Proving Quadratic Reciprocity: Explanation, Disagreement, Transparency and Depth

William D'Alessandro

Abstract

Gauss's quadratic reciprocity theorem is among the most important results in the history of number theory. It's also among the most mysterious: since its discovery in the late 18th century, mathematicians have regarded reciprocity as a deeply surprising fact in need of explanation. Intriguingly, though, there's little agreement on how the theorem is best explained. Two quite different kinds of proof are most often praised as explanatory: an elementary argument that gives the theorem an intuitive geometric interpretation, due to Gauss and Eisenstein, and a sophisticated proof using algebraic number theory, due to Hilbert. Philosophers have yet to look carefully at such explanatory disagreements in mathematics. I do so here. According to the view I defend, there are two important explanatory virtues—*depth* and *transparency*—which different proofs (and other potential explanations) possess to different degrees. Although not mutually exclusive in principle, the packages of features associated with the two stand in some tension with one another, so that very deep explanations are rarely transparent, and vice versa. After developing the theory of depth and transparency and applying it to the case of quadratic reciprocity, I draw some morals about the nature of mathematical explanation.

1 Introduction

C.F. Gauss's quadratic reciprocity theorem is remarkable several times over. For one, the theorem describes a surprising connection between pairs of prime numbers which has inspired and puzzled mathematicians since its discovery. For another, it's among the most re-proved results in mathematics; Gauss himself proved it in eight different ways, and the best current count finds 246 proofs to date ([Lemmermeyer 2019]). Last but not least, attempts to generalize quadratic reciprocity have led to the monumental Langlands Program, one of the most ambitious enterprises in contemporary mathematics.

The reciprocity theorem—hereafter, QR—has had no trouble capturing the attention of mathematicians and historians.[1] But surprisingly little has been written about its rich philosophical dimensions.[2] Among other issues, the theorem and its proofs promise to teach us

[1] [Lemmermeyer 2000] is a thorough and useful entry point to this literature.

[2] See [Gray 2015] (on QR as an example of deep mathematics), [Tappenden 2008] (on the naturalness of the Legendre symbol, and on the importance and mysteriousness of QR generally), and [Yap 2011] (on the usefulness of Gauss's congruence notation for proving QR).

about explanation in mathematics, a topic of much interest to philosophers in recent years.[3] Partly this is because the proofs of QR are so numerous and varied, which invites questions about the relative explanatory value of different approaches. Even more notable, however, is the lack of a clear consensus—and, indeed, the existence of active and longstanding disagreement—about QR's proper explanation. Such controversies can shed a unique kind of light on a discipline's norms, values and goals, but philosophers have yet to look carefully at this aspect of mathematical practice.[4]

My first goal here is to examine this situation and to develop some theoretical tools for understanding it. According to the view I defend, there are two important explanatory virtues—*depth* and *transparency*—which different proofs (and other potential explanations) possess to different degrees. Although not mutually exclusive in principle, the packages of features associated with the two stand in some tension with one another, so that very deep explanations are rarely transparent, and vice versa. (Roughly, a deep explanation derives the explanandum from distant sources and thus places it in a wider theoretical setting, while a transparent explanation makes the explanandum intuitive and clear.) The reciprocity theorem is unusual in that, among its many proofs, there are striking examples of both types. For mathematicians in search of a transparent explanation, QR's deep proofs are likely to seem unwieldy and excessively technical. On the other hand, those who favor depth have sometimes judged QR's transparent proofs to be frivolous, artificial and point-missing. This, I suggest, is at least a large part of the reason why explaining quadratic reciprocity has proven contentious.

While such an analysis is, I hope, worthwhile in its own right, my goal isn't just to present an interesting case study. The conclusions reached here have broader consequences for the theory of explanation, a few of which I deal with below.

There's much ground to cover, so here's my plan. I begin in §2 with some historical background for QR, and I show that the theorem is widely regarded as mysterious and in need of explanation. The goal here is to give nonspecialists a better appreciation for this important piece of mathematics, as well as to display its philosophical interest. §3 develops the general theory of transparency and depth, including their relationship to each other, their links to explanation, and their respective roles in mathematics. §4 presents three proofs of QR: one by induction (due to Gauss), one by counting lattice points (due to Gauss and Eisenstein), and one from algebraic number theory (due to Hilbert). The first proof has been widely judged unexplanatory, while mathematicians have disagreed about whether QR is best explained by proofs like the second or proofs like the third. The theory of §3 can explain this disagreement. As I show, the lattice-point proof is transparent but not deep, while the algebraic proof is deep but not transparent (and Gauss's original proof is neither). Mathematicians who favor one sort of explanatory style are likely to prefer the corresponding

[3] A handful of noteworthy examples, in chronological order: [Steiner 1978], [Resnik & Kushner 1987], [Hafner & Mancosu 2005], [Tappenden 2005], [Mancosu 2008a], [Lange 2009], [Frans & Weber 2014], [Lange 2014], [Pincock 2015], [Inglis & Mejía-Ramos 2019], [D'Alessandro 2020]. Some of these and other works on mathematical explanation are discussed below, especially in §3 to §5.

[4] A recent exception is [Colyvan et al. 2018], which discusses a disagreement about how best to explain the free group theorem. Colyvan et al. suggest that there are at least two fundamentally different ways for a proof to be explanatory; they call the relevant properties "abstractness" and "constructiveness", although neither these qualities nor their connection to explanation are analyzed in much detail.

type of proof.

Finally, I discuss a philosophical consequence in §5. I show that the case of QR is troublesome for certain accounts of explanation, including Marc Lange's influential theory of explanatory proof ([Lange 2014]). The crux of the issue is that it's unclear in advance what kind of proof will furnish a good explanation of QR; the theorem suggests there's something important about prime numbers that awaits a more complete understanding, but we can't say much a priori about what it is that we don't know or what an illuminating proof should look like. This is a problem for theories like Lange's, which identify an explanation with an answer to a specific, determinate why-question provoked by a theorem.

2 On quadratic reciprocity

2.1 What the theorem says

I've said a lot about QR without yet stating the theorem, so let me do so now. The result is about *quadratic residues modulo primes*. (The terminology "b is congruent to a modulo m", symbolized $b \equiv a \pmod{m}$, means that a is the remainder left when b is divided by m. For example, $32 \equiv 8 \pmod{12}$. Here and whenever we talk about congruences, the numbers involved are assumed to be integers.)

A *quadratic residue* modulo m is an integer that's congruent to a perfect square modulo m. In other words, a is a quadratic residue modulo m if we can find an x such that $x^2 \equiv a \pmod{m}$. For instance, 6 and 9 are quadratic residues modulo 10, since $4^2 \equiv 16 \equiv 6 \pmod{10}$ and $3^2 \equiv 9 \pmod{10}$. But 7 isn't a quadratic residue modulo 10; as you can verify, there's no integer x such that $x^2 \equiv 7 \pmod{10}$. (It suffices to check this for $0 \leq x \leq 9$. In fact, the quadratic residues modulo 10 are exactly $0, 1, 4, 5, 6, 9$.)

Just as perfect squares are extremely important in ordinary arithmetic, it's often useful to know whether or not one number is a quadratic residue modulo another number. Adrien-Marie Legendre introduced a clever piece of notation, the *Legendre symbol*, to represent the possible answers to this question. The Legendre symbol is written $\left(\frac{a}{p}\right)$—for a any integer and p a prime—and is defined as follows:

$$\left(\frac{a}{p}\right) = \begin{cases} 0 & \text{if } p \text{ divides } a \\ 1 & \text{if } a \text{ is a quadratic residue modulo } p, \text{ and } p \text{ doesn't divide } a \\ -1 & \text{if } a \text{ isn't a quadratic residue modulo } p, \text{ and } p \text{ doesn't divide } a. \end{cases}$$

We can now state the quadratic reciprocity law, which is often expressed in terms of the Legendre symbol. Stated this way, the theorem takes the form

$$\left(\frac{p}{q}\right)\left(\frac{q}{p}\right) = (-1)^{\frac{p-1}{2}\cdot\frac{q-1}{2}}, \tag{1}$$

where p and q are odd primes. A slightly different but also common statement (which will be useful later) is

$$\left(\frac{p^*}{q}\right) = \left(\frac{q}{p}\right),$$

where p^* is defined as $(-1)^{\frac{p-1}{2}} p$. The following version is the longest on words but the most perspicuous:

> If either p or q is congruent to 1 (mod 4), then p is a quadratic residue modulo q just in case q is a quadratic residue modulo p. If both p and q are congruent to 3 (mod 4), on the other hand, then p is a quadratic residue modulo q just in case q is *not* a quadratic residue modulo p.

This is in fact just a restatement of (1). To see this, note that the exponent $\frac{p-1}{2} \cdot \frac{q-1}{2}$ is odd if both factors are odd, and is even otherwise. In the case that $\frac{p-1}{2} \cdot \frac{q-1}{2}$ is even, we have $\left(\frac{p}{q}\right)\left(\frac{q}{p}\right) = 1$, meaning that p and q are either both or neither quadratic residues modulo each other. And an integer of the form $\frac{n-1}{2}$ is even just in case $n \equiv 1 \pmod 4$. On the other hand, if $\frac{p-1}{2} \cdot \frac{q-1}{2}$ is odd, we have $\left(\frac{p}{q}\right)\left(\frac{q}{p}\right) = -1$, meaning that exactly one of p and q is a quadratic residue modulo the other. And indeed, an integer of the form $\frac{n-1}{2}$ is odd just in case $n \equiv 3 \pmod 4$.

QR itself doesn't say anything about the special cases $\left(\frac{-1}{p}\right)$ and $\left(\frac{2}{p}\right)$; these are dealt with by "supplementary laws" often given alongside the reciprocity theorem. The supplementary laws won't feature much below, but for completeness, they're as follows:

$$\left(\frac{-1}{p}\right) = 1 \iff p \equiv 1 \pmod 4$$

$$\left(\frac{2}{p}\right) = 1 \iff p \equiv 1 \text{ or } 7 \pmod 8.$$

The relationship between $\left(\frac{p}{q}\right)$ and $\left(\frac{q}{p}\right)$ described by QR is unexpected. On the face of things, it's far from obvious that p's status as a quadratic residue modulo q should tell us anything about q's status as a residue modulo p. (Indeed, results like the Chinese Remainder Theorem give us good reason to expect the opposite; cf. §2.3 below.) Yet QR says that these two questions are closely related: if we know the answer to one question, we can immediately deduce the answer to the other.

The sense of mystery surrounding the reciprocity theorem isn't easily dispelled. Even after two centuries of progress in number theory, mathematicians continue to regard QR as a surprising result that demands explanation. The rest of the paper will have much more to say on this issue. It would leave an important part of the story untold, however, to look at possible explanations of QR without mentioning where the theorem came from and why mathematicians find it so interesting. I outline this history briefly in the next section.[5]

2.2 A brief history of reciprocity

The genealogy of QR starts in the mid-1600s, with Pierre de Fermat's investigations in number theory. Fermat was interested in an eminently natural question: Which positive

[5] I mostly follow [Cox 2013] for this history. [Ireland & Rosen 1990], [Frei 1994] and [Baumgart 2015] are also useful guides to various aspects of QR, and the compendious [Lemmermeyer 2000] contains a great many further references.

integers are sums of two squares? The first few numbers on this list, as you can check, are

$$1, 2, 4, 5, 8, 9, 10, 13, 16, 17, 18, 20, 25, 26, 29, 32, 34, 36, 37, 40, 41.$$

Notably, the odd primes appearing here (5, 13, 17, 29, 37, 41) are all and only those of the form $4n + 1$. This observation led Fermat to the statement

$$p = x^2 + y^2 \text{ if and only if } p \equiv 1 \pmod 4,$$

for p an odd prime and for some integers x, y. (Fermat claimed to have proved this, but as usual he didn't provide an argument. Albert Girard had previously announced the same theorem, also without proof, a couple decades earlier.) Fermat then turned to other relationships between congruence conditions and representations of primes. He found, for instance, that

$$p = x^2 + 2y^2 \text{ if and only if } p \equiv 1 \text{ or } 3 \pmod 8,$$
$$p = x^2 + 3y^2 \text{ if and only if } p = 3 \text{ or } p \equiv 1 \pmod 3.$$

Fermat recorded some further conjectures along these lines, but he admitted he didn't know how to prove them. As it turns out, his successors' attempts to do so would eventually lead to the discovery of QR.

Although Fermat was well known in scientific circles of the day, few of his contemporaries shared his zeal for number theory. The next mathematician to make significant progress on these issues was Euler, who read Fermat around 1730 and spent the next several decades proving—or in some cases disproving—his predecessor's claims. (He finally obtained complete proofs of the three theorems mentioned above in 1772.) The techniques Euler devised to solve these problems led him right to the doorstep of the reciprocity theorem, so it's worth briefly explaining his approach.

The general issue raised by Fermat's results is as follows. Given an integer n, we want to find a congruence condition on odd primes that will tell us whether such a prime is of the form $x^2 + ny^2$. The most interesting and difficult part of this problem turns out to be finding a congruence condition that ensures *divisibility*—once we know that $p \mid x^2 + ny^2$, it's not too much harder to show that $p = x^2 + ny^2$, at least for the relatively simple cases mentioned above. So the question that required the largest part of Euler's efforts was this:

Is there a congruence condition on odd primes p that determines when $p \mid x^2 + ny^2$?

Euler took a major step forward—and a major step toward QR—when he realized that the divisibility condition is equivalent to a statement about quadratic residues. Making use of the Legendre symbol, the result is as follows:

$$p \mid x^2 + ny^2 \text{ and } \gcd(x, y) = 1 \iff \left(\frac{-n}{p}\right) = 1. \tag{2}$$

Proposition (2) gives the first glimpse of the relationship between QR and Fermat's question about primes. In fact, we're already fairly close to bridging the gap. Given (2), the

next natural question to ask is: Under what circumstances do we have $\left(\frac{-n}{p}\right) = 1$? Euler's calculations led to conjectures essentially like the following:

$$\left(\frac{3}{p}\right) = 1 \Longleftrightarrow p \equiv \pm 1 \,(\text{mod } 12)$$
$$\left(\frac{5}{p}\right) = 1 \Longleftrightarrow p \equiv \pm 1, \pm 9 \,(\text{mod } 20)$$
$$\left(\frac{7}{p}\right) = 1 \Longleftrightarrow p \equiv \pm 1, \pm 9, \pm 25 \,(\text{mod } 28)\,.$$

The right-hand sides of these biconditionals list exactly the odd squares modulo $4n$. Euler found the same pattern to hold for all the odd prime values of n he examined, though not for composite n. These data suggest the following conjecture, where p and q are odd primes:

$$\left(\frac{q}{p}\right) = 1 \Longleftrightarrow p \equiv \pm a^2 \,(\text{mod } 4q) \text{ for some odd integer } a. \tag{3}$$

Pleasantly, (3) turns out to be true. What's much more surprising is that (3) is equivalent to quadratic reciprocity. This fact is worth dignifying with a line of its own:

$$\text{For distinct odd primes } p \text{ and } q, \text{ statement (3) is equivalent to } \left(\frac{p}{q}\right)\left(\frac{q}{p}\right) = (-1)^{\frac{p-1}{2}\cdot\frac{q-1}{2}}\,. \tag{4}$$

In light of (4), Euler is often credited with discovering QR. Neither statement actually appears in his work, though—what Euler actually found was a general conjecture about conditions of the form

$$\left(\frac{N}{p}\right) = 1 \Longleftrightarrow p \equiv \pm a \;(\text{mod } 4N)\,,$$

for N composite as well as prime. The particular case (4) corresponding to QR, although a consequence of Euler's conjecture, is never singled out for special mention.

The first explicit statement of the reciprocity theorem comes from Legendre some forty years later, in 1785. Legendre's version of QR is quite modern—besides contributing his eponymous symbol, Legendre also gave the theorem its current name. His statement of the "law of reciprocity which exists between two arbitrary prime numbers" is as follows:

$$\left(\frac{n}{m}\right) = (-1)^{\frac{n-1}{2}\cdot\frac{m-1}{2}}\left(\frac{m}{n}\right),$$

where m and n are distinct odd primes.[6]

Legendre was the first to publish an alleged proof of QR, but his argument didn't work: "some of the cases are proved rigorously, some depend on Dirichlet's theorem on primes in arithmetic progressions"—which wouldn't be proved until 1837— "and some are a tangle of circular reasoning" ([Cox 2013], 36). Curiously, Legendre went on repeating his incomplete proof for years, even after its flaws had been pointed out by Gauss and others.

[6] [Legendre 1830], 230.

