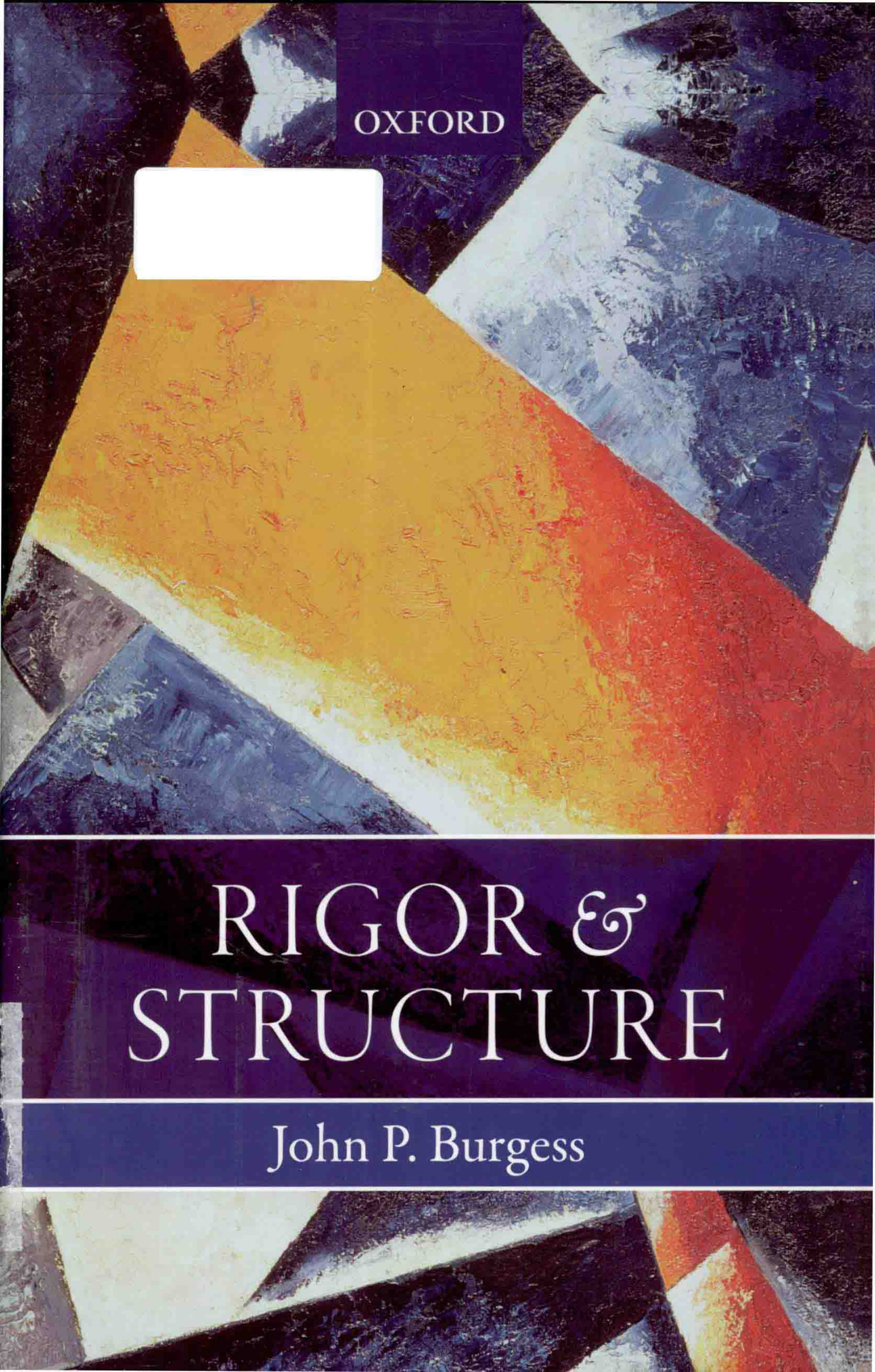

OXFORD

RIGOR & STRUCTURE

John P. Burgess

Rigor and Structure

John P. Burgess

OXFORD
UNIVERSITY PRESS

OXFORD
UNIVERSITY PRESS

Great Clarendon Street, Oxford, OX2 6DP,
United Kingdom

Oxford University Press is a department of the University of Oxford.
It furthers the University's objective of excellence in research, scholarship, and education by publishing worldwide. Oxford is a registered trade mark of Oxford University Press in the UK and in certain other countries

