, To Discover or to Invent

Platonism

> Don't go around saying the world owes you a living; the world owes you nothing; it was here first.
> (Mark Twain, American writer and humorist, 1835-1910)

The standard physics department is staffed by both theoretical researchers and experimentalists. Experiments are meant to be the raw material out of which theories are created. In a math department, you generally won't find experimental mathematicians (although, surprisingly enough, there are exceptions: in a leading Canadian university I found a "Laboratory for Experimental Mathematics"). Mathematicians don't need laboratories. They do their work in an office, on a blackboard or on paper, and all they need is their mind. In this aspect they resemble philosophers, even though, according to the following joke, they aren't completely equal to philosophers:

> The university president sits in on a meeting of the mathematics department. "You know," he says to his hosts, "of all the people in the university, I like mathematicians the best. All they need is paper, a pen, and a wastebasket." Then he thinks for a while, and adds, "but the philosophers are even better. They don't need the wastebasket, either."

Scientists study the world. What do mathematicians study? Are they concerned with something that exists in the world, or with concepts that they created in their feverish minds? In short, do mathematicians discover, or do they invent? As for order, the question is: does mathematics find order in the world, or create it? Does it build something new, like a house, or discover something that already existed, the way Columbus discovered America? If someone invents a concept, such as "even number," does he invent something new, or did the concept pop into his brain because it corresponds to some structure in the world?

This question is supposed to be the subject of dispute among mathematicians. The approach that says that mathematical objects are part of the regular reality, and that they are as real as chairs and tables, is called "Platonism" (incidentally, Platonism is more extreme: Plato claimed that the concept of a table is more real than the table itself). A bitter struggle is assumed be waged between the Platonists and anti-Platonists. In practice, however, this is not the case. There is almost unanimous consent among mathematicians. The twentieth-century American mathematician Ralph Boas declared that he had never met a mathematician who was not a Platonist. Almost all mathematicians believe in the reality of the objects that they explore. They regard the numbers, geometric shapes, and the other objects that engage them - no matter how abstract - as part of the actual world. This means that most mathematicians think that mathematics reveals order that is already out there in the world.

The Sum of an Arithmetic Sequence

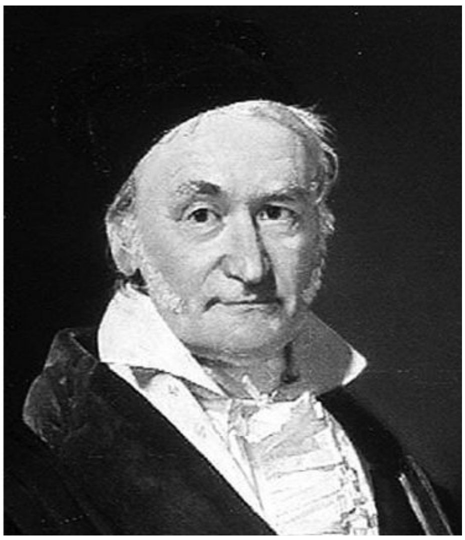
Here is a nice example of a discovery of order in the world: the calculation of the sum of an arithmetic sequence. An "arithmetic sequence" is a sequence of numbers that ascends (or descends) by a fixed number at each step. That is, the difference between an element ("term") and its predecessor is constant, called the "difference" of the sequence. For example, the sequence 2, 5, 8, 11, 14, 17 ascends by jumps of 3, meaning that its difference is 3; the sequence 20, 19, 18, 17 descends by jumps of 1, meaning that its difference is (-1). The source of the name "arithmetic sequence" is the fact that each term in the sequence is the arithmetic average (this is the official name for what is usually just called an "average") of its neighbors: in the first example, for instance, 5 is the average of 2 and 8, and 8 is the average of 5 and 11.

The method of calculating the sum of the terms of an arithmetic sequence was discovered by Karl Friedrich Gauss (1777-1855), the greatest mathematician of the nineteenth century, when he was seven years old. His father, who was a bricklayer, demanded that Karl Friedrich follow in his footsteps. Fortunately for the son, his talent was recognized, and he received a scholarship to study in a university One day, wishing to keep his students from mischief, his teacher asked the pupils of his class to calculate the sum of the numbers between 1 and 100. This is an example of the sum of the terms of an arithmetic sequence: the simple sequence 1, 2, 3, ... 98, 99, 100, whose difference is 1. To the teacher's surprise, Gauss came to him after a few minutes, or perhaps seconds, with the right answer - 5050.

