UNIVERSITY OF TORONTO
MATHEMATICS NETWORK

Answers and Explanations

This is a printed version of the *Answers And Explanations* area of the University of Toronto Mathematics Network web site, `http://www.math.toronto.edu/mathnet/answers/`. It reflects the contents of the site as of September 27, 1999.

This area is designed to be a place where you can delve more deeply into some foundational questions in mathematics, and find expositions on some topics that often seem puzzling when glimpsed from a distance.

(You may also be interested in our Question Corner area, a place where you can ask your own individual questions and see the answers to the questions of others. It can be found at `http://www.math.toronto.edu/mathnet/questionCorner/`)

Contents

1 Is there really such a thing as "infinity"?

It's a tough question, because the word "infinity" can mean different things in different contexts.

In mathematics, whether or not a certain concept exists can depend on the context in which you ask the question. If you want to know more about this, you can refer to a fuller explanation(in section 4) of how a mathematical concept can exist in some contexts but not in others.

Here are some of the contexts in which the question "is there such a thing as infinity" can be asked, and the answers appropriate for each context. The details are given afterwards.

1. In the context of a *number system*, in which "infinity" would mean something one can treat like a number.

   In this context, infinity does not exist.

2. In the context of a *topological space*, in which "infinity" would mean something that certain sequences of numbers converge to.

   In this context, infinity does exist.

3. In the context of measuring sizes of sets, in which "infinity" means a measurement of the size of an infinite set.

   In this context, such "infinity" concepts do exist but there are more than one of them, since not all infinite sets have the same size. So there does not exist any one single "infinity" concept; instead, there exists a whole collection of things called "infinite cardinal numbers".

Here are the details.

1. *In the context of a number system.*

   A *number system* is any collection of objects that has the basic properties (like addition, multiplication, and so on) we normally associate with numbers. More information is available on this in section 5 below.

   Number systems come in many sizes. There is the "natural number system", which is just the set of numbers used in counting: 1, 2, 3, 4, and so on. Or, one can expand this number system to include additional concepts, such as negative numbers, fractions, even the so-called "imaginary" numbers (which are not really imaginary at all). Each of these concepts exists provided we look for it in the context of a large enough number system.

   Now the question is, does infinity exist in the same way that these concepts (negative numbers, fractions, etc.) do?

   In other words, *does there exist any number system which, as well as including the familiar numbers we are used to, also includes an "infinity" concept?*

   The answer is **no**; for more information, see section 1.1.

2. *In the context of a topological space.*

   You've probably never heard the term "topological space" before; it occurs in an advanced branch of mathematics. Don't worry; we don't need to get into that advanced area. All we need is the following rough idea:

   Roughly speaking, a *topological space* is any collection of objects for which there is a definition of which sequences of objects converge to other objects, and which sequences don't.

The real number system is a topological space: there's a definition of what it means for a sequence of numbers to converge. For instance, the sequence 1.1, 1.01, 1.001, 1.0001, etc. converges to the number 1, while the sequence 1, 2, 1, 2, 1, 2, 1, 2, etc. does not converge to anything.

In areas such as calculus, one often speaks of a sequence like 1, 2, 3, 4, ... as "converging to infinity". Is this just a convenient phrase, or can there actually exist an object "infinity" that this sequence is converging to?

In other words, the question is: *does there exist some topological space (that is, a set of objects plus a definition of what convergence means) which, as well as including the familiar real numbers we are used to, also includes an "infinity" concept to which some sequences of real numbers converge?*

The answer is **yes**; for more information, see section 1.2.

It is important to realize, though, that this topological space **is not a number system**. Although it includes an additional object called "infinity" as well as the familiar real numbers, you **cannot** add, subtract, multiply, or divide this additional object the way you can numbers.

Sometimes you will see a statement like "$1/\infty = 0$". This does not really mean what it seems to say. You are not really dividing 1 by infinity. Instead, it is a statement about *sequences*. What it means is that if a sequence $a_1,\ a_2,\ a_3,\ \ldots$ converges to infinity, then the sequence of reciprocals $1/a_1,\ 1/a_2,\ 1/a_3,\ \ldots$ converges to zero.

(Similar looking expressions like "$\infty - \infty$" don't make any sense, for just because two sequences $a_1,\ a_2,\ a_3,\ \ldots$ and $b_1,\ b_2,\ b_3,\ \ldots$ each converge to infinity tells you nothing about what the sequence $a_1 - b_1,\ a_2 - b_2,\ a_3 - b_3,\ \ldots$ does.)

3. *In the context of measuring sizes of sets.*

   The natural numbers (1, 2, 3, etc.) measure sizes of finite sets. They are possible answers to the question "how many things am I holding?"

   Is there a similar concept that can measure the size of an infinite set? For instance, is there any object called "infinity" that can serve as the answer to the question "how many elements are there in the infinite set $\{1, 2, 3, \ldots\}$"?

   the answer is a qualified **yes**; however, we really shouldn't use the word "infinity" for this one object alone, because there are other infinite sets with different sizes. For more information, see section 1.3.

