Chapter 10

Means, sequences, and series

From *Algebra: For the Enthusiastic Beginner* (Draft version, July 2024)

In mathematics, a *sequence* (also called a *progression*) is a list of quantities that follow a particular pattern. There are many types of sequences/progressions, but the two most common ones are *arithmetic* and *geometric* sequences. The goal of this chapter is to study these sequences and learn about their properties and the wide variety of problems you can solve with them.

We'll begin in Section 10.1 by discussing *means*, which serve as a warmup to general sequences. We'll look at the *arithmetic mean* (which is a fancy name for the average) and the *geometric mean*. We'll then see in Section 10.2 how these two means relate to each other via the Arithmetic-Mean/Geometric-Mean (AM-GM) inequality. In Sections 10.3 and 10.4 we'll study arithmetic *sequences* and *series*. (A series is the sum of the terms in a sequence.) We'll then cover geometric sequences and series in Sections 10.5 and 10.6. We'll derive the formulas for the various types of series along the way.

This chapter contains many interesting examples and exercises. In particular, three classics are Exercise 10.31 (Fibonacci sequence), Exercise 10.40 (bicyclists-and-bird problem), and Exercise 10.41 (Archimedes' parabola theorem).

10.1 Means

10.1.1 Arithmetic mean

The *arithmetic mean* (AM) of two numbers a and b is often denoted by the Greek letter μ (mu) and is defined as

$$\text{Arithmetic Mean} \equiv \boxed{\mu \equiv \frac{a+b}{2}} \tag{10.1}$$

For example, the AM of 4 and 10 is $(4+10)/2 = 14/2 = 7$, and the AM of 3 and 8 is $(3+8)/2 = 11/2 = 5.5$. It's fine if one or both of the numbers are negative; the AM of -5 and 11 is $(-5+11)/2 = 6/2 = 3$. In the special case where $a = b$, the AM equals $(a+a)/2 = a$. So the AM of two identical numbers equals their common value.

The arithmetic mean (AM) is just a fancy name for the *average*: The AM (the average) of a and b is the number that is halfway between a and b. Equivalently, the AM is the same distance from both a and b. For example, the AM of 4 and 10 is 7, because 7 is 3 larger than 4, and 10 is also 3 larger than 7. So 7 is the same distance (namely 3) from both 4 and 10. We can demonstrate this equidistance in general. We want to show that

$$\mu - a = b - \mu. \tag{10.2}$$

(We're assuming here that b is larger than a, but that actually doesn't matter.) Using our definition of the AM in Eq. (10.1), this becomes

$$\frac{a+b}{2} - a = b - \frac{a+b}{2} \implies \frac{a+b}{2} - \frac{2a}{2} = \frac{2b}{2} - \frac{a+b}{2} \implies \frac{b-a}{2} = \frac{b-a}{2}, \tag{10.3}$$

which is indeed true.

Alternatively, we could have gone in the other direction. We could have *defined* the AM (again call it μ) of a and b to be the number that is equidistant from a and b. To find the value of μ, we then simply need to write down the two distances and equate them. This just means writing down Eq. (10.2) again, and solving for μ:

$$\mu - a = b - \mu \implies 2\mu = a + b \implies \mu = \frac{a+b}{2}, \tag{10.4}$$

in agreement with our original definition of μ in Eq. (10.1). So the two definitions (the one in Eq. (10.1), and the "equidistant" one) are equivalent.

Another way of stating the equidistant property is to say that whatever number d we add to a to get μ, it's the same number d that we add to μ to get b. For example, if we start with $a = 4$, then we need to add $d = 3$ to get $\mu = 7$, and then we need to add $d = 3$ again to get $b = 10$. In other words, a, μ, and b can be written in the form,

$$a, \qquad \mu = a + d, \qquad b = a + 2d, \tag{10.5}$$

where d is the common distance.

Eq. (10.1) gives the definition of the AM of two numbers. We can also define the AM of any number of numbers; it's again just the average (the sum divided by the number of numbers). So the AM of n numbers is

$$\boxed{\mu \equiv \frac{a_1 + a_2 + \cdots + a_n}{n}} \tag{10.6}$$

For example, the AM of the three numbers -2, 5, and 9 is $(-2+5+9)/3 = 12/3 = 4$. We used the a_k index notation in Eq. (10.6); see the "Index notation" entry in the Glossary.

A note on pronunciation: When the word "arithmetic" is used as a noun (as in "Let's do some arithmetic"), it is pronounced aRITHmetic, with the accent on the second syllable, as you undoubtedly know. However, when "arithmetic" is used as an adjective (as in "arithmetic mean," or in "arithmetic sequence" which we'll encounter in Section 10.3), it is pronounced arithMETic, with the accent on the third syllable.

Exercise 10.1 The AM of 5 and what other number equals 13? What if the AM is instead -2?

Exercise 10.2 Eq. (10.2) can alternatively be written as $(\mu - a) + (\mu - b) = 0$. Show that the analogous statement for n numbers $a_1, \ldots, a_n$ leads to Eq. (10.6).

10.1.2 Geometric mean

The *geometric mean* (GM) of two numbers a and b (which we'll label as g) is defined as

$$\text{Geometric Mean} \equiv \boxed{g \equiv \sqrt{ab}} \tag{10.7}$$

For example, the GM of 3 and 12 is $\sqrt{3 \cdot 12} = \sqrt{36} = 6$, the GM of 4 and 100 is $\sqrt{4 \cdot 100} = \sqrt{400} = 20$, the GM of 5 and 17 is $\sqrt{5 \cdot 17} = \sqrt{85} \approx 9.22$, and the GM of 1/8 and 32 is $\sqrt{(1/8) \cdot 32} = \sqrt{4} = 2$. Since the GM involves a square root, a and b need to be both positive or both negative. But people generally restrict the GM to positive numbers. In the special case where $a = b$, the GM equals $\sqrt{a \cdot a} = a$. So the GM of two identical numbers equals their common value.

The GM of a and b is the number that is "halfway" between a and b, in a *multiplicative* sense (as opposed to an *additive* sense with the AM above). By this we mean that the multiplicative *factor* (as opposed to the additive *difference*) by which g is larger than a is the same as the factor by which b is larger than g. For example, the GM of 3 and 12 is 6. And 6 is larger than 3 by a factor of 2, and 12 is larger than 6 also by a factor of 2.

Equivalently, the "halfway" (or "equal factor") property of the GM says that the ratio of g to a equals the ratio of b to g:

$$\frac{g}{a} = \frac{b}{g}. \tag{10.8}$$

Let's check that the definition of g in Eq. (10.7) satisfies this relation. Plugging $g = \sqrt{ab}$ into Eq. (10.8) gives

$$\frac{\sqrt{ab}}{a} = \frac{b}{\sqrt{ab}} \implies \frac{\cancel{\sqrt{a}}\sqrt{b}}{\cancel{\sqrt{a}}\sqrt{a}} = \frac{\sqrt{b}\cancel{\sqrt{b}}}{\sqrt{a}\cancel{\sqrt{b}}} \implies \sqrt{\frac{b}{a}} = \sqrt{\frac{b}{a}}, \tag{10.9}$$

which is indeed true. Or you could just cross multiply in the first of these equations; the result is $ab = ab$.

As with the AM, we could have gone the other way and *defined* the GM via the "equal ratio" expression in Eq. (10.8), and then solved for g:

$$\frac{g}{a} = \frac{b}{g} \implies g^2 = ab \implies g = \sqrt{ab}, \tag{10.10}$$

in agreement with the definition of g in Eq. (10.7).

Again, as with the AM, another way of stating the "equal factor" property is to say that whatever number r we *multiply* a by to get g, it's the same number r that we multiply g by to get b. For example, if we start with $a = 3$, then we need to multiply it by $r = 2$ to get $g = 6$, and then we need to multiply 6 by $r = 2$ again to get $b = 12$. In other words, a, g, and b can be written in the form,

$$a, \qquad g = ar, \qquad b = ar^2, \tag{10.11}$$

where r is the common ratio.

Eq. (10.7) gives the definition of the GM of two numbers. More generally, for n numbers the GM is defined to be the nth root of the product of all the numbers, instead of the square root of the two numbers we had in Eq. (10.7). That is,

$$\boxed{g = \sqrt[n]{a_1 a_2 \cdots a_n} = (a_1 a_2 \cdots a_n)^{1/n}} \tag{10.12}$$

For example, the GM of the three numbers 18, 20, and 75 is $(18 \cdot 20 \cdot 75)^{1/3} = 27{,}000^{1/3} = 30$.

Remarks:

1. All of the above results for the GM are parallel to the results for the AM, as you can see by comparing the six equations in Eqs. (10.1)–(10.6) with, respectively, the six equations in Eqs. (10.7)–(10.12). The difference between the two sets of equations is that addition in the AM case is replaced with multiplication in the GM case. A general consequence of this replacement is that multiplying a number by n (adding n copies of it) gets replaced with raising a number to the nth power (multiplying n copies of it). Likewise, dividing a number by n gets replaced with raising a number to the $1/n$th power; see Eqs. (10.6) and (10.12).

2. Geometrically, the GM of two numbers, namely $\sqrt{ab}$, is the length of the side of a square that has the same area as a rectangle with side lengths a and b. This is true because the area of the square is $\left(\sqrt{ab}\right)^2 = ab$, which equals the area ab of the rectangle. Similarly, the GM of three numbers, namely $(abc)^{1/3}$, is the length of the side of a cube that has the same volume as a rectangular block with side lengths a, b, and c. This is true because the volume of the cube is $\left((abc)^{1/3}\right)^3 = abc$, which equals the volume abc of the rectangular block; see Section A.5.1 in Appendix A.

3. There is a nice interpretation of the geometric mean in terms of right triangles. Referring back to Fig. A.52 in Appendix A, the geometric mean of the lengths BD and DA is $\sqrt{(a^2/c)(b^2/c)} = ab/c$, which equals CD. So the altitude to the hypotenuse equals the GM of the lengths into which the hypotenuse is divided: $CD = \sqrt{(BD)(DA)}$.

4. Fig. 10.1 gives a visual description of the difference between the AM and GM. The AM is the same *distance* d away from two given numbers, while the GM is the same *factor* r away from the two numbers. For example, the GM of 4 and 36 is 12, because 12 is the same factor (namely 3) away from both 4 and 36: If you divide 12 by 3, you get 4. And if you multiply 12 by 3, you get 36.

   For any given number N, if someone asks you to produce two numbers whose AM is N, that's easy. There is an infinite number of answers. You just pick any arbitrary number d, and then add and subtract it from N to obtain $N + d$ and $N - d$. Likewise for the GM, but

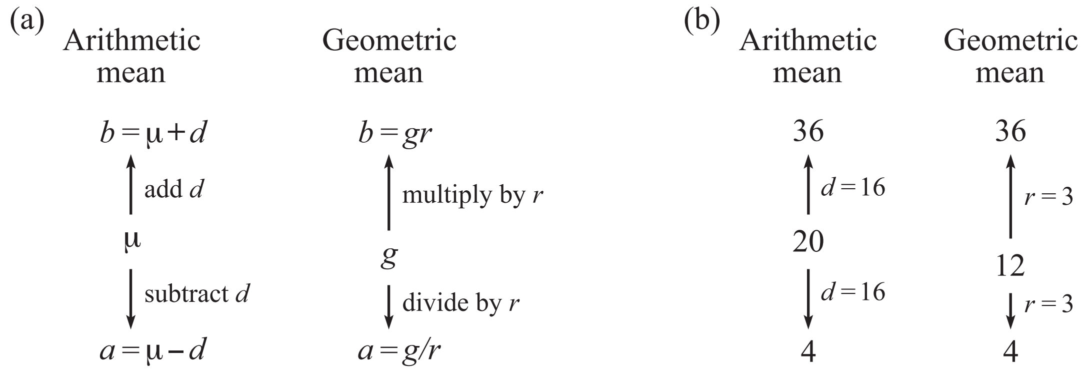

Figure 10.1: (a) The AM (μ) is the same *distance* d from a and b, while the GM (g) is the same *factor* r from a and b. (b) A numerical example for each is shown.

now you pick any arbitrary number r, and then multiply and divide by it to obtain Nr and N/r. From Eq. (10.1) the AM of $N+d$ and $N-d$ is correctly N. And from Eq. (10.7) the GM of Nr and N/r is also correctly N.

5. For two numbers, we have seen that the AM of a and b is the number μ that is halfway between a and b, where we mean halfway in the usual additive sense: Adding d to a gives μ, and then adding another d to μ gives b; see Eq. (10.5). We have also seen that the GM of a and b is the number g that is "halfway" between a and b, in a *multiplicative* sense. Multiplying a by r gives g, and then multiplying g by another r gives b; see Eq. (10.11). These facts are also encapsulated in Fig. 10.1.

   If you ask for the number that is "halfway" between two given numbers, most people will probably assume that you mean additively. But technically the word "halfway" is a little ambiguous, because you could mean additively or multiplicatively. Additively, the number that is halfway between 10 and 1000 is 505 (495 away from each), but multiplicatively it is 100 (a factor of 10 away from each). The meaning of "halfway" is usually clear from the context, but be sure to specify, if there's any chance of confusion. ♣

> What do we mean by the mean,
> This number that's halfway between?
> For 32, 8,
> You can have a debate:
> Is it 20, or only 16?

Exercise 10.3 The GM of 4 and what other number equals 12? What if the GM is instead $\sqrt{2}$?

Exercise 10.4 Eq. (10.8) can alternatively be written as $(g/a)(g/b) = 1$. Show that the analogous statement for n numbers $a_1, \ldots, a_n$ leads to Eq. (10.12).

Exercise 10.5 Show that the log of the GM of a and b equals the AM of the logs of a and b, for any value of the base.