How did little Karl Friedrich do this? He found a useful order in the series (a "series," in mathematical terminology, is the sum of a sequence) 1 + 2 + 3 + ... + 98 + 99 + 100. He combined the 1 with 100, for a total of 101; 2 with 99 - once again 101; and 3 + 98 also equals 101. The 100 numbers divide into 50 pairs, each of which totals 101, and so the sum is 50 times 101, which is 5050.

This method depended on the fact that 100 is an even number, and hence the numbers 1, 2, 3, ... 98, 99, 100 can be divided into pairs. Here is another way to calculate the same sum, that will work also for an odd number of terms: look at the average of the terms of the sequence. This is equal to its "middle," namely, the middle between 1 and 100, which is $50\frac{1}{2}$. The sum of 100 numbers, whose average is $50\frac{1}{2}$ is 100 times $50\frac{1}{2}$, which is 5050.

Did Gauss invent, or discover his method? Obviously, it was a discovery. In fact, If he hadn't discovered this, someone else would have. Actually, it's hard to believe that this method was not known in 1784, the year in which he found it. The mathematics of the time was quite sophisticated, and if, for example, Leonhard Euler (1707-1783), who preceded Gauss, had put his mind to the problem, he undoubtedly would have solved it.



Karl Friedrich Gauss (Germany, 1777-1855), the greatest mathematician of the nineteenth century (and some say, of all time). He contributed to the theory of complex numbers, number theory, and modern algebra. Together with the physicist Wilhelm Eduard Weber, he built the first telegraph. He spent his later years in seclusion in the observatory in Gottingen, and published very little. The biographer Eric Bell estimated that if all of his discoveries had been published in his lifetime, mathematics would have progressed by fifty years

Is Poetry Invented or Discovered?

What about poetry? We would think that there is no need to even ask the question: obviously, poetry is invented. But, let's listen to what a mathematician (and a poet) has to say about this. Sofja Kowalewskaya (1891-1950) was the favorite student of Karl Weierstrass, one of the important mathematicians of the late nineteenth century. In one of her letters she related to Weierstrass's statement (cited above) that a true mathematician must be something of a poet:

> In order to understand this, one must renounce the ancient prejudice that a poet must invent something that does not exist, that imagination and invention are identical [...] The poet must see what others do not see, must look deeper than others look.

Her words strike true. As we already mentioned, poets, like mathematicians, are hunters, engaged in the search for hidden patterns in the world. A metaphor that is on target reveals a similarity that is concealed, but that is out there. After all, "on target" implies that the target was already present. When the poet Yehuda Amichai writes

> Careful angels passed fate within fate,
> Their hands shook not, nothing dropped or fell.
> (Yehuda Amichai, "Twenty New Squares," *Poems*)

he expresses an existing truth: our fate is no more in our hands than the thread is master of its fate; there are forces that direct it, as the seamstress directs the thread. This is beautiful, not because it is an invention, but, mainly, because it is true. As Franz Kafka declared, poetry is always a search for the truth.

Order and Beauty

Saving Energy

When everything works out, and everything falls into its proper place, we don't say "Everything worked out finely," but: "Everything worked out beautifully." The sudden revelation of order leads to a sense of beauty. The question is - why? It is obvious why recognizing order is useful. Familiarity with the order of things around us means less effort in coping with the world. But utility and beauty are not one and the same. Why should the discovery of an orderly pattern cause aesthetic pleasure?

To understand this, we must first realize that we are not talking just about order. Order by itself is not sufficient to arouse a sense of beauty. Nothing is more orderly than a blank sheet of paper, and no collection of sounds is more orderly than absolute silence. Nonetheless, a blank sheet of paper is not a work of art, and silence does not possess the beauty of a Mozart symphony. A series of rhythmic beats is a very orderly and predictable phenomenon - we know exactly which sound will come nexthear, and when, but this is not music. In order to create a sensation of beauty, we need some factor beyond order.