For more information on the above topics, refer to one of the sections below:

1.1 More Information on Why "Infinity" Does Not Exist in the Context of Any Number System

No "infinity" concept exists in the context of any number system, if by number system one means a collection of concepts that have operations like addition and multiplication the way familiar numbers do, operations which obey the usual properties of arithmetic.

One way to see this is to think, what would infinity minus 1 be? It couldn't be a finite number, since no finite number plus 1 equals infinity. So it must be infinite, and this would mean

$$\infty - 1 = \infty.$$

From this one can immediately see that the rules of arithmetic must be violated, since if they held one could subtract infinity from both sides to conclude that $-1 = 0$, which isn't true.

Therefore, there is no number system which possesses the usual rules of arithmetic and in which infinity exists. In other words,

Infinity does not exist, if by "exist" one means in the context of a number system.

A side note:

One should note that these arguments show there's no such thing as a number system with a single "infinity" concept. But there's nothing to stop us constructing a number system containing "infinitely large numbers" which are bigger than all the usual numbers; it's just that there would have to be many of them and no single one of them could be called "infinity".

A fairly straightforward example is the set of all polynomials. They're a number system (you can add, subtract, and multiply any two polynomials), and if you were to try to put an ordering on them (some system for saying which polynomials are larger than which), the most natural way to do so would be to consider the degree 1 polynomials to be "larger" than the degree 0 polynomials. So a polynomial like "$x + 2$" would be considered larger than any constant (no matter how big that constant was). Therefore, in this number system, the polynomial "$x + 2$" would be an "infinitely large number" because it's bigger than all the ordinary numbers that have no x in them. But it would be quite unjustified to call it "infinity" because it's just one of many degree 1 polynomials, and there's no reason to single it out with a special name rather than, say, $x + 1$ or $x + 3$ or $2x - 17$.

1.2 More Information on Why "Infinity" Does Exist in the Context of A Topological Space

Remember, we are trying to see why the answer is "yes" to the following question:

> *Does there exist some topological space (that is, a set of objects plus a definition of what convergence means) which, as well as including the familiar real numbers we are used to, also includes an "infinity" concept to which some sequences of real numbers converge?*

How do we know that a topological space with these properties really exists? Well, we just build one! Here are two ways to do it:

- Start with the real numbers, take any additional object you like, add it to the set of real numbers, and make the resulting set into a topological space. It turns out that all you need to do to make a set into a topological space is come up with an appropriate collection of subsets (which in the case of the real numbers are unions of open intervals). It would take us too far afield to discuss exactly why specifying an "appropriate collection of subsets" is all it takes to completely determine all notions of which sequences converge and which don't. However, that is in fact all it takes, and in our new set (the real numbers together with the additional object) there does turn out to be an appropriate collection of subsets which makes it into a topological space, and the resulting notion of "convergence" is such that sequences like 1, 2, 3, ... converge to the additional object, which justifies calling the additional object "infinity".
- If the previous construction seems too abstract, think about drawing the real numbers on an interval of finite length, instead of the usual infinitely long number line. (It may seem surprising that this can be done, but all you need to do is compress the distances so that larger numbers are drawn closer together. For instance, you could draw the number 1 at

a distance 1 from the end, 2 at a distance 1/2 from the end, 3 at a distance 1/3 from the end, and so on. The following picture illustrates this (not quite to scale):

$$\underset{\cdots\text{-}1 \qquad\qquad\qquad\qquad\qquad 1 \qquad 2\ 3\ 4\ \cdots}{\rule{14cm}{0.4pt}} \leftarrow \text{the interval}$$

Now you can take the endpoints of the interval, call them "infinity" and "−infinity", and say that a sequence of numbers a_1, a_2, a_3, ... "converges to infinity" if and only if the dots where those numbers are drawn converge to the right-hand endpoint of the interval.

No matter which approach you take, you can define a topological space which has the real numbers as a subset, also has an additional object called "infinity", and in which the notion of convergence is such that sequences like 1, 2, 3, ... converge to this additional object. In this sense, therefore, infinity exists.

1.3 Why "Infinity" Does Exist in the Context of Measuring the Size of an Infinite Set

Yes, there does exist an object one could call "infinity" in the sense that it measures the size of the infinite set $\{1, 2, 3, \ldots\}$ in the same way that an ordinary natural number measures the size of a finite set.

One way to see that there does in fact exist an object that can serve as a measure of the size of a set is just to use *the collection of all sets of that size* as your object. For example, the number "1" can be defined as the collection of all the single-element sets in the world, the number "2" can be defined as the collection of all pairs of things in the world, and so on. (This may sound like a circular definition but that's just because I've been oversimplifying a bit).

In a similar way, one can define the "size" of the infinite set $\{1, 2, 3, \ldots\}$ to be the collection of all the sets in the world that have the same size as this set.

So, if one asks the question *does there exist any object that one can call "infinity", in the sense that it measures the size of the infinite set* $\{1, 2, 3, \ldots\}$ *in the same way that an ordinary natural number measures the size of a finite set*, the answer is **yes**. However, we really shouldn't use the word "infinity" for this one object alone, because there are other infinite sets with different sizes. This is an interesting topic, so we will discuss it briefly.