It was Gauss himself, of course, who turned reciprocity into a proper theorem. He famously discovered his first proof at the age of 19, in 1796, without having read Euler or Legendre. (So Gauss didn't use Legendre's term 'reciprocity'; he calls QR "the fundamental theorem" in the *Disquisitiones Arithmeticae* and "the golden theorem" in his diary.) The *Disquisitiones* contains Gauss's first proof and another one using quadratic forms. Gauss found six more proofs over the next twenty years (four published, two not), for a lifetime total of eight. For the most part, each proof uses quite different ideas from the others. Later sections of this paper will say more about Gauss's proofs, his thoughts about their relative merits, and his reasons for returning repeatedly to QR.

2.3 Explaining QR

2.3.1 The need for explanation

Since its earliest days, mathematicians have viewed QR as a remarkable fact that calls out for explanation. Gauss's himself was driven to repeatedly reprove the theorem in part because he sought a better understanding of its sources and significance. On this score, not much has changed since 1796. Unlike many classical results—whose original discoveries were hard-won, but which have come to seem obvious and trivial in light of subsequent advances—QR's reputation hasn't diminished over time. Indeed, if anything the opposite is true. Ireland and Rosen's classic textbook calls it "among the deepest and most beautiful results of elementary number theory" and the forebear of "the very general Artin reciprocity law, perhaps the most impressive theorem in all number theory" ([Ireland & Rosen 1990], 54).

Descriptions of QR as mysterious or surprising also abound in the literature. In his *Guide to Elementary Number Theory*, for instance, Underwood Dudley writes that "there seems to be no reason why square roots (mod p) should be related to square roots (mod q)" ([Dudley 2009], 61). Paul Garrett echoes this thought: "From a naive viewpoint, there is no reason why any such thing [as QR] should be true" ([Garrett 2005], 23). Audrey Yap says that the reciprocity theorem "seem[s] quite non-obvious and surprising" ([Yap 2011], 412). And Jonathan Rogawski calls QR "one of the deepest and most mysterious results of elementary number theory" ([Rogawski 2000], 35).

The number theorist Fernando Gouvêa goes further, offering QR as an exemplar of mathematicians' desire for explanatory proofs:

> It's often said that proofs serve as the criterion for truth in mathematics: we prove things in order to establish that they are true. This is certainly true, but it doesn't explain something else we do, namely, provide new proofs of old results. We already know those theorems are true, so in giving new proofs we are not seeking to establish that. What we are seeking is understanding. We want to know *why* the theorem is true, and a proof can (sometimes) tell us that.
>
> ...A case in point is the quadratic reciprocity theorem. First conjectured by Euler and Legendre and first proved by Gauss, it is a staple of elementary number theory courses. It relates the answers to two yes-no questions about two (distinct) odd prime numbers p and q:
>
> (1) *Does there exist a whole number n such that $n^2 \equiv q \pmod{p}$?*

(2) *Does there exist a whole number m such that $m^2 \equiv p \pmod{q}$?*

There is, a priori, no reason to expect these two questions to be related. ([Gouvêa 2015]; emphasis in original)

What exactly is so mysterious about the reciprocity phenomenon? Some of the above remarks suggest a general answer: on the basis of other evidence from number theory, we'd have expected two arbitrary primes not to "know anything about" each other's congruence properties. But the reciprocity theorem violates this expectation.

The Chinese Remainder Theorem is one fundamental fact that points in the opposite direction from QR. As Gouvêa writes, the CRT "strongly suggests that 'life mod p' and 'life mod q' are completely independent, since it tells us that for any choice of a and b one can always find x such that both $x \equiv a \pmod{p}$ and $x \equiv b \pmod{q}$" ([Gouvêa 2015]). Kiran Kedlaya tells a more detailed version of this story:

> [Quadratic reciprocity] is simple but also mysterious, because it violates our intuition that congruences modulo different primes should act independently. For instance, the Chinese remainder theorem asserts that (in a suitably precise sense) knowing that a random integer is odd or even does not prejudice it toward having any particular remainder modulo 3. ...The Chinese remainder theorem can be interpreted as saying that local [prime-related] phenomena at one point really are local, in that they do not influence local phenomena at another point. However, just as a particle physicist cannot explain the behavior of the universe by analyzing individual particles in isolation, [QR shows that] one cannot hope to understand the behavior of integers by looking at individual primes in isolation. ([Kedlaya 2008], 40)

It's worth unpacking this point. An immediate consequence of the CRT is that, given any primes $p_1, p_2, \ldots, p_n$ and any integers $a_1, a_2, \ldots, a_n$, there exists an integer x such that

$$\begin{aligned} x &\equiv a_1 \pmod{p_1} \\ x &\equiv a_2 \pmod{p_2} \\ &\vdots \\ x &\equiv a_n \pmod{p_n}. \end{aligned}$$

Kedlaya's example, then, involves the case $p_1 = 2$, $p_2 = 3$. Suppose we know that an integer x is odd, i.e. that $x \equiv 1 \pmod{2}$. Does this tell us anything about the residue class of $x \pmod 3$? No. By the CRT, the pairs of congruences

$$\begin{array}{l|l|l} x \equiv 1 \pmod 2 & x \equiv 1 \pmod 2 & x \equiv 1 \pmod 2 \\ x \equiv 0 \pmod 3, & x \equiv 1 \pmod 3, & x \equiv 2 \pmod 3 \end{array}$$

all admit of solutions. In this sense, then, the questions "what is $x \pmod p$?" and "what is $x \pmod q$?" are independent.

One can think of the CRT as a statement about "linear residues", and so it would be natural to expect quadratic residues to exhibit a similar kind of independence. That is, just

as $x \equiv q \pmod{p}$ and $x \equiv p \pmod{q}$ are always simultaneously solvable for any primes p and q, we might expect $x^2 \equiv q \pmod{p}$ and $x^2 \equiv p \pmod{q}$ to be likewise always solvable—or, at least, uncorrelated in their solvability. But QR tells us otherwise.

Unsurprisingly, number theorists haven't been content to linger in a state of puzzlement. Gauss's own prolonged engagement with QR was partly motivated by dissatisfaction with his early proofs, which he considered insufficiently natural. (More on this point in §4 and §5.1 below.) And many of Gauss's successors have followed in his footsteps. As Harold Edwards explains:

> The reason that the law of quadratic reciprocity has held such fascination for so many great mathematicians should be apparent. On the face of it there is absolutely no relation between the questions "is p a square mod λ?" and "is λ a square mod p?" yet here is a theorem which shows that they are practically the same question. ...[Many] great mathematicians have taken up the challenge presented by this theorem to find a natural proof or to find a more comprehensive "reciprocity" phenomenon of which this theorem is a special case. ([Edwards 1977], 177, quoted in [Tappenden 2008], 260)

The fact that QR has been proved in over two hundred ways in as many years tells us something about the importance of this challenge. Indeed, philosophers could hardly ask for a clearer example of the role of explanatory concerns in mathematics. If any theorem has ever needed explaining, and if mathematicians have ever devoted serious resources to the pursuit of an explanation, QR is surely such a case.

The next obvious question is, of course, "Does the reciprocity theorem have a good explanation? (And if so, what is it?)"[7] As I've indicated, the answer turns out not to

[7]A referee worries that these questions are predicated on a doubtful assumption about mathematical explanation: namely, that every mathematical fact has a unique correct explanation. As the referee points out, this claim seems false. Some mathematical facts seem not to have any explanation at all. And others seem to have more than one good explanation.

One could take this worry in two ways. The first interpretation is that I myself appear to be committed to the above assumptions. In fact, I'm not: I reject both the existence claim and the uniqueness claim. In particular, I don't assume at the outset that we have explained or can explain QR, or that QR must have a single correct explanation if it has any. As far as I can tell, I haven't said anything that presupposes or appears to endorse either claim.

The second interpretation is that the doubtful assumptions are made by the mathematicians whose views on QR are at issue. Perhaps the worry is that, if the disagreement about QR is based on mistaken endorsement of a false theory of explanation, then the disagreement ceases to look very interesting: rather than bothering ourselves about the details of these mathematicians' views, we should just note their incorrect starting point and pay the controversy no further mind. I have two replies. First, it's unclear that the QR debate would be uninteresting even if it partly depended on false assumptions. Second and more importantly, I see no reason to think that these mathematicians actually do endorse either the existence or uniqueness claim in general. As I show below, some of them *do* accept one kind of proof as explanatory and reject another kind as unexplanatory. But this isn't inherently problematic. The claim that one has an explanation for QR doesn't imply that every mathematical fact has an explanation. The claim that a certain purported explanation is unacceptable doesn't imply that QR can have only one correct explanation. And even the claim that QR *does* have only one correct explanation doesn't imply that this is true of every mathematical fact. So I don't think the disagreement about QR is based on an obvious mistake of the above sort. Neither

be straightforward. Some proofs have certainly been embraced as explanatory by some mathematicians, but there's significant disagreement about what constitutes an acceptable reason for QR's truth. Fortunately, number theorists' pain here is philosophers' gain. The fact that such disagreements occur, as well as the details of this particular case, can teach us some useful lessons about mathematical explanation.

I defer full discussion of these issues until §4.4, at which point I'll have presented some of the relevant proofs of QR. The next section sets the stage for the analysis of those proofs by introducing the notions of transparency and depth.

3 Transparency and depth as alternative explanatory styles

This part is the philosophical centerpiece of the paper. Its purpose is to describe two general types of explanatory proof, so I need to put quadratic reciprocity temporarily to one side (for the most part, although it appears once or twice as an example). I begin with transparency. My goals are to clarify the notion itself, to describe some major sources of transparency, and to justify the claim that transparent proofs are explanatory. The following part deals with depth in a similar way.

The ideas discussed below aren't totally new. Mathematicians, philosophers, psychologists, education theorists and others have often invoked something like them in their assessments of different types of proof. But it's frequently claimed that concepts like transparency and depth, while perhaps heuristically useful, are too imprecise to do serious theoretical work. (E.g.: "The notion of 'mathematical depth' is used quite frequently in informal contexts by mathematicians, but it is not clear that there is a coherent notion here" ([Urquhart 2015], 1); "perhaps most of the informal evaluative words used by mathematicians have meanings that are too diffuse to be explicated satisfactorily" ([Gowers 2008], 39) One objective of this section is to lay the groundwork for answering such complaints. Although I won't attempt anything like a formalization of my target notions—I don't think that kind of precision is possible here—I will try to provide a usefully clear and detailed account, supported by plenty of examples and easy to apply to other cases.

3.1 Transparency

3.1.1 What transparency is, and what it's not

Some proofs (as well as other kinds of explanations) are explanatory by virtue of identifying a compelling reason, and showing in a simple and clear way that the explanandum follows. Such an explanation renders the fact to be explained more vivid, more obvious, easier to grasp or reason about, or in general more comprehensible and cognitively tractable. I'll call explanations of this kind *transparent*. This section elaborates on transparency, its sources and its relationship to explanation.

version of the worry, then, appears to be a real problem.

Proofs that are transparent in my sense are often described as intuitive, natural, straightforward, simple, lucid, elegant, or clear. Often mathematicians use the word 'transparent' itself, with more or less the sense I'm giving it here. A few examples:

- "I have not attempted to state results under the weakest possible conditions; on the contrary, I have often imposed relatively strong conditions if that allows a simpler and more transparent proof" ([Severini 2005], xii).

- "Anyone already familiar with the subject matter of this book will be surprised that it covers so much ground in its 276 pages. The authors have achieved this partly by finding shortcuts to some complicated proofs in the literature (their short, transparent proof of the Borwein-Press variational principle is particularly welcome)..." ([Ioffe 2000], E60).

- "It is well known that the Riemann curvature tensor satisfies the two Bianchi identities... These have always seemed a bit mysterious, despite their short proofs from an abstract viewpoint. From the work of DeTurck on Ricci curvature, it became clear that the Bianchi identities are intimately related to the group of diffeomorphisms. This led to the following natural and conceptually transparent proof" ([Kazdan 1981], 341).

Here, "transparent" is allied with "simple" and "natural", and opposed to "complicated" and "mysterious". A highly non-transparent proof, then, is one that's poorly motivated, hard to follow, and that fails to give a straightforward reason why the theorem is true.[8]

Another important point also emerges from the above remarks: although transparent proofs are often relatively short, shortness is neither necessary nor sufficient for transparency. Some short proofs, for instance, are "gimmicky", in the sense that they rely on clever but opaque tricks.[9] Others, as in Kazdan's example, invoke powerful abstract machinery that makes little direct contact with the subject matter of the theorem. Neither type of proof is likely to satisfy the criteria for transparency.

Relatedly, although transparent proofs are characteristically simple, the kind of simplicity at issue shouldn't be confused with mere brevity, information-theoretic economy, or sparsity of logical structure. Rather, transparent proofs are simple in the sense of being *facile*—they make their results and the reasons for their truth easier to grasp. Possessing this sort of simplicity is compatible with being relatively long, with containing a lot of information, and with involving numerous inferential steps. (Cf. [Booß-Bavnbek & Wojciechowski 1993]: "We determine the topology of the space of self-adjoint Fredholm operators in Hilbert space. This was done [by Atiyah and Singer], but though transparent, their computations are quite long and complicated" (127).)

I've said that a transparent explanation makes its explanandum "more comprehensible and cognitively tractable". What does this mean? Importantly, the kind of benefit I'm

[8]To be clear, the meaning I'm giving to 'transparent' should be understood as somewhat stipulative. Although it's compatible with (and inspired by) the way many mathematicians use the term, I'm not claiming that my definition exactly captures what they all have in mind. Indeed, there's probably some variation in different authors' understanding of transparency.

[9]An infamous example is Don Zagier's "one-sentence proof" that every prime congruent to 1 modulo 4 is a sum of two squares ([Zagier 1990]).

interested in isn't a mere "sense of understanding", which feels satisfying but which may be illusory and ineffectual (cf. [Trout 2002]). Transparent explanations must *do* something (for an appropriate subject under appropriate conditions). They must actually improve our cognitive situation in a suitable respect, whether this is accompanied by an "aha!" experience or not.

In addition to the subjective "sense of understanding", however, there's plausibly a notion of understanding that involves a kind of objective epistemic or cognitive success, and transparency might seem to be related to understanding in this sense. I don't want to rule out this possibility; indeed, I think it's plausible, as I'll discuss further in §3.1.3 below. But I'd like to avoid wedding my account of transparency too closely to any particular theory about understanding. Philosophers and psychologists have had much to say on the subject lately, but even a consensus on the basics seems far off. (For instance, some authors maintain that understanding is reducible to knowledge ([Kelp 2016]), others go psychological and identify it with the storage of schemas in long-term memory [Inglis & Mejía-Ramos 2019], yet others connect it to the possession of appropriate models [Knuuttila & Merz 2009]), and so on.) It would be tendentious to make assumptions about which of these views is correct, but neither can I engage with the debate in detail here.

Relatively little of an analytical nature has been written about the notion of transparency. One of the few discussions I'm aware of is in [Raman-Sundström & Öhman 2016]. Raman-Sundström and Öhman define a transparent proof as one in which "[t]he structure of the argument is clear. In a proof that is strong on this criterion, it is easy [to] see 'what is going on'. In other words, the structure of the proof is natural for the particular argument, and there is no *deus ex machina* component. ...[I]f a proof is transparent, a reader with the appropriate background should be in an ideal position to grasp the ideas of the proof" (188).

I have no particular objection to this, although one might worry that it trades one vague intuitive notion for others that are even less clear. (What does it mean for "the structure of a proof" to be "natural for a particular argument", for example?) But I agree with Raman-Sundström and Öhman that a transparent proof must be easy to follow and free from unmotivated tricks.

3.1.2 Sources of transparency

Some methods of proof are particularly well-suited to giving clear presentations of compelling reasons. Geometric and combinatorial methods are examples; mathematicians are often happy if they can translate problems into these terms, because the associated arguments are likely to be revealing.

Richard Stanley's *Enumerative Combinatorics* features a nice discussion of combinatorial proof and transparency. One of his examples is as follows. Let n and k be positive integers, and let $f(k,n)$ denote the number of sequences $(X_1, X_2, \ldots, X_k)$ of subsets of $\{1, 2, \ldots, n\}$ whose intersection is empty. Stanley first shows, by a direct but ponderous proof involving exponential functions and binomial coefficients, that $f(k,n) = \left(2^k - 1\right)^n$. As he notes:

> This argument is a flagrant example of a noncombinatorial proof. The resulting answer is extremely simple despite the contortions involved to obtain it, and it cries out for a better understanding. ([Stanley 2012], 14)

He then sketches a counting argument that's much more satisfying. First, it's clear that $\left(2^k-1\right)^n$ is the number of sequences $(Z_1, Z_2, \ldots, Z_n)$ such that each Z_i is a proper subset of $\{1, 2, \ldots, k\}$. So we'll be done if we can find a bijection between this set of sequences and the one in the problem statement. In fact, this is simple to do: given a sequence $(Z_1, Z_2, \ldots, Z_n)$, define $(X_1, X_2, \ldots, X_k)$ by declaring that $i \in X_j$ just in case $j \in Z_i$. "This rule is just a precise way of saying the following: The element 1 can appear in any of the X_i's except all of them, so there are $2^k - 1$ choices for which of the X_i's contain 1; similarly there are $2^k - 1$ choices for which of the X_i's contain $2, 3, \ldots, n$, so there are $\left(2^k-1\right)^n$ choices in all. ...[The fact that this is a bijection] should be intuitively clear" (14).