The secret lies in a concept that has its roots in Sigmund Freud, and even before him: saving energy. Freud was an avowed mechanist, and believed that the psyche could be described in physical terms. He thereby followed the school that had its beginnings in the Industrial Revolution in England and the main proponent of which was Herbert Spencer. In the 1890s, when Freud was still taking his first tentative steps in psychoanalysis, he wrote the draft of a thick book entitled *Physiology for Psychologists*, in which he tried to explain mental phenomena in terms borrowed from the physical world of his time. Freud championed the Spencerian idea that the psyche tries to reduce effort as much as possible, that is, to save energy. When this works well, he claimed, the saved energy is released in the form of pleasure.

Like many before and after him, Freud quickly learned that psychological terms that were effective as metaphors soon become useless when we try and use them in a concrete sense. The concept of "saving energy" is too general to predict the behavior of human beings. As a result, *Physiology for Psychologists* was shelved around the year 1895, but echoes of it would reverberate throughout Freud's career. The idea of saving energy was expressed most clearly in a book he wrote in 1905 on humor, *Jokes and Their Relation to the Unconscious*. The book's thesis was that the pleasure we derive from a joke results from saving the energy of repression. The joke enables us to enjoy forbidden things without having to repress them. Consequently, the energy that was prepared to repress the forbidden idea is unnecessary, and is transformed into pleasure.

Freud himself was not satisfied with this book, and, later on, regarded it as a needless deviation from his main course. Totally out of character with its subject, this is a weighty book, perhaps the most abstruse of all of Freud's books, and its discussions are tortuous and forced. The Freudian theory of the joke was not particularly successful (and, in my mind, rightly so). The idea of saving energy, however, found supporters and there were those who continued what Freud had started, especially as regards pleasure resulting from one aspect of art, the formal aspect. In order to save energy, the proponents of this view argue, energy must, first of all, be harnessed. This happens when a person encounters stimuli that are seen, at first glance, as chaotic. The individual prepares the energy needed to contend with this lack of order. If some hidden order is suddenly revealed, then this energy will no longer be needed, and will be released in the form of aesthetic pleasure. According to this idea, the sensation of beauty arises when unexpected order is suddenly revealed in a phenomenon that seems to be disordered.

Music is one realm in which this explanation was clearly successful. Music is received by the human ear on two levels. On one level, it seems to be disorganized "noise." Our mind constantly attempts to decipher the stimuli that arrive from the outside world, and so it prepares energy to organize the sounds. But then the order underneath the noise reveals itself. We find connections between the sounds that enable us to predict what's coming. This happens in two dimensions: rhythm and harmony. Rhythm is the organization in time, that enables us to anticipate when the note will be heard. Harmony is the connection between the frequencies of the notes. In the next chapter, we will explain a bit about both.

In complex music these links are not straightforward, and cannot be perceived consciously. This means that, on the conscious level, we do not fully understand the order in the musical work. The sensation of beauty is born out of the gap between the perceived lack of order and the hidden order that is unconsciously revealed.

Mathematical Harmonies

Rhythm and Prediction

> The pleasure we obtain from music comes from counting, but counting unconsciously. Music is nothing but unconscious arithmetic.
> (Gottfried Leibniz, German mathematician and philosopher, 1646-1716)

Why do we enjoy music? The first answer is that it saves us the energy of predicting. Like any animal, man, too, is constantly engaged in predicting his surroundings. Man is a future-directed creature. His eyes are fixed in the front of his head, because he cares more about where he will be than where he was. Most of his thoughts are of the future, not his past. Even a historian, when he prepares an omelette, is more interested in where the egg will be in another moment than in where it was. There's a simple reason for this future orientation: this is how living creatures were formed by evolution. Evolution selected those life forms that are capable of directing themselves toward the future, in surviving, and in leaving descendants after them - "after them" in time.

