Mathematical Beauty, Understanding, and Discovery

Carlo Cellucci

Abstract In a very influential paper Rota stresses the relevance of mathematical beauty to mathematical research, and claims that a piece of mathematics is beautiful when it is enlightening. He stops short, however, of explaining what he means by 'enlightening'. This paper proposes an alternative approach, according to which a mathematical demonstration or theorem is beautiful when it provides understanding. Mathematical beauty thus considered can have a role in mathematical discovery because it can guide the mathematician in selecting which hypothesis to consider and which to disregard. Thus aesthetic factors can have an epistemic role *qua* aesthetic factors in mathematical research.

Keywords Mathematics · Beauty · Enlightenment · Understanding

1 Aesthetic Terms and Neuroscience

Hardy claimed that "there is no permanent place in the world for ugly mathematics" (Hardy 1992, 14). When Snow "taxed him and said that if one would present a solution which finally proved Goldbach's theorem, and if it was ugly, would he accept it? Hardy replied, 'This is impossible. It couldn't possibly be a proof of Goldbach's theorem if it were ugly'" (Snow 1973, 812).

Hardy's position was too extreme. There are permanent parts of mathematics that one could hardly say to be beautiful. As an example, Hersh and John-Steiner mention the formula which solves the general quadratic equation. This "is one of the most memorized formulas in math. Not beautiful!" (Hersh and John-Steiner 2011, 61). Nevertheless, from a more moderate point of view, several mathematicians assume that at least some mathematics is beautiful. In particular, they make aesthetic judgments of demonstrations and theorems, qualifying some of them as 'elegant' or 'clumsy', 'beautiful' or 'ugly'.

C. Cellucci (✉)
Department of Philosophy, Sapienza University of Rome, Rome, Italy
e-mail: carlo.cellucci@uniroma1.it

🌀 Springer

Even this more moderate point of view, however, has been disputed. Doubt has been raised whether the apparent aesthetic judgments that mathematicians often make are really aesthetic at all. For example, Todd claims that there are "strong reasons for casting doubt on the aesthetic nature of at least many of the claims made in mathematics and science" (Todd 2008, 63).

Recent findings in the neuroscience of aesthetics are relevant to the question. They have been established using functional magnetic resonance imaging to image the activity in the brains of a number of mathematicians, when they viewed mathematical formulae which they had individually rated as beautiful, indifferent or ugly. The findings indicate that "the experience of mathematical beauty correlates with activity in the same brain area"—the medial orbito-frontal cortex—that is "active during the experience of visual, musical, and moral beauty" (Zeki et al. 2014, 8). Moreover, "the activity there is parametrically related to the declared intensity of the experience of beauty, whatever its source" (ibid.). This does not mean that the medial orbito-frontal cortex alone is responsible for, or underlies, the experience of mathematical beauty. Other brain areas are active during that experience, "distinct from the areas engaged when viewing paintings" or "listening to musical excerpts" (ibid., 10). Nevertheless, the findings suggest that "the activity in a common area of the emotional brain that correlates with the experience of beauty derived from different sources merely mirrors neurobiologically the same powerful and emotional experience of beauty that mathematicians and artists alike have spoken of" (ibid., 8).

However, the findings in the neuroscience of aesthetics only address the question what neural mechanisms allow us to experience beauty. They say nothing about what is mathematical beauty, and in particular about why it is that a demonstration or a theorem is beautiful. The aim of this paper is to try to suggest an answer to these questions.

2 Two Different Traditions about Mathematical Beauty

With respect to the question what is mathematical beauty there are two different traditions.

(1) Mathematical beauty consists of the intrinsic properties of certain mathematical entities, and is therefore independent of the observer and period of history.
(2) Mathematical beauty is a projection of the observer. If certain mathematical entities exhibit properties that are valued by the aesthetic criteria of the observer, such properties will be called 'aesthetic properties' and the observer will project beauty on those mathematical entities and will describe them as beautiful. Thus mathematical beauty is strongly dependent upon the observer and period of history.

The first tradition goes back to antiquity, the second one has been widespread in the modern and contemporary age.

3 Mathematical Beauty as an Intrinsic Property

An eminent representative of the first tradition about mathematical beauty is Plato. He states that "straight lines and circles, and the plane and solid figures which are formed out of them by means of compasses, rulers and squares" are "not, as other things are, beautiful in a relative way, but they are by their very nature forever beautiful by themselves" (Plato, *Philebus*, 51 c 3–d 1). Thus they are "not beautiful at one time and not at another, or beautiful by one standard and ugly by another, or beautiful in one place and ugly in another because" they are

"beautiful to some people but ugly to others" (Plato, *Symposium*, 211 a 3–5). Instead, they are of a kind of beauty which "always is and does not come into to be or perish, nor does it grow or wane" (ibid., 211 a 1–2). We may obtain knowledge of this kind of beauty only by intellectual intuition, by which alone we may arrive at "that particular knowledge which is knowledge solely of the beautiful itself" (ibid., 211 c 7–8).

The first tradition about mathematical beauty has had a large following among mathematicians and scientists generally. For example, Dirac states that, while "beauty does depend on one's culture and upbringing for certain kinds of beauty, pictures, literature, poetry, and so on," mathematical beauty "is of a completely different kind and transcends these personal factors. It is the same in all countries and at all periods of time" (Kragh 1990, 288).

However, that mathematical beauty is the same in all countries and at all periods of time is somewhat problematic, because often what is considered mathematically beautiful in a certain age or culture is not so considered in another age or culture. In particular, philosophers and mathematicians have given different characterizations of mathematical beauty.

Thus Plato states that "measure and proportion manifest themselves in all areas as beauty and virtue" (Plato, *Philebus*, 64 e 6–7). For "nothing beautiful lacks proportion" (Plato, *Timaeus*, 87 c 5).

Aristotle states that "the supreme forms of beauty are order, symmetry, and definiteness, which the mathematical sciences demonstrate in a special degree" (Aristotle, *Metaphysica*, M 3, 1078 a 36–b 2).