Stanley concludes as follows:

> Not only is the preceding combinatorial proof much shorter than our previous proof, but it also makes the reason for the simple answer completely transparent. It is often the case, as occurred here, that the first proof to come to mind turns out to be laborious and inelegant, but that the final answer suggests a simpler combinatorial proof. (15)

Similar remarks apply to geometric proofs, and in general to arguments that lend themselves to visualization. This point probably doesn't need much belaboring: mathematicians since Euclid have enriched their proofs with diagrams, and it's a standard practice to gain intuition for a problem by casting it in a geometric light. Here, for instance, is Richard Swan on a "remarkable theorem" about sums of signs of matrix entries:

> The original [algebraic] proof of the theorem was elementary but very complicated. In attempting to simplify this proof, I found a more transparent proof based on the use of graph theory. One advantage of this approach is that complicated algebraic definitions can be replaced by much simpler geometric definitions merely by drawing a picture of the appropriate graph. ([Swan 1963], 367)

I don't mean to claim that combinatorial, geometric or pictorial proofs are generally highly transparent, or that other kinds of proof rarely are. My point is just that visualizability and the like often function as transparency-promoting features.

The reason for this is no mystery: brains and perceptual systems like ours are better at performing some kinds of operations than others. We can get quite a lot of immediate intuitive knowledge by looking at a graph, but this same isn't often true for a complicated algebraic expression. Similarly, reasoning about the sizes of finite sets is easier and more vivid than contemplating a formula like

$$f(k,n) = \sum_{i=0}^{n} \left(\begin{array}{c} n \\ i \end{array} \right) 2^{n-i} f(k-1, n-i)$$

(cf. [Stanley 2012], 14).

Of course, these advantages have their limits. A geometric proof centering on technical properties of schemes in 6-dimensional affine space may not be very transparent. In general, it matters whether or not other transparency-enhancing features are present—for example, whether the relevant sets are manageably small, whether the pictures involved are simple and clear, whether the geometry is low-dimensional and not too exotic, and so on.

There are other routes to transparency than the ones just mentioned. Another common scenario involves showing that a result depends in a straightforward way on antecedently well-understood facts. Here, for instance, is George Rousseau on his 1991 proof of QR:

> As is known, Euler's criterion and the theorems of Fermat and Wilson can be proved in a very simple manner by determining in two ways the product of the elements of a suitable finite abelian group... We show that the same is true for the quadratic reciprocity law. This law is thus seen to depend on nothing more mysterious than the Chinese Remainder Theorem, without need for special lemmas or auxiliary considerations which go beyond the sphere of simple congruences. ([Rousseau 1991], 423)

Rousseau's proof derives QR "in a very simple manner" from the Chinese Remainder Theorem, one of the most basic and familiar results in number theory. So although there's nothing vividly combinatorial, geometric or pictorial about the proof, it has a good claim to transparency: it makes QR more accessible by relating it to a well-known phenomenon.

3.1.3 Transparency and explanation

I've claimed that some proofs are explanatory by virtue of being transparent. I'll conclude this section by elaborating on this claim and saying some things in its defense.

We've already seen an example of a transparent proof that appears to be explanatory—Stanley's combinatorial proof was supposed to "[make] the reason for the [theorem] completely transparent". It's not too hard to find other cases like this. For instance, Nadler offers "a simple, transparent proof of Darboux's Theorem that we feel shows systematically and conceptually why the theorem is true" ([Nadler 2010], 174). And Di Bucchianico and Loeb present "a new transparent proof of Feinsilver's theorem" ([Di Bucchianico & Loeb 1998], 195). They note that the theorem has been proved at least twice before, but "[t]he merit of our proof is that it explains why the result is true" (205).

Of course, the fact that some transparent proofs are explanatory doesn't yet show that any proofs are explanatory *by virtue of* being transparent—and unfortunately for philosophers, mathematicians rarely venture such helpful metaphysical clarifications. Instead of appealing directly to mathematicians' testimony, then, let me briefly explain why I think this claim is reasonable.

Some philosophers have suggested that the property of being an explanation just is the property of producing understanding (in an appropriate way, for an appropriate subject). [10] As I mentioned earlier, the nature of understanding is controversial. But I think the idea is a plausible one.[11] Moreover, I think that having a compelling answer to a why-question—together with a simple and clear demonstration that the answer is a good one—gets one a pretty long way toward possessing understanding, whatever exactly the latter state turns out to be.

[10] See for instance [Grimm 2010], [Inglis & Mejía-Ramos 2019], [Khalifa 2013], [Wilkenfeld 2014], [Turri 2015], [Waskan et al. 2015].

[11] In a similar vein, I've argued in [D'Alessandro 2020] that some explanations aren't grounded in objective dependence relations, and that some type of epistemic or cognitive account of explanation may be the best alternative.

[Inglis & Mejía-Ramos 2019] is a clear recent defense of a "cognitivist" view of explanatory proof. As Inglis and Mejía-Ramos write:

> Our account suggests that the archetypal explanatory proof would have at least three properties. First, it would have features that make it easy, or at least as easy as possible, to select the information from sensory memory into working memory that is necessary for a successful processing stage. ...Second, it would have features that make it easier to coordinate the new knowledge contained in the proof with existing schemas retrieved from long-term memory, and therefore to reorganise the new and existing information into coherent new schemas. Finally, it would be likely to split the working memory load it gives to its readers between their visual and verbal/auditory channels so that the chances of their working memory capacity being exceeded during the schema-organisation process is minimised. (13)

It's characteristic of transparent proofs that they have the first of these features; since such proofs are straightforward and clear, their key components are easy to extract for further processing. They're likely to satisfy the second criterion for the same reason. Finally, as discussed earlier, transparent proofs often have visual (or visualizable) parts, so that they distribute the load to working memory in the way Inglis and Mejía-Ramos describe.

In summary, then, I think there are good reasons to take understanding-based accounts of explanation seriously. And anyone who accepts such an account should be amenable to the idea that transparency is an explanatory virtue. Transparent proofs are the kinds of things that reliably produce understanding—certainly on Inglis and Mejía-Ramos's cognitivist theory, but also, I think, on any reasonable analysis.

3.2 Depth

3.2.1 What depth is, and what it's not

Some other proofs are explanatory in virtue of deriving their result from a remote, relatively fundamental source. If the byword of transparency is "presenting reasons clearly", that of depth is "exposing hidden reasons".

While philosophers have said little about transparency, mathematical depth and its relationship to explanation are comparatively well-studied.[12] Interestingly, Jeremy Gray traces the concept to the *Disquisitiones* and its reception: "the word 'depth' came into mathematics in the early nineteenth century through the work of Gauss on number theory, and... Gauss's successors agreed that this part of Gauss's work was deep" ([Gray 2015], 177).

Philosophers and others have characterized depth in a few substantially different ways. (More on some of these below.) A useful starting point for my purposes is G.H. Hardy's *A Mathematician's Apology*, which outlines something like the concept I'm concerned with:

> It seems that mathematical ideas are arranged somehow in strata, the ideas in each stratum being linked by a complex of relations both among themselves and

[12]Thanks in part to a 2015 special issue of *Philosophia Mathematica* on mathematical depth (Vol. 23, No. 2).

> with those above and below. The lower the stratum, the deeper (and in general the more difficult) the idea. Thus the idea of an 'irrational' is deeper than that of an integer; and Pythagoras's theorem is, for that reason, deeper than Euclid's [i.e. the infinitude of primes]. ([Hardy 2012], 110)

So a theorem can be deep, and one way to compare the depth of two theorems is in terms of the depth of their constituent concepts. The notion of depth also applies to proofs:

> [We may be able to] recognize and prove, for example, some property of the integers, without any knowledge of the contents of lower strata. Thus we proved Euclid's theorem by consideration of properties of integers only. But there are also many theorems about integers which we cannot appreciate properly, and still less prove, without digging deeper and considering what happens below. ...[For instance, how are the primes distributed?] We can answer [this question], with rather surprising accuracy, but only by boring much deeper, leaving the integers above us for a while, and using the most powerful weapons of the modern theory of functions.([Hardy 2012], 110-11)

So a deep proof, according to Hardy, is one that uses deep ideas. (More specifically, it uses ideas that are deep relative to the associated theorem or problem statement.) Depth in this sense can be contrasted with elementariness—where deep proofs appeal to remote and sophisticated concepts, elementary proofs stick to the same family of ideas as their theorems. This opposition is sometimes invoked explicitly: "A more precise form of [Mordell's finite basis theorem] was given by Weil... who gives an elementary as well as a deep proof of a far-reaching generalization" ([Cassels 1950], 244).

I think Hardy's characterization is on the right track, but the notion of depth I'm interested in (and the one I think most mathematicians have in mind) has a bit more content. Depth in my target sense is a salutary feature, and one that's closely connected to explanation. So not just any proof that happens to journey into lower strata should count as deep. Making such a detour doesn't inherently promote understanding; greater generality and exotic cross-connections can very well obscure, distract and complicate rather than enlighten. A deep proof, on the other hand, should be revealing. It should expose non-obvious reasons, and it should help show how its theorem fits into a larger system of concepts and results. This richer notion of depth is opposed to gimmicks, shortcuts, artificiality, naivete, and other forms of missing the point. As Jeremy Gray writes, "what is deep in the Gaussian sense has to be fruitful, fundamental, explanatory, and important because it is necessarily organizing" ([Gray 2015], 192).

A nice illustration of these points comes from Alexander Givental's "The Pythagorean Theorem: What Is It About?". Givental begins with the theorem's most popular proof (pictured in Figure 1), which shows by rearrangement of triangles that the sum of the squares A and B has the same area as the square C. Givental writes:

> [This proof] is very convincing indeed. Yet it pictures the whole issue as a cut-and-paste puzzle and leaves us with a feeling of disproportion: one of the most fundamental facts of nature is due to an ingenious tiling trick. The vast majority of other proofs are similar in nature. ([Givental 2006], 261)

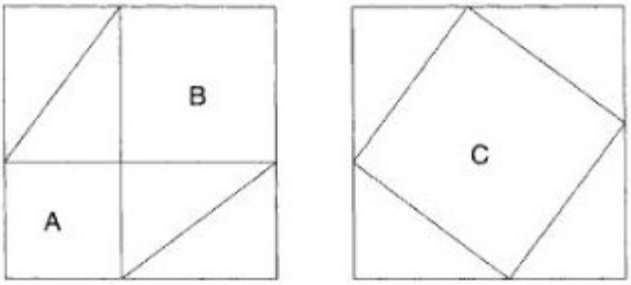

Figure 1: "Tiling" proof of the Pythagorean theorem, from [Givental 2006].

Indeed, the tiling proof is highly transparent—it makes a simple reason for the theorem's truth immediately clear. But there are other possible reasons to be dissatisfied. One is that the proof fails to generalize. In fact, we know (and Euclid knew) that "$A + B = C$ for the areas A, B and C of *similar* figures of *any* shape built on the sides of a right triangle. The Pythagorean theorem is clearly the special case where the shape is the square. However, the tiling argument looks hopeless when the shape is arbitrary" ([Givental 2006], 261).

This gives a clue about what a better, and deeper, proof might look like: it should exploit the notion of similarity, and it should thereby exhibit the Pythagorean theorem as an instance of a more general phenomenon. Here's such a proof. Suppose we have a right triangle with area c. If we now draw the height from the right angle to the hypotenuse, we've created two smaller triangles—similar to each other and to the original—with areas a and b such that $a + b = c$. Next, suppose we construct similar shapes A, B, C of any other kind on the hypotenuses of the three triangles. The area of each of these new shapes will be a multiple k of the area of the corresponding triangle; that is, $A = ka$, $B = kb$, $C = kc$. (This is from "the intuitively obvious fact that the ratio k of the areas of two figures built on the same segment depends only on their shapes and does not change under their simultaneous rescaling" ([Givental 2006], 262). Since $a + b = c$, then, it follows that $ka + kb = kc$, or in other words $A + B = C$. The Pythagorean theorem is the special case where A, B, C are squares. This simple argument isn't new, of course, although it's less common than the square-tiling proof—as Givental notes, this is essentially how Euclid proves *Elements* VI.31.

Givental claims that, through his preferred proof,

> the true nature of the Pythagorean theorem as a statement based on similarity is revealed. What is "similarity" after all? From the abstract point of view it is *conformal isometry*: isometry of a metric space with itself equipped with a rescaled metric. This explains why the Pythagorean theorem is a genuinely *Euclidean* phenomenon (and not only in the historical sense of the word): among all Riemannian metrics of constant curvature, only the Euclidean one admits nontrivial conformal isometries. (263-4)

Is Givental's proof a step forward in depth? It seems so. Although the proof isn't especially profound in absolute terms, its foil—the square-tiling textbook proof—is about as shallow as possible, being a pure picture proof with a minimum of technical content. The similarity proof digs at least one stratum deeper (and probably several more, if we think of similarity as conformal isometry). Indeed, summarizing [Givental 2006] for *Mathematical Reviews*,

Edmund Harriss writes:

> This paper considers the Pythagorean Theorem, starting by rejecting the "cut and paste" nature of the traditional school proof as it is based on a tiling trick. It therefore reveals little about what the theorem means. The paper then gives a deeper proof using the nontrivial conformal isometries that only exist in Euclidean geometry. ([Harriss 2006])

Harriss's remarks highlight a feature of depth mentioned earlier: a deep proof frequently "shows what its theorem means", by situating it in a broader theoretical context and thus displaying its relations to other important facts. Here, by deriving the Pythagorean theorem from isometry considerations, we get a glimpse of its place in the universe of Riemannian geometries. (Manya Raman-Sundström agrees: she writes that the similarity proof "get[s] at... what the theorem is about", and "captures very well the essence of why the theorem is true" ([Raman-Sundström 2016], 275.)

A final remark: depth is sometimes associated with qualities like complexity, difficulty and impurity ([Shanks 1978], 64; [Gray 2015]). Although many deep proofs may have these properties, I doubt that they're essential to depth in general. Impurity, for instance, is specifically about crossing over divisions between subject matters (as in an analytic proof of the Fundamental Theorem of Algebra). But a deep proof might instead draw on more sophisticated parts of the same subject. (As [Arana 2015] points out, the deep original proof of Szemerédi's theorem "has widely been judged pure... as a combinatorial proof of a combinatorial theorem" (169).) Likewise, it's not clear that depth requires difficulty and complexity. Perhaps a deep proof must be initially non-obvious, to the extent that its ideas are far removed from the apparent topic of the theorem. But such a proof might eventually be viewed as natural and elegant once the relevant mathematics is better understood. (This raises a question about the relationship between depth and transparency, which I'll examine more closely in §3.3.)

3.2.2 Comparisons with other work on depth

I now want to consider some recent philosophical accounts of depth and their relationship to the picture presented here. As I've indicated, my notion of depth is fairly close to Jeremy Gray's; his [Gray 2015] focuses in particular on Gauss, quadratic reciprocity and subsequent European attitudes toward deep mathematics. For Gray, "a deep property in mathematics is one that is, or has been, hidden, has significant organizational power, and shapes, guides, and helps explain a large body of ideas" (193). A deep proof, then, seems to be one that exposes and makes essential use of deep properties.

Gray's main example is in fact Gauss's second proof of QR, based on the theory of quadratic forms. It's certainly true that Gauss's work on quadratic forms was deep—the section of the *Disquisitiones* on this subject takes up over half the book, and the ideas it contains are highly important and original—but I'm less sure about the depth of the second proof itself, even compared to Gauss's other proofs of QR. As far as I know, Gauss himself never marked the second proof as especially deep or otherwise valuable. Nor, in general, have subsequent number theorists done so.

I think it's important to distinguish—as Gray's paper could have done more carefully—between a deep proof and a proof that's merely based on deep mathematics. For instance, ZFC set theory should count as deep by any reasonable standard, and one can use ZFC (along with suitable definitions) to prove that $3+4=7$. But this isn't a deep *proof*. It doesn't shed light on any hidden aspect of the sum of 3 and 4; it doesn't "show what the theorem means"; it doesn't reveal important connections with other concepts and results.[13] Gauss's second proof is probably a more complex case, but I suspect the same diagnosis applies, at least to some degree. The theory of quadratic forms is deep, but it doesn't supply a particularly deep proof of QR.