First Edition published in 2015

Impression: 1

Published in the United States of America by Oxford University Press
198 Madison Avenue, New York, NY 10016, United States of America

British Library Cataloguing in Publication Data

Data available

Library of Congress Control Number: 2014953343

ISBN 978-0-19-872222-9

Printed and bound by
CPI Group (UK) Ltd, Croydon, CR0 4YY

Rigor and Structure

Preface

Few buzzwords are more often encountered in philosophical discussions of the nature of mathematics than "rigor" or "structure," and few are more diversely understood. Thus while we are commonly told that the distinctive method of mathematics is rigorous proof, and that the special topic of mathematics is abstract structure, even a superficial look at the literature in which such formulations appear quickly shows that neither mathematicians nor logicians nor philosophers are agreed as to the exact sense in which mathematics is or ought to be concerned with rigor, or with structure. Differing perspectives on such issues can be of practical as well as theoretical importance, leading to divergent views on matters of policy: on the role computer-dependent proofs should be allowed to play, or the relative importance of set theory as opposed to category theory. The aim of the present essay is to clarify our understanding of mathematical rigor and mathematical structure, and especially of the relation between the two, and to sketch how such a clarified understanding can bear on disputed policy questions.

An account of the origin of this work may help explain its organization. Much writing by philosophers about mathematics in recent decades has been concerned with two views about the nature of mathematical objects and mathematical existence, known in the literature as *nominalism* and *structuralism*. Fifteen and more years ago, my colleague Gideon Rosen and I wrote a study of nominalism (Burgess and Rosen 1997). Even at the time I had the feeling that a similar study of structuralism was called for, but other projects intervened, so that only five or so years ago did I begin serious work. My developing views on structuralism were presented in lectures delivered under various titles at Stanford, NYU, Penn, Bristol, Paris, Oxford, and at home in Princeton, and one that would have been delivered in Helsinki but for Eyjafjallajökull. The thesis I was led to was that *the features of mathematical practice that structuralists want to explain in terms of the peculiar nature of mathematical objects are better explained in a different way, as consequences of the way the ancient ideal of rigor is realized in modern mathematics.* More specifically, what that way of working involves is that *the individual mathematician is responsible for the rigor of the derivation of any new results claimed from results in the previous literature, but may remain indifferent to just how the results in the previous literature were derived from first principles.* The exposition and defense of this thesis is to be found in Chapter 3 of the present work.

When I began to work on the present book, my aim was to develop the material in my various lectures into a long paper on the thesis I have just adumbrated. However, when I lectured on my work in progress, I found that in the question-and-answer sessions afterwards, members of the audience always wanted to press a question I had passed over, namely, the question of what the first principles from which the rigorous development of mathematics proceeds should be understood to be. More specifically, the question raised was whether set theory or category theory should be thought of as the starting point. When I was presented with the opportunity to co-teach a graduate seminar on set theory versus category theory, I eagerly accepted, resolving in preparation for the task to think harder about the question of first principles, which I had been inclined to postpone. I was eventually led to a corollary to my main thesis, which in hindsight may appear an obvious consequence of it: *The working mathematician may remain indifferent not only to how results in the existing literature were derived from first principles, but also to what exactly those first principles are—though they* might as well be *those of standard set theory.* The exposition and defense of this corollary is to be found in Chapter 4 of the present work.