Poincaré states that "the mathematical entities to which we attribute" the "character of beauty and elegance" are "those whose elements are harmoniously arranged," and "this harmony is at once a satisfaction to our aesthetic requirements, and an assistance to the mind" (Poincaré 1914, 59).

Hardy states that in beautiful theorems "there is a very high degree of unexpectedness, combined with inevitability and economy" (Hardy 1992, 113).

Hersh and John-Steiner state that "we can point to three important elements of beauty in mathematical content: simplicity, concrete specificity, and unexpected or surprising integration or connection of disparate elements" (Hersh and John-Steiner 2011, 60).

Proportion, order, symmetry, definiteness, harmony, unexpectedness, inevitability, economy, simplicity, specificity, integration are different properties. This explains Hofstadter's claim that "there exists no set of rules which delineate what it is that makes a piece beautiful, nor could there ever exist such a set of rules" (Hofstadter 1999, 555).

4 Mathematical Beauty as a Projection of the Observer

An eminent representative of the second tradition about mathematical beauty is Kant. As Breitenbach argues, contrary to the widespread view that "Kant's aesthetics leaves no room for beauty in mathematics," Kant holds that, "while mathematical properties themselves are not beautiful, it is the demonstration of such properties that can be the object of aesthetic appreciation" (Breitenbach 2013b, 2).

Indeed, according to Kant, "beauty is not a quality of the objects considered for itself" (Kant 2000, 221). The experience of mathematical beauty does not consist, as Plato claims, in an intellectual intuition into properties of mathematical entities, but "contains merely a relation of the representation of the object to the subject" (ibid., 97). It consists in our subjective emotional response to the demonstrations of such properties, which show a "harmony of the two cognitive faculties, the sensibility and the understanding" (ibid., 198). We may "call a demonstration of such properties beautiful, since by means of this the understanding, as the

faculty of concepts, and the imagination, as the faculty of exhibiting them, feel themselves strengthened a priori (which, together with the precision which is introduced by reason, is called its elegance): for here at least the satisfaction" is "subjective" (ibid., 239). Since the representations are "rational but related in a judgment solely to the subject (its feeling)," they "are to that extent always aesthetic" (ibid., 90).

5 Rota's Phenomenology of Mathematical Beauty

The second tradition about mathematical beauty has several representatives. Of course, they need not share all of Kant's views about mathematical beauty, but only the view that beauty is not a quality of the objects considered for themselves and is, rather, a projection of the observer. Some recent representatives of the second tradition are Breitenbach (2013a), McAllister (2005), Rota (1997), Sinclair (2004, 2006, 2011). I will give some details about Rota's position because this will be useful as reference in what follows.

Rota states that "the beauty of a piece of mathematics is strongly dependent upon schools and periods of history. A theorem that is in one context thought to be beautiful may in a different context appear trivial" (Rota 1997, 175). Thus mathematical beauty is dependent upon the context. Appreciating it requires "familiarity with a huge amount of background material" (Rota 1997, 179). For example, "a proof is viewed as beautiful only after one is made aware of previous clumsier proofs" (ibid.). Familiarity with a huge amount of background material "is arrived at the cost of time, effort, exercise and Sitzfleisch" (ibid., 177). Therefore, the appreciation of mathematical beauty cannot be instantaneous. We must avoid the "light-bulb mistake" which consists in believing that mathematical beauty is "appreciated with the instantaneousness of a light bulb being lit" and hence its appreciation is "an instantaneous flash" (ibid., 179).

Beauty plays a positive role in the development of mathematics, because "the lack of beauty in a piece of mathematics is of frequent occurrence, and it is a strong motivation for further mathematical research" (ibid., 178). An even "cursory look at any mathematics research journal will confirm" that "much mathematical research consists precisely of polishing and refining statements and proofs of known results" (ibid.). This arises because a mathematician may have been able "to follow the proof of a statement" and hence "to formally verify its truth in the logical sense of the term, but is still missing something," that is, "the sense of the statement" (ibid., 181). Indeed, "the mere logical truth of a statement does not enlighten us to the sense of the statement. Enlightenment, not truth, is what the mathematician seeks" (ibid.).

Nevertheless, "the phenomenon of enlightenment is seldom explicitly acknowledged among mathematicians" (ibid.). For neither mathematical beauty nor mathematical truth "admits degrees" (ibid., 180). Conversely, "enlightenment admits degrees; some statements are more enlightening than others," and "mathematicians universally dislike any concepts admitting degrees" (ibid., 181). For this reason, mathematicians talk of 'mathematical beauty', but this is only a trick that they "have devised to avoid facing up to the messy phenomenon of enlightenment" (ibid., 182). Mathematicians "say that a proof is beautiful when such a proof finally gives away the secret of the theorem, when it leads" them "to perceive the actual, not the logical inevitability of the statement that is being proved" (ibid.). They say that a theorem is beautiful when they see "how the theorem 'fits' in its place, how it sheds light around itself, like a *Lichtung*, a clearing in the woods" (ibid.). But when they say so, "they really mean to say that the theorem is enlightening" (ibid., 181).

6 Some Limitations of Rota's Views

Despite its suggestiveness, Rota's approach to mathematical beauty has some limitations.

(1) Rota claims that beauty plays a positive role in the development of mathematics because the lack of beauty in a piece of mathematics is a strong motivation for further mathematical research. But sometimes the quest for beauty may be an obstacle to the development of science. For example, the current theory of quantum electrodynamics uses a mathematical procedure, renormalization, which leads to accurate predictions but is mathematically incorrect. Thus Dirac states that, although renormalization "has been very successful in setting up rules for handling the infinities and subtracting them away," the "resulting theory is an ugly and incomplete one," and hence "cannot be considered as a satisfactory solution of the problem of the electron" (Dirac 1951, 291). The theory "does involve neglecting infinities which appear in its equations, neglecting them in an arbitrary way. This is just not sensible mathematics. Sensible mathematics involves neglecting a quantity when it is small—not neglecting it just because it is infinitely great and you do not want it!" (Dirac 1978, 36). Therefore Dirac developed an alternative theory, but the theory was unsatisfactory since it did not lead to accurate predictions.