Other means

In addition to the arithmetic and geometric means, there are other means, including:

$$\text{Root Mean Square (RMS):} \quad \boxed{\sqrt{\frac{a^2+b^2}{2}}}$$

$$\text{Harmonic Mean (HM):} \quad \boxed{\frac{2ab}{a+b}} \tag{10.13}$$

The "root mean square" name comes from the fact that $\sqrt{(a^2+b^2)/2}$ is the square *root* of the *mean* (the average) of the *squares* of a and b.

The harmonic mean has the property that the reciprocal of the HM is the AM (the average) of the reciprocals of a and b:

$$\frac{1}{\text{HM}} = \frac{1}{2}\left(\frac{1}{a}+\frac{1}{b}\right). \tag{10.14}$$

If you take the reciprocal of the HM in Eq. (10.13), you will indeed obtain the expression in Eq. (10.14).

We've defined the RMS and HM for two numbers a and b, but as with the AM and GM, they both generalize to an arbitrary number n of numbers. For example, the RMS of three numbers is $\sqrt{(a^2+b^2+c^2)/3}$. And the HM generalizes via Eq. (10.14); the HM of three numbers is given by $1/\text{HM} = (1/3)(1/a+1/b+1/c)$. The RMS is ubiquitous in statistics, and the HM makes appearances in all sorts of subjects (travel times, optics, finance, etc.). Both the RMS and HM have interesting geometric interpretations, which you are encouraged to investigate on your own.

10.2 AM-GM inequality

10.2.1 Derivation

It turns out that for two positive (or zero) numbers a and b, their arithmetic mean is always greater than or equal to their geometric mean:

$$\boxed{\frac{a+b}{2} \geq \sqrt{ab}} \qquad \text{(AM-GM inequality)} \tag{10.15}$$

This is called the "Arithmetic-Mean/Geometric-Mean inequality," or "AM-GM inequality" for short, since the full name is a bit of a mouthful. Both a and b need to be positive (or zero) in this inequality, because if one of them is negative, then the product ab is negative, so we can't take the square root (without introducing imaginary numbers). And if they're both negative, then the inequality isn't true, since the lefthand side is negative while $\sqrt{ab}$ is positive by definition.

As a numerical example, let's pick $a = 3$ and $b = 12$. Eq. (10.15) then says that $(3+12)/2 \geq \sqrt{3\cdot 12} \implies 7.5 \geq 6$, which is indeed true. As another example, if $a = 1$ and $b = 9$ then Eq. (10.15) says that $(1+9)/2 \geq \sqrt{1\cdot 9} \implies 5 \geq 3$, which is again true. The numbers in the limerick on page 587 provide a third example. And consistent with Eq. (10.15), we made sure to place g lower than μ in Fig. 10.1; g is always less than (or equal to) μ.

Equality holds in Eq. (10.15) if $a = b$. This is true because if $a = b$, then Eq. (10.15) becomes $(a+a)/2 \geq \sqrt{aa} \implies a \geq a$, so the two sides are equal.

This neat inequality's claim
Is conveyed by its rather long name.
And equality holds
If by chance it unfolds
That your a and your b are the same.

It is also true (as we'll see in the proofs below) that equality holds in Eq. (10.15) *only if* $a = b$. That is, if a and b are *not* equal, then equality does *not* hold; the lefthand side of Eq. (10.15) (the AM) is strictly greater than the righthand side (the GM). (This is the case in the two examples above that resulted in $7.5 \geq 6$ and $5 \geq 3$.) If we combine this "only if" statement with the "if" statement in the preceding paragraph, we can say that equality holds in Eq. (10.15) *if and only if* $a = b$. See the "If and only if" section following the exercises below.

For more than two numbers, instead of just the a and b in Eq. (10.15), the AM-GM inequality also holds and takes the general form:

$$\boxed{\frac{a_1 + a_2 + \cdots + a_n}{n} \geq \big(a_1 a_2 \cdots a_n\big)^{1/n}} \qquad \text{(General AM-GM inequality)} \qquad (10.16)$$

You should verify this for some arbitrarily chosen numbers. For example, if you pick the three numbers 3, 5, and 10, then Eq. (10.16) says that $(3+5+10)/3 \geq (3\cdot 5\cdot 10)^{1/3}$, which is (as you can check) equivalent to $6 \geq 5.31$, which is indeed true. Similar to Eq. (10.15), equality in Eq. (10.16) is achieved if and only if all of the n numbers are equal.

The formal proof of the general case in Eq. (10.16) is somewhat involved, so we won't present that here. But Exercises 10.6–10.10 present six proofs of the $n = 2$ case in Eq. (10.15).

Exercise 10.6 (Proof 1) The square of any number is greater than or equal to zero. In particular, if we judiciously pick the number to be $\sqrt{a} - \sqrt{b}$, then $(\sqrt{a} - \sqrt{b})^2 \geq 0$. Use this to derive the AM-GM inequality.

Exercise 10.7 (Proof 2) Start with the true relation $(a-b)^2 \geq 0$, and then add a helpful quantity to both sides.

Exercise 10.8 (Proofs 3 and 4) This exercise gives two proofs.

(a) From Eq. (10.5) or Fig. 10.1, we can write a and b as $a = \mu - d$ and $b = \mu + d$. Use these forms to prove the AM-GM inequality.

(b) Similarly, from Eq. (10.11) or Fig. 10.1, we can write a and b as $a = g/r$ and $b = gr$. Use these forms to prove the AM-GM inequality. You will need to use the result from Example 8.9.

Exercise 10.9 (Proof 5) If we square both sides of Eq. (10.15) and multiply by 4, we see that proving the AM-GM inequality is equivalent to showing that $(a+b)^2 \geq 4ab$. Algebraically, this relation is true from the reasoning in the second proof above. Show geometrically that $(a+b)^2 \geq 4ab$ is true, by drawing four helpful rectangles inside a square with side length $a+b$.

Exercise 10.10 (Proof 6) Prove the AM-GM inequality by finding the lengths of the two dashed lines in the semicircle in Fig. 10.2. You will need to use the fact that the h dashed line divides the large right triangle into two similar smaller right triangles; see Exercise A.12 in Appendix A.

You can accept the fact that the top angle in the large right triangle is a right angle, which means that the result in Exercise A.12 is in fact applicable here. The right angle follows from the fact that the angle is inscribed in a semicircle.

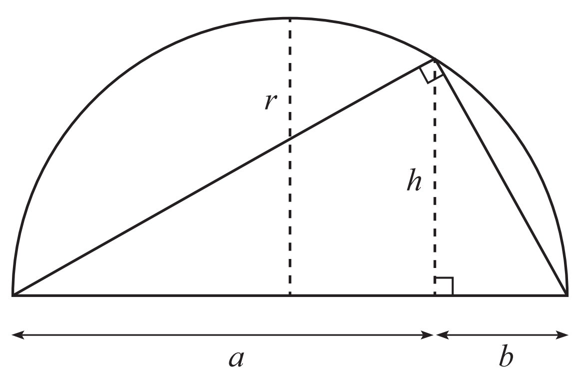

Figure 10.2: The setup for a geometric proof of the AM-GM inequality.

Exercise 10.11 At the end of Section 10.1 we introduced the RMS and the HM. It turns out that the four types of means we have encountered satisfy

$$\text{RMS} \geq \text{AM} \geq \text{GM} \geq \text{HM}. \tag{10.17}$$

We have already established the middle AM-GM inequality. Prove the other two inequalities.

Exercise 10.12 The Pythagorean theorem (see Section A.6) says that the sides of a right triangle satisfy $a^2 + b^2 = c^2$. For a given length c of the hypotenuse, apply the AM-GM inequality in a helpful manner to find the maximum possible value of the product ab of the lengths of the legs.

10.2.2 If and only if

[*This section is an aside on some logic terminology. It isn't critical at the moment, so it can be skipped on a first reading. But it's very important for math in general (and life too), so be sure to come back to it at some point.*]

We stated above (in the paragraph following the limerick) that equality holds in Eq. (10.15) *if and only if* $a = b$. What exactly does the phrase "if and only if" mean? The quick answer is that the general statement, "A if and only if B," means that A and B go hand in hand. Whenever you have A, you also have B, and vice versa. You can't have one without the other. To understand why this is the case, we'll first consider an "if" statement, then an "only if" statement, and then we'll combine the two.

IF: Consider the statement, "You open your umbrella *if* it is raining." This says that whenever it is raining, you always open your umbrella. However, it doesn't say anything about situations where it isn't raining. You might very well also open your umbrella when it's sunny out, to give yourself some shade. This is perfectly consistent with the statement, "You open your umbrella if it is raining." You might open your umbrella many other times too.

This is summarized in Fig. 10.3(a). The small circle represents cases where it rains, and the large circle represents cases where you open your umbrella. (So all points in the "rain" circle are in both the "rain" and "open" circles.) Since the "rain" circle is completely contained within the "open" circle, we see that whenever it rains, it is necessarily also the case that you open your umbrella. (No points in the "rain" circle lie outside the "open" circle.) We can therefore say, "If it is raining, then you open your umbrella." This is the same as the original statement, "You open your umbrella if it is raining," simply rearranged. Note that there are many points in the large "open" circle that aren't in the small "rain" circle. These correspond to cases where you open your umbrella when it isn't raining (such as when you want some shade).

In Fig. 10.3 we're using circles to represent things that happen, such as "it rains." But the regions don't need to be circles. They can have any arbitrary shape, as long as one is completely inside the other. The regions are just abstract representations of happenings. Their exact shape is meaningless.

ONLY IF: Now consider the statement, "You open your umbrella *only if* it is raining." The "only if" here makes this statement have a very different meaning from the one above. The present statement says that you never open your umbrella when it isn't raining. Equivalently, the only times you open your umbrella are when it is raining. In other words, the statement

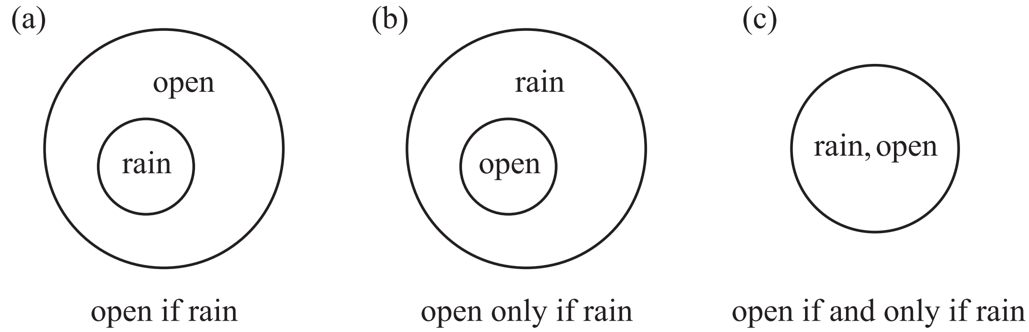

Figure 10.3: Visualizing "if," "only if," and "if and only if" statements in terms of circles. In (a), for example, the "rain" circle is completely contained within the "open" circle, which means that whenever it rains, you also open your umbrella. That is, "If it rains, then you open your umbrella."

says, "If you open your umbrella, then it must be raining." (So you *don't* open it just to get some shade.) However, the statement does *not* say that you open your umbrella *every* time it rains. There may be times when it rains and you don't open it.

This is summarized in Fig. 10.3(b), which has its circles reversed from Fig. 10.3(a). The "open" circle is now completely contained within the "rain" circle, which means that whenever you open your umbrella, it is necessarily also the case that it is raining. The previous sentence can be stated in two equivalent ways, as we noted in the preceding paragraph:

"You open your umbrella *only if* it is raining."
is equivalent to
"If you open your umbrella, then it must be raining." (10.18)

These statements both say that you never open your umbrella when it isn't raining. (No points in the "open" circle in Fig. 10.3(b) lie outside the "rain" circle.) But again, similar to Fig. 10.3(a), there are many points in the large "rain" circle that aren't in the small "open" circle. These correspond to cases where it rains but you don't open your umbrella.

IF AND ONLY IF: Finally, consider the statement, "You open your umbrella *if and only if* it is raining." The "if" part of this means that Fig. 10.3(a) describes the situation. And the "only if" part means that Fig. 10.3(b) also describes it. So the "rain" circle must be completely contained within the "open" circle; that's Fig. 10.3(a). And the "open" circle must be completely contained within the "rain" circle; that's Fig. 10.3(b). The only way both of these things can be true is if the two circles are actually the *same* circle, as shown in Fig. 10.3(c). That is, the instances where you open your umbrella are *exactly the same* as the ones where it rains.

The "if and only if" phrase therefore means that the two things go hand in hand; you can't have one without the other. Every time it rains, you open your umbrella. And every time you open your umbrella, it is raining. Said in another way, it never rains without you opening your umbrella, and you never open your umbrella when it isn't raining.

General result: The above results are summarized in Fig. 10.4, which is Fig. 10.3 with "rain" replaced with A, and "open" replaced with B, to be general. The two equivalent statements in Eq. (10.18) are now "B only if A" and "If B then A." These say the same thing. They both say that the B circle is completely contained within the A circle in Fig. 10.4(b). We can also switch A and B, of course, which corresponds to Fig. 10.4(a):

"A only if B" means the same thing as "If A then B."