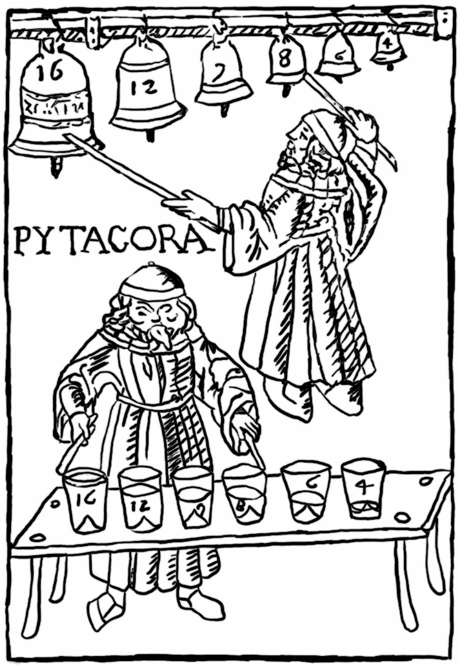
Order is related to the formation of the future. Our knowledge of the order in the world is useful for anticipating the future, and for molding our future environment for our benefit. This is the reason for the pleasure we derive from musical rhythm. An expected rhythm saves the investment of energy in the deciphering of the stimuli that reach our ears. But the rhythm must not be too predictable, because in order to save energy, it must first be rallied. If the rhythm is sufficiently complex, and we are incapable of consciously deciphering it, we prepare energy in order to guess the next note. When the order is revealed to us (usually unconsciously), this energy is no longer necessary - we know what to expect. Then, the energy that was saved turns into pleasure.

Pythagoras

Now we have an explanation for the pleasure we derive from rhythm. What about the second element of music, its harmony? This is more of a puzzle. We all know that some combinations of notes are pleasing to hear, while others are less so. For example, the *do* note closely complements the *do* note that is one octave higher. Actually, when we hear both at the same time, we are hardly aware that these are two different notes. The *do-sol* and *do-mi* combinations, as well, are pleasing to the ear. The notes *do*, *mi*, *sol* are the basic chord of the *do major* scale, which is probably the best-known and most basic of all the scales, since its notes are played on the white piano keys. A composition in *do major* frequently begins with notes close to *do*, *mi*, *sol* (in whatever order), strays from them, wanders about, before finally returning to them. Music is built on the tension between the digressions from the original harmony and this harmony itself.

But why is one combination of notes pleasing, while another grates on our ears? Surprisingly enough, the answer to this question is mathematical, and it was discovered by one of the most fascinating individuals in the history of mathematics, Pythagoras. He was the founder and leader of a most rare band of people: a mathematical religious sect. The sect numbered about 600 men and women, who lived in the Greek colony of Crotona in the south of the Apennine peninsula (modern-day Italy), and engaged in study and research. They donated all their possessions to the community, and they swore to keep their discoveries secret. According to the legend, one day Pythagoras was passing by a blacksmith's workshop, and he realized that when the smith struck rods with a simple ratio between their lengths - for example, one was twice as long as the other, or one was one and a half times as long, the combination of the two sounds was pleasant, but when this ratio was not simple, the combination was less pleasent.

In modern terminology, if two sounds are pleasant together, then the ratio between their frequencies is simple, that is, it is expressed by small numbers (3:2 is simpler than 11:5). The frequency of a sound is the number of times the air vibrates per second when the sound is produced, or, in more precise language: the number of peaks per second of the sound waves. If the note is produced by a string, this is the number of vibrations per second of the string. A difference of a single octave between notes (like that between a *do* and the *do* an octave higher) means a ratio of 2 between their frequencies: the frequency of a high *do* is twice that of the *do* below it. The frequency of the note *sol*, the fifth in the octave (when beginning with *do*) is frac{3}{2} times that of the low *do* of that octave. In other words, for every 2 vibrations of the *do* piano note, there will be 3 vibrations of the *sol* note. The ratio between them, as well, is a simple one: 3:2. The ratio between the *mi* and *do* notes is 5:4, which also is quite simple. This is why *do*, *mi*, and *sol* sound well together.