(2) Rota claims that, when mathematicians say that a piece of mathematics is beautiful, they really mean to say that it is enlightening. But Rota stops short of explaining what he means by 'enlightening'.

(3) Rota claims that, unlike enlightenment, beauty does not admit degrees, and mathematicians dislike any concepts admitting degrees. But beauty, including mathematical beauty, does admit degrees, as it is clear from the fact that we commonly say that something is more beautiful than something else. Further evidence for the fact that mathematical beauty admits degrees is provided by a poll of readers of *The Mathematical Intelligencer* who ranked 24 theorems, on a scale from 0 to 10, for beauty (see Wells 1990).

7 Beauty and Perception

Before trying to suggest how to avoid the limitations of Rota's approach to mathematical beauty, a preliminary objection must be considered. The objection is that perceptual experience is central to beauty, while the recognition of something that might count as mathematical beauty has nothing to do with perception.

Thus Zangwill states that, "as the etymological origins of the word 'aesthetic' suggest, aesthetic properties are those that we appreciate in perception. Lovers of beauty are 'lovers of sight and sounds'" (Zangwill 1998, 81). Indeed, "aesthetic properties are properties which something has only if it has sensory properties" (ibid., 66). Conversely, "proofs, theories" do "not necessarily have any sensory embodiment" (ibid., 78). Therefore, when we say that a proof is elegant or beautiful, this "is not genuine aesthetic appreciation," but "aesthetic terms are metaphorically applied in these cases" (ibid., 79).

Similarly, van Gerwen states that "beauty centrally involves a perceptual experience" (van Gerwen 2011, 250). For "beauty is proven to exist in perception" (ibid., 257). Conversely, "the recognition of something that might count as mathematical beauty" has "nothing to do with perception" (ibid., 259). Therefore, "to speak of mathematical beauty is to speak in a loose manner" (ibid., 264).

The objection, however, is unjustified because a perceptual experience need not be central to beauty. As Kant states, "that is beautiful which pleases in the mere judging (neither in

sensation nor through a concept)" (Kant 2000, 185). If that is beautiful which pleases in the mere judging, a perceptual experience need not be central to beauty.

The objection would imply not only that there is no mathematical beauty, but also that there is no literary beauty, since the recognition of something that might count as literary beauty has little to do with perception. Admittedly, to read a novel, I must be able to see the text in front of me, but one could hardly say that my aesthetic experience in reading a novel consists in deciphering the words on the page in front of me. Nor do I literally see the facts narrated in the novel, I only imagine them, and imagination is the capacity to represent something even when it is not itself present before my eyes.

Zangwill claims that "if a literary work has aesthetic properties, they derive from the particular choice of words, because of the way they sound," and "if a literary work has values which are not linked" to "the sonic properties of words, then they are not aesthetic values" (Zangwill 1998, 75). But, since there is no special problem in translating a literary work in other languages so that its values are preserved, a literary work has values which are not linked to the sonic properties of words. Thus, if Zangwill were right, no literary work would have aesthetic values. But it would be difficult to deny that literary works have aesthetic values. Therefore, the assumption that a perceptual experience is central to beauty does not seem to be justified.

Of course, sculptural beauty, painterly beauty, musical beauty, literary beauty, mathematical beauty, etc., are all different kinds of beauty. But this means that there are different kinds of beauty even in the realm of beauty derived from more perceptually based sources.

8 From Enlightenment to Understanding

As already stated, a limitation of Rota's approach to mathematical beauty is that Rota stops short of explaining what he means by 'enlightening'. In fact, 'enlightening' is not the most appropriate word to use in this context. For 'enlighten' means 'to give insight', and 'insight' suggests that mathematical beauty is appreciated with the instantaneousness of a light bulb being lit. This contrasts with Rota's warning that we must avoid the light bulb mistake.

It seems more appropriate to state that, when mathematicians say that a piece of mathematics is beautiful, they mean to say that it gives understanding. In fact a relation, not between mathematical beauty and understanding, but between beauty in natural science and understanding, is established by Kosso and Breitenbach. Thus Kosso argues that there is "a link between beauty and scientific understanding" (Kosso 2002, 40). Breitenbach argues that, "in searching for beauty, scientists aim for theories that provide understanding" (Breitenbach 2013a, 85).

It may seem odd to state that when mathematicians say that a piece of mathematics is beautiful they mean to say that it gives understanding, because understanding is not commonly associated with beauty. But here by 'understanding' I mean the recognition of the fitness of the parts to each other and to the whole. Understanding thus meant is an aesthetic property because, according to a tradition going back to antiquity, one aspect of the appreciation of beauty in a work of art is the recognition of the fitness of the parts to each other and to the whole.

As Heisenberg points out, in antiquity "there were two definitions of beauty, which stood in a certain opposition to one another" (Heisenberg 2001, 57). The "one describes beauty as the proper conformity of the parts to one another, and to the whole. The other, stemming from Plotinus, describes it, without any reference to the parts, as the translucence of the eternal splendor of the 'one' through the material phenomenon" (ibid.). Heisenberg supports "the

first and more sober definition of beauty" (ibid., 69). He also establishes a relation between beauty and understanding. For he states that, "if the beautiful is conceived as a conformity of the parts to one another and to the whole, and if, on the other hand, all understanding is first made possible by means of this formal connection, the experience of the beautiful becomes virtually identical with the experience of connections either understood or, at least, guessed at" (ibid., 59).

Understanding, meant as the recognition of the fitness of the parts to each other and to the whole, is context-dependent. In particular, it is dependent upon the mathematician's background knowledge, so it is usually recognizable only to the well trained. Appreciating mathematical beauty requires apprehending the ideas involved in a way that reveals the fitness of the parts to each other and to the whole, and this is something that a poor background knowledge cannot provide.

Of course, the fitness of the parts to each other and to the whole means different things in art, music, literature or mathematics, and, within mathematics, in demonstrations or theorems. Therefore, each case requires special consideration.