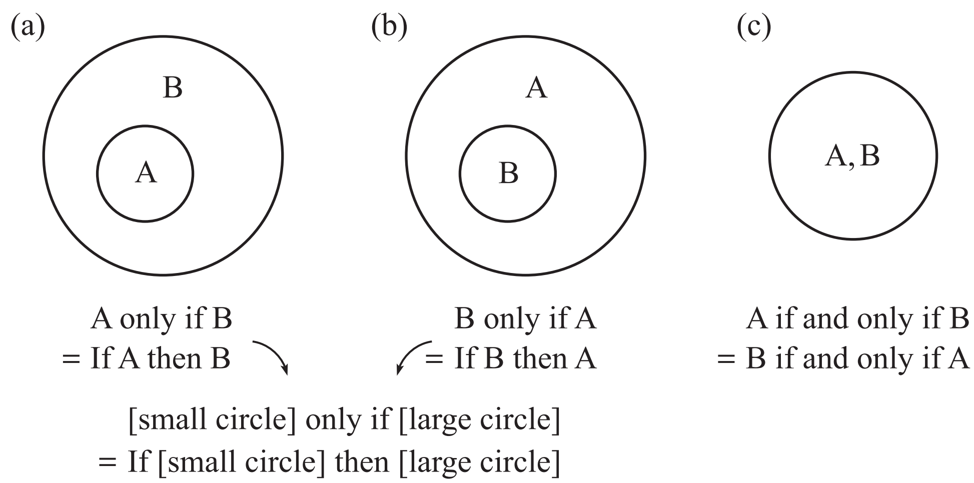

Figure 10.4: The general version of Fig. 10.3.

If you want to get right to the point, you can forgo the A and B letters and just talk in terms of small and large circles, as we have done at the bottom of Fig. 10.4. The two statements written there follow from the fact that the small circle is completely contained within the large circle. If you're in the small circle, you're automatically also in the large circle.

We see that to convert from an "only if" statement to an "if" statement, we can't just erase the word "only." But what we *can* do is erase the "only" while also moving the "if" to be in front of the other letter (and adding a "then" between the letters). So "A only if B" becomes "If A then B." Both of these statements say that the A circle is completely contained within the B circle in Fig. 10.4(a). That is, every A point is also a B point (but not necessarily the other way around; there are some B points that aren't A points). Every time you have A, you must also have B.

Fig. 10.4(c) corresponds to the statement, "A if and only if B," and also to the statement, "B if and only if A." These statements are equivalent. They both say that the circles representing A and B are the same circle. A and B go hand in hand. You can't have one without the other.

Notation: To save some writing, "iff" is used as shorthand for "if and only if." So "A iff B" means the same thing as "A if and only if B."

Returning to our earlier AM-GM-inequality statement, "Equality holds in Eq. (10.15) *if and only if* $a = b$," we see that equality goes hand in hand with $a = b$. Equality holds if $a = b$, and it doesn't hold if $a \neq b$.

10.2.3 Maxima and minima

[*This section is more abstract and can be skipped on a first reading.*]

The standard way of finding the maximum or minimum of a function involves calculus. However, we saw in Section 8.4.2 that the quadratic formula (in particular the discriminant) can also be used to find maxima and minima in certain circumstances (ones that generate a quadratic equation). It turns out that the AM-GM inequality can also be used to find maxima and minima. The method is most easily understood through some examples:

Example 10.1 Assuming that x is positive, what is the minimum value of the function $f(x) = x + 1/x$? What value of x produces the minimum? We already found the minimum value back in Example 8.9 by using the quadratic formula, but let's now use the AM-GM inequality.

Solution With $a = x$ and $b = 1/x$, the AM-GM inequality in Eq. (10.15) becomes

$$\frac{x + 1/x}{2} \geq \sqrt{x \cdot \frac{1}{x}} \implies x + \frac{1}{x} \geq 2. \tag{10.19}$$

So 2 is the desired minimum value. And we know that equality is achieved in the AM-GM inequality when $a = b$, which implies $x = 1/x \implies x^2 = 1 \implies x = 1$. ($x = -1$ is also a solution to this equation, but we're assuming that x is positive. And a and b are always assumed to be positive in the AM-GM inequality anyway.)

Example 10.2 Here is a slightly trickier example: Assuming that x is positive, what is the minimum value of $x + 1/x^2$? What value of x produces the minimum?

Solution We'll use the general AM-GM inequality in Eq. (10.16), with three numbers a, b, and c. We'll choose $a = b = x/2$ and $c = 1/x^2$. (See below for the reason why we chose these a and b values, which at the moment are admittedly a bit out of the blue.) Eq. (10.16) then gives

$$\frac{1}{3}\left(\frac{x}{2} + \frac{x}{2} + \frac{1}{x^2}\right) \geq \sqrt[3]{\frac{x}{2} \cdot \frac{x}{2} \cdot \frac{1}{x^2}} \implies x + \frac{1}{x^2} \geq 3\sqrt[3]{\frac{1}{4}} = \frac{3}{2^{2/3}}. \tag{10.20}$$

So the minimum value of $x + 1/x^2$ is $3/2^{2/3} = 1.89$. This is achieved when all three quantities in the inequality are equal, that is, when $x/2 = 1/x^2 \implies x^3 = 2 \implies x = 2^{1/3} = 1.26$. You can check that plugging this value of x into $x + 1/x^2$ yields $3/2^{2/3}$.

You might think that the minimum value of $x + 1/x^2$ should be 2, and be achieved at $x = 1$, just like in the preceding example with $x + 1/x$. However, in the present case, if you imagine starting x at 1 and then increasing its value, the value of $1/x^2$ will decrease faster than the value of x increases (at least up to a point, which apparently is $x = 1.26$). You can test this with, say, $x = 1.05$ and $x = 1.1$. The sum $x + 1/x^2$ decreases, at least for a little while, as you increase x above 1. The minimum value of $x + 1/x^2$ is therefore less than 2 (and happens to be 1.89).

If you plot the function $x + 1/x^2$ in Desmos and zoom in to the bottom of the curve (in the positive-x region), you can verify that it is located at (approximately) the point $(1.26, 1.89)$. You should perform the analogous Desmos verification for all the other examples and exercises in this section, too.

REMARK: Eq. (10.20) gave us the correct answer, but how did we know to choose a and b to be $x/2$? The first thing to note is that using Eq. (10.15) with the two quantities x and $1/x^2$ doesn't produce a constant number on the righthand side, like x and $1/x$ did in Eq. (10.19). So Eq. (10.15) doesn't help us. The minimum value we're looking for must be a fixed number (with no x's).

The goal is to write the given expression $x + 1/x^2$ as the sum of some number of terms whose product is a *constant* (involving no x's), so that we have a pure number on the righthand side of Eq. (10.20). And we quickly see that the three terms $x/2$, $x/2$, and $1/x^2$ get the job done, because the x's cancel in their product. If we were instead given the expression $x + 1/x^3$, then we would want to write it as the sum of the *four* terms $x/3, x/3, x/3$, and $1/x^3$, because the x's cancel in their product; see Exercise 10.14.

Technically, any a and b values that are multiples of x, and that add up to x, will produce a constant number on the righthand side of Eq. (10.20). For example, if we choose $a = 3x/4$ and $b = x/4$ (and $c = 1/x^2$), then Eq. (10.20) becomes $x + 1/x^2 \geq 3(3/16)^{1/3} = 1.72$, as you can verify. Does this mean that 1.72 is the correct minimum value, instead of the 1.89 result we obtained above?

No, because *equality* in the AM-GM inequality in Eq. (10.16) is achieved only if *all n numbers are equal*. And there is no way for $3x/4$ to be equal to $x/4$ (unless $x = 0$, which definitely doesn't yield a minimum). That's why we picked a and b in Eq. (10.20) to be equal (and hence equal to $x/2$, since they must add up to x). Picking other values for a and b (with $a + b = x$) will yield a perfectly valid inequality for $x + 1/x^2$, such as the $x + 1/x^2 \geq 1.72$ one above. But the point is that *equality will never occur* in such cases, because we can't make all of the numbers be equal. So we can't say that the minimum value equals 1.72. The function never actually gets that low. All we can conclude is that the actual minimum value is some number that is larger than 1.72. And it happens to be 1.89 from our earlier reasoning. ♣

Example 10.3 Assuming that x is positive, what is the maximum value of $(1 - x)x^2$? What value of x produces the maximum?

Solution Letting $a = 1 - x$ and $b = x^2$ in Eq. (10.15) doesn't give us anything useful, because although the given product appears on the righthand side, the lefthand side still involves x's. We want the lefthand side to be a constant, so that we obtain an actual number (with no x's) for the maximum value.

In the preceding example, we needed to find three quantities whose *sum* was $x + 1/x^2$, and whose *product* was a pure number. The goal now is to find three quantities whose *product* is $(1-x)x^2$ (or a constant multiple of this, that's fine), and whose *sum* is a pure number.

A little fiddling tells us that we want to pick the three quantities to be $1-x$, $x/2$, and $x/2$, because then the general AM-GM inequality in Eq. (10.16) gives

$$\frac{1}{3}\left((1-x)+\frac{x}{2}+\frac{x}{2}\right) \geq \sqrt[3]{(1-x)\cdot\frac{x}{2}\cdot\frac{x}{2}} \implies \frac{1}{3} \geq \sqrt[3]{\frac{(1-x)x^2}{4}}. \tag{10.21}$$

Cubing both sides leads to $4/27 \geq (1-x)x^2$. The desired maximum value is therefore $4/27$. This is achieved when all three quantities in the inequality are equal, that is, when $1-x = x/2 \implies x = 2/3$. You can quickly check that plugging this value of x into $(1-x)x^2$ yields $4/27$. You should verify with Desmos that the maximum of the function $(1-x)x^2$ (for positive x) is located at the point $(2/3, 4/27) \approx (0.67, 0.15)$.

Note that because the sums in Eqs. (10.15) and (10.16) are on the lefthand sides (the "greater than" sides), and the products are on the righthand sides (the "less than" sides), the above strategy for producing maxima and minima will always produce a *minimum* value for a *sum* (as in Examples 10.1 and 10.2), and a *maximum* value for a *product* (as in Example 10.3).

For a sum, your goal is to write the given expression as the sum of things whose *product* is a constant (involving no x's), as we did in Example 10.2. For a product, your goal is to write the given expression as the product of things whose *sum* is a constant (involving no x's), as we did in Example 10.3. It doesn't ruin anything to introduce an overall factor into the product, as we did with the 4 in Eq. (10.21).

The $x + 1/x$ case in Example 10.1 was fairly simple, because we didn't need to figure out how to rewrite the expression in order to produce a constant on one side of the AM-GM inequality; we could just use the given terms. Likewise for Exercise 10.13 below. In other cases, however, it takes some thought to figure out how you should rewrite the given expression and break the terms up into the correct combination that yields a constant on one side of the AM-GM inequality. But that's a small price to pay for a maximization/minimization procedure that doesn't require any calculus!

The AM and GM relation
Has a nice max and min application.
You can safely ignore
All that calculus lore,
But you *do* need the right combination.

Having said all this, the fact of the matter is that people rarely use the present AM-GM method to find the maxima and minima of functions. Calculus is the standard way. But the AM-GM method is a nice amusement. It falls into the category of things that you do simply because you *can*.

In the following four exercises, remember to check your answers by plotting the given functions in Desmos.

Exercise 10.13 Assuming that x is positive, what is the maximum value of $x(1-x)$? What value of x produces the maximum?

Exercise 10.14 Assuming that x is positive, what is the minimum value of $x + 1/x^3$? What value of x produces the minimum?

Exercise 10.15 Assuming that x is positive, what is the maximum value of $(1-x)(1+2x)$? What value of x produces the maximum? *Hint:* It's fine to temporarily introduce an overall multiplicative factor.

Exercise 10.16 Assuming that x is positive, what is the minimum value of $x^2 + 1/x^3$? What value of x produces the minimum?

10.3 Arithmetic sequences

A *sequence* (or *progression*) is an ordered list of quantities that fit a particular pattern or rule. For example, in the following list,

$$7, \quad 11, \quad 15, \quad 19, \quad 23, \quad 27, \quad 31, \quad 35, \tag{10.22}$$

the rule is that each number is obtained by adding 4 to the preceding one. This type of sequence, where the difference between successive numbers is constant, is called an *arithmetic sequence*. More generally, if the first term in an arithmetic sequence is a, and if the constant difference is d, then the terms in the sequence take the form of

$$\boxed{a, \quad a+d, \quad a+2d, \quad a+3d, \quad a+4d, \ldots} \tag{10.23}$$

The sequence in Eq. (10.22) has $a = 7$ and $d = 4$, and we have halted the sequence after the 8th term (the 35). It's perfectly fine if the difference d is negative. The arithmetic sequence $8, 5, 2, -1, -4, \ldots,$ has $a = 8$ and $d = -3$. And a can be negative too, of course. A sequence can have any number of terms, even an infinite number.

If we label the nth term in a sequence as a_n, then the sequence in Eq. (10.23) can be written more succinctly as

$$\boxed{a_n = a + (n-1)d} \tag{10.24}$$

For example, if $n = 4$ then this equation says that $a_4 = a + (4 - 1)d = a + 3d$, which is correctly the 4th term in the sequence in Eq. (10.23). Note that Eq. (10.24) can equivalently be written as

$$a_n = a_1 + (n - 1)d, \tag{10.25}$$

where we have replaced a with a_1, since the first term a_1 equals a.

Instead of a_n, you can alternatively label the nth term as $a(n)$ or $f(n)$, which treats the terms as a function of n. That is, $f(n) = a + (n - 1)d$, where a and d are given quantities. Although it is perfectly reasonable to do this, the a_n notation is more compact, so that's generally what people use.

The coefficient in front of the d in Eq. (10.24) is in fact $n - 1$ and not n, because in Eq. (10.23), the 2nd term involves a single d, the 3rd term involves $2d$, and so on. So in general the nth term involves $(n - 1)d$. The pictorial representation of this fact is shown in Fig. 10.5. The 5th term, a_5, for example, is obtained by adding on only $4d$, instead of $5d$, to the first term, a. This is known as the "fencepost theorem":

FENCEPOST THEOREM: If you have n fenceposts, then there are only $n - 1$ horizontal boards between them. Equivalently, n numbers have only $n - 1$ spacings between them.