Pythagoras discovering the link between harmony and numbers
(from "The theory of music" by Franccino Gaforio, Milan 1492

Helmholtz

Why do simple ratios between frequencies cause pleasure? Pythagoras discovered the phenomenon, but was incapable of explaining it. Another 2,400 years would have to pass before this question would be answered. The enigma was solved by the German Hermann von Helmholtz (1821-1894), a true Renaissance man: a mathematician, physicist, and physiologist, who also studied aesthetics. His explanation of the pleasure from harmony was based on the phenomenon of "high notes." When a chord vibrates at a certain frequency, it also vibrates, at the same time, at frequencies 2, 3, 4, ... times higher. These secondary vibrations are weaker, and become weaker the greater the ratio to the original frequency (usually, there is hardly any vibration at a frequency 11 times higher), but they are audible. In other words, when the note *do* is played for us, most times we will also hear the *do* that is an octave higher, with a frequency exactly twice that of the original *do* note, and also the *sol* in the higher octave, which has a frequency precisely three times that of the original *do*. The simple ratio between the two frequencies means that they share high notes. For example: the *do* and the *sol* in the same octave share the *sol* high note in a higher octave. It is 3 times as high as the *do*, and twice as high as the lower *sol*; it also is not far from both, and therefore adds significantly to what we hear. And so, when we hear such notes together, we reveal a hidden order. The notes are different, but, without our being conscious of this, we reveal a factor common to both. Instead of chaos, order emerges. Is this a complete explanation for the pleasure we derive from music? Of course not. This can hardly explain the feelings that music arouses in people. We will have to look for these feelings in the tensions between disharmony and harmony, and between disorder and order. Helmholtz's explanation, however, was a first step to our understanding of this phenomenon.

Mystical Numbers

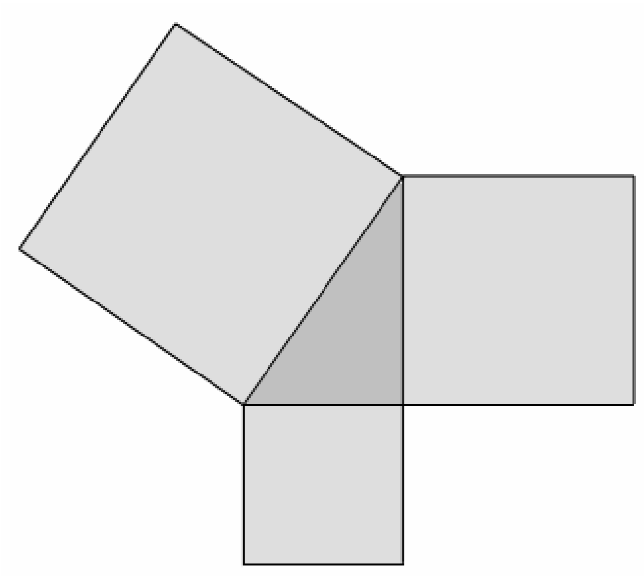
All this was beyond the knowledge of the ancient Greeks, who knew nothing of the frequencies of notes. And when people don't know, they fantasize. In order to explain harmony, Pythagoras and his school invented fanciful theories of the magical powers of numbers and the ratios between them. "The world is a number" was their slogan. That is, the world is ruled by simple numerical ratios. The Pythagoreans believed that every important natural phenomenon has to obey numerical laws. They maintained that there are simple ratios between the diameters of the planetary orbits, and that the planets consequently emanate "celestial music." Going beyond this, they claimed that every significant size in the world could be expressed as a number that is a ratio between whole numbers.

A number that is the quotient of two whole numbers is called a "rational number" (from the word "ratio"). Every whole number is rational: 4, for example, is rational, because it is the ratio between itself and 1, that is, $4:1 = 4$. Every fraction with a whole-number numerator and denominator is rational, because the fraction bar is actually a division sign: frac{17}{3} is really the ration between 17 and 3. The Pythagoreans believed that every important size in nature had to be expressed as a rational number.

Sobering Up

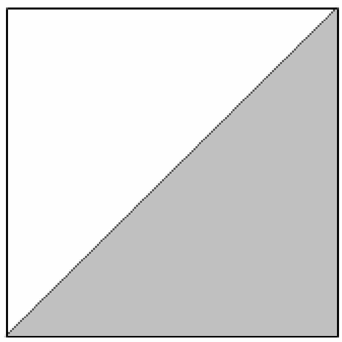
The intellectual achievements of the ancient Greeks were nothing short of miraculous. A very small people, numbering no more than a few hundred thousand, developed conceptual systems whose fruits we continue to enjoy to the present. They were motivated by their infinite respect for abstract ideas. For the Greeks, abstractions held magical power, and, for them, abstractions were more important than the real world. The Greeks were the first to study abstract concepts for their own sake, without regard for their usefulness in the world around them. The Egyptians and the Babylonians, too, studied numbers, but they did so for practical ends. The Greeks were the first to see numbers as a world worthy to be explored for its beauty and inner harmony.