As my preparation for the seminar began, it became clear to me that before launching into making the case for my thesis I had better begin, by way of background, with some general discussion of rigor. At first, I expected that some brief prefatory remarks would suffice, but as I went deeper into the matter I found there was more that needed explaining about the ideal of rigor, about how it is realized in present-day mathematics, and about how it came to be realized in that way, than I had originally anticipated. Such preliminary explanations are now to be found in Chapter 1 of the present work.

As my preparation continued, it also became clear to me that if I were to discuss category theory versus set theory, and the corollary to my main thesis, I would need to explain not only what rigor is, but also how set theory came to occupy the position some category theorists would now like to challenge. I mean the position of being in some sense the accepted foundation or starting point for rigorously building up the rest of mathematics. Such further preliminary explanations are given in Chapter 2 of the present work.

Thus, what I originally expected to be a longish paper has turned into the not-so-short book I now lay before the reader. Lest the prospective reader be daunted by the sheer number of pages to follow, let me add an explanatory remark. There is an old, and doubtless false, story about the mathematician Euler confronting the philosopher Diderot at the court of Catherine the Great with a nonsensical mathematical refutation of atheism: "Sir, $(a + b)^n/n = x$, hence God exists. Reply!" I dislike this kind of bullying by invoking technicalities, and have accordingly tried in the present work to keep all mathematics at as elementary a level as I knew how, for the sake of philosophical readers. Equally, I have tried to explain all philosophical background in as elementary and jargon-free way as I knew how, for the sake of mathematical readers. Such is and has long been the state of lack of communication between the two fields that no other course seemed to me feasible, if I wanted to have any hope of not utterly putting off readers of one sort or the other. The result is that mathematical readers will find a fair amount of elementary mathematical exposition intended for philosophers that they will want to skip or skim over, while philosophical readers will similarly find a fair amount of elementary philosophical exposition intended for mathematicians that they will want to read over quickly, if they read it over at all. Thus, for either class of readers the present work will effectively be shorter than it appears to be.

Princeton, November 2014

Acknowledgments

I now turn to the pleasant duty of acknowledging intellectual debts. My oldest and deepest are to two of my teachers: Arnold E. Ross, in whose summer program in number theory for high-school students I first learned what a proof is, and Ivo Thomas, from whom I learned in the same program the elements of classical and intuitionistic formal logic. I am also much indebted to Stewart Shapiro, once a fellow student with me of Ross and Thomas, and now the structuralist philosopher whose work has most concerned me. He was present and participated in the discussion at more than one of my talks on structuralism, and besides engaged in an informative correspondence with me about some of the issues. I owe much also to other questioners from the audiences after my talks, and to two other structuralist philosophers with whom I have been in communication over the years, Geoffrey Hellman and Charles Parsons.

Closer to home, to Hans Halvorson, the Princeton colleague with whom I jointly conducted the graduate seminar on set theory and category theory, and later an undergraduate seminar on philosophy of mathematics, I owe among other things much of such knowledge of category theory as I possess. I must hasten to emphasize, however, that he is not in any way responsible for the opinions I express, and especially not for any that may be uncongenial to his friends among category enthusiasts. I also profited from attending my colleague Benjamin Morison's seminar on Euclid's *Elements*, book 1, and to a greater degree than will be apparent from my scattered remarks on Euclid in this work.

I received useful comments on earlier drafts from Oystein Linnebo, from my emeritus colleague Paul Benacerraf, from my student Jack Woods, from my friends Juliette Kennedy and Jouko Väänänen, and on typographical matters from Michael Scanlan. I am grateful also to OUP's two anonymous referees, who convinced me that some of the things I was inclined to say were open to objections, and others to misinterpretations, which I had not anticipated. The many members of the staff of OUP and affiliates who became involved with the project were exceptionally efficient and helpful throughout, from the first inception of the idea for such a volume to the final proofreading.