If any of Gauss's original proofs is worthy of that title, it's probably his sixth, using "Gauss sums" (certain sums of roots of unity). In Gauss's own words: "the sixth proof calls upon a completely different and most subtle subtle principle, and gives a new example of the wonderful connection between arithmetic truths that at first glance seem to lie very far from one another" ([Gray 2018], 333). This language is evocative of depth, and the same sentiment appears in modern sources. In their *Proofs from the Book*—chronicling the best known demonstrations of various noteworthy results—Aigner and Zeigler write:

> With so many proofs [of QR] present the question which of them belongs in the Book can have no easy answer. Is it the shortest, the most unexpected, or should one look for the proof that had the greatest potential for generalizations to other and deeper reciprocity laws? We have chosen two proofs (based on Gauss' third and sixth proofs), of which the first may be the simplest and most pleasing, while the other is the starting point for fundamental results in more general structures. ([Aigner & Zeigler 2010], 23)

(§5 below presents the third proof and its claim to transparency.) I won't discuss the sixth proof in detail, since algebraic number theory has gone well beyond this "starting point" and I want to sketch a more modern type of deep proof. But if one is looking for depth in Gauss's original arguments, I believe the sixth rather than the second is the most compelling candidate.

Marc Lange's account of depth differs more substantially from mine. He holds that, "in at least some cases, one proof of a given theorem is deeper than another by virtue of supplying a deeper explanation of the theorem" ([Lange 2015a], 196). A deeper explanation, in turn, "answer[s] not only the why questions answered by the shallower explanations, but also some more why questions besides—especially why questions that were prompted but left unanswered by the shallower explanations" (203).

Lange's examples are characteristically interesting and well-chosen, and I agree with his claim that deep proofs are explanatory. But it's hard to accept the details of Lange's view. One issue is that, taken at face value, the account seems to involve some dubious assumptions about why-questions and answers. (Does every proof answer a definite collection of why-questions? If so, which ones, and how do we tell? Is it plausible that different proofs of a given theorem often stand in strict inclusion relations with respect to the sets of why-questions they answer? At the very least, Lange needs to say more here.) Perhaps the more

[13] See [D'Alessandro 2018a] for a sustained argument that the reduction of arithmetic to set theory isn't explanatory.

fundamental point, though, is that depth doesn't seem to be a merely quantitative property. A proof isn't deep solely on account of *how many* things it explains; it also matters, and probably matters more, *what kind* of explanation it offers. As we've seen, characterizations of depth in the mathematical literature almost invariably invoke remoteness, hiddenness, or something like Hardy's "strata" metaphor. I don't believe one can pin down these qualities just by counting answers to why-questions. (Perhaps it's true that deeper proofs tend to answer more why-questions on average. But this is a symptom of depth, on my view, rather than a defining feature.)

Some other authors use the term 'deep' to denote a quite different property than the one I'm interested in. For example, [Urquhart 2015] proposes to analyze depth in terms of the size and structure of proof trees; this is a fine project, but it doesn't reflect a notion of depth that's closely related to explanation or other epistemic goods. Other authors have discussed deep theorems without saying much about proofs (e.g. [Arana 2015], [Stillwell 2015]), and it's not clear whether or how these accounts can be generalized.

3.3 The relationship between transparency and depth

Finally, I want to examine the relationship between transparency and depth. Doing so is interesting for theoretical purposes, but more to the point, it will help set the stage for my later analysis of the disagreement over proofs of QR.

It's useful to start with the obvious analogy, which turns out to be a pretty good one. Where lakes and other bodies of water are concerned, transparency and depth aren't inherently incompatible; some very deep reservoirs are also very clear. (Lake Baikal in Russia is one example. Incidentally, this is why it was chosen to host the GVD underwater neutrino detector.[14]) Nevertheless, the two properties tend to work against each other. In particular, greater depth means more room for occluding material; so in the absence of some kind of special conditions, a deep pool won't be easy to see through.

Much the same goes for proofs. There's nothing inherently impossible about a proof that's both transparent and deep, but the combination is rare. Since deep proofs draw on distant sources in non-obvious ways, they tend to be longer, more complex, and hence less facile. (Here, too, greater depth means more room for "occluding material", in the form of unwieldy machinery and obscure argumentative maneuvers.) On the other hand, since transparent proofs make their results clear and intuitive, they're typically simpler, more elementary, and hence less profound.

Are there any familiar examples of deep but transparent proofs? Perhaps Gödel's first incompleteness theorem provides a candidate. The main ideas of the standard proofs (arithmetization of syntax, diagonalization, etc.) are profound, but the structure of the proofs is easy enough to grasp and leaves no confusion about why the theorem is true. ("[T]he method of Gödel's proof explicitly produces a particular sentence that is neither provable nor refutable in F; the 'undecidable' statement can be found mechanically from a specification of F. The sentence in question is a relatively simple statement of number theory, a purely universal arithmetical sentence" ([Raatikainen 2018]).)

[14] "The choice of this lake—the largest and deepest freshwater reservoir in the world—was determined by the high transparency of its water, its depth, and the ice cover that allows the installation of deep-water equipment during two months in winter" ([Domogatsky 2015], 23).

The analogy between pools and proofs is useful, but it has its limits. In the world of reservoirs, shallowness promotes transparency: other things being equal, a less deep pool will be easier to see through. Not so with mathematics. Shallow proofs don't have any particular tendency to be cognitively tractable. As Gauss's first proof of QR illustrates (see §4 below), a proof can be entirely elementary yet unilluminating and hard to follow.

Transparency and depth represent alternative explanatory styles, and I've claimed that some mathematicians prefer one style over the other. Why is this? What reasons might one have for valuing transparent proofs and deprecating deep proofs, or vice versa? Probably these differences are partly a matter of general mathematical taste, and partly due to variations in goals, prior knowledge and other circumstantial features.

As for the former point, some of the best mathematicians have famously been system-builders who strive to get to the bottom of things, while others are problem-solvers who prove difficult theorems with deftness and insight. Alexander Grothendieck seems to have taken himself and Jean-Pierre Serre as representatives of these two styles, as Colin McLarty describes:

> Grothendieck says Serre generally uses the hammer and chisel. He calls Serre "Super Yang" against his own "Yin"—but not at all in the sense of being heavy-handed—rather Serre is "the incarnation of elegance"... That is the difference. Serre cuts elegantly to an answer. Grothendieck creates truly massive multi-volume books with numerous coauthors, offering set-theoretically vast yet conceptually simple mathematical systems adapted to express the heart of each matter and to dissolve the problems. ([McLarty 2008], 302-3)

Timothy Gowers writes in a similar vein about "two cultures of mathematics", consisting respectively of problem-solvers and theory-builders ([Gowers 2000]). The first kind of mathematics revolves around intuitive heuristic principles "that allow proofs to be condensed in the mind, and therefore more easily memorized and more easily transmitted to others" (72)—i.e., practices associated with transparency—while the second kind prizes "deep theorems of great generality" (72) which "suddenly [place] a large number of existing results in their proper context" (68)—i.e., practices associated with depth. (Gowers places Erdős and his combinatorialist followers in the first camp, and much of abstract, category-powered contemporary mathematics in the second.)

Of course, it's also possible to prefer one proof style over the other for situational reasons, apart from one's general tastes or the values of one's mathematical culture. Different proofs may be appropriate for students as opposed to seasoned mathematicians, for non-specialists as opposed to specialists, for developing intuition and a problem-solving repertoire as opposed to assimilating a network of theory.

Transparent proofs can be very useful for beginners in a subject, since they promote understanding without requiring extensive prior knowledge or extreme cognitive effort. Hence, as Gila Hanna notes, teachers and education theorists are mostly interested in mathematical explanations to the extent that they're transparent—they "consider a proof to be explanatory when it helps convey mathematical insights to an audience in a manner that is pedagogically appropriate. This latter view brings cognitive factors very much into play" ([Hanna 2018], 3).

It's tempting to suppose that most experts, by contrast, must prefer deep proofs. (And one might take this to show that deep proofs represent a more mature, informed, and hence in some sense better type of explanation.) But the first of these claims is doubtful, and the second wouldn't obviously follow even if the first were true. Achieving expertise on a topic may make some deep proofs more accessible and useful, if one happens to be interested in the relationship between the topic and the relevant deeper mathematics. But gaining the ability to appreciate deep proofs needn't obviate or negate the cognitive benefits associated with transparency. Having command of a simple, intuitive explanation is useful for specialists too, and a more tractable proof is sometimes desired even when a deep one is already available: "Since Weil's proof of the Riemann conjecture for curves was very deep... Stepanov's elementary proof for the case of hyperelliptic curves was quite welcome" ([Lemmermeyer 2000], 345).

Interestingly, although Gauss is often (and rightly) viewed as a deep mathematician *par excellence*, he himself seems to have valued simplicity, elegance and directness at least as much as depth. Looking back in 1808 on his first six proofs of quadratic reciprocity, Gauss writes: "Although these proofs leave nothing to be desired as regards rigor, they are derived from sources much too remote, except perhaps the first, which however proceeds with laborious arguments and is overloaded with extended operations. I do not hesitate to say that till now a *natural* proof has not been produced" ([Smith 1959], 113; emphasis in original). Derivation from remote sources is characteristic of depth, but Gauss apparently didn't see this trait as a proof's highest virtue. His longstanding goal was to find an elementary argument which wouldn't be too hard to follow—a transparent proof, in other words, rather than a deep one. We should hesitate to rank transparency as a lesser explanatory virtue if Gauss didn't do so himself.

I'll have more to say later about the relationship between transparent and deep proofs of QR (and the reasons for favoring one over the other in this particular case). My goal here has been to note a few general facts: that transparency and depth are in substantial tension with one another but not in principle mutually exclusive, that there are a variety of individual, sociological and contextual reasons why a mathematician might prefer one type of proof over the other, and that we should resist the temptation to think of depth as a superior kind of explanatory quality.

This concludes my general discussion of transparency and depth. The next part will put the theory to work. I'll consider three interesting and illustrative proofs of QR: one by induction (due to Gauss), one by counting lattice points (due to Gauss and Eisenstein), and one from algebraic number theory (due to Hilbert). The first proof has been widely judged unexplanatory, while mathematicians have disagreed about whether QR is best explained by proofs like the second or proofs like the third.

The theoretical tools developed here can explain this disagreement. As I'll show, the lattice-point proof is transparent but not deep, while the algebraic proof is deep but not transparent. Mathematicians who favor one sort of explanatory style are likely to prefer the corresponding type of proof.

4 Three proofs of quadratic reciprocity

4.1 An unexplanatory proof: Gauss, by induction

Gauss discovered his first proof of the reciprocity theorem at the age of 19, after a year of "torment" and "greatest effort". The proof uses an inductive argument that's quite different from Legendre's earlier attempt (which, at any rate, Gauss was unfamiliar with at the time). It appears as the first of two proofs of QR in the *Disquisitiones*, comprising articles 135-144 ([Gauss 1966], 92-97).

Although he was delighted by QR itself and relieved to have found a conclusive argument, Gauss found his first proof very unsatisfying. As discussed above, Gauss described it as "proceed[ing] with laborious arguments" and "overloaded with extended operations", and he claimed that a "natural proof" had not yet been found ([Smith 1959], 113). (As we'll see in the next section, Gauss considered his third proof to be the only one worthy of that title.)

Gauss's successors have generally echoed his negative assessment. Jeremy Gray describes the argument as "a bull-at-a-gate proof that splits the problem into eight cases" ([Gray 2015], 179). Others have called it "awkward" ([Cox 2013], 58), "elementary but quite complicated" ([Frei 1994], 76), "unreadable" ([Rowe 1988], 15), and "very repulsive to any but the most laborious students" ([Smith 1894], 59).

I won't present the proof in full here, but the idea is as follows (with some slight changes to Gauss's notation). The base case of QR involves the primes $p = 3$ and $q = 5$, and it's easily checked—neither is a quadratic residue modulo the other, in conformity with the fact that $5 \equiv 1 \pmod 4$. We can then assume that QR holds for all primes less than some prime T. Does QR necessarily hold for T itself? If not, then one of the following counterexamples has to exist. (Here p is some prime less than T, and Gauss's notations aRb and aNb mean "a is a quadratic residue modulo b" and "a is a nonresidue modulo b", respectively.)

1. p and T are both of the form $4n + 1$, and we have TNp even though pRT.
2. p and T are both of the form $4n + 1$, and we have pNT even though TRp.
3. p and T are both of the form $4n + 3$, and we have both pRT and TRp.
4. p and T are both of the form $4n + 3$, and we have both pNT and TNp.
5. p is of the form $4n + 3$ and T is of the form $4n + 1$, and we have TNp even though pRT.
6. p is of the form $4n + 3$ and T is of the form $4n + 1$, and we have pNT even though TRp.
7. p is of the form $4n + 1$ and T is of the form $4n + 3$, and we have TNp even though pRT.
8. p is of the form $4n + 1$ and T is of the form $4n + 3$, and we have pNT even though TRp.

"If it can be shown that none of these eight cases can occur," Gauss writes, "it will be certain likewise that the truth of the fundamental theorem is circumscribed by no limits" ([Gauss 1966], 94). This turns out to require substantial argumentation, as Gauss splits each of 1-8 into further subcases requiring separate analysis. The reasoning isn't unmanageably difficult, but it is mostly laborious and unenlightening. For instance, the first case proceeds as follows:

Suppose that p and T are of the form $4n+1$ and that $p \equiv c^2 \pmod{T}$. We can assume that c is even and less than T. There are then two subcases to consider: either c is divisible by p or not.

- First, suppose c isn't divisible by p. Write $c^2 = p + dT$. We can see that d has to be of the form $4n+3$, that $d < T$, and that $p \nmid d$. Also we have $c^2 \equiv p \pmod{d}$, and so$p$ is a residue modulo d. Since p and d are less than T, the induction hypothesis implies that the reverse is also true, i.e. d is a residue modulo p. But dT is a residue modulo p by assumption. Since the product of a residue and a nonresidue would be a nonresidue, T must then be a residue modulo p.

- Second, suppose c is divisible by p. Write $c = gp$ and $c^2 = p + hpT$. Again, we can see that h has to be of the form $4n+3$, and that $\gcd(h, p) = \gcd(h, g^2) = 1$. A little algebra shows that $pg^2 = 1 + hT$, so $pg^2 \equiv 1 \pmod{h}$ and hence pg^2 is a residue modulo h. It follows that p itself is a residue modulo h, and so by the induction hypothesis, h is conversely a residue modulo p. Finally, again from the fact that $pg^2 = 1 + hT$, we have $-hT \equiv 1 \pmod{p}$ and so $-hT$ is a residue modulo p. By the same reasoning as the previous case, we can conclude that $-T$ is a residue modulo p. Since T is of the form $4n+1$, however, this implies that T itself is a residue modulo p.

This shows that counterexamples of type 1 can't occur, so we can conclude that TRp whenever pRT. □

The flavor of the reasoning here is characteristic of the whole proof, although some of the other seven parts are more complicated. (For instance, case 2 splits into two subcases, the first of which divides into four further subsubcases.) It's not hard to see why Gauss immediately started looking for an alternative argument.

There's no doubt that Gauss's first proof is cumbersome and aesthetically unappealing, but why think that it's unexplanatory? First, it's worth noting that Gauss himself seems to have held this view. As we saw, he deprecated his first few proofs of QR as not "natural". Gauss's use of the term suggests that, for him, a natural proof is not just one that's short or easy (for instance). Rather, naturalness has to do with finding appropriate reasons and presenting them clearly. These criteria are at least closely allied with explanatoriness. So it's reasonable to read Gauss as denying that his first proof is explanatory.

In any case, there are other reasons to consider the proof unexplanatory. For one, it's a proof by induction, and it's been argued that such proofs generally don't explain their results (cf. [Lange 2009]). What's more, the proof requires separate analysis of at least a dozen distinct subcases. As Mark Colyvan writes, "[proofs of this kind] lack unity. There are often different reasons offered in the different cases and it looks like the theorem itself holds merely by accident. What we would like is a proof that offers the same reason in each case; that would provide an explanation of the theorem in question" ([Colyvan 2012], 81).

Finally, Gauss's inductive proof has neither of the explanation-making features described in §3 above. A transparent proof is one that makes its result more cognitively accessible or tractable; Gauss's case analysis doesn't do this. Showing that QR can't have any of eight varieties of counterexample makes it no more clear or obvious why the theorem should be true. One can grasp the proof completely without getting a better understanding of the relationship between quadratic residues and the reciprocity phenomenon.