But even in comparison with the Greeks' achievements as a whole, geometry enjoys a special pride of place. It was in this field that the Greeks developed the concepts of "axiom" and "proof," and it was here that they reached the highest level of abstraction. Pythagoras was one of the founders of Greek geometry. The theorem that, to this day, is regarded (and rightly so) as the most important and useful geometric theorem is named after him (not rightly so): the "Pythagorean theorem." This theorem states that the sum of the areas of the two squares based on the legs of a right triangle equals the area of the square based on the hypotenuse. This theorem is so important because it enables us to calculate distances. Given the lengths of the legs of a right triangle, we can calculate the length of the hypotenuse. This means that if you know how to calculate east-west and north-south distances, you can calculate the distance between any two points.



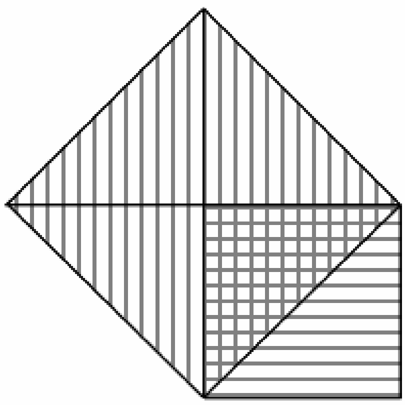
The Pythagorean theorem: the sum of the areas of the two squares on the legs of a right triangle equals the area of the square on the hypotenuse

An interesting special case of this theorem is that in which the legs of a right triangle are of the same length. Let's look at a square whose side is 1 unit long. The diagonal of the square is the hypotenuse of a right triangle, with equal legs, each 1 unit long.



According to the Pythagorean theorem, the length of the diagonal of a square 1 unit long is $\sqrt{2}$ units

According to the Pythagorean theorem, the diagonal, squared, is equal to $1^2 + 1^2 = 2$. Therefore, the length of the diagonal, not squared, is $\sqrt{2}$. There is a simple and especially beautiful proof for this individual case of the theorem, that appears in a somewhat surprising place: one of Plato's dialogues (in which, as in all of Plato's dialogues, the hero is Socrates) entitled *Menon*. Now, let's look at the next drawing:



The area of the square based on the diagonal (with vertical lines) is twice the area of the small square (with horizontal lines), because it contains 4 triangles, while the small square contains only 2. The side of the larger square is therefore $\sqrt{2}$ the length of the side of the small square

Let's assume that the length of the side of the small square (with horizontal lines) is 1. The area of this square is therefore 1 X 1, that is, 1. The large square, that is on a diagonal (and marked with vertical lines), is composed of 4 triangles, while the small square contains only 2 (all of the triangles are congruent, that is, they are capable of perfectly fitting one over the other). Therefore, the area of the large square is twice that of the small square, that is, 2. The length of any square's side is the square root of its area, and so the length of the large square's side is *sqrt*2. But look: the side of the large square is exactly the same length as the diagonal of the small square! And so, this diagonal is $\sqrt{2}$ long.

Geometry played a central role in the lives of the Pythagoreans, and for them the diagonal of a square was an everyday object. The Pythagoreans believed that this diagonal should therefore be expressed by a simple ratio between whole numbers. That is, it has to be a rational number. For many years the Pythagoreans tried to express $\sqrt{2}$ as a rational number. In the end, they realized the surprising - and for them, destructive - fact that this is impossible. $\sqrt{2}$ is not rational. This was such a severe blow to the Pythagoreans that they vowed to each other not to reveal this fact to the outside world. Due to the sect's secretiveness, most of their history remains a mystery to us, and the continuation of the story might very well be spurious: legend has it that Hipassus, the sect member who revealed the secret to the world, was put to death for doing so. This is almost certainly not true, Hipassus drowned, and his death might very well have been an accident, that the sect accredited to punishment by the gods.