Contents

1

Rigor and Rigorization

What Is Rigor?

Go to any university biology library. Pick at random two or three recent volumes of different mainstream journals. Sit down and flip through the pages of several articles in each, not hoping to absorb their substantive content, but merely observing their gross form. Now go over to the mathematics library, and repeat the process. The reader who performs this exercise cannot fail to be struck by an enormous difference in format between papers in a natural science like biology and papers in mathematics. In biology one finds descriptions of experimental design and statistical analyses of data collected; in mathematics, a series of definitions, theorems, and proofs. The distinctive definition–theorem–proof format of professional publications is the single most conspicuous feature of mathematical practice.

In any science, a published paper may turn out to be less innovative or less insightful than it appeared to be to the editors and referees responsible for its acceptance by the journal in which it appears. In such a case, there will simply be a dearth of subsequent citations of the paper, which will fall into oblivion. It may also happen, though less frequently, that there is something so wrong with the paper that professional ethics would call for the publication of a retraction, whether or not one actually appears. In biology, some error in data collection or analysis may be uncovered. In mathematics, some flaw or gap in a proof may be recognized, so that it has to be acknowledged that what appeared to be a proof was not really a proof. In some cases of this kind, the retraction may offer a counterexample to what was originally published as a "theorem"; in others, it may merely be indicated that the previously published result must now be regarded as a mere "conjecture."

The quality whose presence in a purported proof makes it a genuine proof by present-day journal standards, and whose absence makes the proof spurious in a way that if discovered will call for retraction, is called *rigor*. Or at least, that is one of several uses of that term, and the one that will be of concern here. Now the definition of rigor just given, as that which is demanded of proofs in contemporary mathematics, identifies rigor by the role it plays in the mathematical enterprise; but what one would like to have, and what we will be seeking in the pages to follow, is a more intrinsic characterization: an account of what rigor, thus identified, amounts to.

In examining this question, it is natural to begin with what mathematicians themselves may have to say about the matter. They are the experts, after all. For mathematicians, the ability to judge an argument's rigor in our sense, to evaluate whether a proof is up to the standard required in professional publications, is a crucial skill. Mathematicians need to be able to evaluate the rigor of their own work before submitting it, and to judge the rigor of the work of others when called on to serve as referees.

And mathematicians are generally quite able in such matters. Evaluating the rigor of a journal submission, though often tedious in the way that proofreading for typographical errors is tedious, may be the most routine part of refereeing, much more cut and dried than evaluating how innovative or insightful a paper is. For rigor is only a preliminary qualification for publication, as lacking a criminal record may be a preliminary qualification for certain kinds of employment. Failure to meet the requirement may mean rejection, but success in meeting it does not necessarily mean acceptance.

Mathematicians learn what rigor requires, learn to distinguish genuine from bogus proofs, during their apprenticeship as students. This skill is acquired mainly by observation and practice, by example and imitation. Typically, the student "learns what a proof is" in connection with a course on some substantive mathematical topic, perhaps "arithmetic" in the sense of higher-number theory, or more often "analysis" in the sense of the branch of mathematics that begins with calculus. The assessment of purported proofs for rigor is generally *not* the topic of a separate course. In particular, it is *not* generally learned by studying formal logic, courses which tend to be taken more by students of philosophy or computer science than of mathematics.

Mastery of a practical skill does not automatically bring with it an ability to give a theoretical account thereof. Even a champion racing cyclist may find it difficult to explain how to ride a bike, if tied hand and foot and required to indicate what to do using words only, without any physical demonstration. In the same way, expert mathematicians are not automatically equipped to give an accurate theoretical description of the nature of rigorous proof. In any case, they nowadays very seldom attempt to do so. Most mathematicians seem to stand to rigor as Justice Stewart stood to obscenity: They can't define it, though they know it when they see it.

