manifold Q (the configuration space), then its possible dynamical states (positions and momenta) are parameterized by the associated cotangent bundle T^*Q (the phase space), and the dynamics are determined by specifying a Hamiltonian function, $H : T^*Q \to \mathbb{R}$, assigning to each dynamical state its total energy. The Hamiltonian encodes information about the forces acting on the system, and in concert with the natural symplectic structure on T^*Q, it determines a vector field on T^*Q whose integral curves correspond to the dynamical trajectories of the theory.[39] Without invoking any physical interpretation, we can employ the same mathematics to associate with our principal symbol P_0, a vector field Ξ_{P_0}, on T^*X. The integral curves of this vector field are called the *bicharacteristic strips*.

So given X and P, we construct T^*X, P_0, and Ξ_{P_0}. The study of the

35. The following account is based upon Bates and Weinstein 1997, Chapters 1–4; Colin de Verdière 1998; Duistermaat 1974, Section 1; and Guillemin and Sternberg 1977, Chapter 2.

36. I write out the formulae for dim $X = 1$. For higher dimensions, you mess around with multi-indices in the obvious way. See Colin de Verdière 1998, Section 4.3.

37. ω is in fact a nondegenerate closed two-form on T^*X. If $\{x^i\}$ are coordinates on X, and if at each $x \in X$ the $\{\xi^i(x)\}$ are the elements of the cotangent space at X dual to the basis $\{\partial/\partial x^i\}$ of the tangent space at x, then (x, ξ) is a set of coordinates for the manifold T^*X. In terms of these coordinates, $\omega = \sum dx^i \wedge d\xi^i$.

38. The text gives a coordinate-based characterization of P_0. But the recipe leads from P to an intrinsically defined function on T^*X.

39. For $f \in C^{\infty}(T^*X)$, the associated vector field, Ξ_f, is defined by $\omega(\Xi_f, \cdot) = df$. This is just a recasting of Hamilton's equations in a coordinate-independent form.

https://doi.org/10.1086/428072 Published online by Cambridge University Press

associated bicharacteristic strips is the auxiliary mathematical problem that plays a role in the study of the asymptotic behavior of our original differential equation, $Pu = 0$. The account of the role of this auxiliary problem is somewhat involved. I focus on the construction of approximate solutions in the asymptotic regime, and break the procedure into three steps.

First Step: Find a Lagrangian manifold that solves the Hamilton-Jacobi equation. Because the symplectic form ω on T^*X is antisymmetric, if $(x, \xi) \in T^*X$ and $v \in T_{(x,\xi)}T^*X$, then $\omega(v, v) = 0$—i.e., the symplectic product of a tangent vector to T^*X with itself always vanishes. But for fixed v in the tangent space at (x, ξ), there will in fact be many w in that tangent space such that $\omega(v, w) = 0$. We look for an n-dimensional submanifold L of T^*X with the property that at any point of L, any two vectors v and w in the tangent space to L at that point satisfy $\omega(v, w) = 0$ (n is the maximum possible dimension of such a submanifold). Such an L is called a *Lagrangian submanifold*. We are looking for a Lagrangian submanifold that satisfies the Hamilton-Jacobi equation—this just means that P_0 is constant on L.[40]

Second Step: Find a function on L invariant under the flow induced by P_0. Suppose that we have such a submanifold, L. Our next step in constructing our approximate solution to our equation is to look for a scalar function, α, defined on L with the property that the Lie derivative of α in the direction Ξ_{P_0} vanishes.[41]

Third Step: Project the structure (L, α) *down to X*. When all goes well, the structure (L, α) can be projected down to X to yield a second-order approximate solution to our differential equation $Pu = 0$.[42] Let $\pi : T^*X \to X$ be the projection map (so $\pi(x, \xi) = x$). We call a point of L *regular* if the restriction of π to L is a diffeomorphism on some neighborhood of that point; otherwise we call that point *singular*. We call the image under π of the bicharacteristic strips the *bicharacteristic curves*. For any regular point of L there is a sufficiently small neighborhood U, so that the bicharacteristic curves in $\pi(U)$ do not cross; but at the image under π of any singular point of L the bicharacteristic curves will intersect one another. If U is a sufficiently small open set of regular points of L

40. We require, in fact, that L lie in a level set of a regular value of P_0—i.e., $\nabla P_0(x, \xi) \neq 0$. For convenience, we restrict attention to connected Lagrangian submanifolds.