A deep proof, on the other hand, derives its result from remote sources, and thus shows how the result fits into a wider theoretical framework. Gauss's first proof doesn't do this either. The concepts, facts and techniques appearing in the argument, like QR itself, all live on the plane of elementary number theory. This feature of the early proofs, in fact, was another source of Gauss's dissatisfaction: he was interested in higher reciprocity laws, and he hoped to find a route to QR that would point the way to greater generality. As he wrote in 1818:

> From 1805 onwards I have investigated the theory of cubic and biquadratic residues... Theorems were found by induction... which had a wonderful analogy with the theorems for quadratic residues. On the other hand, for a long time all attempts at complete proofs have been futile. This was the motive for endeavoring to add yet more proofs to those already known for quadratic residues, in the hope that of the many different methods given, one or the other would contribute to the illumination of the related arguments [for higher reciprocity]. ([Gauss 1863], 50, quoted and translated in [Cox 2013], 78).

So the first proof is neither deep nor transparent, and hence isn't explanatory in either of these ways. (Incidentally, the argument would be even more unwieldy if not for Gauss's new congruence notation, which lets him write '$a \equiv b \pmod{c}$' in place of '$a = b + cx$ for some x'. See [Yap 2011] for an account of the fruitfulness of this notation in the context of Gauss's first proof.)

4.2 A transparent proof: Gauss-Eisenstein, by counting lattice points

In this and the next section, I turn to two explanatory proofs of QR. Examining these cases serves a couple purposes. First, it will lend further clarity to the distinction between transparency and depth, introduced in §3 above. Second, it will support my main claim. The claim, recall, is that the ongoing disagreement about how best to explain QR is largely based on a difference of explanatory preferences. Some mathematicians favor transparent proofs, while others favor deep proofs. Since different proofs of QR turn out to exemplify depth and transparency to different degrees, it's no surprise that the explanation question has proven contentious.

In this section I present a transparent (though not very deep) proof. While it's probably correct that "[t]here is no truly simple explanation for why quadratic reciprocity works" ([Narins 2001], 361), this one comes respectably close. On account of its appealing intuitiveness and non-technical nature, it's widely used in elementary number theory textbooks and other beginner-friendly sources. The argument is essentially due to Gauss—in which form

it's counted as his third proof of QR[15]—but the presentation here, as is now usual, includes important refinements of Eisenstein's.

We've seen that Gauss was unhappy with his initial proofs, which he found insufficiently natural. In 1807, six years after the publication of the *Disquisitiones*, he finally worked out a more satisfying argument. This third proof has been "considered by Gauss and many others to be the most direct and elegant of his eight demonstrations" ([Smith 1959], 112). Aigner and Zeigler's *Proofs from the Book* calls it "the simplest and most pleasing" of the Gaussian proofs ([Aigner & Zeigler 2010], 23).

We've already seen some of Gauss's remarks on the third proof and its relationship to the others. Here's a more complete version of the passage quoted previously; Gauss's general concern with transparency is clear here, as is his belief that the third proof alone succeeds on this score:

> [Arithmetical truths] are frequently of such a nature that they may be arrived at by many distinct paths and that the first paths to be discovered are not always the shortest. It is therefore a great pleasure after one has fruitlessly pondered over a truth and has later been able to prove it in a round-about way to find at last the simplest and most natural way to its proof. ...For a whole year [the reciprocity theorem] tormented me and absorbed my greatest efforts until at last I obtained a proof given in the fourth section of [the *Disquisitiones*]. Later I ran across three other proofs which were built on entirely different principles. One of these I have already given... the others, which do not compare with it in elegance, I have reserved for future publication. Although these proofs leave nothing to be desired as regards rigor, they are derived from sources much too remote, except perhaps the first, which however proceeds with laborious arguments and is overloaded with extended operations. I do not hesitate to say that till now a *natural* proof has not been produced. I leave it to the authorities to judge whether the following proof which I have recently been fortunate enough to discover deserves this description. ([Smith 1959], 113; emphasis in original).

Gauss's version of the third proof is purely arithmetical. It revolves around Gauss's lemma[16] and calculations relating the floor function $\lfloor x \rfloor$ to quadratic residues. In 1844 Eisenstein published a new proof, showing how Gauss's argument can be understood in terms of counting lattice points inside certain plane figures. Eisenstein's version is generally viewed as a transparency-enhancing improvement. For instance, Lauenbacher and Pengelley call it a "particularly beautiful and economical adaptation of Gauss's third proof" [Lauenbacher & Pengelley 1994a] 29). And it's been claimed that "Eisenstein translated Gauss's arithmetic language of the third and the fifth proof in a very intuitive way into the language of geometry" ([Baumgart 2015], 103). Brett Tangedal is even more effusive:

[15]The argument was Gauss's third published proof, but the fifth proof he discovered. The conventional numbering of Gauss's proofs is based on the order of publication.

[16]Gauss's lemma is the following statement, which is used in quite a few proofs of quadratic reciprocity. Let p be an odd prime and let a be an integer coprime to p. Consider the residues $a \pmod p$, $2a \pmod p$, $3a \pmod p$, $\ldots$, $\left(\frac{p-1}{2}\right) a \pmod p$. (These are all distinct.) Suppose that n of these residues are greater than $p/2$. Then, writing $\left(\frac{a}{p}\right)$ for the Legendre symbol, we have $\left(\frac{a}{p}\right) = (-1)^n$.

> [The lattice-point argument is] a remarkably direct and insightful proof of the classical law of quadratic reciprocity. ...Eisenstein's proof simplifies and improves upon Gauss's third proof at every step and truly deserves to replace the standard proof in the textbooks. ([Tangedal 2000], 130)

Here, then, is the proof. Let ABC be a right triangle with $|AB| = p$ and $|BC| = q$, where p and q are distinct odd primes. As in the diagram below, draw ABC in the Cartesian plane, with the vertex A at the origin and the side AB lying on the x-axis. Then we can ask: how many lattice points with even x-coordinate lie strictly inside ABC? (A *lattice point* is a point whose coordinates are both integers.)

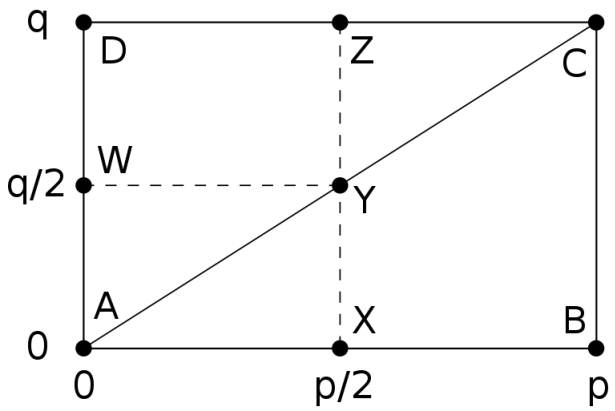

Figure 2: The triangle ABC.

This question has a simple answer, which is easy to see after examining a particular case. Consider the set of points inside ABC with x-coordinate 2; call this set L_2. A point $(2, y)$ belongs to L_2 just in case y is an integer strictly between 0 and $2q/p$. Clearly there are $\lfloor 2q/p \rfloor$ such integers (where $\lfloor x \rfloor$ denotes the floor function). So L_2 contains $\lfloor 2q/p \rfloor$ points. Extending this reasoning, we see that the number of lattice points in the interior of ABC is

$$\left\lfloor \frac{2q}{p} \right\rfloor + \left\lfloor \frac{4q}{p} \right\rfloor + \cdots + \left\lfloor \frac{(p-1)\,q}{p} \right\rfloor = \sum_u \left\lfloor \frac{uq}{p} \right\rfloor,$$

where u ranges over the even integers $2, 4, \ldots, p-1$.

What does this have to do with QR? As it turns out, the quantity $\sum_u \left\lfloor \frac{uq}{p} \right\rfloor$ is related to the Legendre symbol by the following result, known as *Eisenstein's lemma*:

$$\left(\frac{q}{p}\right) = (-1)^{\sum_u \left\lfloor \frac{uq}{p} \right\rfloor}. \tag{5}$$

The idea of the rest of the proof, then, is to interpret QR as a statement about counting lattice points. Specifically, our goal will be to show that the exponent $\left(\frac{p-1}{2}\right)\left(\frac{q-1}{2}\right)$ appearing in the reciprocity theorem is equal to the number of lattice points inside the rectangle $AXYW$.

For brevity, let an *even point* be a lattice point with even x-coordinate (and similarly for an *odd point*). If we consider even points in the interior of the rectangle $XBCZ$, we find

several "columns" each consisting of $q-1$ such points. Some of these points are contained in the lower region $XBCY$ and the rest lie in the upper region YZC. Since $q-1$ is even, these two sets of points have the same parity. (Note that there are no points lying on AC, since this would require uq/p to be an integer.) The following diagram illustrates this for the case $p = 11$, $q = 7$:

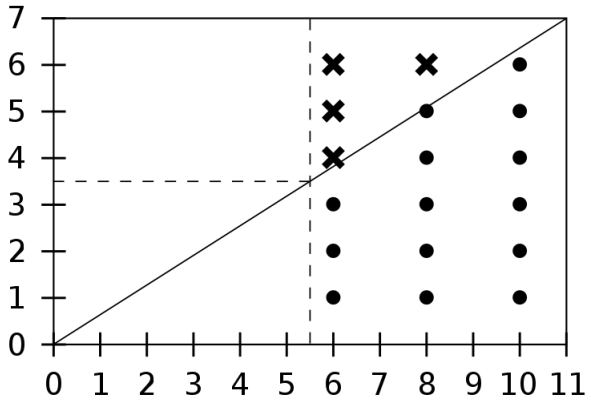

Figure 3: Columns of lattice points.

It's clear that the triangles YZC and AXY are congruent, and moreover that the number of *even* points in YZC equals the number of *odd* points in AXY:

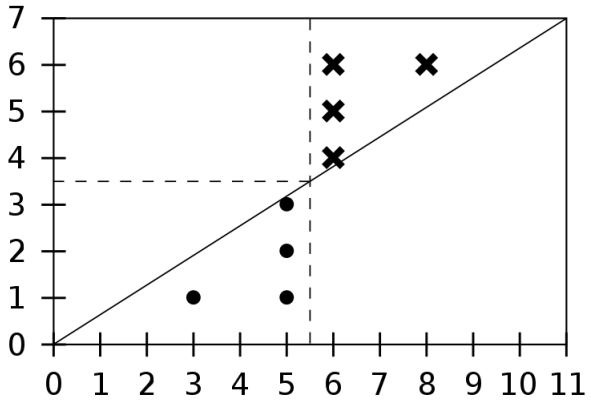

Figure 4: Lattice points with odd x-coordinates in AXY, and with even x-coordinates in YZC.

Letting $E(R)$ (respectively $O(R)$) denote the parity of the number of even (respectively odd) points lying inside a region R, and interpreting addition of parities in the obvious way[17],

[17] I.e., so that the sum of an even and an odd parity is odd, and the sum of two odd or even parities is even.

we've now shown the following:

$$\begin{aligned} E(ABC) =& E(AXY) + E(XBCY) \\ =& E(AXY) + E(YZC) \\ =& E(AXY) + O(AXY). \end{aligned}$$

In other words, the parity of the number of even points in ABC is equal to the parity of the *total* number of points in AXY. By Eisenstein's lemma, it follows that

$$\left(\frac{q}{p}\right) = (-1)^{\lambda},$$

where λ is the total number of points in AXY. Running the same argument for the opposite triangle ADC, we get

$$\left(\frac{p}{q}\right) = (-1)^{\mu},$$

where μ is the total number of points in AWY. Finally, observe that $\lambda + \mu$ is the number of points in the rectangle $AXYW$, which is $\left(\frac{p-1}{2}\right)\left(\frac{q-1}{2}\right)$. Putting these last few facts together, we have the statement of QR:

$$\left(\frac{q}{p}\right)\left(\frac{p}{q}\right) = (-1)^{\lambda+\mu} = (-1)^{\frac{p-1}{2}\cdot\frac{q-1}{2}}. \quad \square$$

I claim that the Gauss-Eisenstein proof is explanatory by virtue of being transparent. In case this claim isn't obvious, let me say a few things in its defense.

First, as we've already seen, the Gauss-Eisenstein approach is often described using the language of transparency. Gauss called his original proof simple, natural and elegant, and contemporary mathematicians tend to agree. Daniel Shanks describes it as "the best of Gauss's many proofs" and "his simplest proof" ([Shanks 1978], 65). Eisenstein's modifications are a further improvement; we've seen this version of the proof praised as intuitive, beautiful and economical. These kinds of terms indicate a proof that's easy to digest and presents a comprehensible reason for its result.

Second, the proof uses methods that promote transparency. In particular, it's geometric, it's readily visualizable, and it turns on a straightforward finite counting argument. Although the argument still demands some cognitive effort, both the general structure and the individual steps are intuitive and relatively easy to grasp. This is all a far cry from Gauss's forbidding inductive proof.

Finally, the proof gives a clear reason for the truth of QR. The theorem holds because $\left(\frac{q}{p}\right)\left(\frac{p}{q}\right)$ is equal to $(-1)^{\sum_u \left\lfloor \frac{uq}{p} \right\rfloor + \sum_u \left\lfloor \frac{up}{q} \right\rfloor}$ (by Eisenstein's lemma), and the sum $\sum_u \left\lfloor \frac{uq}{p} \right\rfloor + \sum_u \left\lfloor \frac{up}{q} \right\rfloor$ is the number of lattice points in a $\frac{p}{2} \times \frac{q}{2}$ rectangle, which is pretty easily seen to be $\frac{p-1}{2} \cdot \frac{q-1}{2}$. This sort of reason might or might not be entirely satisfying, depending on one's broader goals and interests. And one might like to have a further explanation for certain parts of the argument (e.g. Eisenstein's lemma). But this isn't a problem—a good explanation need not be perfectly complete and need not meet the needs of every inquirer. (Compare: general relativity explains the tides, even though this explanation presupposes Einstein's

equations without justification, and even though the simpler Newtonian explanation may be preferable under certain circumstances.)

So the Gauss-Eisenstein proof is transparent, and it's plausibly explanatory for this reason. Note, though, that the proof isn't particularly deep. Although it invokes some novel mathematics—e.g., geometric notions like *triangle* and *lattice point*—these ideas aren't appreciably more abstract or profound than the ones in the statement of QR. Indeed, commentators universally regard the proof as elementary. Moreover, counting lattice points does little to reveal the meaning of QR, in the sense that the approach doesn't obviously lead to interesting generalizations or display QR's place in a larger family of results. It's for this reason, as we'll see in §4.4 below, that the Gauss-Eisenstein proof is sometimes deprecated as superficial, gimmicky and point-missing. Like the tiling proof of the Pythagorean theorem discussed in §3.2, the lattice-point proof is "very convincing indeed"—but it seems to be based on "an ingenious trick" that makes no contact with the fundamental underlying facts. The next section shows a proof of this latter type.

4.3 A deep proof: Hilbert, by algebraic number theory

QR was a landmark discovery, but it raises as many questions as it answers. Are there also interesting rules governing the solvability of $x^3 \equiv p \pmod{q}$, or higher powers of x? If so, do they involve some sort of reciprocity phenomenon? Is there anything to say about the general case, $x^n \equiv p \pmod{q}$?

The answer to all these questions is yes, although this isn't trivial to show—the pursuit of higher reciprocity laws was a major focus of number theory for over a hundred years, from Gauss's work up to Emil Artin's in the 1920s. QR lays the groundwork for this important area of mathematics. Indeed, in their classic textbook, Ireland and Rosen characterize QR as "among the deepest and most beautiful results of elementary number theory and the beginning of a line of reciprocity theorems that culminate in the very general Artin reciprocity law, perhaps the most impressive theorem in all number theory" ([Ireland & Rosen 1990], 54).

The area of mathematics most intimately associated with these results is *algebraic number theory*. This discipline uses the tools of abstract algebra—groups, rings, fields, Galois theory and so on—to address questions about the integers, and its development went hand in hand with progress on higher reciprocity laws. Artin's general reciprocity law uses the language of *class field theory*, a sophisticated branch of algebraic number theory. The large and technical Langlands program represents an ongoing effort to push these ideas still further.

This section presents a proof of QR based on algebraic number theory. The proof doesn't invoke the full power of Artin's law—it would be difficult to even state that theorem here in a way that's enlightening and reasonably concise[18]—but the approach gives a good sense for the modern algebraic viewpoint on reciprocity. I claim that the proof is deep (though not very transparent).

Before giving the proof, let me give a bit of background on higher reciprocity and algebraic number theory. The basic insight here, already known to Gauss, is that going beyond QR

[18]Artin reciprocity says that two algebraic objects—the abelianization of a certain kind of Galois group, and a certain kind of idele class group—are isomorphic. Defining these groups in a precise way would take quite a few layers of definitions.

requires working in an extension of the rational numbers $\mathbb{Q}$. More specifically, the proper setting for proving the nth reciprocity law turns out to be the ring of integers $\mathcal{O}_n$ of the cyclotomic field $\mathbb{Q}(\zeta_n)$, where ζ_n is a primitive nth root of unity. (For instance, ζ_4 is the imaginary unit $i = \sqrt{-1}$, so $\mathcal{O}_4$ is the familiar ring of Gaussian integers of the form $a + bi$.) Within this framework, Gauss and his followers Jacobi and Eisenstein were able to prove laws for cubic and quartic reciprocity. They did so by adapting the technique of "Gauss sums", which had originally appeared in Gauss's last published proof of QR.