Mathematicians' views on the nature of rigor and proof may more often be expressed in aphorisms and epigrams, in anecdotes and jokes, for instance, of the common "an engineer, a physicist, and a mathematician . . ." genre,[1] than in formal theoretical pronouncements. Browsing on the shelves of the library, or perhaps more conveniently using a search engine, will turn up many crisp sayings and droll stories from mathematicians present and past, major and minor, pure and applied. Most of the sayings are semi-jocose and not to be taken completely literally; many if so taken would be rather obviously false: Though there may be a grain of truth in them, considerable threshing and winnowing may be needed to separate it from the chaff. All together, this material represents something less than a full-blown theoretical account, and though mathematicians' witticisms and humor about rigor and proof contain many pointers, these don't all point in the same direction.

To give at least a couple of examples, a search will turn up many versions, variously formulated and variously attributed, of the dictum whose simplest form is just "a proof is what convinces." (A couple of more refined versions will be taken up later.) A search may also turn up different retellings from different authors of the story about the young gentleman and his tutor. In case the reader is unfamiliar with this tale, here is the version, set in eighteenth-century France, given by Isaac Asimov:

> [A] poor scholar was hired to teach geometry to the scion of one of the nation's dukedoms. Painstakingly, the scholar put the young nobleman through one of the very early theorems of Euclid, but at every pause, the young man smiled amiably and said, "My good man, I do not follow you." Sighing, the scholar made the matter simpler, went more slowly, used more basic words, but still the young nobleman

[1] The oldest may be the following. A mathematician, a physicist, and an engineer travel on a train in Scotland. They catch a glimpse of a sheep. "Sheep in Scotland are black," says the engineer. "There is at least one black sheep in Scotland," says the physicist. "There is at least one sheep in Scotland with at least one black side," says the mathematician.

said, "My good man, I do not follow you." In despair, the scholar finally moaned, "Oh, monseigneur, I give you my word what I say is so." Whereupon the nobleman rose to his feet, bowed politely, and answered, "But why, then, did you not say so at once so that we might pass on to the next theorem? If it is a matter of your word, I would not dream of doubting you." (Asimov 1971, 456)

The unstated moral of the story is precisely that, contrary to the dictum that a proof is what convinces, *not* everything that convinces is a proof.

If several Fields medalists, often described as the nearest mathematical analogue of Nobel laureates, were to pledge their word of honor that a certain mathematical conjecture is true, few of us would dream of doubting them. But publication of an assemblage of such testimonials in favor of some outstanding conjecture, say the one known as the *Riemann hypothesis*, though it might well produce widespread conviction that the conjecture is true, would not qualify for the million-dollar prize offered by the Clay Foundation for the solution of one of its "Millennium Problems" (see Carlson et al. 2006). For to qualify for the prize one must publish, in a peer-reviewed journal, a *proof*, and whatever exactly is needed for a proof, a collection of testimonials is not enough. So at the very least, conviction derived from testimony must be distinguished from conviction derived from proof: The supporting arguments for a piece of rigorous mathematics cannot just be appeals to authority, or tradition, or revelation, or faith. The further clarification of the nature of rigor will require characterizing what more is excluded.

Before launching into that project, it should be mentioned that the dictum that "a proof is what convinces" sometimes fails for the opposite of the reason we have been considering so far: Genuine proofs do not always convince, at least not on first exposure. Anyone who has taught undergraduate mathematics knows that proofs do not always immediately convince students, but indeed they do not always immediately convince even the proof's own author. The skeptical Scottish philosopher David Hume put it well:

> There is no . . . Mathematician so expert in his science, as to place entire confidence in any truth immediately upon his discovery of it, or regard it as any thing, but a mere probability. Every time he runs over his proofs, his confidence increases; but still more by the approbation of his friends; and is rais'd to its utmost perfection by the universal assent and applauses of the learned world. (Hume 1739, part IV, section I)

Such are the effects of an awareness of human fallibility.