41. In fact, it is better to take α to be a half-density rather than a function, and likewise for a (introduced below). For half-densities see, e.g., Bates and Weinstein 1997, Appendix A.

42. Indeed, given such a solution, there is a procedure for producing approximate solutions of arbitrary order. This procedure involves solving only *ordinary* differential equations.

https://doi.org/10.1086/428072 Published online by Cambridge University Press

on which π acts as a diffeomorphism, then we can find a function S and a function a (actually a half-density) defined on $\pi(U) \subset X$ such that: (i) $U = \{(x, dS(x)) : x \in \pi(U)\}$; and (ii) $a = (dS)^*\alpha$ (that is, we pull α back to X using the one-form dS, thought of as a map from $\pi(U) \subset X$ to $U \subset T^*X$). So we can break the regular part of L up into subsets, U_j, projectable to X, such that the structure (L, α) can be encoded in functions S_j and a_j living on the $\pi(U_j)$. Roughly speaking, $u(x) = \sum e^{i\tau S_j(x)} a_j(x)$ is our second-order approximate solution to $Pu = 0$. But we still have subtleties to worry about—the phases of the terms in this sum need to be made to mesh, and we have to take the singular points of L into account. The final upshot is: (i) only certain L will be *admissible*—that is, embedded in T^*X in such a way that the whole process can be carried out consistently; (ii) the final form for an asymptotic solution to $Pu = 0$ corresponding to (L, α) will consist of a finite sum of integrals of the form

$$\left(\frac{\tau}{2\pi}\right)^{k/2} \int_K e^{i\tau\phi(x,y)} a(x, y, \tau) dy$$

(where K is a compact subset of $\mathbb{R}^k$, and the y are to be thought of as auxiliary variables).[43]

What does all this have to do with Batterman's examples? Well, for the differential operator of the reduced wave equation the corresponding principal symbol is the kinetic energy deriving from the metric on X (where here X represents physical space). In this case, the bicharacteristic strips correspond to the geodesic flow on T^*X, and their images on X, the bicharacteristic curves, are just the geodesics of X. These last correspond to light rays in the physical situation described by (L, α). The images of the singular points of L under π form the caustic corresponding to this family of rays. It follows from our recipe for pulling back (L, α) to X and from the principle of stationary phase that our approximate solution will have asymptotic behavior at the caustic of the sort we are familiar with from the rainbow case.

In the quantum mechanical case we are interested in the operator corresponding to the time-independent Schrödinger equation—and the corresponding principal symbol is just the classical Hamiltonian. Our recipe

43. Why integrals here? To get an intuitive idea, consider the simplest case, where $X = \mathbb{R}$ and $T^*X = \mathbb{R}^2$. A Lagrangian submanifold, L, is just a level set of P_0 and a point (x, ξ) lying on such a set is singular iff $T_{(x,\xi)}L = \{(0, t) : t \in \mathbb{R}\}$. To take care of such a point, we project some neighborhood of it onto the ξ-axis (rather than the x-axis as we do for regular points), carry out our procedure sketched in the text, then Fourier transform—ending up with an integral expression for the contribution to the shadow on X of (L, α) corresponding to our chosen neighborhood of our singular point of L.