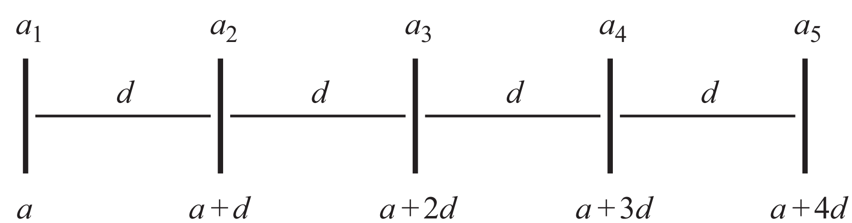

Figure 10.5: The "fencepost theorem" says that n fenceposts have only $n - 1$ horizontal boards between them. A consequence of this is that a_5 involves adding on only $4d$.

You need to remember this $n - 1$ fact if you want to determine the number of numbers in a list like this: 42, 43, 44, . . . , 75, 76, 77. You can't just subtract 42 from 77 to obtain 35, because 35 is the number of *spacings* (horizontal fence boards), and not the number of *numbers* (fenceposts). You need to remember that there is always one more number than spacings. So the number of numbers in the list is 36 instead of 35.

The number of numbers is certainly one more than the number of spacings if we look at a special case where there are only two numbers in the list, say, 42 and 43. And this statement is basically the same as the statement that a hamburger has two buns but only one hamburger patty between them (well, a normal hamburger, at least). And there's no shortage of other analogies along these lines. . .

> This $n - 1$ that we've seen
> Boils down to a fact that's routine:
> The theorem on posts
> Says the two US coasts
> Have just one single country between.

REMARKS:

1. As we mentioned at the beginning of this section, another name for "sequence" is "progression." Technically, when used by themselves, these two words have slightly different definitions. But that won't concern us here, because with the word "arithmetic" in front of them, they have the same meaning. So an arithmetic progression means exactly the same thing as an arithmetic sequence.

2. The three quantities in Eq. (10.5) (the two numbers a and b, along with the arithmetic mean μ between them) are a special case of the arithmetic sequence in Eqs. (10.23) and (10.24). The arithmetic sequence in Eq. (10.5) has only three terms, but that's fine.

3. We stated above that a sequence can have any number of terms, which technically means any number from 1 to infinity. However, a sequence with only one term isn't much of a sequence. Likewise for a sequence with two terms, because it's hard to tell what the pattern/rule is. For example, if a sequence consists of the two numbers 4 and 12, then the rule might be that 12 is obtained from 4 by adding 8. Or maybe 12 is obtained from 4 by multiplying by 3. Or perhaps it's obtained by squaring the 4 and then subtracting 4. These operations all yield 12, so we don't know which of them actually applies. But with three or more numbers, the rule is usually clear from the context (however, see Exercise 10.20). In any case, there's rarely much to be gained from thinking about one or two numbers as a sequence. ♣

Exercise 10.17 In an arithmetic sequence, show that every number is the arithmetic mean between its two neighbors.

Exercise 10.18 In a given arithmetic sequence, the first term is 5, the last term is 59, and there are seven terms in all. What are they?

Exercise 10.19

(a) Two arithmetic sequences start with the same number. The 10th term in the first sequence equals the 5th term in the second. What is the ratio of their spacings (the d's)?

(b) If the common first term is 2, and the d spacing in the first sequence is 4, write out the first ten and five terms of the two sequences.

Exercise 10.20 Consider the sequence: 4, 8, 12. This can be considered to be a (short) arithmetic sequence with first term $a = 4$ and difference $d = 4$. It can also be considered to be a sequence governed by the strange rule, $a_n = Aa_{n-1}^2 + Ba_{n-1}$, for certain values of A and B. That is, each term is obtained by multiplying the square of the previous number by A, and then adding on the previous number times B. Find the particular values of A and B that make this rule work.

10.4 Arithmetic series

Given an arithmetic sequence, is there a quick way to calculate the sum of all the terms? If the number of terms is reasonably small, as in Eq. (10.22), it's simple enough to just plug them into a calculator and add them up that way. Or perhaps you can do the sum in your head, if the numbers are nice. But what if the sequence has 100 terms? You're not going to want to plug 100 numbers into a calculator, and you're certainly not going to want to do the sum in your head. Furthermore, what about the sequence in Eq. (10.23), which involves letters instead of numbers? Even if you wanted to plug 100 things into a calculator, you can't plug in letters (at least not with a standard run-of-the-mill calculator). Is there a way to find a general formula for the sum?

Indeed there is. And the key is to use the technique from back in Exercise 4.18, where we found the sum of the first n positive integers. In the second solution to that exercise, we wrote out the list of integers, and then beneath them we wrote out the list again, in reverse order. We then noted that the sum of the two numbers in every column was always the same. Now, for the sequences above in Eqs. (10.22) and (10.23), the differences between successive numbers are, respectively, 4 and d. These aren't as simple as the difference of 1 in Exercise 4.18, but that doesn't matter. The strategy works exactly the same way here.

To be general, we'll apply the strategy to the sum of the first n terms in an arithmetic sequence whose first term is a and whose separation is d, that is, the sequence defined by Eq. (10.23) or Eq. (10.24). Let's label the sum as S_n, where the subscript n indicates how many of the terms in the sequence we're adding up. So S_n equals

$$\begin{aligned} S_n &= a_1 + a_2 + a_3 + \cdots + a_n \\ &= a + (a+d) + (a+2d) + \cdots + \big(a + (n-1)d\big). \end{aligned} \tag{10.26}$$

Remember that a_n has an $n-1$ in it, as opposed to an n. Listing the n terms in the forward and backward orders yields the two rows in Table 10.1.

a	$a+d$	$a+2d$	$\cdots$	$a+(n-2)d$	$a+(n-1)d$
$a+(n-1)d$	$a+(n-2)d$	$a+(n-3)d$	$\cdots$	$a+d$	a

Table 10.1: Writing an arithmetic sequence forward, and also backward, stacked on top of each other.

In this pair of lists, the sum of every pair of numbers stacked on top of each other is $2a + (n-1)d$; every column of two terms has this same sum. Therefore, since there are n columns, the sum of all the terms in the table is $n\big(2a + (n-1)d\big)$. However, we have counted every number twice. So to obtain the sum of all the numbers counted just once (that is, the desired sum S_n of the terms in the sequence), we need to divide by 2. Therefore,

$$\boxed{S_n = \frac{n\big(2a + (n-1)d\big)}{2}} \tag{10.27}$$

This can be written in the more compact form,

$$S_n = \frac{n\big[a + \big(a + (n-1)d\big)\big]}{2} = \boxed{n\left(\frac{a_1 + a_n}{2}\right)} \tag{10.28}$$

where we have used the expression for a_n in Eq. (10.24), along with the fact that $a_1 = a$.

In retrospect, this $n(a_1 + a_n)/2$ result makes sense, because $(a_1 + a_n)/2$ is the average of all n numbers, and the sum S_n of all n numbers is n times this average. The average is indeed $(a_1 + a_n)/2$, because the average of the two numbers in every column in Table 10.1 takes on this value. The average can also be written as $(a_2 + a_{n-1})/2$, or $(a_3 + a_{n-2})/2$, etc. So the sum S_n can be written as n times any of these other expressions. But the $n(a_1 + a_n)/2$ formula for S_n is easier to remember and apply.

The sum of the terms in an arithmetic *sequence* (or equivalently, an arithmetic *progression*) is called an arithmetic *series*.

A *series* is the sum of a *sequence*.

So a sequence or a progression is just a list of terms, whereas a series is a sum. Said in another way, in a sequence (or progression) you put commas between the terms, whereas in a series you put plus signs between them. For example:

$$\begin{aligned} \text{Arithmetic sequence:}&\quad 7,\ \ 11,\ \ 15,\ \ 19,\ \ 23,\ \ 27,\ \ 31,\ \ 35 \\ \text{Arithmetic series:}&\quad 7 + 11 + 15 + 19 + 23 + 27 + 31 + 35 = 168 \end{aligned} \tag{10.29}$$

Though both have a similar spread,
There's a difference in how they are read.
A sequence, we've seen,
Has some commas between,
While a series has plusses instead.

There is a slight ambiguity in the usage of the word "series." Looking at the arithmetic series in the second line of Eq. (10.29), some people might say, "The series equals 168," while others might say, "The sum of the series equals 168." In the first of these usages, the word "series" means the actual *value* of the sum. In the second, the "series" means the *act* of adding up the numbers (that is, a string of numbers with plus signs between them), but not the actual value of the sum itself.

Which is the correct usage? Both of them are acceptable. The usage is generally clear from the context, and there is little chance of confusion. In this chapter, we'll usually go with "The sum of the series equals 168."

Grammar note: The plural of the word "series" is again "series."

Example 10.4 If we split up the fraction in Eq. (10.27), we can rewrite it as

$$S_n = na + \frac{n(n-1)}{2} \cdot d. \tag{10.30}$$

Explain why the two terms here make sense.

Solution Each of the n terms in the series in Eq. (10.26) contains an a, so this yields a contribution of na to the sum S_n. This is the first term in Eq. (10.30). The d parts of the terms in the series yield a sum of $d\big(1 + 2 + \cdots + (n-1)\big)$. From Eq. (4.55) we know that the sum of the first $n-1$ integers is $(n-1)n/2$. (This is the righthand side of Eq. (4.55) with n replaced by $n-1$.) So the sum of the d parts of the terms is $d \cdot (n-1)n/2$, in agreement with the second term in Eq. (10.30).

Exercise 10.21 In the special case of $a = 1$ and $d = 1$, verify that Eq. (10.27) correctly yields the sum of the first n integers given in Eq. (4.55). Likewise for Eq. (10.28).

Exercise 10.22 With a calculator, verify that the sum of the terms in the sequence in Eq. (10.22) equals 168, as Eq. (10.29) states. Then calculate the sum by using Eq. (10.27), and then again by using Eq. (10.28).

Exercise 10.23 The first term in an arithmetic sequence is 5, and the 9th term is 29. If the sequence has 23 terms, what is the sum?

The $\sum$ summation notation

It gets to be a pain to write out long sums like the ones in Eqs. (10.26) and (10.29), so fortunately there exists a more compact notation. The Greek letter $\sum$ (sigma) is used to denote a sum, as the following four sums illustrate. You should stare at these for a minute to figure out what the notation means, before looking at the commentary that follows.

$$\begin{aligned}\sum_{k=3}^{8} k &= 3+4+5+6+7+8 = 33, \\ \sum_{k=2}^{5} k^3 &= 2^3+3^3+4^3+5^3 = 224, \\ \sum_{-2}^{3}(2k+1)^2 &= (-3)^2+(-1)^2+1^2+3^2+5^2+7^2 = 94, \\ \sum_{1}^{n} k(k+1) &= 1\cdot 2+2\cdot 3+3\cdot 4+\cdots+n(n+1).\end{aligned} \tag{10.31}$$

Here is what the various parts of the notation mean:

1. The expression to the right of the $\sum$ (the k, k^3, etc.) gives the general form of the terms in the sum. Common letters that are used for this purpose are i, j, k, m, and n, but other letters work just as well. The letter is called the *index*.

2. The numbers below and above the $\sum$ give the minimum and maximum values of k (often referred to as the lower and upper limits). So in the first sum above, k runs from 3 to 8. It is understood that k is an integer and that it increases in steps of 1 from the min to the max. You then just need to plug in all of the relevant k values into the expression to the right of the $\sum$, and then add up the results.

3. In the first two sums above, we wrote "$k = 3$" and "$k = 2$" below the $\sum$. However, the "$k =$" part of this isn't really necessary, because it's clear that we're talking about the letter k. So people often drop the "$k =$" and just write the minimum value of k, as we did in the third and fourth sums above.

4. In the fourth sum, the upper limit (the max value of k) is the letter n. This is perfectly fine; the limits don't need to be actual numbers. Of course, if the limits involve letters, then there's no way you can produce a numerical value for the sum until someone gives you the value of n. But you can (usually) find a formula for the sum in terms of n (as in Eq. (4.55), for example). And then once you're given a numerical value of n, you can just plug that into the formula.

Constant factors can be taken outside (that is, factored out of) a sum. For example,

$$\sum_{1}^{4} 3k^2 = 3\sum_{1}^{4} k^2. \tag{10.32}$$

This relation is true because it means the same thing as

$$3\cdot 1^2 + 3\cdot 2^2 + 3\cdot 3^2 + 3\cdot 4^2 = 3\big(1^2 + 2^2 + 3^2 + 4^2\big), \tag{10.33}$$

which is indeed true, by the distributive law. The two sides here are both equal to 90.

Sums can be broken up into other sums. For example, the third sum in Eq. (10.31) can be written as

$$\begin{aligned}\sum_{-2}^{3}(2k+1)^2 = \sum_{-2}^{3}(4k^2+4k+1) &= \sum_{-2}^{3}4k^2 + \sum_{-2}^{3}4k + \sum_{-2}^{3}1\\ &= 4\sum_{-2}^{3}k^2 + 4\sum_{-2}^{3}k + \sum_{-2}^{3}1.\end{aligned} \tag{10.34}$$

Note that in the last sum here, the "expression" to the right of the $\sum$ is simply a 1. That's fine; it doesn't actually need to involve k. This $\sum 1$ sum is shorthand for $1+1+1+1+1+1$, which equals 6. The first 1 here is the $k=-2$ term, the second is the $k=-1$ term, and so on, up to the last 1, which is the $k=3$ term. We do in fact need to write down a 1 for every value of k, and then add them up, even though the 1 itself doesn't involve k.

If you have any doubts about the steps in Eq. (10.34), you should explicitly write out all of the sums in terms of the specific numbers (that is, plug in all the k values), and then verify that the equalities do indeed hold. The task of Exercise 10.24 is to show that the last line of Eq. (10.34) reproduces the 94 that we found in Eq. (10.31).