This technique depends on the unique factorization property, however, which holds in the rings $\mathcal{O}_3$ and $\mathcal{O}_4$ but not for arbitrary $\mathcal{O}_n$. So an amendment was needed to reach a more general reciprocity law. To this end, Kummer introduced his new "ideal numbers". Although unique factorization doesn't always hold for *elements* of $\mathcal{O}_n$, it does turn out to hold for (what we now call) *ideals* of $\mathcal{O}_n$, and this fact led Kummer and Eisenstein to higher reciprocity laws via a generalization of the Gauss sums technique. The farthest-reaching of these results, called Eisenstein reciprocity, describes an ℓth reciprocity law for all odd primes ℓ. Since Eisenstein's time, reciprocity laws have been generalized considerably further—to number fields other than $\mathbb{Q}$, for instance—but the algebraic tools developed by Gauss's successors continue to find use. Many are on display in this section's proof, which I'll now present. (As far as I can tell, the proof first appears in David Hilbert's landmark *Theory of Algebraic Number Fields* of 1896 (in English as [Hilbert 1998]); the proof is in §122, where Hilbert calls it "a new proof of the quadratic reciprocity law" (217)). Proofs of this style can now be found in many places, especially in texts on algebraic number theory: see for instance [Weyl 1968], §3.11; [Samuel 1970], §VI.5; [Janusz 1996], §I.11.)

As usual, let p and q be odd primes, and define $p^* = (-1)^{\frac{p-1}{2}} p$. (This means that $p^* = p$ if $p \equiv 1 \pmod 4$, and $p^* = -p$ otherwise. Note that $p^* \equiv 1 \pmod 4$ in either case; this fact will be useful later.) Our approach will center around the cyclotomic field $\mathbb{Q}(\zeta_p)$, where ζ_p is a primitive pth root of unity. We'll prove QR in the form $\left(\frac{p^*}{q}\right) = \left(\frac{q}{p}\right)$.

Recall that the *Galois group* of $\mathbb{Q}(\zeta_p)/\mathbb{Q}$ is the group of automorphisms of the upper field $\mathbb{Q}(\zeta_p)$ which fix the lower field $\mathbb{Q}$. By basic Galois theory, $\mathrm{Gal}(\mathbb{Q}(\zeta_p)/\mathbb{Q})$ is cyclic of order $p-1$, and hence is isomorphic to $(\mathbb{Z}/p\mathbb{Z})^\times$ (the multiplicative group of integers modulo p). As a cyclic group of even order, $(\mathbb{Z}/p\mathbb{Z})^\times$ has a unique subgroup of index 2, corresponding to the set of squares modulo p. By the Fundamental Theorem of Galois Theory, there exists a corresponding (and also unique) quadratic field extension $K/\mathbb{Q}$.

To proceed, we need to determine the identity of the field K, which equals $\mathbb{Q}\left(\sqrt{d}\right)$ for some squarefree integer d. This is the first appearance of algebraic number theory in the proof: in order to identify K, we have to talk about the *ramification* of primes in an algebraic number field. Since ideals in rings of algebraic integers have the unique factorization property (as noted earlier), we can take a standard prime $q \in \mathbb{Z}$, get an ideal of $\mathcal{O}_p$ by multiplication, and factor this ideal in just one way:

$$q\mathcal{O}_p = (q) = \prod_i P_i^{e_i},$$

where the P_i are prime ideals of $\mathcal{O}_p$ and the exponents e_i are positive integers. The prime q is said to *ramify* in $\mathbb{Q}(\zeta_p)$ if any of the e_i are greater than 1.

It isn't hard to show that p is the only prime that ramifies in $\mathbb{Q}(\zeta_p)$. One can do this by considering the *discriminant*, an important numerical invariant of an algebraic number field:

the discriminant of $\mathbb{Q}(\zeta_p)$ is always a power of p, and it turns out that a prime ramifies in a given field just in case that prime divides the discriminant. Moreover, if a prime ramifies in the lower field $\mathbb{Q}\left(\sqrt{d}\right)$, then it also ramifies in the extension $\mathbb{Q}(\zeta_p)$. So we're looking for a quadratic field $\mathbb{Q}\left(\sqrt{d}\right)$ where p is the only ramified prime.

We can find it with the help of the discriminant again. Another standard fact is that, since $\mathbb{Q}\left(\sqrt{d}\right)$ is a quadratic field, its discriminant is equal to d itself if $d \equiv 1 \pmod 4$ and to $4d$ otherwise. In the latter case, the discriminant's prime divisors would include 2 as well as our odd prime p, and hence both primes would be ramified in $\mathbb{Q}(\zeta_p)$, a contradiction. So we know that $d \equiv 1 \pmod 4$, that $p \mid d$, and that d is squarefree; the only choice is $d = p^*$. Therefore the mystery field K is actually $\mathbb{Q}(\sqrt{p^*})$.

To complete the proof of QR, we take up our second odd prime q and examine its behavior in $\mathbb{Q}(\sqrt{p^*})$. We've already discussed one possible situation, ramification, where there's a term with $e_i > 1$ in the prime ideal factorization $(q) = \prod_i P_i^{e_i}$. There are two other scenarios to consider. In fact, however, things are simpler when the extension is a quadratic field: in this case, the factorization has at most two non-unit terms, i.e. we have $(q) = P_1^{e_1} P_2^{e_2}$. With this proviso, the remaining possibilities are as follows. First, it could happen that $(q) = P$ for some single prime ideal $P \subset \mathbb{Q}(\sqrt{p^*})$. In this case q is said to be *inert*. Second, we could have $(q) = P_1 P_2$, where P_1 and P_2 are distinct prime ideals. Here q is said to *split*.

Since p is the only ramified prime in $\mathbb{Q}(\sqrt{p^*})$, q must be either split or inert. Our next goal is to show that $\left(\frac{p^*}{q}\right) = \left(\frac{q}{p}\right) = 1$ if q splits, while $\left(\frac{p^*}{q}\right) = \left(\frac{q}{p}\right) = -1$ if q is inert. This fact will complete the proof of QR.

Step 1: q splits iff $\left(\frac{p^*}{q}\right) = 1$. By elementary algebraic number theory, the minimal polynomial of $\mathbb{Q}(\sqrt{p^*})$ is $f_{p^*}(x) = x^2 - x + \frac{1-p^*}{4}$, and q splits in $\mathbb{Q}(\sqrt{p^*})$ just in case $f_{p^*}(x)$ factors into two distinct linear polynomials modulo q. But the roots of $f_{p^*}(x)$ are $\frac{1\pm\sqrt{p^*}}{2}$, so we have the factorization $\left(x - \frac{1+\sqrt{p^*}}{2}\right)\left(x - \frac{1-\sqrt{p^*}}{2}\right)$ just in case the square root of p^* exists modulo q, i.e., just in case $\left(\frac{p^*}{q}\right) = 1$.

Step 2: $\left(\frac{p^*}{q}\right) = 1$ iff $\left(\frac{q}{p}\right) = 1$. First, recall that the *Frobenius automorphism* $\mathrm{Frob}_{q,\mathbb{Q}(\zeta_p)} \in \mathrm{Gal}(\mathbb{Q}(\zeta_p)/\mathbb{Q})$ is defined by the map $\zeta_p \mapsto \zeta_p^q$. The restriction of this map to the subfield $\mathbb{Q}(\sqrt{p^*})$ is in fact the Frobenius automorphism $\mathrm{Frob}_{q,\mathbb{Q}(\sqrt{p^*})} \in \mathrm{Gal}(\mathbb{Q}(\sqrt{p^*})/\mathbb{Q})$. Since the Galois group of a quadratic extension only has two elements—the identity map, and the automorphism sending $\sqrt{p^*}$ to $-\sqrt{p^*}$—we can identify it with the group $\{+1, -1\}$. So we can write $\mathrm{Frob}_{q,\mathbb{Q}(\sqrt{p^*})} = \pm 1$. In fact, it turns out that the order of the Frobenius is determined by the splitting behavior of q in $\mathbb{Q}(\zeta_p)$: if q splits, then $\mathrm{Frob}_{q,\mathbb{Q}(\sqrt{p^*})} = +1$, and if q is inert, then $\mathrm{Frob}_{q,\mathbb{Q}(\sqrt{p^*})} = -1$. (This isn't hard to show, but doing so requires a few extra definitions.) In other words, by the result of Step 1, we have $\mathrm{Frob}_{q,\mathbb{Q}(\sqrt{p^*})} = \left(\frac{p^*}{q}\right)$.

The final maneuver is to show that $\mathrm{Frob}_{q,\mathbb{Q}(\sqrt{p^*})} = \left(\frac{q}{p}\right)$. Note that $\mathrm{Frob}_{q,\mathbb{Q}(\sqrt{p^*})} = +1$ occurs just in case $\mathrm{Frob}_{q,\mathbb{Q}(\zeta_p)}$ fixes $\mathbb{Q}(\sqrt{p^*})$, since the former Frobenius is the restriction of the latter. We can use Galois theory to get a clearer idea of what this means. As noted earlier, we have an isomorphism $\mathrm{Gal}(\mathbb{Q}(\zeta_p)/\mathbb{Q}) \cong (\mathbb{Z}/p\mathbb{Z})^\times$. Correspondingly, $\mathrm{Gal}(\mathbb{Q}(\zeta_p)/\mathbb{Q}(\sqrt{p^*}))$ is isomorphic to the subgroup of $(\mathbb{Z}/p\mathbb{Z})^\times$ consisting of the quadratic residues modulo p. From this viewpoint, the condition that $\mathrm{Frob}_{q,\mathbb{Q}(\zeta_p)}$ fixes $\mathbb{Q}(\sqrt{p^*})$ means that q belongs to

$\mathrm{Gal}\left(\mathbb{Q}\left(\zeta_p\right)/\mathbb{Q}\left(\sqrt{p^*}\right)\right)$, that is, q is a quadratic residue modulo p. Putting all this together, we've shown that $\mathrm{Frob}_{q,\mathbb{Q}\left(\sqrt{p^*}\right)} = +1$ if and only if q is a quadratic residue modulo p, or in other words $\mathrm{Frob}_{q,\mathbb{Q}\left(\sqrt{p^*}\right)} = \left(\frac{q}{p}\right)$.

This completes the proof of QR: from Steps 1 and 2, we have $\left(\frac{p^*}{q}\right) = \left(\frac{q}{p}\right)$, as needed.

I claim that this last proof is explanatory by virtue of being deep. The depth of the proof is, I hope, not in question: clearly it makes much use of ideas from algebraic number theory, which are remote from (and of a significantly "lower stratum" than) the ideas in the statement of the theorem. It also gives a satisfying reason why QR is true: it turns out that the Legendre symbols $\left(\frac{p^*}{q}\right)$ and $\left(\frac{q}{p}\right)$ can be interpreted as giving the order of the Frobenius automorphism $\mathrm{Frob}_{q,\mathbb{Q}\left(\sqrt{p^*}\right)}$, which is trivial iff q splits in $\mathbb{Q}\left(\zeta_p\right)$, iff $\left(\frac{p^*}{q}\right) = \left(\frac{q}{p}\right) = 1$. Additionally, the proof can be said to "show what the theorem means". As Yuri Manin and Alexei Panchishkin write:

> Unfortunately, in most modern texts devoted to elementary number theory one cannot find any hint of explanation as to why quadratic reciprocity is anything more than just a curiosity. The point is that primes, the traditional subject matter of arithmetic, have another avatar as Frobenius elements in the Galois group. Acting as such upon algebraic numbers, they encode in this disguise of symmetries much more number-theoretical information than in their more standard appearance as elements of $\mathbb{Z}$. ([Manin & Panchishkin 2005], 4)

Indeed, this approach is closely related to the Artin reciprocity law and other modern generalizations of QR, which also rely on Galois theory, Frobenius elements and the splitting behavior of primes in algebraic number fields. In this sense, the above proof helps situate QR in the broader context of contemporary class field theory.

In light of these qualities, it's unsurprising that many mathematicians consider the proof explanatory. Hermann Weyl is explicit about this in his classic *Algebraic Theory of Numbers*. After giving a version of our proof, Weyl remarks: "There certainly exist more elementary proofs of the reciprocity law, but hardly one that is less artificial and goes as straight to the root of the phenomenon" ([Weyl 1968], 127). Citing the rules governing the splitting of primes in $\mathbb{Q}\left(\zeta\right)$, Jürgen Neukirch agrees that "[t]he law of decomposition in the cyclotomic field provides the correct explanation of Gauss's reciprocity law" ([Neukirch 1999], 63). As we'll see in the next section, Weyl and Neukirch are hardly alone in thinking that algebraic number theory furnishes the best explanation of QR.

The algebraic proof is lacking in transparency, though. Even in the relatively congenial form in which I've tried to present it, the reasoning is complex, and it depends on many non-obvious definitions, facts, sub-arguments and other bits of unwieldy technical machinery. (I've elided some of these in the interest of length and clarity; a version of the proof with all the details completely spelled out would be a good deal more involved.) I think the proof possesses a kind of intellectual beauty, and the connection between QR and the theory of algebraic number fields is certainly remarkable. But I think one would be hard pressed to argue that the argument is illuminatingly simple, intuitive or clear.

4.4 The disagreement about explaining QR

I pointed out at the outset that mathematicians disagree deeply about which proof best explains QR. I also promised that this paper would provide some tools for understanding the disagreement. Now that the relevant proofs are at our disposal, it's time to make good on this promise.

We saw earlier that the Gauss-Eisenstein lattice-point proof is highly regarded by many mathematicians. Gauss himself considered it (in its original form) his best candidate for a natural proof of QR, and by "natural" he seems to have meant "explanatory". Shanks agrees that it's Gauss's best proof, and Tangedal calls it "remarkably direct and insightful". Many others have praised it in a variety of ways. Moreover, it's the most widely used textbook proof. Given the dozens of other options to choose from, this probably says something about its overall merits, including mathematicians' perception of its explanatory value.

But not everyone takes this view. For a sizable and vocal contingent of mathematicians, the Gauss-Eisenstein proof is definitely not an acceptable explanation of QR. For example, here's the number theorist Keith Conrad: "Among proofs I have read that I don't like, two I can recall are the one in Serre's *Course in Arithmetic*, which is a method involving trigonometric identities, and the first proof of QR that I learned, which is essentially the one involving counting lattice points... I do not mean to say lattice-point counting is unimportant, but the lattice-point proof of QR feels like a one-time trick compared to other proofs I have seen" (personal correspondence, 7/29/2018). Helmut Koch dismisses all of QR's early proofs as unexplanatory: "Altogether [Gauss] gave seven proofs of this theorem, however they should all be regarded as verifications, which give no insight into the background of the law" ([Koch 1991], 5). Fernando Gouvêa expounds on this sentiment at length:

> The proof that is usually given in elementary courses goes like this: relate the answers to the questions [about quadratic residues] to counting, in such a way that an even count means the answer is "yes" and an odd count means the answer is "no." ...Then set up a way to relate the two counts, and show that the difference between the answers is odd or even as required by the theorem.
>
> The proof works, but it is remarkable in the fact that it gives us *no insight at all* into why the theorem is true. In particular, it does not yield any direct connection between "life mod p" and "life mod q". Every time I present the proof to students, I point out the feeling that yes, it comes out right, but it comes out right *because the theorem is true*. It's hard to claim (and I do not believe) that counting points in a rectangle explains *why* the theorem is true. ([Gouvêa 2015]; emphasis in original)

What, then, might count as a better explanation? In the view of some authors, elementary methods can't give a satisfactory proof of QR; rather, the theorem properly "belong[s] to the realm of algebraic number theory" ([Lemmermeyer 2000], v). (In fact, Lemmermeyer says that "the most transparent proofs of the quadratic reciprocity law are embedded into the theory of algebraic number fields" (vi). The context suggests that Lemmermeyer means something like "illuminating" or "explanatory"; the statement is pretty clearly untrue if "transparent" is understood in my sense.) Here's a forceful statement of this view from the German mathematician Erich Hecke:

> The development of algebraic number theory has now actually shown that the content of the quadratic reciprocity law only becomes understandable if one passes to general algebraic numbers and that a proof appropriate to the nature of the problem can be best carried out with these higher methods. ...[I]t must be said of the elementary proofs that they possess rather the character of supplementary verification. For this reason we will dispense entirely with a presentation of an elementary proof. Rather we set ourselves the problem of carrying over the concepts of rational number theory, in particular the concept of integer, to other domains of numbers, where new relations between rational integers will also be obtained, e.g., the reciprocity law itself will be presented as a side result. ([Hecke 1981], 53)

Of course, algebraic number theory is the domain of ramified primes, Frobenius elements and the like. So these latter authors are expressing a preference for proofs like the one from the previous section.