Many infinite sets have the same size. For instance, the set of even integers has the same size as the set of all integers. In one sense that may seem surprising, since the second set has all the elements of the first set in it plus all the odd integers as well, but it's fact of life about infinite sets that you can add elements to an infinite set without altering its size. But in another sense it's not surprising at all that, since both sets are infinite, their sizes should each be the same "infinity" concept.

What is really surprising is that there are other infinite sets which do *not* have the same size as the set of integers! For instance, the set of all real numbers is a much *bigger* set. So the "infinity" concept that describes the size of the set of real numbers is a different one from the one which describes the size of the set of integers. In other words, in the context of measuring sizes of sets there is no one single "infinity" concept, but lots of different ones appropriate to the different possible sizes.

These concepts that describe the sizes of infinite sets are called *infinite cardinal numbers.* The size of the set of integers (which is the same as the size of the set of even integers, which is the same as the size of the set of odd intgers, which is the same as the size of the set $\{1, 2, 3, \ldots\}$ of natural numbers, and so on) is called $\aleph_0$ (pronounced "aleph-null"); it's the Hebrew letter Aleph with a zero subscript. There are also different cardinal numbers that describe the sizes of sets (such as the set of real numbers) that are not the same size as the set of integers. There are in fact infinitely many different infinite cardinal numbers!

2 Do "Imaginary Numbers" Really Exist?

An "imaginary number" is a multiple of a quantity called "i" which is defined by the property that i squared equals -1. This is puzzling to most people, because it is hard to imagine any number having a negative square. The result: it is tempting to believe that i doesn't really exist, but is just a convenient mathematical fiction.

This isn't the case. Imaginary numbers do exist. Despite their name, they are not really imaginary at all. (The name dates back to when they were first introduced, before their existence was really understood. At that point in time, people were imagining what it would be like to have a number system that contained square roots of negative numbers, hence the name "imaginary". Eventually it was realized that such a number system does in fact exist, but by then the name had stuck.)

Before discussing why imaginary numbers exist, it's helpful to think about why we're even asking the question. Why is it so hard to accept that there could be numbers with negative squares? One has to come to terms with the things that seem so puzzling and confusing about this concept and see that they are not really so unreasonable after all, before one can move on to accept the existence of imaginary numbers. Having done that, we can move on to seeing why they exist, and what relevance they have.

Therefore, we will address the following questions:

2.1 Why is the existence of imaginary numbers not really as unreasonable as it seems?

It may seem hard to believe that imaginary numbers could possibly exist. The source of this difficulty stems from what one means by "existence". In mathematics, whether or not a certain concept exists can depend on the context in which you ask the question.

When talking about numbers, there are many **very different contexts** that one could have in mind.

Here are the four most familiar ones:

- *The Natural Numbers.* These are the counting numbers 1, 2, 3, ... that are possible answers to the question "how many?" They are abstract concepts that describe sizes of sets.
- *The Integers.* These are abstract concepts that describe, not sizes of sets, but the *relative sizes* of two sets. They are the possible answers to the question "how many more does A have than B has?" They include both positive numbers (meaning A has more than B) and negative numbers (meaning B has more than A).
- *The Rational Numbers.* These are abstract concepts that describe *ratios of sizes of sets.* They do not model sizes of sets the way that natural numbers do. If you say "I ate 3/4 of a pie", you are not saying that the set of things you ate had 3/4 elements. Instead, you are expressing a ratio of two integer quantities: 3, the number of pie-quarters that you ate, and 4, the number of pie-quarters that make up a whole pie.

- *The Real Numbers.* These are abstract concepts that describe *measurements of continuous quantities*, such as length, weight, quantity of fluid, etc. (Don't let the word "real" fool you; the real numbers are no more "real" in the ordinary English sense of the word than are any other kind of numbers.)

Concepts that exist in one of these contexts may not exist in another. The question "does there exist a number between 1 and 2?" has the answer **no** in the first two contexts (you cannot go to the beach and pick up more than one but fewer than two pebbles), but **yes** in the last two contexts (you could eat three cookie halves, which is in between one whole cookie and two whole cookies).

Although in the first two contexts there does not exist a number between 1 and 2, most people are quite comfortable with the fact that such numbers do exist in other contexts. For instance, people don't usually have trouble accepting the existence of the fraction $3/2$. Why then is it so hard to believe that the concept of "a number whose square is -1", though it does not exist in any of the four contexts mentioned above, might nonetheless exist in some other context?

It is because we usually forget the fact that we already have four quite different meanings for the word "number". We have become so familiar with each of the four contexts that we have jumbled them together in our mind as if they were a single concept. When we encounter a notion like "square root of -1" which does not exist in any of these four contexts, we think that it cannot exist at all, because we think the word "number" is a single concept that embodies just these four contexts.

Instead, what we *should* be thinking is something like this:

> Okay, I know about four different number systems: one in which "number" means a measurement of how many items are in a set, a second one in which "number" means a relative measurement of the sizes of two sets, a third one in which "number" means a ratio of sizes of two sets, and a fourth one in which "number" means a measurement of a continuous quantity.
>
> In neither of these four number systems does there exist a square root of -1.
>
> Might there be a fifth context, a number system (where "number" means something different from any of the above four things) in which there *does* exist a square root of -1?