https://doi.org/10.1086/428072 Published online by Cambridge University Press

for constructing approximate solutions is a generalization of the WKB method, and the admissible Lagrangian submanifolds lie in the level sets corresponding to the points of the spectrum of the quantum Hamiltonian. So the auxiliary mathematical problem coincides with a mathematical treatment of the classical limit of our quantum problem. In this case, the bicharacteristic strips are the classical dynamical trajectories in T^*X. So it is no longer so surprising that these play a special role in the asymptotic analysis of the energy spectrum of our quantum system—since they are definable from knowledge of the quantum Hamiltonian alone, and play a key role in available techniques for constructing approximate solutions. Indeed, the proof of the semiclassical trace formula turns upon the ingredients introduced above: continuous superpositions of Hamilton-Jacobi-like solutions, and application of the principle of stationary phase (see Uribe 2000, 72 ff.).

5.2. Empty Formalism? The foregoing serves to make clear the asymptotic results that Batterman focuses on are indeed contained in the fundamental equations he considers. A great intuitive analyst ignorant of mathematical physics could be handed a linear partial differential equation (defined on a manifold X and depending on a parameter) and asked to construct asymptotic approximate solutions or to investigate the asymptotic properties of the spectrum of the corresponding operator. This analyst might well rediscover the material sketched above—rediscovering in the process the symplectic geometry of T^*X and the mathematical theory of Hamiltonian mechanics, considered now only as props for the study of our differential equation. Indeed, this is more or less how symplectic geometry and Hamiltonian mechanics figure in some mathematical treatments of the results sketched above (see, e.g., Guillemin and Sternberg 1977, Chapters 1–3).

The equation studied by our untutored genius might or might not have a physical interpretation. Either way, our analyst will be able to prove results about asymptotic behavior of solutions of the equation, or the spectrum of the corresponding differential operator, in the limit where the parameter grows large. These results will be tied up with the details of the principal symbol on T^*X, and the associated bicharacteristic strips and curves. This further structure is definable in terms of X and our differential equation alone—and so is conceptually available as soon as we set out to study the equation.

Now suppose that it is the reduced wave equation that our analyst has been set to work on. She has, unbeknownst to her, proved the results that we need to explain the occurrence of the rainbow: Handed a certain set of initial and boundary conditions, she has shown that in the high frequency limit, Airy's integral provides a good approximation for the be-

https://doi.org/10.1086/428072 Published online by Cambridge University Press

havior of the quantity called the *intensity of light* for a certain region of space; she has shown, perhaps via numerical integration, that the pattern of the 'intensity of light' matches the curves illustrated in the textbooks for the 'rainbow'; finally, she has shown that certain qualitative aspects of this pattern of the 'intensity of light' are invariant under a certain family of perturbations of the mathematical problem that she has been set. So far, the analyst has only a mathematical understanding of the problem. In order to transform her work into an explanation with physical content, we need: (i) to impart to her the standard sense of the 'intensity of light'; (ii) to explain why the given initial and boundary conditions correspond to a situation in which a cloud of spherical water droplets is illuminated by white light; and (iii) to explain why the perturbations studied correspond to changes in the shape of the drops. None of this would appear to require reference to the concepts of geometric optics.[44] A similar story can be told about quantum mechanical examples: The output of the semiclassical analysis is a characterization of the energy spectrum of the system—which one does not require the concept of a classical mechanical particle in order to understand.

Batterman claims, very plausibly, that much of our detailed understanding of physical phenomena turns upon semiclassical considerations, in which we consider asymptotic behavior of solutions of the equations of a given theory as a parameter that they involve goes to infinity. Noticing that the asymptotic behavior of solutions of wave optics is sensitive to the structure of caustics, and that the asymptotic features of quantum mechanics are sensitive to the structure of the set of classical periodic trajectories, Batterman concludes that our physical understanding of the phenomena in question has one leg in the physics of the more fundamental theory and one leg in the physics of the less fundamental theory. But this is too hasty: The mathematics of the less fundamental theory is definable in terms of that of the more fundamental theory; so the requisite mathematical results can be proved by someone whose repertoire of interpreted physical theories includes only the latter; and it is far from obvious that the physical interpretation of such results requires that the mathematics of the less fundamental theory be given a physical interpretation. Putting it another way: In the examples considered, it can be shown that given