This is an interesting situation. Some authors consider the lattice-point proof to be "remarkably insightful", while others insist that it gives "no insight at all". Those in the latter camp frequently hold up algebraic-style proofs as the correct route to explaining QR. All parties involved are knowledgeable and capable mathematicians who have spent more than a little time thinking about the subject. What should we make of this? My diagnosis should be clear by now. The Gauss-Eisenstein proof and the algebraic proof are both explanatory, but they exhibit very different kinds of explanatory virtue: the former is transparent (but "shallow"), while the latter is deep (but "opaque"). Those who accept the Gauss-Eisenstein proof as a good explanation value transparency, while the others demand depth.

As evidence that this diagnosis is right, consider the kinds of reasons given by each side for their views. Those who accept the lattice-point proof as explanatory praise it as direct, simple, elegant and intuitive: qualities associated with transparency. Their opponents, importantly, don't deny that the proof has *these* qualities. Rather, they complain that it misses the point: it "feels like a one-time trick" that fails to make contact with the deep ideas of algebraic number theory. (Note that this is the very same criticism made by Givental of the shallow tiling proof of the Pythagorean theorem.) Only by taking this more expansive viewpoint, Hecke says, does the content of QR become understandable. This agrees with the earlier observation that deep proofs "show what a theorem means".

Is there anything to say about why the particular authors in question might have the preferences that I've ascribed to them? I'm hesitant to engage in psycho-sociological speculation, but there are a few points worth noting. As discussed in §3.3, a mathematician might prefer either a transparent or a deep explanation for any of several reasons: their personal style, the norms of their subdisicplinary "culture", or the interests of their audience, to name a few. I think we can see each of these factors at work in the case of QR.

Gauss presents an interesting case of distinctive individual style. While he's often (rightly) praised for his deep ideas and results, Gauss's explanatory ideal seems to have been close to transparency rather than depth. His long-sought natural proof of QR, as we saw in §4.2, was to be simple, clear, elementary and not unduly cognitively taxing. (Although he was also interested in approaches that would generalize to higher reciprocity laws, Gauss seems to have considered such proofs more instrumentally useful than explanatorily

illuminating.) Of course, Gauss didn't live to see the blossoming of algebraic number theory in the later 19th century and beyond, and there's no doubt he would have appreciated its power, elegance and generality. But it seems probable that, even if Gauss had had the resources to discover an algebraic-style derivation of QR, this type of proof wouldn't have met his criteria for naturalness.

The differing norms of mathematical subcultures may also be a factor. §3.3 discussed Gowers' distinction between a "theory-building" culture that values depth and generality and a "problem-solving" culture that prizes transparency and directness. Gowers explicitly names algebraic number theory as an example of the former type of discipline ([Gowers 2000], 66), and the Hecke quote above is from a book on that subject. (But the rest of the evidence isn't totally straightforward: Tangedal is an algebraic number theorist, for instance, while Lemmermeyer and Gouvêa both also work in the history of mathematics. So cultural considerations may not be extremely important here.)

Finally, these authors are writing for different audiences with varying interests and areas of expertise. It's not surprising that introductory textbooks often prove QR via counting lattice points, since it (like other transparent explanations) demands relatively little in the way of cognitive effort and special knowledge. Works like [Lauenbacher & Pengelley 1994a], [Tangedal 2000] and [Aigner & Zeigler 2010] are also directed at fairly wide, non-specialist audiences. By contrast, the books of Weyl, Neukirch, Lemmermeyer and Hecke demand significant mathematical sophistication. It's not unreasonable for works like this to focus on deeper proofs, since their readers will be in a better position to appreciate such explanations.

As I hope is now clear, the existence of disagreements like this one shouldn't be too surprising or disturbing. It's perfectly consistent, and likely appropriate in some contexts, to value one type of explanation more than another. So the case of QR doesn't show that mathematicians are confused or irrational, or that their explanatory practices are incoherent. It's plausible, rather, that different authors are legitimately seeking different types of explanation, and the proofs they prefer are fitting responses to their divergent interests.

5 QR and the theory of explanatory proof

I promised at the outset of the paper that the case of QR would yield interesting philosophical conclusions. Some of these goods were delivered in §3, but at least one more important point awaits unpacking. This section does the job.

The marquee issue facing theorists of mathematical explanation is, of course, the general question about the nature of the phenomenon. What makes a proof (or other piece of mathematics) explanatory or not? Since we have in hand a couple of explanatory proofs as well as an unexplanatory control, it's worth considering what we can say on this subject.

The first wave of the recent literature on mathematical explanation appeared in the 1970s and 80s, centering around the work of Mark Steiner (especially [Steiner 1978]) and Phillip Kitcher (especially [Kitcher 1989]). Both authors proposed general theories of explanatory proof. But both have taken numerous beatings in work from the last couple decades and, as far as I know, are no longer actively defended by anyone. So I won't discuss them further.[19] Among ideas that remain seriously on the table, the most ambitious and well-

[19]For the record, Steiner's theory claims that a proof is explanatory when it's "deformable" in a certain

developed account is surely Marc Lange's ([Lange 2014], [Lange 2016]). Lange's theory is young, however, and as of yet it's received relatively little critical scrutiny. The case of QR provides a good opportunity to put it to the test.

On Lange's view, a proof is explanatory when it "exploits a certain kind of feature in the problem: the same kind of feature that is outstanding in the result being explained" ([Lange 2014], 489). Thus it only makes sense to talk about explanatory proof when a why-question has been prompted by some specific noteworthy quality of a theorem. One such quality is *symmetry*: when a result shows some phenomenon to be surprisingly symmetrical, we find ourselves wondering why, and we consider a proof explanatory only if it makes use of a similar symmetry. Another feature that demands explanation is *unity*: when a theorem shows different cases to exhibit a striking commonality, an explanatory proof will be one that exposes their underlying sameness. (This is why "brute-force" methods like proof by exhausation often seem unsatisfying.) Finally, the *simplicity* of a result is often salient, and in this case we deem a proof explanatory when it reveals some correspondingly simple feature of the problem situation. Although Lange holds that symmetry, unity and simplicity are among the features that most often call out for explanation, he allows that other qualities can also be salient in this way ([Lange 2014], 524).

If Lange's theory is right, we ought to expect two things. First, we should be able to identify a specific why-question prompted by some salient aspect of QR. Second, a proof of QR should count as explanatory if and only if it exploits the same type of feature.

Mathematicians clearly view QR as mysterious and in need of explanation, but what is its relevant outstanding quality? I don't think this is very obvious. Although one could probably shoehorn the case into any or all of Lange's three main categories, none seems particularly natural, and none leads to the right verdicts about proofs.

Consider symmetry. The reciprocity relationship is perhaps symmetrical in a sense, insofar as the values of $\left(\frac{p}{q}\right)$ and $\left(\frac{q}{p}\right)$ each partly determine one another. But this is a somewhat strained usage of "symmetrical". And in any case there's nothing in the proofs of QR that looks a definitive exploitation of a relevant symmetry. The lattice-point proof, for example, features a diagram that's symmetrical in certain ways, but it has no evident symmetry corresponding to the reciprocal relationship of $\left(\frac{p}{q}\right)$ and $\left(\frac{q}{p}\right)$. The other proofs don't seem to involve any meaningful symmetries at all.

Unity seems like a non-starter. What's noteworthy about QR isn't that it unifies a set of seemingly disparate facts. (Of course it does have infinitely many special cases, like any general theorem about prime numbers, but it isn't especially surprising or impressive that these cases fall under a common rule.)

Finally, consider simplicity. I'm not inclined to say that QR is very simple. Thanks to the ingeniousness of the Legendre symbol, it's possible to express QR in a compact form, but it takes a nontrivial amount of work to explain what the equation $\left(\frac{p}{q}\right)\left(\frac{q}{p}\right) = (-1)^{\frac{p-1}{2}\cdot\frac{q-1}{2}}$ means. And I don't think the theorem seems simple even once one does understand. (Odd

respect, while Kitcher's account is a mathematical application of his well-known unificationist theory of scientific explanation. Some philosophers remain sympathetic to modified or restricted versions of these views. For instance, [Weber & Verhoeven 2002] develop an amended Steinerian account that's only intended to handle certain kinds of explanatory proof. And [Tappenden 2005] is a sympathetic exploration of the idea that unification promotes explanation in mathematics. See [D'Alessandro 2019] for more on Steiner's and Kitcher's views and their respective receptions.

primes p and q are either *both* or *neither* quadratic residues modulo each other if at least one is congruent to 1 (mod 4), and if both are congruent to 3 (mod 4), then exactly *one* is a quadratic residue modulo the other? A far cry from the elegance of $e^{i\pi} + 1 = 0$.) What's more, the algebraic number theory proof of QR isn't simple, and yet many mathematicians consider it explanatory.

So I think symmetry, unity and simplicity are unpromising candidates for QR's outstanding feature. Is there a better one? As we saw in §2.3, when mathematicians describe what they find salient about QR, they sometimes mention the surprising contrast with results like the Chinese Remainder Theorem, which suggest that residuosity modulo p should be independent of residuosity modulo q. So one might frame the question here as "Why is the quadratic residuosity relationship one of reciprocity rather than independence?"

In order to answer this question, I take it that a proof would have to somehow address the linear case and the respect in which it differs from the quadratic case. Notably, however, neither of the explanatory proofs from §4 does so. These proofs shed no light at all on the contrast between QR and facts like the CRT. So either Lange's account is wrong about this aspect of explanatory proofs, or neither proof is really explanatory, or else the QR-CRT contrast isn't the uniquely salient outstanding feature here.

Actually, I'd like to offer something of a fourth option: although QR calls out loudly for explanation, this isn't because it prompts any one sharp and distinctive why-question. The theorem is peculiar, and it seems somehow profound, but it's not immediately clear what the real issue is. Rather, QR exhibits a kind of inchoate mysteriousness that reflects our meta-ignorance about the situation—there's apparently something important about quadratic residues that we haven't fully understood, but the shape and extent of that something aren't obvious in advance. In part, the need for explanation is a need to better delineate the very questions that need answering.

I think this goes a long way toward explaining why QR has been reproved so many times. To some extent, we won't know what to expect from an explanation until we have it, and it's hard to tell when the job is completely done—there's always a chance that a new proof will reveal a further aspect of the phenomenon that we didn't previously know to look out for.

This kind of situation isn't uncommon (in mathematics, science or elsewhere), but it's hard to reconcile with Lange's theory. As Lange notes with admirable candor: "My proposal predicts that if [a] result exhibits no noteworthy feature, then to demand an explanation of why it holds, not merely a proof that it holds, makes no sense. There is nothing that its explanation over and above its proof would amount to until some feature of the result becomes salient" ([Lange 2014], 507).

The case of QR, and others like it, seem to me like plausible counterexamples. It certainly makes sense to demand an explanation of quadratic reciprocity. And indeed we have some good candidates for explanatory proofs. But it's difficult to point to any single, uniquely salient feature of QR which we feel pressed to explain, or which these proofs do explain. Rather, the question "Why is QR true?" seems to largely amount to the question "What's going on with quadratic residues? Why do they exhibit this (odd and apparently significant) behavior, and what is it that we don't understand about them?" A proof will count as explanatory according to this standard when it manages to roll back the fog of our meta-ignorance in some suitable way; we can't say anything much more specific about what we

need until after we've gotten it. In short, then, Lange's theory may work well for *known unknowns*, but it has trouble dealing with *unknown unknowns*, like the case of QR.

6 Conclusion

This paper has covered a lot of ground. To summarize, my main claims have been the following. First, the quadratic reciprocity theorem has much to teach us about mathematical explanation, including the neglected issue of explanatory disagreements and their philosophical significance. I've argued that such disagreements arise in part because—to a degree matched by few other important theorems—QR admits of both deep and transparent explanatory proofs, and mathematicians may legitimately prefer one type of explanation over the other. In the course of making this case, I've developed the notions of transparency and depth in some detail, and I've tried to exhibit their usefulness. Finally, I've suggested that QR puts some interesting constraints on an acceptable theory of mathematical explanation.

These conclusions raise some further questions. Notably: (1) Are depth and transparency the only kinds of explanatory virtues that a proof can have? If not, which others are there? (2) What does this tell us about the nature of (mathematical) explanation in general? My conjectures: (1) Any explanatory proof has some explanation-making feature that can be assimilated to one of these two notions. For instance, I think the "abstract explanations" described in [Pincock 2015] are more perspicuously understood as deep explanations. (2) Theories of explanation that require ontic or counterfactual dependence relations are too narrow. As far as mathematical explanation is concerned, at least, we need an account that lets cognitive considerations count.

I leave these conjectures sketchy and unsupported for now, but I hope to revisit them in future work.

Acknowledgments

An earlier version of this paper was my MS thesis in mathematics at the University of Illinois at Chicago. I'm grateful to my advisor, Ramin Takloo-Bighash, and my other committee members, Nathan Jones and Jamie Tappenden, for their help and support. Thanks especially to Ramin, who introduced me to quadratic reciprocity and whose faith, advice and good cheer were invaluable over the many months I worked on this project. Thanks to Dick Gross, Jordan Ellenberg and Keith Conrad for correspondence with Ramin and I about proofs of QR. Many thanks to Marc Lange for his characteristically perceptive and generous comments and for his ongoing support. Many thanks as well to Kenny Easwaran and the Texas A&M philosophy and math departments for their hospitality. As usual, Daniel Sutherland and Lauren Woomer gave me helpful feedback. Finally, thanks to the 2019 Midwest Philosophy of Mathematics Workshop participants—including John Baldwin, Paddy Blanchette, Bernd Buldt, Curtis Franks, Doug Marshall, Rebecca Morris, Daniel Nolan, Patrick Ryan, and Susan Vineberg—for their comments, discussion and encouragement.

References

[Aigner & Zeigler 2010] Aigner, Martin and Günter M. Zeigler. 2010. *Proofs from the Book* (4th edition). Berlin: Springer.

[Arana 2015] Arana, Andrew. 2015. "On the depth of Szemerédi's theorem." *Philosophia Mathematica* 23, 163-176.

[Baker 2009] Baker, Alan. 2009. "Mathematical accidents and the end of explanation." In Otávio Bueno and Øystein Linnebo (eds.), *New Waves in Philosophy of Mathematics*, Palgrave Macmillan: New York, 137-159.

[Baumgart 2015] Baumgart, Oswald. 2015. *The Quadratic Reciprocity Law: A Collection of Classical Proofs* (ed. and trans. by Franz Lemmermeyer). Birkhäuser: Heidelberg.

[Booß-Bavnbek & Wojciechowski 1993] Booß-Bavnbek, Bernhelm and Krzysztof P. Wojciechowski. 1993. *Elliptic Boundary Problems for Dirac Operators*. Springer: New York.

[Cassels 1950] Cassels, J.W.S. 1950. "The rational solutions of the diophantine equation." *Acta Mathematica* 82, 243-273.

[Colyvan 2012] Colyvan, Mark. 2012. *An Introduction to the Philosophy of Mathematics*. Cambridge University Press: Cambridge.

[Colyvan et al. 2018] Colyvan, Mark, John Cusbert and Kelvin J. McQueen. 2018. "Two flavours of mathematical explanation." In Alexander Reutlinger and Juha Saatsi (eds.), *Explanation Beyond Causation: Philosophical Perspectives on Non-Causal Explanations*, Oxford University Press: New York, 231-249.

[Cox 2013] Cox, David A. 2013. *Primes of the Form $x^2 + ny^2$: Fermat, Class Field Theory, and Complex Multiplication* (2nd edition). John Wiley & Sons: New York.

[D'Alessandro 2018a] D'Alessandro, William. 2018. "Arithmetic, set theory, reduction and explanation." *Synthese* 195, 5059-5089.

[D'Alessandro 2018b] D'Alessandro, William. 2018. "Mathematical explanation beyond explanatory proof." *British Journal for the Philosophy of Science*. DOI: 10.1093/bjps/axy009.

[D'Alessandro 2019] D'Alessandro, William. 2019. "Explanation in mathematics: Proofs and practice." *Philosophy Compass*. DOI: 10.1111/phc3.12629.

[D'Alessandro 2020] D'Alessandro, William. 2020. "Viewing-as explanations and ontic dependence." *Philosophical Studies* 177, 769-792.

[Detlefsen 1988] Detlefsen, Michael. 1988. "Fregean hierarchies and mathematical explanation." *International Studies in the Philosophy of Science* 3, 97-116.

[de Regt 2004] de Regt, Henk W. 2004. "Visualization as a tool for understanding." *Perspectives on Science* 22, 377-396.