In any situation involving the $\sum$ notation, if you're ever unsure what the sum means, you should simply start plugging in the k values. That should make things clear.

The lefthand sides of the equations in Eq. (10.31) admittedly look a bit scary at first sight, with the large and serious-looking $\sum$'s staring at you. But with a little practice, you'll soon get used to the fact that the $\sum$ notation is just a concise way of writing down the explicit sums that appear on the righthand sides of the equations in Eq. (10.31). If you have a sum with 100 terms, I think you'll quickly embrace the $\sum$ sigma notation, because it certainly beats writing out 100 terms!

> This fancy new sigma notation
> Is a helpful compact innovation.
> Though at first it seems frightening,
> It's really enlightening –
> It shortens a lengthy summation!

Exercise 10.24 Show (by explicitly plugging in the k values and adding up the numbers) that the last line of Eq. (10.34) equals 94.

Exercise 10.25 Write the sum in Eq. (10.29) in terms of the $\sum$ notation.

Exercise 10.26 Write the arithmetic series in Eq. (10.26) in terms of the $\sum$ notation, and then show how your answer leads to Eq. (10.30).

Exercise 10.27 There are two plausible ways to interpret the sum below. What are the two ways, and what results do they give?

$$\sum_{1}^{4} k + 1$$

10.5 Geometric sequences

We saw in Section 10.3 that in an arithmetic sequence, each term is obtained by *adding* a fixed number d to the preceding one. Another type of sequence is one where each term is obtained by *multiplying* the preceding one by a fixed number r. That is, the ratio of any two successive terms always has the same value r. (We've chosen the letter r for ratio, instead of d for difference.) This kind of sequence is called a *geometric sequence* (or *geometric progression*). An example is:

$$4, \quad 12, \quad 36, \quad 108, \quad 324, \quad 972. \tag{10.35}$$

The first term in this geometric sequence is 4, and each successive term is obtained by multiplying the previous one by $r = 3$.

More generally, if the first term in a geometric sequence is a, and if the ratio is r, then the terms in the sequence take the form of

$$\boxed{a, \quad ar, \quad ar^2, \quad ar^3, \quad ar^4, \ldots} \tag{10.36}$$

As with an arithmetic sequence, a geometric sequence can have any number of terms, even an infinite number. The ratio r can have any value (larger than 1, equal to 1, less than 1, 0, negative; anything). However, when $r = 0$ the sequence takes the unexciting form of a, 0, 0, 0, etc. And when $r = 1$ the sequence is also unexciting: a, a, a, a, etc.

If we label the nth term in a geometric sequence as a_n, then the sequence in Eq. (10.36) can be written concisely as

$$\boxed{a_n = ar^{n-1}} \qquad \text{or equivalently} \quad a_n = a_1 r^{n-1}, \tag{10.37}$$

since the first term a_1 equals a.

As with Eq. (10.24), the a_n in Eq. (10.37) involves an $n - 1$ on the righthand side, and not an n. This is consistent with the fact that the 2nd term in Eq. (10.36) involves a single r, the 3rd term involves r^2, and so on. So in general the nth term involves r^{n-1}.

As in the arithmetic case, a *geometric progression* means the same thing as a *geometric sequence*. Also similar to the arithmetic case, the three quantities in Eq. (10.11) (the two numbers a and b, along with the geometric mean g between them) are a special case of the geometric sequence in Eqs. (10.36) and (10.37). The geometric sequence in Eq. (10.11) has only three terms, but that's fine.

The name "geometric sequence" presumably comes from the fact that this type of sequence appears in geometric pictures like Figs. 10.6 and 10.8 below, as you'll demonstrate.

An important difference between arithmetic and geometric sequences is that geometric ones eventually grow much faster (if $r > 1$). Consider, for example, the following arithmetic and geometric sequences, both of which have 1 and 2 as their first two terms:

$$\begin{aligned} \text{Arithmetic:}\quad & 1,\ 2,\ 3,\ 4,\ \ 5,\ \ \ 6,\ \ \ 7,\ \ \ \ 8, \ldots \\ \text{Geometric:}\quad & 1,\ 2,\ 4,\ 8,\ 16,\ 32,\ 64,\ 128, \ldots \end{aligned} \tag{10.38}$$

The geometric sequence clearly grows much faster than the arithmetic one.

Remark: We already encountered this difference in growth rates back in Tables 9.2 and 9.3, and Fig. 9.16. When interest is compounded (which leads to multiplicative, or equivalently geometric, behavior), your money grows faster than when the interest isn't compounded (which leads to additive, or equivalently arithmetic, behavior).

Fig. 9.16 shows that the difference between the geometric and arithmetic behaviors is more pronounced as time goes by. This means that if you start investing late, the difference doesn't have enough time to become large. For short times, the geometric behavior isn't much better than the arithmetic behavior, as Table 9.2 and the left part of Fig. 9.16 show. So start early if you want to reap the benefits of geometric growth! ♣

It's not just a trite hypothetical
Or a pointless abstract theoretical.
Start young and you'll see
That the growth will then be
Geometric, and not arithmetical!

Example 10.5 Fig. 10.6 shows a chain of growing circles inside a given angle. Let the radii of three successive circles be a, b, and c, and consider the dark-shaded right triangle shown. The hypotenuse is the sum $c+b$ of the two radii, and the short leg is the difference $c - b$, as shown. By considering the analogous (similar) light-shaded right triangle shown for the circles with radii a and b, show that b is the geometric mean of a and c, that is, $b = \sqrt{ac}$. (Equivalently, a, b, and c form a geometric sequence; see Exercise 10.28.) You will need to use a ratio result for similar triangles from Section A.3 in Appendix A.

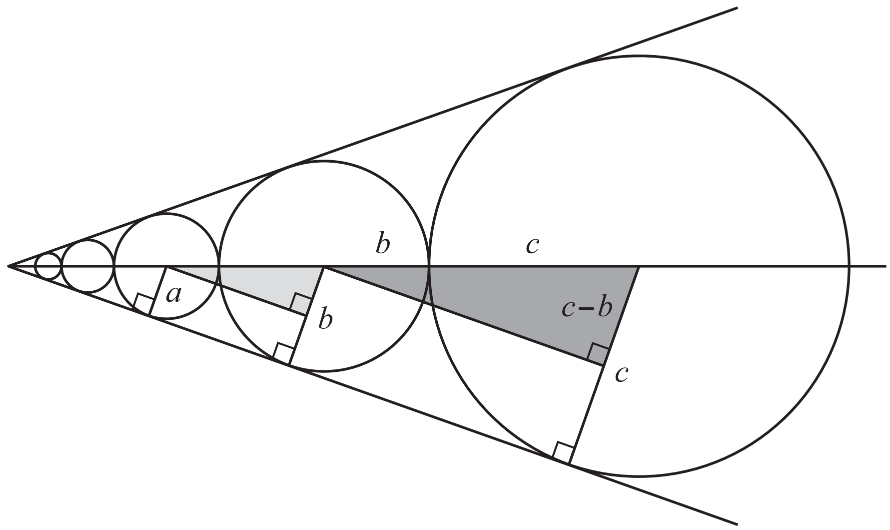

Figure 10.6: In this chain of growing circles, $b = \sqrt{ac}$. More generally, each radius is the geometric mean of the two adjacent radii. Equivalently, the radii form a geometric sequence.

Solution The two shaded right triangles are indeed similar, because the two tilted b and c radii are parallel (because they both make a right angle with the lower arm of the given angle), which means that they make the same angle with the horizontal line. So the larger angles in the two triangles are equal, which means that the smaller angles are also equal (since the other angle is 90°). The two triangles are therefore similar, since they have the same three angles.

As with the dark-shaded right triangle, the hypotenuse and short leg of the light-shaded right triangle associated with the circles with radii a and b are likewise the sum $b + a$ and difference $b - a$. From Section A.3 we know that the ratios of corresponding sides of two similar triangles are equal. So the ratio of the short leg to the hypotenuse is the same in both triangles. This gives

$$\begin{aligned}
\frac{\text{short leg}}{\text{hypotenuse}}: \quad \frac{b-a}{b+a} = \frac{c-b}{c+b} \quad &\Longrightarrow \quad (b-a)(c+b) = (b+a)(c-b) \\
&\Longrightarrow \quad \cancel{bc} + b^2 - ac - \cancel{ab} = \cancel{bc} - b^2 + ac - \cancel{ab} \\
&\Longrightarrow \quad 2b^2 = 2ac \\
&\Longrightarrow \quad b = \sqrt{ac}. \qquad (10.39)
\end{aligned}$$

So b is the geometric mean of a and c, as desired. Equivalently, the $b^2 = ac$ relation in the third line can be rewritten as $b/a = c/b$. This says that the ratio of the radii of the adjacent b and a circles equals the ratio of the radii of the adjacent c and b circles. Since this is the defining characteristic of a geometric sequence (equal ratios of successive terms), the three radii therefore from a geometric sequence.

If we repeated the above process with the b, c, and (the next, not shown) d radii, we would similarly find that $c/b = d/c$. (There was nothing special about the three a, b, c circles we chose to work with.) Therefore, $b/a = c/b = d/c$. And so on. We can also work our way leftward from the a circle. We conclude that the ratio of the radii of any two adjacent circles always takes on the same value. In other words, the radii all form a geometric sequence. If we let the common ratio be r, and if we arbitrarily give one of the circles a radius of 1, then the radii of all the circles take the form of: $\ldots 1/r^3, 1/r^2, 1/r, 1, r, r^2, r^3, \ldots$

Remark: There is a nice geometric way of seeing why the radii in Fig. 10.6 form a geometric sequence, without going through the algebra in Eq. (10.39) (even though the algebra worked out very cleanly!). Consider two adjacent circles, such as B and C in the top picture in Fig. 10.7. Let the ratio of the radii of C and B be r (which happens to be $r = 1.5$ in Fig. 10.7, although the specific value isn't important). Assume that we don't know anything about any of the other ratios of adjacent radii.

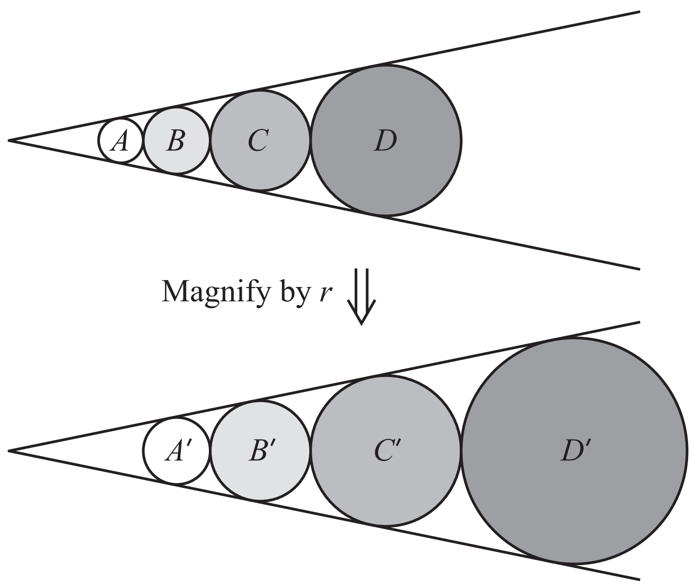

Figure 10.7: A geometric way of showing that the radii form a geometric sequence.

Imagine magnifying the whole picture by a factor of r, to obtain the bottom picture in Fig. 10.7. The new B circle (call it B') is now as large as the C circle was; we have aligned B' and C in the figure. Since B' and C have the same size, it follows that C' and D do also, because there is only one way to add a circle on the right side of B' and C while having it be tangent to the two lines that form the angle. So C' has the same size as D. But C' is r times as large as C, since the bottom picture in Fig. 10.7 is obtained by magnifying the top one by r. Hence D is also r times as large as C. The D-to-C ratio is therefore the same as the C-to-B ratio. They both equal r.

In the same manner, we can work our way rightward or leftward to show that the ratio of the radii of any two adjacent circles equals r. The radii therefore form a geometric sequence with ratio r, as we wanted to show. ♣

Exercise 10.28 In a geometric sequence, show that every number is the geometric mean between its two neighbors.

Exercise 10.29 In a given geometric sequence, the first term is 3, the last term is 24, and there are seven terms in all. What are they?

Exercise 10.30 Fig. 10.8 shows a chain of growing squares inside a given angle. Let the sides of three successive squares be a, b, and c, and consider the dark-shaded right triangle shown. The horizontal leg is b, and the vertical leg is $(c-b)/2$, because twice this distance plus b must equal c. By considering the analogous (similar) light-shaded right triangle shown for the squares with sides a and b, show that b is the geometric mean of a and c, that is, $b = \sqrt{ac}$. (Equivalently, a, b, and c form a geometric sequence.)

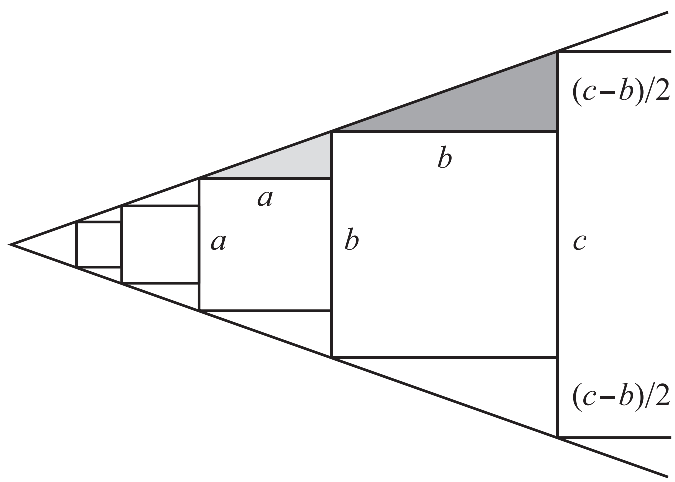

Figure 10.8: In this chain of growing squares, $b = \sqrt{ac}$. More generally, the side of each square is the geometric mean of the sides of the two adjacent squares. Equivalently, the sides form a geometric sequence.