The answer to that final question is "yes, there is". It is called the *Complex Number System.* Although it will involve a notion of "number" that is something different from what we are used to, the difference is not fundamentally any greater than is the difference between the concepts of "number of elements in a set" (natural number) and "ratio of sizes of two sets". In other words, the complex numbers are not that much more different from familiar numbers than rational numbers (fractions) are from natural numbers.

2.2 How can one show that imaginary numbers really do exist?

One does it in exactly the same way one would show that fractions exist.

Let's look at a way to show that fractions exist. Of course, that's something you know already; you don't need a mathematical exposition to prove it to you. But the point of going through it is that *exactly the same argument* can be used to show that imaginary numbers exist. Having become convinced that the argument is a legitimate one by seeing it work in a familiar context, you should be more willing to accept it in the somewhat mysterious context of "imaginary" numbers.

2.2.1 Argument that Fractions Exist

Suppose the only things you knew about were the natural numbers (1, 2, 3, etc.), and you had to show that "three halves" exists. In other words, you need to show that there exists some number which, when doubled, gives you 3. You could argue as follows:

> Granted, no such thing exists within the Natural Number System.
>
> However, there is a different number system in which such a thing does exist: the Rational Number System. The "numbers" in this different number system will be *fractions*: totally different objects from the natural numbers (they won't represent sizes of sets; instead, they'll represent ratios of sizes), but that doesn't make them any less real.
>
> Do fractions really exist? Yes. Do they really form a number system? Yes. Within this number system, is there a number which, when doubled, gives 3? Yes. Therefore, "three halves" exists.

2.2.2 Validity of the Argument

To see that the three key answers (in the last paragraph of the argument) really are "yes", let's look at the questions one by one.

Do fractions really exist? Yes; they're just pairs of natural numbers. (Let's just talk about positive fractions here, to make the discussion as simple as possible and avoid having to worry about things like the denominator being zero). Pairs of natural numbers certainly exist, so fractions exist. We write such a pair by writing the first number over the second number, e.g. a/b.

Do fractions really form a number system? Yes. A *number system* is just a collection of objects for which

- there's a definition of what it means for two objects to be equal,
- there is a rule for how to add two objects together and a rule for how to multiply two objects together (subtraction and division can be deduced from these, provided that all objects have corresponding negatives and some objects have corresponding reciprocals), and
- these rules for addition and multiplication satisfy the familiar properties of arithmetic, such as *commutativity* (order doesn't matter), *associativity* (in a sum of three of more terms, it doesn't matter which two you add first, and likewise for products), and *distributivity* ($a(b+c) = ab + ac$).

Roughly speaking, any collection of objects that satisfies these properties is, by definition, a number system. (Strictly speaking, some of these properties need to be stated a little more precisely, but the rough statement is quite enough for our purposes!)

These properties are all satisfied by fractions. We have a definition of when two fractions are to be considered equal:

> $a/b = c/d$ if and only if $ad = bc$.

We have a rule for adding two fractions:

> $a/b + c/d = (ad + bc)/(bd)$

and a rule for multiplying two fractions:

$$(a/b)(c/d) = (ac)/(bd).$$

One can check that these rules do indeed satisfy the familiar properties of arithmetic.

Therefore, fractions form a number system.

Within this number system, is there an object which, when doubled, gives 3? Yes. It is the fraction 3/2 . When you double it, you get the fraction 3/1.

Strictly speaking, 3/1 is something different from the natural number 3. After all, it's a *pair* of natural numbers, 3 and 1 (representing the ratio "3 to 1"), not a single natural number.

However, fractions of the form $a/1$ behave identically to the way ordinary natural numbers a behave. They add and multiply in exactly the same way that ordinary natural numbers do:

$$a/1 + b/1 = (a + b)/1$$
$$(a/1)(b/1) = (ab)/1.$$

The "/1" just "comes along for the ride".

Since numbers are just abstract concepts anyway, and since natural numbers a and fractions of the form $a/1$ are completely identical as far as their arithmetic behaviour is concerned, it is perfectly legitimate to view them as just two different representations of the same underlying concept.

With this in mind, we can consider the fraction 3/1 (the ratio "3 to 1") and the natural number "3" to be the same thing. This enables us to say that 3/2, when doubled, gives 3.

This completes the argument that "three halves" exists. Of course, that's something you knew already; it's obvious that fractions exist. But even though you already knew that fractions exist, and didn't need this long argument proving it, the point of going through the details of the argument is that **exactly the same argument** can be used to show that imaginary numbers exist.

The argument that "imaginary" numbers exist is almost word-for-word identical to the above argument. So, being convinced that the above argument is a valid one, you should be better able to accept the argument that imaginary numbers exist.

2.2.3 Argument that Imaginary Numbers Exist

This argument is patterned after the above argument that fractions exist; you'll probably find it helpful to view the two of them side by side.

The issue is the existence of the mysterious quantity "i", since imaginary numbers are just multiples of i. In other words, we want to see that there exists some number which, when squared, gives you -1. Here is such an argument:

> Granted, no such thing exists within any of the four familiar number systems (the Natural Number System, the Integers, the Rational Number System, or the Real Number System).
>
> However, there is a different number system in which such a thing does exist: the Complex Number System. The "numbers" in this different number system will be totally different objects from the familiar real numbers (they will in fact be pairs of real numbers), but that doesn't make them any less real.
>
> Do complex numbers really exist? Yes. Do they really form a number system? Yes. Within this number system, is there a number which, when squared, gives -1? Yes. Therefore, i exists.