9 Beauty in Works of Art

In works of art, a kind of fitness of the parts to each other and to the whole is expressed by the golden ratio. Two quantities are said to be in the golden ratio (φ) if the sum of the two quantities is to the larger quantity as the larger quantity is to the smaller. That is, a and b, with $a > b$, are in the golden ratio if $\frac{a+b}{a} = \frac{a}{b} = \varphi$. Since $\frac{a+b}{a} = 1 + \frac{b}{a} = 1 + \frac{1}{\varphi}$, we have $1 + \frac{1}{\varphi} = \varphi$, so $\varphi + 1 = \varphi^2$, hence $\varphi^2 - \varphi - 1 = 0$. The positive solution of this equation is $\varphi = \frac{1+\sqrt{5}}{2} = 1.61805$, which gives the value of the golden ratio to the first five decimal places.

The golden ratio is observed in many works of art. Thus, in Leonardo da Vinci's *Mona Lisa*, the ratio of the height to the width of the canvas is the golden ratio. The proportions of Mona Lisa's body exhibit several golden ratios. For example, the ratio of the height to the width of Mona Lisa's face—from the top of the forehead to the base of the chin, and from left cheek to right cheek, respectively—is the golden ratio.

In Botticelli's *The Birth of Venus*, the ratio of the width to the height of the canvas is very nearly the golden ratio. The proportions of Venus' body exhibit several golden ratios. For example, the ratio of Venus' height overall to the height of her navel is the golden ratio.

10 Beauty in Demonstrations

Of course, the beauty of a demonstration cannot be expressed in terms of the golden ratio, it involves another kind of relationship between the parts of the demonstration. There is fitness of the parts of a demonstration to each other and to the whole when it is clear what the whole idea of the demonstration is, what the contribution of each part of the demonstration to the whole idea is, and why such contribution is essential. Then the demonstration provides understanding. Indeed, it shows why the statement that is being demonstrated holds and so, as Rota states, it leads us to perceive the actual, not the logical inevitability of the statement that is being demonstrated.

This notion of fitness is significant, as it is apparent from the statements of several outstanding mathematicians.

Thus Poincaré states that, even when the mathematician has cut up "each demonstration into a very great number of elementary operations" and has "ascertained that each is correct," he may not "have grasped the real meaning of the demonstration" since he may not have seized that "which makes the unity of the demonstration" (Poincaré 1958, 21–22).

Mordell states that, "even when a proof has been mastered" and "may be strictly logical and convincing," the "reader may feel that something is missing. The argument may have been presented in such a way as to throw no light on the why and wherefore of the procedure or on the origin of the proof or why it succeeds" (Mordell 1959, 11).

Rota, as we have seen, states that a mathematician may have been able to follow the proof of a statement and hence to formally verify its truth in the logical sense of the term, but is still missing something.

What is missing is the fitness of the parts of the demonstration to each other and to the whole, in the sense explained above. Only a demonstration which shows such fitness provides understanding. (There is a strict relation between demonstrations providing understanding and demonstrations providing explanation; see Cellucci 2014b).

11 An Example of a Beautiful Demonstration

To give a simple example of a beautiful demonstration we consider the problem: Construct a square which has half the area of a given square. A solution is given by the following diagram, where the given square is the larger square, and the middle square—the one looking like an equilateral kite—is the square which has half the area of a given square.
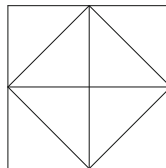
From the diagram it is apparent that the given square consists of eight equal right triangles, while the middle square consists of four equal right triangles. This gives a demonstration that the middle square has half the area than the larger square. The demonstration is beautiful because it provides understanding, in the sense explained in the previous section.

Essentially the same demonstration is obtained by folding the corners of the larger square towards the center of the square.

The same diagram is used in Plato's *Meno* to solve the inverse problem: Construct a square which has twice the area of a given square. Let the given square be one of the four smallest squares of which the larger square consists. From the diagram it is apparent that the given square consists of two equal right triangles, while the middle square, the one looking like an equilateral kite, consists of four equal right triangles. This provides a demonstration that the middle square has twice the area of the given square. The demonstration is beautiful because it provides understanding.

It might be objected that demonstrations based on diagrams are not genuine demonstrations. But it may be argued that this objection is unwarranted (see Cellucci 2013, Sect. 21.11).

12 Differences in Beauty Between Geometrical Demonstrations

There are significant differences between demonstrations in terms of understanding. This has been pointed out by several mathematicians. Thus Atiyah states: "I remember one theorem that I proved, and yet I really couldn't see why it was true," but "five or six years later I understood why it had to be true. Then," using "quite different techniques," I "got an entirely different proof" that made "quite clear why it had to be true" (Atiyah 1988, I, 305). Auslander states: "A proof is supposed to explain the result. Now, it must be admitted that not all proofs meet this standard" and this has "often led to the development of new, more understandable, proofs" (Auslander 2008, 66). Nielsen states: "Sometimes, finding new proofs can give us significant new insights that help us understand why a result is true in the first place" (Nielsen 2011, 198).

Since there are significant differences between demonstrations in terms of understanding, between them there are significant differences in beauty. To give an example of this let us consider the problem: Show that the square on the hypotenuse of a right triangle is equal to the sum of the squares on the other two sides (the Pythagorean Theorem).

Demonstration 1 (Euclid, *Elements*, I.47). Describe the square BDEC on BC, and the squares GB and HC on BA and AC, respectively. Draw AL parallel to either BD or CE, and draw AD and FC.
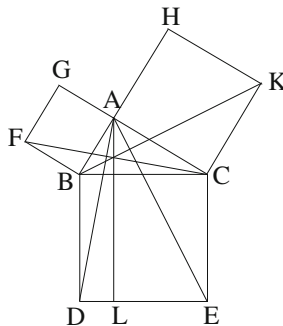
The triangle ABD is equal to the triangle FBC because DB=BC, FB=BA and angle ABD=angle FBC (since both angles consist of angle ABC plus a right angle).