[Di Bucchianico & Loeb 1998] Di Bucchianico, A. and D. E. Loeb. 1998. "Natural exponential families and umbral calculus." In Bruce E. Sagan and Richard P. Stanley (eds.), *Mathematical Essays in Honor of Gian-Carlo Rota*, Birkhäuser: Boston, 195-212.

[Domogatsky 2015] Domogatsky, Grigory. 2015. "A new neutrino telescope for Lake Baikal." *CERN Courier* 55, 23-24.

[Dudley 2009] Dudley, Underwood. 2009. *A Guide to Elementary Number Theory*. Mathematical Association of America: Washington, D.C.

[Edwards 1977] Edwards, Harold M. 1977. *Fermat's Last Theorem: A Genetic Introduction to Algebraic Number Theory*. Springer-Verlag: New York.

[Frans & Weber 2014] Frans, Joachim and Erik Weber. 2014. "Mechanistic explanation and explanatory proofs in mathematics." *Philosophia Mathematica* 22, 231-248.

[Frei 1994] Frei, Günther. 1994. "The reciprocity law from Euler to Eisenstein." In Chikara Sasaki, Mitsuo Sugiura and Joseph W. Dauben (eds.), *The Intersection of History and Mathematics*, Birkhäuser Verlag: Basel, 67-90.

[Garrett 2005] Garrett, Paul. 2005. "Cryptographic primitives." In Paul Garrett and Daniel Lieman (eds.), *Public-Key Cryptography*, American Mathematical Society: Providence, RI, 1-62.

[Gauss 1863] Gauss, Carl Friedrich. 1863. *Werke*, Vol. II. Königlichen Gesellschaft der Wissenschaften: Göttingen.

[Gauss 1966] Gauss, Carl Friedrich. 1966. *Disquisitiones Arithmeticae* (English edition, trans. by Arthur A. Clarke). Springer-Verlag: New York.

[Givental 2006] Givental, Alexander. 2006. "The Pythagorean theorem: What is it about?" *American Mathematical Monthly* 113, 261-265.

[Gouvêa 2000] Gouvêa, Fernando Q. 2000. Review of *Reciprocity Laws: From Euler to Eisenstein*, by Franz Lemmermeyer. *MAA Reviews*, June 23, 2000, online at <https://www.maa.org/press/maareviews/reciprocity-laws-from-euler-to-eisenstein>.

[Gouvêa 2015] Gouvêa, Fernando Q. 2015. Review of *The Quadratic Reciprocity Law: A Collection of Classical Proofs*, by Oswald Baumgart, ed. and trans. by Franz Lemmermeyer. *MAA Reviews*, June 30, 2015, online at <https://www.maa.org/press/maareviews/the-quadratic-reciprocity-law-a-collection-of-classical-proofs>.

[Gowers 2000] Gowers, Timothy. 2000. "The two cultures of mathematics." In V. Arnold, M. Atiyah, P. Lax and B. Mazur (eds.), *Mathematics: Frontiers and Perspectives*, American Mathematical Society: Providence, RI, 65-78.

[Gowers 2008] Gowers, Timothy. 2008. "Mathematics, memory and mental arithmetic." In Mary Leng, Alexander Paseau and Michael Potter (eds.), *Mathematical Knowledge*, Oxford University Press: Oxford, 33-58.

[Gray 2004] Gray, Jeremy. 2004. "An introduction to Gauss's mathematical diary." In G. Waldo Dunnington, *Carl Friedrich Gauss: Titan of Science*, Mathematical Association of America: Washington, D.C., 449-468.

[Gray 2015] Gray, Jeremy. 2015. "Depth—A Gaussian tradition in mathematics." *Philosophia Mathematica* 23, 177-195.

[Gray 2018] Gray, Jeremy. 2018. *A History of Abstract Algebra: From Algebraic Equations to Modern Algebra*. Cham, Switzerland: Springer.

[Grimm 2010] Grimm, Stephen R. 2010. "The goal of explanation." *Studies in the History and Philosophy of Science* 41, 337-344.

[Hafner & Mancosu 2005] Hafner, Johannes and Paolo Mancosu. 2005. "The varieties of mathematical explanation." In Paolo Mancosu, Klaus Frovin Jørgensen and Stig Andur Pedersen (eds.), *Visualization, Explanation and Reasoning Styles in Mathematics*, Springer: Berlin, 215-250.

[Hanna 2018] Hanna, Gila. 2018. "Reflections on proof as explanation." In Andreas G. Stylianides and Guershon Harel (eds.), *Advances in Mathematics Education Research on Proof and Proving: An International Perspective*, Springer: Cham, Switzerland, 3-18.

[Hanna & Mason 2014] Hanna, Gila and John Mason. 2014. "Key ideas and memorability in proof." *For the Learning of Mathematics* 34, 12-16.

[Harriss 2006] Harriss, Edmund O. 2006. Review of "The Pythagorean theorem: What is it about?", by Alexander Givental. *Mathematical Reviews*, online at <https://mathscinet.ams.org./mathscinet-getitem?mr=2204490>.

[Hardy 2012] Hardy, G.H. 2012. *A Mathematician's Apology*. Cambridge University Press: New York.

[Hecke 1981] Hecke, Erich. 1981. *Lectures on the Theory of Algebraic Numbers* (trans. by George U. Brauer and Jay R. Goldman). Springer-Verlag: New York.

[Hilbert 1998] Hilbert, David. 1998. *The Theory of Algebraic Number Fields*, trans. by Iain T. Adamson. Springer: Berlin.

[Inglis & Mejía-Ramos 2019] Inglis, Matthew and Juan Pablo Mejía-Ramos. 2019. "Functional explanation in mathematics." *Synthese*. DOI: 10.1007/s11229-019-02234-5.

[Ioffe 2000] Ioffe, Alexander D. 2000. Review of A. A. Milyutin and N. P. Osmolovskii, *Calculus of Variations and*

Optimal Control. In Donald G. Babbitt and Jane E. Kister (eds.), *Featured Reviews in* Mathematical Reviews*, 1997-1999*, American Mathematical Society: Providence, RI, E63-E64.

[Ireland & Rosen 1990] Ireland, Kenneth and Michael Rosen. 1990. *A Classical Introduction to Modern Number Theory* (2nd edition). New York: Springer-Verlag.

[Janusz 1996] Janusz, Gerald J. 1996. *Algebraic Number Fields*, 2nd edition. American Mathematical Society: Providence, RI.

[Kazdan 1981] Kazdan, Jerry L. 1981. "Another proof of Bianchi's identity in Riemannian geometry." *Proceedings of the American Mathematical Society* 81, 341-342.

[Kedlaya 2008] Kedlaya, Kiran S. 2008. "From quadratic reciprocity to class field theory." In Timothy Gowers, June Barrow-Green and Imre Leader (eds.), *The Princeton Companion to Mathematics*, Princeton: Princeton University Press, 39-42.

[Kelp 2016] Kelp, Christoph. 2016. "Towards a knowledge-based account of understanding." In S. R. Grimm, C. Baumberger and S. Ammon (eds.), *Explaining Understanding: New Perspectives from Epistemology and Philosophy of Science*, London: Routledge, 251-271.

[Khalifa 2013] Khalifa, Kareem. 2013. "The role of explanation in understanding." *British Journal for the Philosophy of Science* 64, 161-187.

[Kim 2004] Kim, Sey Y. 2004. "An elementary proof of the quadratic reciprocity law." *American Mathematical Monthly* 111, 48-50.

[Kitcher 1989] Kitcher, Phillip. 1989. "Explanatory unification and the causal structure of the world." In Phillip Kitcher and Wesley Salmon (eds.), *Scientific Explanation* (Minnesota Studies in the Philosophy of Science, Volume XIII), Minneapolis: University of Minnesota Press, 410-505.

[Koch 1991] Koch, Helmut. 1991. *Introduction to Classical Mathematics I: From the Quadratic Reciprocity Law to*

the Uniformization Theorem. Trans. by John Stillwell. Dordrecht: Springer.

[Knuuttila & Merz 2009] Knuuttila, Tarja and Martina Merz. 2009. "Understanding by modeling: An objectual approach." In H. W. De Regt, S. Leonelli and K. Eigner (eds.), *Scientific Understanding: Philosophical Perspectives*, Pittsburgh, PA: University of Pittsburgh Press, 146-168.

[Lange 2009] Lange, Marc. 2009. "Why proofs by mathematical induction are generally not explanatory." *Analysis* 69, 203-211.

[Lange 2010] Lange, Marc. 2010. "What are mathematical coincidences (and why does it matter)?" *Mind* 119, 307-340.

[Lange 2014] Lange, Marc. 2014. "Aspects of mathematical explanation: Symmetry, unity, and salience." *Philosophical Review* 123, 485-531.

[Lange 2015a] Lange, Marc. 2015. "Depth and explanation in mathematics." *Philosophia Mathematica* 23, 196-214.

[Lange 2015b] Lange, Marc. 2015. "Explanation, existence and natural properties in mathematics—A case study: Desargues' theorem." *Dialectica* 69, 435-472.

[Lange 2016] Lange, Marc. 2016. *Because Without Cause: Non-Causal Explanations in Science and Mathematics*. New York: Oxford University Press.

[Lange 2017] Lange, Marc. 2017. "Mathematical explanations that are not proofs." *Erkenntnis*, DOI: 10.1007/s10670-017-9941-z.

[Lauenbacher & Pengelley 1994a] Lauenbacher, Reinhard C. and David J. Pengelley. 1994. "Eisenstein's misunderstood geometric proof of the Quadratic Reciprocity Theorem." *College Mathematics Journal* 25, 29-34.

[Lauenbacher & Pengelley 1994b] Lauenbacher, Reinhard C. and David J. Pengelley. 1994. "Gauß, Eisenstein, and the 'third' proof of the Quadratic Reciprocity Theorem: Ein kleines Schauspiel." *Mathematical Intelligencer* 16, 67-72.

[Legendre 1830] Legendre, Adrien-Marie. 1830. *Théorie des Nombres* (vol. 1), 3rd edition. Duprat: Paris.

[Lemmermeyer 2000] Lemmermeyer, Franz. 2000. *Reciprocity Laws: From Euler to Eisenstein*. Springer: Berlin.

[Lemmermeyer 2019] Lemmermeyer, Franz. 2019. "Proofs of the quadratic reciprocity law." Online at <http://www.rzuser.uni-heidelberg.de/~hb3/rchrono.html>, accessed 6-18-2019.

[Mancosu 2008a] Mancosu, Paolo. 2008. "Mathematical explanation: Why it matters." In Paolo Mancosu (ed.), *The Philosophy of Mathematical Practice*, New York: Oxford University Press, 134-150.

[Mancosu 2008b] Mancosu, Paolo (ed.). 2008. *The Philosophy of Mathematical Practice*. New York: Oxford University Press.

[Manin & Panchishkin 2005] Manin, Yuri and Alexei Panchishkin. 2005. *Introduction to Modern Number Theory: Fundamental Problems, Ideas and Theories*, 2nd edition. Springer: Berlin.

[McLarty 2008] McLarty, Colin. 2007. "The rising sea: Grothendieck on simplicity and generality". In Jeremy J. Gray and Karen Hunger Parshall (eds.), *Episodes in the History of Modern Algebra (1800-1950)*, American Mathematical Society: Providence, RI.

[Nadler 2010] Nadler, Sam B. 2010. "A proof of Darboux's Theorem." *American Mathematical Monthly* 117, 174-175.

[Narins 2001] Narins, Brigham (ed.). 2001. *World of Mathematics* (vol. 1). Gale Group: New York.

[Neukirch 1999] Neukirch, Jürgen. 1999. *Algebraic Number Theory*, trans. by Norbert Schappacher. Springer: Berlin.

[Pincock 2015] Pincock, Christopher. 2015. "The unsolvability of the quintic: A case study in abstract mathematical explanation." *Philosophers' Imprint* 15 (3), 1-19.

[Raatikainen 2018] Raatikainen, Panu. 2018. "Gödel's incompleteness theorems." In Edward N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy (Fall 2018 Edition)*, URL =

<https://plato.stanford.edu/archives/fall2018/entries/goedel-incompleteness/>.

[Raman-Sundström 2016] Raman-Sundström, Manya. 2016. "The notion of fit as a mathematical value." In Brendan Larvor (ed.), *Mathematical Cultures: The London Meetings 2012-2014*, Springer: Cham, Switzerland, 271-286.

[Raman-Sundström & Öhman 2016] Raman-Sundström, Manya and Lars-Daniel Öhman. 2016. "Mathematical fit: A case study." *Philosophia Mathematica* 26, 184-210.

[Resnik & Kushner 1987] Resnik, Michael D. and David Kushner. 1987. "Explanation, independence and realism in mathematics." *British Journal for the Philosophy of Science* 38, 141-158.

[Rogawski 2000] Rogawski, Jonathan. 2000. "The nonabelian reciprocity law for local fields." *Notices of the AMS* 47, 35-41.

[Rousseau 1991] Rousseau, G. 1991. "On the quadratic reciprocity law." *Journal of the Australian Mathematical Society* (*Series A*) 51, 423-425.

[Rowe 1988] Rowe, David E. 1988. "Gauss, Dirichlet, and the law of biquadratic reciprocity." *Mathematical Intelligencer* 10, 13-26.

[Samuel 1970] Samuel, Pierre. 1970. *Algebraic Theory of Numbers*, trans. by Allan J. Silberger. Hermann: Paris.

[Severini 2005] Severini, Thomas A. 2005. *Elements of Distribution Theory*. Cambridge University Press: New York.

[Shanks 1978] Shanks, Daniel. 1978. *Solved and Unsolved Problems in Number Theory* (2nd edition). Chelsea Publishing Company: New York.

[Smith 1959] Smith, David Eugene. 1959. *A Source Book in Mathematics, Volume One*. Dover: New York.

[Smith 1894] Smith, Henry J. S. 1894. *The Collected Mathematical Papers of Henry John Stephen Smith*, *Volume I*, ed. J.W.L. Glaisher. Oxford: Clarendon Press.

[Stanley 2012] Stanley, Richard P. 2012. *Enumerative Combinatorics, Volume I* (2nd edition). Cambridge University Press: New York.

[Steiner 1978] Steiner, Mark. 1978. "Mathematical explanation." *Philosophical Studies* 34, 135-151.

[Stillwell 2015] Stillwell, John. 2015. "What does 'depth' mean in mathematics?" *Philosophia Mathematica* 23, 215-232.

[Swan 1963] Swan, Richard G. 1963. "An application of graph theory to algebra." *Proceedings of the American Mathematical Society* 14, 367-373.

[Tangedal 2000] Tangedal, Brett A. 2000. "Eisenstein's lemma and quadratic reciprocity for Jacobi symbols." *Mathematics Magazine* 73, 130-134.

[Tappenden 2005] Tappenden, Jamie. 2005. "Proof style and understanding in mathematics I: Visualization, unification and axiom choice." In P. Mancosu, K. Jørgensen and S. Pedersen (eds.), *Visualization, Explanation and Reasoning Styles in Mathematics*, Springer: Berlin.

[Tappenden 2008] Tappenden, Jamie. 2008. "Mathematical concepts and definitions." In Paolo Mancosu (ed.), *The Philosophy of Mathematical Practice*, Oxford University Press: New York, 256-275.

[Trout 2002] Trout, J.D. 2002. "Scientific explanation and the sense of understanding." *Philosophy of Science* 69, 212-233.

[Turri 2015] Turri, John. 2015. "Understanding and the norm of explanation." *Philosophia* 43, 1171-1175.

[Urquhart 2015] Urquhart, Alasdair. 2015. "Mathematical depth." *Philosophia Mathematica* 23, 233-241.

[Waskan et al. 2015] Waskan, Jonathan, Ian Harmon, Zachary Horne, Joseph Spino and John Clevenger. 2014. "Explanatory anti-psychologism overturned by lay and scientific case classifications." *Synthese* 191, 1013-1035.

[Weber & Verhoeven 2002] Weber, Erik and Liza Verhoeven. 2002. "Explanatory proofs in mathematics." *Logique at Analyse* 179-180, 299-307.

[Weyl 1968] Weyl, Hermann. 1968. *Algebraic Theory of Numbers*. Princeton University Press: Princeton, NJ.

[Wilkenfeld 2014] Wilkenfeld, Daniel A. 2014. "Functional explaining: A new approach to the philosophy of explanation." *Synthese* 191, 3367-3391.

[Yap 2011] Yap, Audrey. 2011. "Gauss' quadratic reciprocity theorem and mathematical fruitfulness." *Studies in History and Philosophy of Science* 42, 410-415.

[Zagier 1990] Zagier, Don. 1990. "A one-sentence proof that every prime $p \equiv 1 \pmod 4$ is a sum of two squares." *American Mathematical Monthly* 97, 144.