2.2.4 Validity of the Argument

To see that the three key answers (in the last paragraph of the argument) really are "yes", let's look at the questions one by one.

Do complex numbers really exist? Yes; we just define a complex number to be a pair of real numbers. Real numbers certainly exist, so pairs of them exist.

Do complex numbers really form a number system? Yes. Remember that any collection of objects for which

- there is a definition of what the objects are and when two objects are equal,
- there is a rule for how to add two objects,
- there is a rule for how to multiply two objects, and
- these rules obey familiar arithmetic laws like commutativity, associativity, and distributivity,

is, by definition, a number system.

These properties are all satisfied by complex numbers.

We have a definition of when two complex numbers are to be considered equal: they are equal if and only if they are the same pair of real numbers.

We have a rule for adding two complex numbers (which, remember, are nothing more than pairs of real numbers):

$$(a, b) + (c, d) = (a + c, b + d)$$

and a rule for multiplying two complex numbers:

$$(a, b)(c, d) = (ac - bd, ad + bc)$$

The rule for multiplication may look very strange, but there's nothing wrong with that; one can still verify that these rules do indeed satisfy the familiar properties of arithmetic.

Therefore, complex numbers form a number system.

Within this number system, is there an object which, when squared, gives -1? Yes. It is the pair (0,1). When you square it using the above rule of multiplication, you get

$$(0, 1)(0, 1) = ((0)(0) - (1)(1), (0)(1) + (1)(0)) = (-1, 0).$$

Strictly speaking, the complex number $(-1, 0)$ is something different from the real number -1. After all, it's a *pair* of real numbers, -1 and 0, not a single real number.

However, complex numbers of the form $(a, 0)$ behave identically to the way ordinary real numbers a behave. They add and multiply in exactly the same way that ordinary real numbers do:

$$(a, 0) + (b, 0) = (a + b, 0)$$
$$(a, 0)(b, 0) = (ab, 0).$$

The ",0" just "comes along for the ride".

Since numbers are just abstract concepts anyway, and since real numbers a and complex numbers of the form $(a, 0)$ are completely identical as far as their arithmetic behaviour is concerned, it is perfectly legitimate to view them as just two different representations of the same underlying concept.

With this in mind, we can consider the complex number $(-1, 0)$ and the real number -1 to be the same thing (this may seem a little hard to swallow, but remember it is no different from saying that the fraction $3/1$ and the natural number 3 are the same thing, something that we do all the time; it may be helpful to re-read the corresponding paragraph for fractions to see just how similar the two cases are).

This enables us to say that $(0, 1)$, when squared, gives -1. Therefore, i exists; it is merely the pair of numbers $(0, 1)$ under the above rules for adding and multiplying.

2.3 What relevance do imaginary numbers have to the real world?

Okay, now we've seen that imaginary numbers exist. However, they exist in the context of a different number system, something different from the number systems we are used to. The "complex numbers" that make up this system are *pairs of numbers*; do they really deserve to be called "numbers" in their own right?

Well, remember that fractions are pairs of numbers also. They clearly deserve to be called numbers in their own right, since they can measure "how much" in some contexts (for instance, "I ate three quarters of a pie"). So, the principle of considering a pair of numbers (in this example, 3 and 4) as a number in its own right is well established.

The fact remains, though, that complex numbers have much less *direct* relevance to real-world quantities than other numbers do. An imaginary number could not be used as a measurement of how much water is in a bottle, or how far a kite has travelled, or how many fingers one has.

Nonetheless, there are a few real world quantities for which complex numbers are the natural model. The strength of an electromagnetic field is one example. The field has both an electric and a magnetic component, so it takes a pair of real numbers (one for the intensity of the electric field, one for the intensity of the magnetic field) to describe the field strength. This pair of real numbers can be thought of as a complex number, and it turns out that the strange rule of multiplication of complex numbers has relevance to the physics of an electromagnetic field.

Although such direct applications of complex numbers to the real world are few, their *indirect* applications are many. Many properties related to real numbers only become clear when the real numbers are thought of as sitting inside the complex number system. Therefore, complex numbers aid in the understanding even of things that are described by ordinary, familiar real numbers.

It's like trying to understand a shadow. The shadow lives in a two-dimensional world, so only two-dimensional concepts are directly applicable to it. However, thinking of the three-dimensional object casting the shadow can aid in understanding it, even though three-dimensional concepts don't have any direct application to the two-dimensional world of the shadow. Likewise, complex numbers may not be directly applicable to a real world measurement any more than a three-dimensional object is directly applicable to a 2-dimensional shadow, but they can still help us understand it.

Here's another analogy that may be helpful. Consider the following argument:

> Population A has 236 people, 48 of them children.
>
> Population B has 1234 people, 123 of them children.
>
> The fraction of children in population A is $48/236$ while the fraction of children in population B is $123/1234$.
>
> $48/236$ (approx. 0.2) is much greater than $123/1234$ (approx. 0.1).