On the other hand, the triangle ABD is half the parallelogram BL because they have the same base BD and are in the same parallels BD and AL. In the same way, the triangle FBC is half the square GB because they have the same base FB and are in the same parallels FB and GC.
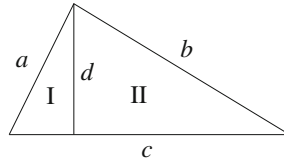
Since the triangle ABD is equal to the triangle FBC, it follows that the parallelogram BL is equal to the square GB.

Similarly, it can be shown that the parallelogram CL is equal to the square HC.

Therefore, the whole square BDEC is equal to the sum of the two squares GB and HC.

Demonstration 2. (Euclid, *Elements*, VI.31). Let c be the length of the hypotenuse of the right triangle, and let a, b be the lengths of the other two sides. We must show that $c^2 = a^2 + b^2$.

Drop a perpendicular d from the right angle to the hypotenuse. The perpendicular divides the right triangle into two right triangles. Let I be the right triangle with hypotenuse a and side d, let II be the right triangle with hypotenuse b and side d, and let III be the whole right triangle.

The triangle I is similar to the triangle III because they have two equal angles, that is, a right angle and the angle formed by a and c. The triangle II is similar to the triangle III because they have two equal angles, that is, a right angle and the angle formed by b and c. Since the relation of similarity is symmetric and transitive, from this it follows that the triangles I, II, and III are all similar.

Now, the areas of similar triangles are proportional to the square of the corresponding sides. Then the areas of the triangles I, II, and III are proportional to the square of their hypotenuses, that is, for some k, area of III $= kc^2$, area of I $= ka^2$, and area of II $= kb^2$. On the other hand, area of III = area of I + area of II. Thus $kc^2 = ka^2 + kb^2$, and hence $c^2 = a^2 + b^2$.

Demonstration 1 is rather clumsy. Thus Schopenhauer states: "Euclid's stilted, indeed underhand, proof leaves us without an explanation of why" (Schopenhauer 2010, 98). In that proof, "lines are often drawn without any indication of why" and the reader "must admit in astonishment what remains completely incomprehensible in its inner workings" (ibid., 96). Similarly, Rav states: "Euclid's proof is a tour de force to fit a preset methodology of a purely geometric, deductive argument," so it "hides the heuristic path to the discovery and a more intuitive proof of the Pythagorean theorem" (Rav 2005, 52).

Conversely, Demonstration 2 is beautiful because it provides understanding. Indeed, it shows why the square on the hypotenuse of a right triangle is equal to the sum of the squares on the other two sides—the right triangle on its own hypotenuse is similar to the right triangles on the other two sides, and the areas of similar triangles are proportional to the square of their corresponding sides.

13 Differences in Beauty Between Non-Geometrical Demonstrations

There are significant differences also between non-geometrical demonstrations in terms of understanding. To give an example of this let us consider the problem: Show that $\sqrt{2}$ is not a fraction.

Demonstration 1 (Euclid, *Elements*, X, Appendix 27). Suppose that $\sqrt{2} = p/q$ where p and q are integers. We may assume that p and q have no common factor. (If they do, we cancel it). Then $2 = p^2/q^2$, so $p^2 = 2q^2$. Thus p^2 is even. Therefore p itself is even (for the square of an odd number is odd). Then $p = 2r$ for some r, so $(2r)^2 = 2q^2$, that is, $4r^2 = 2q^2$, hence $q^2 = 2r^2$. Thus q^2 is even, and therefore, as above, q itself is even. So p and q are both even, hence they have a common factor. But by our assumption p and q have no common factor. Contradiction. Therefore $\sqrt{2}$ cannot be a fraction.

Demonstration 2. Suppose that $\sqrt{2} = p/q$ for two integers p and q. Then $2 = p^2/q^2$, so $p^2 = 2q^2$. Since every integer > 1 can be represented as a product of primes, and this representation is unique, apart from the order of the factors, we may assume that p and q have been so represented. Since $p^2 = p \cdot p$, in the representation of p^2 every prime will occur an even number of times. Similarly in the representation of q^2. Then in the representation of $2q^2$ the prime number 2 will occur an odd number of times. Since $p^2 = 2q^2$ and the representation is unique, this is impossible. Therefore $\sqrt{2}$ cannot be a fraction.

As Davis and Hersh state, Demonstration 2 "exhibits a higher level of aesthetic delight" because it "exposes the 'real' reason" and hence "seems to reveal the heart of the matter," while Demonstration 1 "conceals it" (Davis and Hersh 1981, 299–300). In fact Demonstration 2 is beautiful because it provides understanding. It shows why $\sqrt{2}$ cannot be a fraction—the parities of the exponent of 2 in the representations of p^2 and q^2 as a product of primes will be different.

Note that this is not in contrast with the fact that Hardy cites the theorem that $\sqrt{2}$ is not a fraction as an example of a 'serious' theorem, that is, "a theorem which connects significant ideas" (Hardy 1992, 16). Then Hardy cites the theorem as an example of a beautiful theorem, because "the beauty of a mathematical theorem depends a great deal on its seriousness" (ibid., 17). He does not cite Demonstration 1 as an example of a beautiful demonstration.

14 An Example of a Beautiful Theorem

The beauty of a theorem involves another kind of relationship between the parts of the theorem. There is fitness of the parts of a theorem to each other and to the whole when the concepts involved in the theorem are basic and the theorem establishes some basic relation between them in a transparent way.

To give an example of a beautiful theorem let us consider Euler's identity, $e^{i\pi} + 1 = 0$, or equivalently $e^{i\pi} = -1$. The poll of readers of *The Mathematical Intelligencer* mentioned in Sect. 6, named Euler's identity the most beautiful theorem in mathematics (Wells 1990, 38). In the functional magnetic resonance imaging by which the activity in the brains of a number of mathematicians has been imaged, "the formula most consistently rated as beautiful," both "before and during the scans, was Leonhard Euler's identity" (Zeki et al. 2014, 4). Nahin even claims that Euler's identity sets "the gold standard for mathematical beauty" and "will still appear, to the arbitrarily advanced mathematicians ten thousand years hence, to be beautiful and stunning and untarnished by time" (Nahin 2006, xxxii).