44. This is entirely clear, I think, for (i) and (iii). It is perhaps less obvious in the case of (ii). Indeed, Wilson (2003) suggests that it is just here that geometrical-optical concepts play a crucial role. But note that if this point is correct, it would lead to a picture of the explanatory importance of less-than-fundamental theories rather different from that developed by Batterman—such theories would be needed in showing that particular solutions of the equations of the fundamental theory represent a given phenomenon as occurring, rather than in showing that its occurrence is robust under perturbations of such solutions.

https://doi.org/10.1086/428072 Published online by Cambridge University Press

a physically interpreted fundamental theory, one is able to prove the results required to explain the phenomena in question without recourse to less fundamental theories, except as mathematical crutches that are in any case definable from the mathematics of the more fundamental theories; so the less fundamental theories, as robust, interpreted physical theories, do not make an ineliminable contribution to our understanding in these examples.

REFERENCES

Bates, Sean, and Alan Weinstein (1997), *Lectures on the Geometry of Quantization*. Providence, RI: American Mathematical Society.

Batterman, Robert (2002), *The Devil in the Details: Asymptotic Reasoning in Explanation, Reduction, and Emergence*. Oxford: Oxford University Press.

Berry, Michael, and C. Upstill (1980), "Catastrophe Optics: Morphologies of Caustics and Their Diffraction Patterns", in Emil Wolf (ed.), *Progress in Optics 18*. Amsterdam: North-Holland, 257–346.

Boyer, Carl (1987), *The Rainbow: From Myth to Mathematics*. Princeton, NJ: Princeton University Press.

Colin de Verdière, Yves (1998), "Une Introduction à la Méchanique Semi-Classique", *L'Enseignement Mathématique* 44: 23–51.

Duistermaat, Johannes (1974), "Oscillatory Integrals, Lagrange Immersions, and Unfolding of Singularities", *Communications on Pure and Applied Mathematics* 27: 207–281.

——— (1978), "The Light in the Neighborhood of a Caustic", in Albrecht Dold and Beno Eckmann (eds.), *Séminaire Bourbaki vol. 1976/77. Exposés 489–506.* Berlin: Springer-Verlag, 19–29.

Feyerabend, Paul (1962), "Explanation, Reduction, and Empiricism", in Herbert Feigl and Grover Maxwell (eds.), *Scientific Explanation, Space, and Time*. Minneapolis: University of Minnesota Press, 28–97.

——— (1965), "Reply to Criticism: Comments on Smart, Sellars, and Putnam", in Robert Cohen and Marx Wartofsky (eds.), *Boston Studies in the Philosophy of Science, vol. 2: In Honor of Philipp Frank.* New York: Humanities Press, 223–261.

Guillemin, Victor, and Shlomo Sternberg (1977), *Geometric Asymptotics*. Providence, RI: American Mathematical Society.

Helffer, Bernard (1997), "*h*-Pseudodifferential Operators and Applications: An Introduction", in Jeffrey Rauch and Barry Simon (eds.), *Quasiclassical Methods*. Berlin: Springer-Verlag, 1–49.

Hempel, Carl (1965), *Aspects of Scientific Explanation and Other Essays in the Philosophy of Science*. New York: Free Press.

——— (1966), *Philosophy of Natural Science*. Englewood Cliffs, NJ: Prentice-Hall.

——— ([1969] 2001), "Reduction: Ontological and Linguistic Facets", in James Fetzer (ed.), *The Philosophy of Carl G. Hempel: Studies in Science, Explanation, and Rationality*. Oxford: Oxford University Press, 189–207. Reprint. Originally published in Sidney Morgenbesser, Patrick Suppes, and Morton White (eds.), *Philosophy, Science, and Method: Essays in Honor of Ernest Nagel*. New York: St. Martin's Press, 179–199.