Therefore, population A is, on the whole, a much younger population than population B.

This argument has used *fractions*, non-integer numbers, in a problem where they have no physical relevance! You can't measure populations in fractions; you can't have "half a person", for example. The kind of numbers that have direct relevance to measuring numbers of people are the natural numbers; fractions are just as alien to this context as the complex numbers are alien to most real-world measurements. And yet, despite this, allowing ourselves to move from the natural numbers to the larger set of rational numbers enabled us to deduce something about the real world situation, even though measurements in that particular real world situation only involve natural numbers.

In the same way, being willing to think about what happens in the larger set of complex numbers allows us to draw conclusions about real world situations even when actual measurements in that particular real world situation only involve the real numbers. This happens all the time in engineering applications. If you've seem some calculus, you may be able to understand the following example: in engineering and other applications, one very often has to solve an equation of the form $ay'' + by' + cy = 0$ (*) for the unknown function y.

It turns out that there's a way to get the solutions provided one can solve the quadratic equation $ar^2 + br + c = 0$ for the variable r. In the real numbers, there may not be any solutions. However, in the complex numbers there are, so one can find all complex-valued solutions to the equation (*), and then finally restrict oneself to those that are purely real-valued. The starting and ending points of the argument involve only real numbers, but one can't get from the start to the end without going through the complex numbers. Since equations like (*) need to be solved all the time in real-life applications such as engineering, complex numbers are needed.

3 Does the number e have any real physical meaning, or is it just a mathematical convenience?

Yes, the number e does have physical meaning. It occurs naturally in any situation where a quantity increases at a rate proportional to its value, such as a bank account producing interest, or a population increasing as its members reproduce.

Obviously, the quantity will increase more if the increase is based on the total current quantity (including previous increases), than if it is only based on the original quantity (with previous increases not counted). How much more? The number e answers this question.

To put it another way, the number e is related to the how much more money you will earn under *compound interest* than you would under simple interest.

Therefore, we will address the following topics:

- Simple and Compound Interest:

  a discussion of the meaning of simple and compound interest, and related concepts in more general situations such as population growth.

- A Physical Meaning for the Number e:

  how the concepts of simple and compound interest lead to a physical interpretation of the number e in one particular case (the case in which your money would double under simple interest).

- The General Situation:

how the number e enters in to the picture when we aren't in this one particular case.

- The Number e as a Limit:

  what happens when interest is compounded only at specific times instead of continually, and how this yields a mathematical definition for the number e as a certain limit.

- The Number e in Calculus:

  why the number e is an especially important and natural one in calculus.

.

3.1 Simple and Compound Interest

If a bank account produces "simple interest", the interest is paid directly and does not become part of the account's interest-bearing balance. That balance remains constant, and it's easy to figure out how much money you earn. For instance, if a \$100 account earns 10% simple interest a year, the balance remains at \$100 and each year's interest payment is \$10.

The situation is the same for population growth if the offspring produced are not capable of further reproduction: the fertile population, and hence the rate of growth, remain constant.

But what if the offspring *can* reproduce further? Or, to return to the bank example, what if the interest becomes part of the interest-bearing balance as soon as it is earned, so that you earn interst on the interest, and interest on the interest on the interest, etc. (this is called "compound interest")? How much money will you end up with? How different is it from what you would end up with under simple interest alone?

The answer depends on how large the simple interest is compared to the original amount. If it's only a small fraction, then including it with the original amount won't change that amount by much, so the compound interest (interest on the original amount plus previously earned interest) won't be much different from the simple interest on the original amount alone. The greater the simple interest is compared to the original amount, the more of an effect compounding will have.

Let's ask the question in the case when the ratio is 1 (simple interest = 100% of original amount):

Question: *If you would earn 100% interest (i.e., your money would double) under simple interest, how much money would you end up with under compound interest*?

Answer: *You would have e times your original amount.*

This gives us the following physical meaning for the number e:

3.2 A Physical Meaning for the Number e

The number e is the factor by which a bank account earning continually compounding interest (or a reproducing population whose offspring are themselves capable of reproduction, or any similar quantity that grows at a rate proportional to its current value) will increase, if, without the compounding (or without the offspring being capable of further reproduction) it would have doubled (increased by 100%).

3.3 The General Situation

We've seen that e is the factor by which a continually-compounding bank account will increase if under simple interest it would have doubled (increased by 100%).

What if the simple interest isn't 100%? In other words, suppose your money wouldn't have doubled under simple interest, but increased by some other factor?

Suppose for instance the simple interest is 200%. Then we can split the time period up into two halves, with the simple interest being 100% for each half (for example, if you earn 200% simple interest per year, you're earning 100% interest for each six-month period).

At the end of the first half and the beginning of the second half, the balance under compound interest will be e times the original balance.

At the end of the second half, the balance under compound interest will be e times the balance at the beginning of the second half. In other words, it will be

e times (e times original amount)

which is e^2 times the original amount. So, if the simple interest earned is 200% (2 times original amount), the final balance under compound interest is e^2 times the original amount.

This turns out to be true in general. If the simple interest earned is R times the original balance, then the final balance under compound interest is e^R times the original balance.