Why *sqrt*2 Is Not a Rational Number

Actually, why isn't $\sqrt{2}$ a rational number? Why can't it be expressed as $\frac{m}{n}$ for any pair of whole numbers *m* and *n*? Let's assume that $\frac{m}{n} = \sqrt{2}$, and attempt to reach a contradiction based on this assumption. First, we can assume that $\frac{m}{n}$ is a reduced fraction, and if not, we will reduce it. Now, by the definition of $\sqrt{2}$, $\frac{m}{n} = \sqrt{2}$ means that frac{m^2}{n^2} = (frac{*m*}{*n*})^2 = 2. If we multiply both sides by n^2, we get:

(*) $2n^2 = m^2$

Now, we will divide our discussion into two cases: in one, *m* is an odd number, and in the other, it is even. Each of these two cases will lead to a contradiction, thereby yielding the desired conclusion that frac{*m*}{*n*} = *sqrt*2 is impossible. If *m* is an odd number, the right side of the (*) equation is the square of an odd number, which itself is an odd number, while the left side is a multiple of 2, and therefore is an even number. Since an odd number cannot be equal to an even one, the right side cannot be equal to the left, and this is our contradiction. (As an example of this case, we could guess that $\sqrt{2} = \frac{7}{5}$, which is quite close to the actual value of $\sqrt{2}$. But then 2 = (frac{7}{5})^2, which would mean that 2 X $5^2 = 7^2$, and then the left side, which is 50, is even, while the right side, which is 49, is odd.)

Now for the case in which *m* is an even number. Since, by assumption, the fraction frac{*m*}{*n*} is reduced, *n* must be odd (if it were even, then the entire fraction could be reduced by 2). Since an odd number squared is odd, on the left side of (*) there is the product obtained by multiplying 2 by an odd number. Such a product cannot be divisible by 4. But on the right side we have an even number squared, which is divisible by 4. Since a number divisible by 4 cannot be equal to a number not divisible by 4, once again, the two sides of (*) cannot be equal. (An example of this case: the number frac{10}{7}, too, is very close to $\sqrt{2}$, since (frac{10}{7}^{2} = frac{100}{49}, which is very close to frac{100}{50}, which equals 2. But if $\sqrt{2} = \frac{10}{7}$, then (frac{10}{7})^2 = 2, that is, 10^2 = 2 X 7^2. But 10^2, which is 100, is divisible by 4, while 2 X 7^2 is not. (Obviously, 2 X 7^2 = 98, and not 100. The purpose of this example, however, is to show that it is not divisible by 4, and so these two expressions **could not** be equal.)

In each of the two cases, we have shown that the equation (*) is impossible. This is the contradiction that we sought, in order to refute our original assumption that $\sqrt{2} = \frac{m}{n}$

There is also a shorter way to present the same argument, one that is also a bit more abstract. Let's take another look at (*). The number of factors of 2 that m^2 contains is even. (For example, 40, which is 2 X 2 X 2 X 5, contains 3 factors of 2, and therefore its square, 40^2, which is 2 X 2 X 2 X 2 X 2 X 2 X 5 X 5, contains 6 factors of 2, namely, an even number of 2-factors. n^2 contains an even number of factors of 2. But then $2n^2$ contains an even number + 1 of factors of 2 (an even number from n^2, and an additional 1 from it being multiplied by 2), that is, an odd number of factors of 2. (For example, if 2^3 X 3 X 5 = 120 = n, then 2^6 X 3^2 X 5^2 = n^2, and 2^7 X 3^2 X 5^2 = $2n^2$, in which the power of 2 namely 7, is odd.) Then $2n^2$ contains a different number of factors of 2 than m^2, and therefore $m^2 = 2n^2$ is an impossibility.

How do we then express $\sqrt{2}$? One way is as an infinite decimal fraction: $\sqrt{2}$ = 1.4142135623..., which means that $\sqrt{2}$ can be approached by the rational numbers 1, 1.4, 1.41, 1.414, and so on. Let's compare the infinite decimal expansion with the approximations 1.4 = frac{7}{5} and 1.4285714285... = frac{10}{7}. $\sqrt{2}$ is almost exactly in the middle between them! Almost, but of course not exactly in the middle: the middle is a rational number, and we know that *sqrt*2 is not rational.