Euler's identity is beautiful because it provides understanding. Indeed, it involves five of the most important numbers in mathematics, 0, 1, i, π, e, along with the fundamental concepts of addition, multiplication and exponentiation, and establishes a basic relation between them in a transparent way. More generally, it establishes a relation between what Halmos calls "the three major parts of mathematics," namely, "algebra (a symbolic outgrowth of arithmetic), geometry (the study of shape), and analysis (the abstract version of calculus)" (Halmos 1992, 1325). For 0, 1, i come from algebra, π comes from geometry, and e comes from analysis—since $e = \lim_{n \to \infty} (1 + \frac{1}{n})^n$.

To see how Euler's identity comes about, observe that we can represent powers of the complex number $1 + \frac{i\pi}{n}$, $(1 + \frac{i\pi}{n})^2$, $(1 + \frac{i\pi}{n})^3$, ..., $(1 + \frac{i\pi}{n})^n$ by drawing similar right triangles that share an edge, where 1 is the length of the first spoke and $\frac{\pi}{n}$ the length of first edge. For example, for $n = 6$, this yields the spiral:
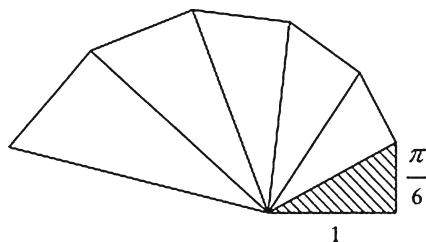
As we increase the number n of similar right triangles, the length of the hypotenuse of each right triangle becomes very close to the length of the previous spoke. So the spiral becomes closer to circular. For example, for $n = 24$, we have:

1
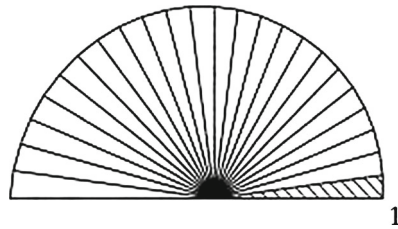
Taking n to the limit, we get an infinite number of spokes that are all the same size, taking us around the unit circle to the opposite side, -1. Thus $\lim_{n \to \infty} (1 + \frac{i\pi}{n})^n = -1$.

-1   1

On the other hand, from the definition of e it follows that $e^{i\pi} = \lim_{n \to \infty} (1 + \frac{i\pi}{n})^n$. Therefore $e^{i\pi} = -1$. This gives away the secret of Euler's identity. (For a similar explanation of Euler's identity, see Conway and Guy 1996, 254–256).

15 Beauty and Discovery

Although the quest for beauty may sometimes be an obstacle to the development of science, it may also have a positive role in the development of mathematics, not only as a motivation for finding better demonstrations of known results, but more importantly as a part of the process of discovery.

In the last century it was widely held that one must distinguish sharply between the context of discovery and the context of justification, because the context of discovery "evades distinct analysis" (Reichenbach 1947, 1). Therefore "the context of discovery is left to psychological analysis," and epistemology "is concerned with the context of justification" (ibid., 2).

This view, however, is unjustified because the context of discovery does not evade distinct analysis, it can be analyzed in terms of the analytic method—a method first used by Hippocrates of Chios to solve the problem of the duplication of the cube and the problem of the quadrature of certain lunules, and first explicitly stated by Plato.

According to the analytic method, to solve a problem, one looks for some hypothesis from which a solution to the problem can be derived. The hypothesis is obtained from the problem, and possibly other data already available, by some non-deductive rule, it need not belong to the same field as the problem, and must be plausible, that is, compatible with the existing data. But the hypothesis is in its turn a problem that must be solved, and is solved in the same way, and so on. In the analytic method there are no principles, everything is a hypothesis. The problem and the other data already available are the only basis for solving the problem. Solving a problem is both a process of discovery and a process of justification, because it involves finding hypotheses by non-deductive rules and checking that the hypotheses thus

found are plausible. (I speak of plausibility rather than truth because, as I have argued in Cellucci 2014a, the goal of science is not truth but plausibility. For details about the analytic method, see Cellucci 2013, Chap. 4).

Since in the analytic method the hypotheses are obtained by non-deductive rules, this shows the limit of the widespread belief that "the two main pillars of mathematics are deductive reasoning and abstraction" and the Greeks "essentially created mathematics, as we know it today, based on these two pillars" (Ó Cairbre 2009, 42). As it is apparent from Hippocrates of Chios, the Greeks essentially created mathematics, as we know it today, based on the analytic method.

However, the non-deductive rules by which hypotheses are obtained in the analytic method may yield so many hypotheses that it could be unfeasible to check all of them for plausibility. The sense of beauty may guide us in selecting which hypothesis to check and which to disregard, thus acting as a shortcut.

Avigad states that "mathematics presents us with a combinatorial explosion of options; understanding helps us sift through them, and pick out the ones that are worth pursuing" (Avigad 2008, 320). Indeed, "when we do mathematics, we are like Melville's sailors, swimming in a vast expanse. Just as the sailors cling to sides of their ship, we rely on our understanding to guide us and support us" (ibid.). However, Avigad views understanding as based on a "moment of insight" which occurs when "I am working through a difficult proof" and "I struggle to make sense of it," then "all of a sudden, something clicks, and everything falls into place—now I understand" (ibid., 322). Now, saying that understanding is based on a moment of insight amounts to committing the light bulb mistake denounced by Rota.

If we agree that understanding is the non-instantaneous recognition of the fitness of the parts to each other and to the whole, then understanding, hence beauty, helps us sift through the hypotheses generated by non-deductive rules and pick out the ones that are worth pursuing. As Breitenbach states, "because of the link to our capacities of understanding, following beauty can provide a heuristic means for choosing between theories" (Breitenbach 2013a, 94). Thus beauty may have a role in the context of discovery. (This is an aspect of the role emotion may have in that context; see Cellucci 2013, Chapter 15). That beauty may have a role in the context of discovery is already pointed out by Poincaré. According to him, "mathematical discovery" consists "in making new combinations" with concepts "that are already known" and in selecting "those that are useful" (Poincaré 1914, 50–51). The selection is made based on the "feeling of mathematical beauty," which is "a real aesthetic feeling that all true mathematicians recognize" (ibid., 59).