Exercise 10.31 The task of this exercise is to derive the general formula for the nth term of the *Fibonacci sequence*, which is the sequence whose first two terms are both 1, and which has the property that every term is the sum of the two preceding terms. That is, $a_n = a_{n-1} + a_{n-2}$. You can verify that the first few numbers in the sequence are:

$$1, \quad 1, \quad 2, \quad 3, \quad 5, \quad 8, \quad 13, \quad 21, \quad 34, \quad 55, \quad 89, \ldots \tag{10.40}$$

The following derivation is a wonderful application of many topics we've covered in this book, including geometric sequences, quadratic equations, and systems of equations.

(a) If a geometric sequence has the property that each term is the sum of the two preceding terms (that is, $a_n = a_{n-1} + a_{n-2}$), what is the ratio r?

(b) You should have obtained two possible answers in part (a). Call them r_+ and r_- (they come from the quadratic formula). Consider a sequence whose nth term is given by $a_n = Ar_+^{n-1} + Br_-^{n-1}$, where A and B are arbitrary numbers of your choosing. (Having the exponents be n would work just as well, but $n - 1$ makes the math cleaner.) Show that this sequence has the property that $a_n = a_{n-1} + a_{n-2}$, for any A and B.

(c) Find the special values of A and B that make the first two terms in the sequence both be equal to 1. These values yield the desired formula for the nth term of the Fibonacci sequence. Check that the formula works for $n = 1, 2, 3$.

(The solution to this exercise contains a long list of remarks about the Fibonacci numbers, if you want to study them further.)

Exercise 10.32 Consider a sequence governed by the rule $a_n = 2a_{n-1} + 3a_{n-2}$, with the first two terms both equalling 1. So it starts off as 1, 1, 5, 13, 41, 121, etc. Follow the procedure in Exercise 10.31 to find the general formula for a_n.

10.6 Geometric series

10.6.1 Infinite geometric series

In Eq. (10.27) we found a formula for the sum of the terms in an arithmetic sequence (with n terms, first term a, and difference d). That is, we found a formula for the sum of an arithmetic series. We'll now derive a formula for the sum of the terms in a geometric sequence. That is, we'll derive a formula for the sum of a geometric series.

We'll start off here with *infinite* series (ones with an infinite number of terms), and then in Section 10.6.2 we'll look at finite series. Although the reasoning below works for any first term a, we'll pick the first term to be 1, just to make the algebra a little cleaner. Our goal then is to find the sum of the following series, where the terms go on forever:

$$S = 1 + r + r^2 + r^3 + r^4 + \cdots . \tag{10.41}$$

This geometric *series* is the geometric *sequence* in Eq. (10.36), but with $a = 1$ and with plus signs instead of commas between the terms.

Depending on the value of r, the sum in Eq. (10.41) might *diverge* (get larger and larger and head off to infinity), or it might *converge* (approach a definite finite number). If we want the sum to converge, then we need r to be less than 1. (More precisely, we need $|r| < 1$, but let's just look at positive r for now.) This is true because if r is larger than 1, the powers of r in the sequence get larger and larger. So if we include an infinite number of terms, the sum will be infinitely large. Even if r is just equal to 1, all of the powers of r are then equal to 1, so the series in Eq. (10.41) is $1 + 1 + 1 + \cdots$, which equals infinity if we include an infinite number of terms.

Therefore, $r < 1$ is our only hope for having the series converge to a finite number. Of course, for all we know at the moment, the sum might *still* end up being infinite even if $r < 1$. For example, if $r = 0.9$ the series is

$$1 + 0.9 + (0.9)^2 + (0.9)^3 + (0.9)^4 + \cdots . \tag{10.42}$$

It's not obvious whether this sum converges to a finite number or heads off to infinity. On one hand, the powers of 0.9 get smaller and smaller, which might make you think that the sum is finite. But on the other hand, there is an infinite number of terms, so even though they get smaller, you might think that the sum is infinite. It's not obvious which effect (the smallness of the terms, or the largeness of the number of terms) wins.

It turns out that the numbers get small quickly enough, so the sum is *finite*. We'll demonstrate this by explicitly calculating the sum and showing that it is finite. However, see Exercise 10.35 for another kind of series where the numbers get small, but not quickly enough, so the sum is *infinite*.

As a first attempt at finding a formula for the geometric series in Eq. (10.41), we can try writing the terms in the reverse order, as we did in Table 10.1 for an arithmetic series. However, this doesn't help because (a) the series is infinite, which means that it has no last term (which would be the first term in the reverse order), and (b) even if the series had a finite number of terms, the sums of the two terms in the various columns are *not* the same, as they were in Table 10.1. For example, if the series is $1 + r + r^2 + r^3$, then the reverse series is $r^3 + r^2 + r + 1$, so the column sums are $1 + r^3$ and $r + r^2$; each sum occurs twice. These aren't equal (except in the special case where $r = 1$).

We therefore need another strategy. The key is to note that if we multiply the series S in Eq. (10.41) by r, we end up with nearly the same series, with the only difference being that the first term is now r instead of 1. Both the S and rS series have an infinite number of terms, so there's no difference on the "far" end, because there is no end; the terms go on forever. The series differ only in where they start. (However, see the first remark below for a caveat to this. We're assuming $r < 1$ in the present discussion.)

Let's write the rS series below the S series, and align the terms like this:

$$\begin{aligned} S &= 1 + r + r^2 + r^3 + r^4 + \cdots \\ rS &= \phantom{1 + {}} r + r^2 + r^3 + r^4 + \cdots \end{aligned} \tag{10.43}$$

If we now *subtract* rS from S (instead of *adding* the two lines, as we did in Table 10.1), all of the terms (an infinite number of them!) cancel except for the 1, so we end up with

$$S - rS = 1 \quad \Longrightarrow \quad S(1-r) = 1 \quad \Longrightarrow \quad \boxed{S = \frac{1}{1-r}} \tag{10.44}$$

This formula is valid for $|r| < 1$; see the first remark below.

> For sums with an infinite run,
> There's an elegant way that they're done.
> It's quick and exact –
> Write down S and subtract
> r times S, and you're left with just 1!

An equivalent way of stating the above reasoning is to say that the rS series in Eq. (10.43) equals the S series, but with the first term (the 1) missing. That is, $rS = S - 1$. This can be rewritten as $1 = S(1 - r)$, which yields Eq. (10.44).

Having shown that the sum of the general geometric series in Eq. (10.41) equals $1/(1-r)$, we can now apply this result to the specific series in Eq. (10.42). With $r = 0.9$, Eq. (10.44) gives $S = 1/(1 - 0.9) = 1/(0.1) = 10$. This value is plausible, although if you start adding up the powers of 0.9 with your calculator, you'll need to add a rather large number of them in order to be convinced that the sum is in fact 10. The sum of the first 10 terms in the series is 6.51, the sum of the first 30 terms is 9.58, and the sum of the first 60 terms is 9.98. So the series converges rather slowly to 10.

In contrast, if we have the series $1 + 0.1 + (0.1)^2 + (0.1)^3 + \cdots$, with $r = 0.1$, then the sum of just the first four terms is 1.111. (No calculator needed! Each term in the series produces a 1 in a different position in the decimal.) And Eq. (10.44) gives the true sum as $S = 1/(1 - 0.1) = 1/(0.9) = 1.111\ldots$, where the 1's go on forever. (Again, we already knew that the sum consists of an infinite string of 1's, even without using the formula for S.) Since this result is only a tiny bit larger than 1.111, we see that the series converges very quickly. That is, not many terms are needed if you want to produce a reasonable estimate of the true sum. Putting everything together, we see that when r is close to 1, the series converges slowly; and when r is close to 0, the series converges quickly.

Remarks:

1. Eq. (10.44) is valid only if $r < 1$ (more precisely, $|r| < 1$). It certainly isn't valid for, say, $r = 2$ because it would give a sum of $1/(1-r) = 1/(1-2) = -1$. This is clearly incorrect, because it is negative, and also because it is finite. The actual sum $1 + 2 + 2^2 + 2^3 + \cdots$ is infinite, because there is an infinite number of terms, each of which is larger than the preceding one.

   Which step in the above derivation makes Eq. (10.44) invalid when $r \geq 1$? The flawed step is the $S - rS = 1$ one, or equivalently the statement that all of the terms in the two sets of "$\cdots$" dots in Eq. (10.43) cancel when we subtract the equations. The rS series is technically one power of r "ahead" of the S series. This means that if the S series were instead just a very long (rather than infinite) series that went up to, say, r^{1000}, then the rS series would go up to r^{1001}. This leftover term would then not cancel when we subtract the two equations.

   Fortunately this complication doesn't matter when $r < 1$, since the r^{1001} term is very small in this case. This is true because for any $r < 1$ (more precisely, $|r| < 1$), if we raise it to a large enough power (and we're actually raising it to an infinite power, since we have an infinite series), the result is essentially zero.

However, the leftover r^{1001} term *does* matter if $r > 1$ (and if $r = 1$ too), because the powers of r get larger and larger, which means the leftover term *cannot* be ignored. It is technically infinitely large if we have an infinite series. This makes the $S - rS = 1$ relation be (highly) incorrect. We'll see the precise consequence of the leftover term when we discuss *finite* geometric series below in Section 10.6.2.

2. Eq. (10.44) works perfectly fine for negative r too, as long as $|r| < 1$. That is, the full range of validity of Eq. (10.44) is $-1 < r < 1$. If $r = -1/2$, for example, then large powers of r are very small. They might be positive or negative, depending on whether the power is even or odd, but that doesn't matter. The numbers will be close to zero in any case. So from the reasoning in the preceding remark, we can ignore the leftover term, which means that the $S - rS = 1$ relation holds. The result in Eq. (10.44) is therefore valid, so the sum for $r = -1/2$ equals $1/\big(1 - (-1/2)\big) = 2/3$, which is plausible (if you add up the first few terms) but by no means obvious.

3. Another way of deriving Eq. (10.44) (which is really the same method in the end) is to note that multiplying S by $1 - r$ gives

$$\begin{aligned}(1-r)S &= (1-r)(1 + r + r^2 + r^3 + r^4 + \cdots) \\ &= (1 + r + r^2 + r^3 + r^4 + \cdots) - (r + r^2 + r^3 + r^4 + r^5 + \cdots) \\ &= 1.\end{aligned} \tag{10.45}$$

All of the terms except the 1 cancel. So $(1-r)S = 1$, which yields Eq. (10.44). Of course, the above multiplication by $1 - r$ is a little out of the blue if you haven't already seen the original derivation in Eqs. (10.43) and (10.44). ♣

In the more general case where the first term in a geometric series is an arbitrary a instead of 1, the series is $a + ar + ar^2 + ar^3 + \cdots$. This is simply a times the series in Eq. (10.41), so the sum is a times the sum in Eq. (10.44):

$$\boxed{a + ar + ar^2 + ar^3 + ar^4 + \cdots = \frac{a}{1-r} = \frac{\text{first term}}{1 - \text{ratio}}} \tag{10.46}$$

When working with infinite geometric series, you will often find yourself repeating the mantra for the sum: "First term over 1 minus the ratio." As with Eq. (10.44), this result is valid only for $|r| < 1$.

In terms of the $\sum$ summation notation from Eq. (10.31), the result in Eq. (10.46) takes the form of

$$\boxed{\sum_0^\infty ar^n = \frac{a}{1-r}} \tag{10.47}$$

Alternatively, you can have the sum go from 1 to ∞, with an exponent of $n - 1$.

Example 10.6 You go for a walk and start by walking one mile east. From this point on, the next stage of your walk is always halfway back to where you just were. So after walking one mile east, you walk $1/2$ mile west, and then $1/4$ mile east, and then $1/8$ mile west, and so on. What is the total distance you walk? What location do you converge on?

Solution You walk distances of 1, then $1/2$, then $1/4$, and so on. So from Eq. (10.44) with $r = 1/2$, the total distance you walk is

$$1 + \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \frac{1}{16} + \cdots = \frac{1}{1 - 1/2} = \frac{1}{1/2} = 2. \tag{10.48}$$

To find your final position, let's take eastward to be the positive x direction. Then the eastward stages increase your x coordinate, while the westward ones decrease it. So from Eq. (10.44) with $r = -1/2$, your final position is

$$1 - \frac{1}{2} + \frac{1}{4} - \frac{1}{8} + \frac{1}{16} - \cdots = \frac{1}{1 - (-1/2)} = \frac{1}{3/2} = \frac{2}{3}. \tag{10.49}$$

The total *distance* you walk is therefore 2 miles, whereas your final *position* is the $2/3$-mile mark. This is correctly less than the total distance of 2 miles, due to the partial cancellation that occurs because of the minus signs in Eq. (10.49). Every step you take always contributes a positive amount to the *distance* you walk; there is never any cancellation there. But it can (and will, if a step is leftward) contribute a negative amount to your *position*. If you walk a total distance of 2 miles, then no matter what path you take, you can never end up farther than 2 miles from where you started. And that occurs only if you walk in an exact straight line in the same direction for the entire time.

Example 10.7 Consider the following two series with ratios r and $-r$:

$$S_r = 1 + r + r^2 + r^3 + r^4 + \cdots, \tag{10.50}$$
$$S_{-r} = 1 + (-r) + (-r)^2 + (-r)^3 + (-r)^4 + \cdots = 1 - r + r^2 - r^3 + r^4 + \cdots.$$

Find the sum $S = S_r + S_{-r}$ of these two series in two different ways:

(a) Add all of their terms together, and then apply Eq. (10.46) to their sum.

(b) Apply Eq. (10.44) to each series separately, and then add the two results.