This is one of the reasons why exponentials of the form e^x occur more frequently in practice than do exponentials with other bases such as 10^x, and why one usually uses exponentials and logarithms base e rather than base 10 or any other base. (The other, more important, reason has to do with the special role that base e exponentials play in calculus).

3.4 The Number e as a Limit

So far we've talked about "continually compounding" interest: interest is assumed to be earned continually. In many bank accounts, however, interest only starts becoming eligible to earn further interest on specific dates (such as the end of each month, or the end of each year) when it is officially added in to your balance. The calculation of compound interest is different in this case, and this different calculation gives a nice mathematical definition of the number e.

Suppose the simple interest rate is 100% for a certain period. Under simple interest alone, the money will double during that period. Under continually compounding interest, it will be multiplied by e. But what happens if interest is compounded at n equally spaced dates during that period (such as at the and of each month), rather than continually?

In the interval before the first compounding date, you are just earning simple interest on the original balance. Let's call the original balance B. The interest rate for this interval is $1/n$, since the total interest rate for all n intervals is 1 (100% is just another way of writing 1). So, on the first compounding date, one gets paid interest in the amount of $(1/n)B$. The balance after this payment is therefore $B + (1/n)B$ which equals $(1 + (1/n))B$. In other words:

(balance at end of first interval) $= (1+(1/n))$ times (balance at start of first interval)
$= (1 + (1/n))B$.

During the second interval (between the first and second compounding dates), one is again earning interest at a rate of $1/n$, but this time on this new balance $(1 + (1/n))B$. Therefore,

(balance at end of second interval) $= (1 + (1/n))$ times (balance at start of second interval)

$= (1 + (1/n))$ times $(1 + (1/n))B$
$= (1 + (1/n))^2 B.$

If you do this same calculation again, you find that the balance at the end of the third interval is $(1 + (1/n))^3 B$ and so on. The balance at the end of the nth and final interval is $(1 + (1/n))^n B$.

Therefore: if the interest is compounded n times during the period, the final balance is $(1 + (1/n))^n$ times the original balance.

Compounding at more and more frequent intervals means that one is approximating more and more closely the idea of *continual compounding* which we talked about earlier, in which the final balance is e times the original balance. What this means is that

The number e is the limit of the quantity $(1 + (1/n))^n$ as n goes to infinity.

This gives a nice mathematical definition of the number e.

3.5 The Number e in Calculus

In calculus there is the notion of a *derivative* of a function, which is a measure of its rate of change with respect to changes in its input. When one differentiates an exponential function of the form $f(x) = a^x$ one gets

$$\begin{aligned} f'(x) &= \lim_{h \to 0} \frac{f(x+h) - f(x)}{h} \\ &= \lim_{h \to 0} \frac{a^{x+h} - a^x}{h} \\ &= \lim_{h \to 0} \frac{a^x a^h - a^x}{h} \\ &= a^x \lim_{h \to 0} \frac{a^h - 1}{h} \end{aligned}$$

which is a^x times a number (the limit of $(a^h - 1)/h$ as h goes to 0) which is a constant independent of x, depending only on the base a of the exponential.

Therefore, the derivative of an exponential function is just a constant times the function value. What the constant is depends on what the base of the exponential function is.

The number e is that value of the base which yields the constant 1, so that the derivative of the function $f(x) = e^x$ is actually equal to e^x itself.

Moreover, the derivative of the more general function $f(x) = Be^{Rx}$ is RBe^{Rx}. If we were using some other number a as the base instead of e, there would be an additional constant out front.

Not only does this make it more natural and convenient to use e as the base in exponentials and logarithms, but it relates back to the compound interest interpretation given earlier. If a bank account is growing under compound interest, with an interest rate of R per unit time, that means that at any instant in time the growth rate is R times the current balance. We've already seen that the balance at the end of the period is Be^R where B is the beginning balance. More generally, to figure out the balance somewhere in the middle of the time period, let t denote the fraction of the time period that has passed; it turns out that the balance at that time is Be^{Rt}. The derivative of this function is BRe^{Rt} which is R times the function value; this is just another way of saying that the growth rate is R times the current balance.

4 What Does "Existence" Mean in Mathematics?

Understanding what *existence* means in mathematics is the key to understanding what it means for concepts like "infinity" or "imaginary numbers" to exist—something that puzzles a lot of people when they first encounter these weird ideas!

Mathematical objects do not exist in the same sense that a physical object exists; nobody has ever bumped their elbow on a number, for instance.

Instead, mathematical objects are abstract concepts (often abstracted from a real world situation, by isolating just the part of the situation that is relevant for a particular discussion).

When we ask whether or not a mathematical object exists, we must have in mind an appropriate context: a particular, precisely defined collection of concepts. Then we ask, "among these concepts, is there one which matches the object we are looking for?" If so, we say that the object exists; if not, it doesn't exist.

For example, the *natural numbers* (that is, the numbers 1, 2, 3, 4, and so on) are the concepts obtained by abstracting the property of "size" from collections of objects. The number 2 is the abstract concept that expresses what the following collections of objects have in common: the eyes on a person's face, the occurrences of the letter "b" in the word "bib", the wheels on a bicycle, and so on.