That the sense of beauty guides us in selecting which hypothesis to check and which to disregard is because the hypotheses thus selected provide understanding. On the other hand, that they provide understanding does not guarantee that they are plausible. The hypotheses may let us recognize the fitness of the concepts involved to each other and to the whole, but the concepts involved may overlook or confuse some aspect of the problem the hypotheses are intended to solve. Then the hypotheses may not be plausible. Such was the case of Frege's hypothesis that two functions have equal extensions if and only if they map every object to the same value. Frege's hypothesis was beautiful because it provided understanding of "what people have in mind, for example, where they speak of the extensions of concepts" (Frege 1964, 4). But it was not plausible, as it is apparent from the fact that it led to Russell's paradox.

Indeed, selecting hypotheses is only one part of the analytic method, another essential part is ascertaining that the hypotheses thus selected are plausible. This requires a separate argument. Dirac claims that "if one is working from the point of view of getting beauty in one's equations, and if one has really a sound insight, one is on a sure line of progress" (Dirac 1963, 47). But it is not so, for neither beauty nor a sound insight can guarantee that the

hypotheses selected on the basis of them are plausible. Establishing that they are plausible essentially involves a comparison with the existing data. Poincaré too recognized that, once useful combinations have been selected on the basis of the feeling of mathematical beauty, "it is necessary to verify them" (ibid., 56).

On the other hand, this by no means diminishes the role of the sense of beauty in selecting hypotheses. As already stated, this role depends on the fact that the non-deductive rules by which hypotheses are obtained in the analytic method may yield so many hypotheses that it could be unfeasible to check all of them for plausibility.

16 An Example of the Role of Beauty in Discovery

To give a simple example of the role of mathematical beauty in the context of discovery, I consider a likely story of how an unknown Egyptian mathematician found that the area of a triangle is half the base times the height. A likely story, because, as it is generally the case with discoveries, there is no report as to how discovery did happen. Nevertheless, the story seems to be authorized by the description of Problem 51 in the Rhind Papyrus.
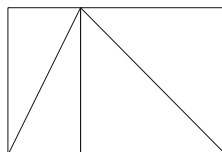
The Egyptian mathematician observes that, from a right-angled triangle, one can "get its rectangle" (Clagett 1999, 163). That is, from a right-angled triangle one can get a rectangle with the same base and the same height.

Then the Egyptian mathematician halves the area of the rectangle. Since the area of the rectangle is the base times the height, the area of the right-angled triangle will be half the base times the height.

From this several hypotheses might be inferred, for example, that the area of the isosceles triangle will be half the base times the height. But the Egyptian mathematician selects the more general hypothesis—obtained by induction from a single case—that the area of any triangle is half the base times the height. He selects this hypothesis because it is the aesthetically most appealing one, since it establishes a basic relation between three basic parameters of any triangle, base, height, and area in a transparent way, and thus provides understanding.

Not only such hypothesis is the aesthetically most appealing one, but it is easy to see that it is plausible. To this purpose the Egyptian mathematician observes that, dropping a perpendicular, any triangle can be divided into two right-angled triangles. On the other hand, as he has already observed, from each of these two right-angled triangles one can get a rectangle with the same base and the same height. Therefore, from the whole triangle itself one can get a rectangle with the same base and the same height.

Then the Egyptian mathematician halves the area of such rectangle. Since the area of the rectangle is the base times the height, the area of the whole triangle will be half the base times the height.

This gives a demonstration of the hypothesis. The demonstration is beautiful because it provides understanding. Indeed, it shows why the area of a triangle is half the base times the height.

17 Epistemic Role of Aesthetic Factors

The quest for beauty has often been a motivation for doing research in mathematics. Indeed, von Neumann even states that the mathematician's "criteria of selection, and also those of success, are mainly aesthetical" (von Neumann 1961, 2062).

This does not mean that in mathematics the quest for beauty is an end in itself. On the contrary, it is instrumental to the development of mathematics. This fact is often overlooked or denied. For example, Todd states that "aesthetic judgements and the evaluation of scientific theories are odd bedfellows, and their conjunction a just object of suspicion" (Todd 2008, 62). Indeed "aesthetic appreciation and epistemic satisfaction are distinct," and hence one must "avoid collapsing them into each other" (ibid., 75).

But this amounts to confining epistemology to the evaluation of scientific theories, hence to the context of justification, assuming that aesthetic factors cannot have any epistemic role *qua* aesthetic factors. Indeed, if epistemology is confined to the context of justification, then aesthetic factors cannot have any epistemic role at all, not only *qua* aesthetic factors. But, as argued above, epistemology need not be confined to the context of justification. Mathematical beauty can have a role in the context of discovery, because it can guide us in selecting which hypothesis to consider and which to disregard. Therefore, the aesthetic factors can have an epistemic role *qua* aesthetic factors.

References

Atiyah, M. (1988). *Collected works*. Oxford: Oxford University Press.

Auslander, J. (2008). On the roles of proof in mathematics. In B. Gold & R. A. Simon (Eds.), *Proof and other dilemmas. Mathematics and philosophy* (pp. 62–77). Washington: The Mathematical Association of America.

Avigad, J. (2008). Understanding proofs. In P. Mancosu (Ed.), *The philosophy of mathematical practice* (pp. 317–353). Oxford: Oxford University Press.

Breitenbach, A. (2013a). Aesthetics in science. A Kantian proposal. *Proceedings of the Aristotelian Society*, *113*(1), 83–100.

Breitenbach, A. (2013b). Beauty in proofs. Kant on aesthetics in mathematics. European Journal of Philosophy, 1–23. doi:10.1111/ejop.12021.