Solution

(a) If we add all the terms in the two series together, the odd powers of r cancel, so we obtain $2 + 2r^2 + 2r^4 + 2r^6 + \cdots$. The ratio in this series is r^2, so Eq. (10.46) (with $a = 2$, and with r replaced by r^2) yields $S = 2/(1 - r^2)$.

(b) Applying Eq. (10.44) to each sum separately gives

$$1 + r + r^2 + r^3 + r^4 + \cdots = \frac{1}{1-r},$$
$$1 - r + r^2 - r^3 + r^4 + \cdots = \frac{1}{1-(-r)} = \frac{1}{1+r}. \tag{10.51}$$

Adding these two results by getting a common denominator yields

$$\frac{1}{1-r} \cdot \frac{1+r}{1+r} + \frac{1}{1+r} \cdot \frac{1-r}{1-r} = \frac{2}{(1+r)(1-r)} = \frac{2}{1-r^2}, \tag{10.52}$$

in agreement with the result in part (a).

Exercise 10.33 When $r = 1/2$, Eq. (10.44) says that

$$1 + \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \cdots = \frac{1}{1-1/2} = 2. \tag{10.53}$$

Can you think of a way of visualizing this sum on a number line, which makes it clear why the sum equals 2?

Exercise 10.34 One infinite geometric series has twice the first term, and half the ratio, compared with a second infinite geometric series. If the two series have the same sum, what are the two ratios?

Exercise 10.35 In an infinite geometric series with $|r| < 1$, the terms decrease quickly enough, so that even though there is an infinite number of them, the sum is finite instead of infinite. However, not all decreasing series have this property. Sometimes the terms don't decrease quickly enough, and the sum ends up being infinite. Such is the case for the sum of the reciprocals of the positive integers:

$$1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \frac{1}{5} + \cdots. \tag{10.54}$$

This sum is called the *harmonic series*. Show that the sum is infinite. *Hint*: It can be grouped as

$$1 + \frac{1}{2} + \left(\frac{1}{3} + \frac{1}{4}\right) + \left(\frac{1}{5} + \frac{1}{6} + \frac{1}{7} + \frac{1}{8}\right) + \cdots. \tag{10.55}$$

Exercise 10.36 In contrast with the preceding exercise, the sum of the reciprocals of the *squares* of the positive integers is finite:

$$1 + \frac{1}{2^2} + \frac{1}{3^2} + \frac{1}{4^2} + \frac{1}{5^2} + \cdots. \tag{10.56}$$

These terms happen to decrease fast enough to make the sum be finite. Demonstrate this by grouping the terms in a manner that is (somewhat) similar to the grouping in the preceding exercise.

Exercise 10.37 Fig. 10.9 provides a geometric method for deriving the result in Eq. (10.44). The sides of the squares form a geometric sequence with ratio r. Make use of similar triangles (see Section A.3 in Appendix A) to calculate the length of CB, which equals the desired sum $S = 1 + r + r^2 + r^3 + \cdots$. To be rigorous, you should also explain why the line AB is in fact a straight line, as drawn.

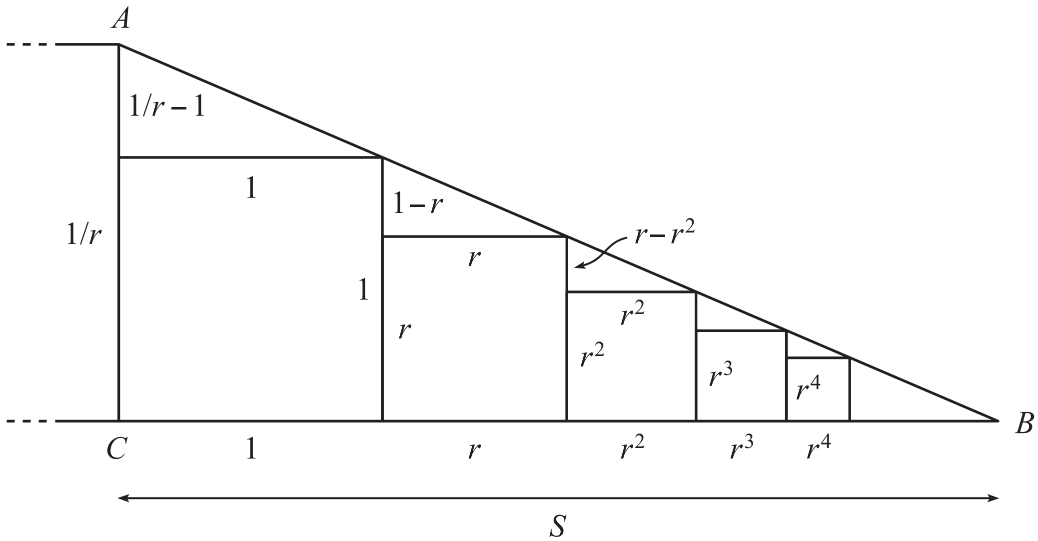

Figure 10.9: The setup for a similar-triangle proof of Eq. (10.44).

Exercise 10.38

(a) The side lengths of the squares shown in Fig. 10.10 are 1, 1/2, 1/4, 1/8, etc., and they go on forever. Find the total area of all the squares by summing the appropriate geometric series.

(b) Find the total area of the squares again, in a different way. Let A_{sq} be the desired total area of the squares, and let A_{tri} be the total area of the little triangles above the squares. Determine the question marks in the following two equations:

$$A_{\text{sq}} + A_{\text{tri}} = (?) \qquad \text{and} \qquad A_{\text{sq}} = (?) \cdot A_{\text{tri}}. \tag{10.57}$$

You now have two equations in the two unknowns, A_{sq} and A_{tri}. Solve for A_{sq}.

Exercise 10.39

(a) From Eq. (10.46), the sum of the geometric series with first term $a = 1/4$ and ratio $r = 1/4$ is

$$\frac{1}{4} + \frac{1}{4^2} + \frac{1}{4^3} + \frac{1}{4^4} + \cdots = \frac{1/4}{1 - 1/4} = \frac{1/4}{3/4} = \frac{1}{3}. \tag{10.58}$$

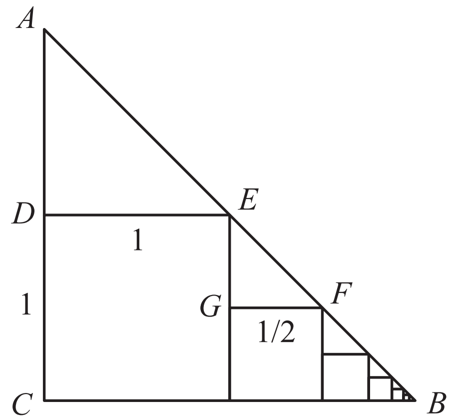

Figure 10.10: A geometric way of finding the total area of all the squares.

Explain how this $1/3$ result can be obtained from each of the two figures in Fig. 10.11(a).

(b) More generally, the sum of the geometric series with first term $a = 1/N$ (we'll assume N is an integer here) and ratio $r = 1/N$ is

$$\frac{1}{N} + \frac{1}{N^2} + \frac{1}{N^3} + \frac{1}{N^4} + \cdots = \frac{1/N}{1 - 1/N} = \frac{1/N}{(N-1)/N} = \frac{1}{N-1}\,. \tag{10.59}$$

Explain how this $1/(N-1)$ result can be obtained from the figure in Fig. 10.11(b), which is drawn for $N = 5$.

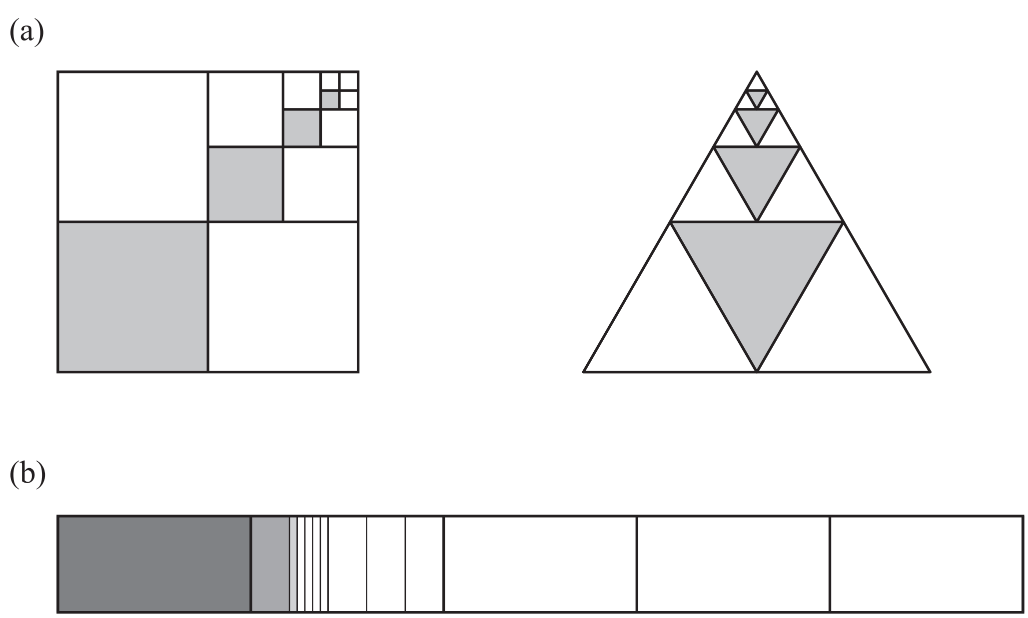

Figure 10.11: Geometric methods for finding sums of the form $1/N + 1/N^2 + 1/N^3 + \cdots$.

Exercise 10.40 Two bicyclists, A and B, start a distance d apart and travel toward each other, both at a speed of 10 mph. Right when they start, a bird departs from A and travels to B. Upon reaching B, it immediately turns around and heads back to A. And then upon reaching A, it turns around and heads back to B, and so on. If the bird flies at 30 mph, what is the total distance it travels by the time the bicyclists meet? Solve this in two different ways:

(a) The long way: Determine how far the bird travels during each stage of the flight (find the first few distances, and you'll see a pattern). Then sum the appropriate series.

(b) The quick way: You know how far each bicyclist travels by the time they meet. Think about how you can get the bird's distance from that.

Exercise 10.41 The task of this exercise is to prove an interesting theorem about parabolas, first demonstrated by Archimedes. The method of proof is a wonderful application of geometric series. Fig. 10.12 shows a $y = x^2$ parabola. Segment AB is drawn between two arbitrary points on the parabola, and then triangle ABC is formed, with point C lying on the parabola with its x coordinate being halfway between the x coordinates of A and B. The three x values of A, C, and B can then be written as $x - a$, x, and $x + a$ for some values of x and a.

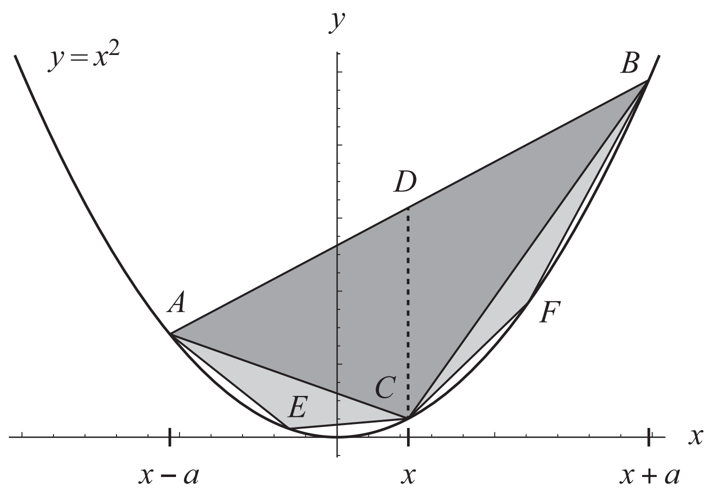

Figure 10.12: The setup for proving Archimedes' theorem, which says that the area of the parabolic section associated with AB is $4/3$ times the area of triangle ABC.

The *parabolic section* associated with segment AB is defined as the "half moon" region lying below AB, bounded by AB and the curved parabola. So the parabolic section includes all of the shaded regions in Fig. 10.12, plus the four thin unshaded slivers bordering the parabola. Archimedes' parabola theorem says:

> ARCHIMEDES' THEOREM: The area of the parabolic section associated with AB is $4/3$ times the area of triangle ABC (the dark-shaded area).

This theorem holds for arbitrary points A and B on the parabola. But the point C must be halfway between them (with regard to the x coordinates). The theorem can be proved as follows:

(a) Find the length of the vertical dashed line CD Fig. 10.12, by using the fact that the height of D is the average of the heights of A and B. Then find the areas of triangles CDA and CDB by tilting your head sideways and considering CD to be their base. You should find that the area of the dark-shaded triangle ABC is a^3, which interestingly is independent of x.

(b) Now find the area of the parabolic section associated with AB. You can do this by successively filling in (partially at each stage) the missing curved pieces of the section. The light-shaded triangles in Fig. 10.12 show the next stage, where E is halfway between A and C (with regard to their x coordinates), and F is halfway between C and B. Use the result from part (a) to find the sum of the areas of the two light-shaded triangles.

In the same manner, imagine partially filling in the remaining four little unshaded slivers of the section with triangles, and find their area. And then partially fill in the remaining eight little slivers, and so on. The entire section will eventually be filled in. You should find that the sum of the areas of all the triangles (that is, the total area of the section) is a geometric series that sums to $4a^3/3$, which is $4/3$ times the area of triangle ABC, as desired.

> If you start with the shape ABC
> And then fill in the slivers, you'll see
> The triangles' sum
> Will converge and become
> Simply a cubed times 4 over 3!