If we were to ask "does there exist a natural number between 1 and 2", we mean, "among the collection of natural numbers, is there one (say x) such that $1 < x < 2$?" The answer to this question is "No". You cannot, for example, go to the beach and pick up more than one pebble but fewer than two pebbles.

If we were to ask "does there exist a number between 1 and 2", the answer would depend on what we meant by the word *number*.

If we were using the word "number" to mean "natural number" (that is, measurement of the size of a set), then the answer would be "No; no such number exists."

However, there are other contexts in which the answer might be "Yes". For example, we might be in a context where *number* is meant to refer, not to a natural number, but to a *rational number*: that is, a fraction.

Rational numbers are something quite different from natural numbers: instead of being measurements of sizes of sets, they are *ratios* of the sizes of two sets. For example, the fraction 3/4 is expressing the ratio "3 to 4".

In this context, where *number* refers to a ratio not to the size of a single set, and where "2" and "1" really mean the fractions "2/1" and "3/1" respectively, then the answer to our question is "yes": there does exist a number between 1 and 2, for instance the fraction 3/2.

This illustrates that, before one can say whether a concept exists or not, we have to be quite clear about the context in which the question is being asked.

If you are still puzzled by this, you might want to read more (in a discussion of whether or not "imaginary numbers" exist) about how there are many quite different meanings for the word "number", and how whether or not a concept exists can depend on the meaning you have in mind.

5 Number Systems with Different Kinds of Numbers

In mathematics, the phrase *number system* refers to any collection of objects that has the following properties:

- there is a rule for how to add two objects together,
- there is a rule for how to multiply two objects together,
- the rules for addition and multiplication satisfy the familiar properties of arithmetic, such as
  - *commutativity* (order doesn't matter),
  - *associativity* (in a sum of three of more terms, it doesn't matter which two you add first, and likewise for products), and
  - *distributivity* ($a(b+c) = ab + ac$), and
- all objects have corresponding *negatives* (from which we can define subtraction as the addition of one object to the negative of another), and some objects have reciprocals (from which we can define division as the multiplication of one object by the reciprocal of another). The properties of subtraction and division follow automatically from the corresponding properties of addition and multiplication, so we need not mention them explicitly.
- In some cases we might consider two objects in the collection to be equal even if they are not the same object; for instance, the fractions 1/2 and 2/4 are equal, even though they are different pairs of numbers. When this is the case, we also need, in addition to the properties above, a precise definition of when two objects are equal.

Roughly speaking, *any collection of objects that satisfies these properties is, by definition, a number system.*

Note: this is only a rough statement. If you wanted to say things more accurately, you'd have to take into account the fact that

- strictly speaking, some of these properties need to be stated a little more precisely than I stated them;
- the phrase "number system" is not often used; instead, one uses more fancy language such as "field" or "commutative ring"; and
- usually the rules for addition and multiplication satisfy various other properties besides the ones I have listed, and one uses terminology that precisely expresses which of these properties are present.

However, the rough statement is enough for our purposes.

5.1 Examples of Number Systems

The most common number systems are the following:

- *The Natural Numbers.* These are the counting numbers 1, 2, 3, ... that are possible answers to the question "how many?" They are abstract concepts that describe sizes of sets.
- *The Integers.* These are abstract concepts that describe, not sizes of sets, but the *relative sizes* of two sets. They are the possible answers to the question "how many more does A have than B has?" They include both positive numbers (meaning A has more than B) and negative numbers (meaning B has more than A).

- *The Rational Numbers.* These are abstract concepts that describe *ratios of sizes of sets.* They do not model sizes of sets the way that natural numbers do. If you say "I ate $3/4$ of a pie", you are not saying that the set of things you ate had $3/4$ elements. Instead, you are expressing a ratio of two integer quantities: 3, the number of pie-quarters that you ate, and 4, the number of pie-quarters that make up a whole pie.

- *The Real Numbers.* These are abstract concepts that describe *measurements of continuous quantities*, such as length, weight, quantity of fluid, etc. (Don't let the word "real" fool you; the real numbers are no more "real" in the ordinary English sense of the word than are any other kind of numbers.)

- *The Complex Numbers.* These are pairs of real numbers, with pairs of the form $(x, 0)$ behaving the same as ordinary real numbers x, but with other pairs (whose second entry is non-zero) behaving differently. The rule for multiplication is

  $$(a, b) \text{ times } (c, d) = (ac - bd, ad + bc)$$

  which means that the pair $(0, 1)$, when squared, gives $(-1, 0)$ which in this context is considered to be the same as the real number -1 (since a complex number of the form $(x, 0)$ is indistinguishable by its arithmetic properties from the real number x, we can consider it to be describing the same underlying concept, much as we consider a fraction of the form $x/1$ to be describing the same underlying concept as the integer x).

  Thus, the complex numbers are a number system in which there is an object whose square is -1. This object is denoted by i, and the pair (a, b) is written as "$a + bi$". When $a = 0$, such a number is called an "imaginary number".

  Complex numbers are extremely important mathematically. They have less direct relevance to real-world situations, since they aren't measurements of single quantities. However, they relate directly to a few real-world situations (such as the strength of an electromagnetic field), and they relate indirectly to many more since many properties of real numbers can be more easily understood in the larger context of complex numbers.

  More information on this is available in section 2.