Kemeny, John, and Paul Oppenheim (1956), "On Reduction", *Philosophical Studies* 7: 6–19.

Kuhn, Thomas (1996), *The Structure of Scientific Revolutions*, 3rd edition. Chicago: University of Chicago Press.

Nagel, Ernest ([1949] 1960), "The Meaning of Reduction in the Natural Sciences", in Arthur Danto and Sidney Morgenbesser (eds.), *Philosophy of Science.* New York: Meridian Books, 288–312. Reprint. Originally published in R. Stauffer (ed.), *Science and Civilization*. Madison, WI: University of Wisconsin Press, 99–145.

https://doi.org/10.1086/428072 Published online by Cambridge University Press

——— (1961), *The Structure of Science: Problems in the Logic of Scientific Explanation*. New York: Harcourt.

——— ([1970] 1979), "Issues in the Logic of Reductive Explanations", in Ernest Nagel, *Teleology Revisited and Other Essays in the Philosophy and History of Science*. New York: Columbia University Press, 95–117. Reprint. Originally published in Howard Keifer and Milton Munits (eds.), *Mind, Science, and History*. Albany, NY: SUNY Press, 117–137.

Nussenzveig, Herch (1977), "The Theory of the Rainbow", *Scientific American* 236 (4): 116–127.

Olscamp, Paul (ed.) (1965), *Descartes: Discourse on Method, Optics, Geometry, and Meteorology*. Indianapolis: Bobbs-Merrill.

Popper, Karl (1957), "The Aim of Science", *Ratio* 1: 24–35.

Pruppacher, Hans, and James Klett (1997), *Microphysics of Clouds and Precipitation*. Dordrecht: Kluwer.

Redhead, Michael (2004), "Asymptotic Reasoning", *Studies in History and Philosophy of Modern Physics* 35: 527–530.

Reichenbach, Hans (1951), *The Rise of the Scientific Philosophy*. Berkeley: University of California Press.

Robert, Didier (1998), "Semiclassical Approximation in Quantum Mechanics: A Survey of Old and Recent Mathematical Results", *Helvetica Physica Acta* 71: 44–116.

Sellars, Wilfred (1961), "The Language of Theories", in Herbert Feigl and Grover Maxwell (eds.), *Current Issues in the Philosophy of Science*. New York: Holt, Rinehart, and Winston, 57–77.

——— (1965), "Scientific Realism or Irenic Instrumentalism", in Robert Cohen and Marx Wartofsky (eds.), *Boston Studies in the Philosophy of Science, vol. 2: In Honor of Philipp Frank.* New York: Humanities Press, 171–204.

Sklar, Lawrence (1967), "Types of Inter-Theoretic Reduction", *British Journal for the Philosophy of Science* 18: 109–124.

——— (2003), "Comments on R. Batterman: *The Devil in the Details*", unpublished paper given at the Pacific APA, 2003.

Tricker, R. A. R. (1970), *Introduction to Meteorological Optics*. London: Mills and Boon.

Uribe, Alejandro (2000), "Trace Formulae", in Salvador Pérez-Esteva and Carlos Villegas-Blas (eds.), *First Summer School in Analysis and Mathematical Physics*. Providence, RI: American Mathematical Society, 61–90.

Varadarajan, Veeravalli (1997), "The Method of Stationary Phase and Applications to Geometry and Analysis on Lie Groups", in Bent Ørsted and Henrik Schlichtkrull (eds.), *Algebraic and Analytic Methods in Representation Theory*. New York: Academic Press, 167–242.

Wilson, Mark (2003), "Comments on *The Devil in the Details*", unpublished paper given at the Pacific APA, 2003.

https://doi.org/10.1086/428072 Published online by Cambridge University Press