Cellucci, C. (2013). *Rethinking logic. Logic in relation to mathematics, evolution, and method*. Berlin: Springer.

Cellucci, C. (2014a). Knowledge, truth, and plausibility. *Axiomathes*. doi:10.1007/s10516-014-9238-7.

Cellucci, C. (2014b). Explanatory and non-explanatory demonstrations. In P. E. Bour, G. Heinzmann, W. Hodges, & P. Schroeder-Heister (Eds.), *Logic, methodology and philosophy of science*. Proceedings of the fourteenth international congress London: College Publications (in press).

Clagett, M. (1999). *Ancient Egyptian science. A source book, vol. 3: Ancient Egyptian mathematics*. Philadelphia: American Mathematical Society.

Conway, J. H., & Guy, R. K. (1996). *The book of numbers*. New York: Copernicus.

Davis, P. J., & Hersh, R. (1981). *The mathematical experience*. Boston: Birkhäuser.

Dirac, P. A. M. (1951). A new classical theory of electrons. *Proceedings of the Royal Society of London, Series A*, *209*, 291–296.

Dirac, P. A. M. (1963). The evolution of the physicist's picture of nature. *Scientific American*, *208*(5), 45–53.

Dirac, P. A. M. (1978). *Directions in physics*. New York: Wiley.

Frege, G. (1964). *The basic laws of arithmetic. Exposition of the system*. Berkeley: University of California Press.

Halmos, P. (1992). Mathematics. In C. G. Morris (Ed.), *Academic Press dictionary of science and technology* (p. 1325). San Diego: Academic Press.

Hardy, G. H. (1992). *A mathematician's apology*. Cambridge: Cambridge University Press.

Heisenberg, W. (2001). Science and the beautiful. In K. Wilber (Ed.), *Quantum questions* (pp. 56–69). Boston: Shambhala.

Hersh, R., & John-Steiner, V. (2011). *Loving + hating mathematics. Challenging the myths of mathematical life*. Princeton: Princeton University Press.

Hofstadter, D. R. (1999). *Gödel, Escher, Bach. An eternal golden braid*. New York: Basic Books.

Kant, I. (2000). *Critique of the power of judgment*. Cambridge: Cambridge University Press.

Kosso, P. (2002). The omniscienter. Beauty and scientific understanding. *International Studies in the Philosophy of Science*, *16*(1), 39–48.

Kragh, H. S. (1990). *Dirac. A scientific biography*. Cambridge: Cambridge University Press.

McAllister, J. W. (2005). Mathematical beauty and the evolution of the standards of mathematical proof. In M. Emmer (Ed.), *The visual mind II* (pp. 15–34). Cambridge: The MIT Press.

Mordell, L. J. (1959). *Reflections of a mathematician.* Montreal: Canadian Mathematical Congress. McGill University.

Nahin, P. J. (2006). *Dr. Euler's fabulous formula: Cures many mathematical ills*. Princeton: Princeton University Press.

Nielsen, M. (2011). *Reinventing discovery. The new era of networked science*. Princeton: Princeton University Press.

Ó Cairbre, F. (2009). The importance of being beautiful in mathematics. *IMTA Newsletter*, *109*, 29–45.

Poincaré, H. (1914). *Science and method*. London: Nelson.

Poincaré, H. (1958). *The value of science*. Mineola: Dover.

Rav, Y. (2005). Reflections on the proliferous growth of mathematical concepts and tools. Some case histories from mathematicians' workshops. In C. Cellucci & D. Gillies (Eds.), *Mathematical reasoning and heuristics* (pp. 49–69). London: College Publications.

Reichenbach, H. (1947). *Elements of symbolic logic*. New York: Macmillan.

Rota, G.-C. (1997). The phenomenology of mathematical beauty. *Synthese*, *111*(2), 171–182.

Schopenhauer, A. (2010). *The world as will and representation* (Vol. 1). Cambridge: Cambridge University Press.

Sinclair, N. (2004). The roles of the aesthetic in mathematical inquiry. *Mathematical Thinking and Learning*, *6*(3), 261–284.

Sinclair, N. (2006). *Mathematics and beauty. Aesthetic approaches to teaching children*. New York: Teachers College Press.

Sinclair, N. (2011). Aesthetic considerations in mathematics. *Journal of Humanistic Mathematics*, *1*(1), 2–32.

Snow, C. P. (1973). The classical mind. In J. Mehra (Ed.), *The physicist's conception of nature* (pp. 809–813). Dordrecht: Reidel.

Todd, C. S. (2008). Unmasking the truth beneath the beauty. Why the supposed aesthetic judgments made in science may not be aesthetic at all. *International Studies in the Philosophy of Science*, *22*(1), 61–79.

van Gerwen, R. (2011). Mathematical beauty and perceptual presence. *Philosophical Investigations*, *34*(3), 249–267.

von Neumann, J. (1961). The mathematician. In J. R. Newman (Ed.), *The world of mathematics* (Vol. IV, pp. 2053–2063). New York: Simon & Schuster.

Wells, D. (1990). Are these the most beautiful? *The Mathematical Intelligencer*, *12*(3), 37–41.

Zangwill, N. (1998). Aesthetic/sensory dependence. *British Journal of Aesthetics*, *38*(1), 66–81.

Zeki, S., Romaya, J. P., Benincasa, D. M. T., & Atiyah, M. F. (2014). The experience of mathematical beauty and its neural correlates. *Frontiers in Human Neuroscience*. doi:10.3389/fnhum.2014.00068.

Carlo Cellucci is emeritus professor at the Sapienza University of Rome, Italy. His recent publications (2013-2014) include the book Rethinking Logic: Logic in Relation to Mathematics, Evolution, and Method, Springer, New York 2013, and the papers "Rethinking Philosophy", "Knowledge, Truth, and Plausibility", "Why Should the Logic of Discovery be Revived", "Philosophy of Mathematics: Making a Fresh Start", "Top-Down and Bottom-Up Philosophy of Mathematics", "Explanatory and Non-Explanatory Demonstrations".