Arithmetic-geometric series

Here's a fun extension of the geometric-series result in Eq. (10.46). What is this sum?

$$1 + 2r + 3r^2 + 4r^3 + 5r^4 + \cdots . \tag{10.60}$$

This series would be a nice arithmetic one if it didn't have the powers of r. And it would be a nice geometric series if it didn't have the coefficients of 2, 3, 4, etc. The series is a combination of an arithmetic one and a geometric one, in that the nth term is the product of the nth term in an arithmetic series (namely $1 + 2 + 3 + \cdots$) and the nth term in a geometric series (namely $1 + r + r^2 + \cdots$). A series with this property is called (quite reasonably) an *arithmetic-geometric series*.

How do we calculate the sum in Eq. (10.60)? It turns out that a sneaky use of the geometric-series formula in Eq. (10.46) will get the job done. The key is to write the series in Eq. (10.60) in the following form:

$$\begin{aligned} 1+2r+3r^2+4r^3+\cdots = 1+r+r^2+r^3+r^4+\cdots \\ +r+r^2+r^3+r^4+\cdots \\ +r^2+r^3+r^4+\cdots \\ +r^3+r^4+\cdots \\ +r^4+\cdots \\ \vdots\quad \end{aligned} \tag{10.61}$$

This has the correct number of terms of each power of r (two r's, three r^2's, etc.). The above form is helpful because each line is a geometric series, with different starting terms (1, then r, then r^2, etc.), but all with the same ratio r. Using Eq. (10.46) for each line tells us that the sum in Eq. (10.61) equals

$$\frac{1}{1-r}+\frac{r}{1-r}+\frac{r^2}{1-r}+\frac{r^3}{1-r}+\cdots . \tag{10.62}$$

But this itself is a geometric series, with first term $r/(1-r)$ and ratio r. Equivalently, we can just factor out the $1/(1-r)$. So the sum in Eq. (10.60) equals

$$\frac{1}{1-r}\left(1+r+r^2+r^3+\cdots\right)=\frac{1}{1-r}\left(\frac{1}{1-r}\right)=\boxed{\frac{1}{(1-r)^2}} \tag{10.63}$$

In short, the sums of the geometric series in the various lines in Eq. (10.61) themselves form a geometric series. The sum in Eq. (10.60) comes up often in the subject of probability, in particular when calculating "expected values" (the average value you expect to obtain in a probabilistic process).

Exercise 10.42 In Eq. (10.45) we derived Eq. (10.44) by multiplying the series by $1-r$. Use the same technique to give another derivation of Eq. (10.63).

Exercise 10.43 In the spirit of the preceding exercise, find the sum of $1+3r+6r^2+10r^3+15r^4+\cdots$, where the coefficient of the nth term is n more than the previous coefficient.

10.6.2 Finite geometric series

Let's now find the sum of a *finite* geometric series (one with a finite number of terms). We'll use exactly the same strategy that we used above for an infinite series, where we multiplied the series by r and then subtracted that from the original series; see Eqs. (10.43) and (10.44)

For concreteness, let's look at the series $1+r+r^2+r^3+r^4+r^5$. Then instead of Eq. (10.43), we now have

$$\begin{aligned} S &= 1 + r + r^2 + r^3 + r^4 + r^5 \\ rS &= \quad\;\; r + r^2 + r^3 + r^4 + r^5 + r^6 \end{aligned} \tag{10.64}$$

There are only six terms in each series here, instead of an infinite number. So we don't have the "$\cdots$" dots like we had in Eq. (10.43). If we subtract rS from S, what we had in Eq. (10.44) now becomes

$$S - rS = 1 - r^6 \quad\Longrightarrow\quad S(1-r) = 1 - r^6 \quad\Longrightarrow\quad S = \frac{1-r^6}{1-r}. \tag{10.65}$$

When taking the difference, there are *two* uncanceled terms: the 1 (as in the infinite case) and now also the r^6. This is the "leftover" term discussed in the first remark on page 612. We can't ignore it now, as we did in the infinite case (when $|r| < 1$), because now the exponent is only 6, which means that r^6 isn't infinitely small.

In general, the above procedure tells us that if the geometric series goes from 1 up to r^n (instead of r^5 in Eq. (10.64)), then the exponent 6 in the result in Eq. (10.65) becomes an $n+1$. The exponent in the result is 1 more than the highest exponent in the series. So the sum is

$$\boxed{1 + r + r^2 + \cdots + r^{n-1} + r^n = \frac{1-r^{n+1}}{1-r}} \tag{10.66}$$

In the more general case where the first term is a instead of 1, every term is multiplied by a, which means the sum is also. The general formula for the sum of a *finite* geometric series is therefore

$$\boxed{a + ar + ar^2 + \cdots + ar^{n-1} + ar^n = \frac{a - ar^{n+1}}{1-r}} \tag{10.67}$$

Note that since the powers of r run from r^0 up to r^n, there are $n+1$ terms in the series. So the $n+1$ exponent in the result is the number of *terms*, and *not* the highest exponent of r that appears in the series (which is n). It is 1 more than that.

Finite geometric series and finite arithmetic series both have finite sums. But if we want an *infinite* series to have a finite sum, then it has to be a geometric series (and furthermore we need $|r| < 1$). An infinite arithmetic series has no hope of having a finite sum, because the terms far out in the series don't go to zero. Instead, they march off to infinity (positive or negative, depending on the sign of the difference d). So there is no way the sum can be finite (although technically one exception is $0 + 0 + 0 + \cdots = 0$).

REMARKS:

1. Whereas the result in Eq. (10.46) is valid only for $|r| < 1$, the result in Eq. (10.67) is valid for all values of r (greater than 1, less than 1, positive, negative), with one exception, namely $r = 1$, because that makes the $1 - r$ denominator be zero. But the $r = 1$ case is simple enough anyway. All of the terms are equal to a, and since there are $n+1$ of them, the sum is simply $(n+1)a$.

2. In the case where $|r| < 1$, the finite-series result in Eq. (10.67) correctly reduces to the infinite-series result in Eq. (10.46) when $n \to \infty$. This is true because if $n \to \infty$, the ar^{n+1} term in Eq. (10.67) is infinitely small when $|r| < 1$. Eq. (10.67) is therefore the more general of the two results, reducing to Eq. (10.46) when $n \to \infty$ (if $|r| < 1$). Logically, then, it might have made more sense to derive Eq. (10.67) for finite series first, and then produce Eq. (10.46) for infinite series as a special case when $n \to \infty$. But pedagogically we chose to discuss infinite series first, because the result is cleaner since we don't need to worry about the "leftover" ar^{n+1} term.

3. If we multiply both sides of Eq. (10.66) by $1 - r$ (analogous to the multiplication by $1 - r$ in Eq. (10.45)), there are now two uncanceled terms, so we obtain

$$(1 - r)(1 + r + r^2 + \cdots + r^{n-1} + r^n) = 1 - r^{n+1}. \tag{10.68}$$

We actually already knew this result from our factoring work in Chapter 4. Eq. (10.68) is simply Eq. (4.32) with some changes in notation: $a \to r$ and $n \to n + 1$ (and the equation is multiplied through by -1, and the sides are reversed). In retrospect, we could have just written down Eq. (10.66), and likewise Eq. (10.67), by invoking Eq. (4.32). But it was instructive to derive things from scratch here via the method in Eqs. (10.64) and (10.65).

4. In the special case where $r = 2$ (remember that $r > 1$ values are allowed in finite series), Eq. (10.66) gives

$$1 + 2 + 2^2 + \cdots + 2^{n-1} + 2^n = \frac{1 - 2^{n+1}}{1 - 2} = 2^{n+1} - 1. \tag{10.69}$$

In words: The sum of the powers of 2 up to 2^n is 1 less than the next power of 2. For example, $1 + 2 + 4 + 8 + 16 = 31$, which is 1 less than 32. This fact is basically the right half of Fig. 10.16 (from 1 to 2) in the solution to Exercise 10.33, with all of the lengths multiplied by 32. ♣

Exercise 10.44 The finite series on the lefthand side of Eq. (10.67) can be thought of as an *infinite* series starting with a, minus another *infinite* series starting with ar^{n+1}. This is true because all of the terms from ar^{n+1} onward will cancel in the subtraction, leaving you with the finite series in Eq. (10.67).

Explain how the result in Eq. (10.67) quickly follows from the above observation, combined with the $a/(1 - r)$ result in Eq. (10.46) for an infinite series.

Exercise 10.45 Eq. (10.66) gives the sum of the series $1 + r + r^2 + \cdots + r^n$ as $(1 - r^{n+1})/(1 - r)$. If you read this series backward, you can alternatively consider it to have a first term of $a = r^n$ and a ratio of $1/r$. Apply Eq. (10.67) to this backward series and show that the sum is again $(1 - r^{n+1})/(1 - r)$.

Exercise 10.46 The series S in Eq. (10.64) can be factored in these two ways, as you can check:

$$S = (1 + r^3)(1 + r + r^2) \qquad \text{and} \qquad S = (1 + r)(1 + r^2 + r^4). \tag{10.70}$$

Each of the trinomials here is a geometric series. Use Eq. (10.66) to find their sums, and then verify that the above two expressions for S agree with the result in Eq. (10.65)

Exercise 10.47 Consider the following product of a finite geometric series and an infinite one:

$$(1 + r + r^2 + r^3)(1 + r^4 + r^8 + r^{12} + \cdots). \tag{10.71}$$

Find the value of this product in two different ways:

(a) Multiply the two factors, and then sum the resulting series.

(b) Sum each series first, and then multiply the results.

Exercise 10.48 Find the sum of the *finite* arithmetic-geometric series,

$$1 + 2r + 3r^2 + 4r^3 + \cdots + nr^{n-1}. \tag{10.72}$$

Use the same strategy we used in Eqs. (10.61)–(10.63), where we found the sum of the *infinite* arithmetic-geometric series in Eq. (10.60). (The answer here is a little messy.)

10.7 Exercise solutions

1. If the AM is $\mu = 13$, then we want to find the x for which $(5 + x)/2 = 13$. Solving for x quickly gives $x = 21$. As a double check, $\mu = 13$ is the same distance (namely 8) from both 5 and 21.

   If the AM is $\mu = -2$, then we want to find the x for which $(5 + x)/2 = -2$. Solving for x quickly gives $x = -9$. As a double check, $\mu = -2$ is the same distance (namely 7) from both -9 and 5.

2. With n numbers, the statement analogous to $(\mu - a) + (\mu - b) = 0$ is

$$\begin{aligned}(\mu - a_1) + (\mu - a_2) + \cdots + (\mu - a_n) = 0 \quad &\Longrightarrow \quad n\mu - (a_1 + a_2 + \cdots + a_n) = 0 \\ &\Longrightarrow \quad \mu = \frac{a_1 + a_2 + \cdots + a_n}{n},\end{aligned} \tag{10.73}$$

   in agreement with Eq. (10.6).

   REMARK: The first equation in Eq. (10.73) says that the sum of the *signed* differences from μ to all of the n numbers equals zero. Some of these differences are positive, and some are negative, because some of the a's are less than μ, and some are larger.

   Equivalently, if you take all of the terms in Eq. (10.73) that have an a_i that is larger than μ and put them on the righthand side, then the equation looks like (where a_k is less than μ, and a_{k+1} is larger than μ)

$$(\mu - a_1) + (\mu - a_2) + \cdots + (\mu - a_k) = (a_{k+1} - \mu) + \cdots + (a_n - \mu). \tag{10.74}$$

   All of the terms on both sides of this equation are positive. The equation says that the sum of the differences from μ to all of the a's that are less than μ (that's the lefthand side) equals the sum of the (magnitudes of the) differences from μ to all of the a's that are greater than μ (that's the righthand side). For example, the AM of the five numbers $-6, -2, 1, 5, 12$ is $\mu = 2$, as you can check. The differences from 2 to each of the three numbers smaller than 2 (namely $-6, -2, 1$) are 8, 4, and 1. And the (magnitudes of the) differences from 2 to each of the two numbers larger than 2 (namely 5, 12) are 3 and 10. Both of these sets of differences add up to 13. So the sums are equal, as Eq. (10.74) claims. In this specific case, Eq. (10.74) takes the form of

$$\begin{aligned}&\big((2 - (-6)\big) + \big(2 - (-2)\big) + (2 - 1) = (5 - 2) + (13 - 2) \\ &\qquad\Longrightarrow \quad 8 + 4 + 1 = 3 + 10. \quad \clubsuit\end{aligned} \tag{10.75}$$

3. If the GM is $g = 12$, then we want to find the x for which $\sqrt{4 \cdot x} = 12$. Solving for x quickly gives $x = 36$. As a double check, $g = 12$ is the same factor (namely 3) from both 4 and 36.

   If the GM is $g = \sqrt{2}$, then we want to find the x for which $\sqrt{4 \cdot x} = \sqrt{2}$. Solving for x quickly gives $x = 1/2$. As a double check, $g = \sqrt{2}$ is the same factor (namely $2\sqrt{2}$) from both $1/2$ and 4.

4. With n numbers, the statement analogous to $(g/a)(g/b) = 1$ is (raising both sides to the $1/n$ power to obtain the last line)

$$\begin{aligned} \frac{g}{a_1} \cdot \frac{g}{a_2} \cdot \cdots \cdot \frac{g}{a_n} = 1 \quad &\Longrightarrow \quad \frac{g^n}{a_1 a_2 \cdots a_n} = 1 \\ &\Longrightarrow \quad g = (a_1 a_2 \cdots a_n)^{1/n}, \end{aligned} \tag{10.76}$$

in agreement with Eq. (10.12).

The first equation in Eq. (10.76) says that the product of the ratios between g and all of the n numbers equals 1. Some of these ratios are greater than 1, and some are less than 1, because some of the a's are less than g, and some are larger.

Equivalently, the product of the ratios of g to the a's (where a is less than g) equals the product of the ratios of the a's to g (where a is greater than g). The reasoning behind this statement is similar to the reasoning in the remark in the solution to Exercise 10.2, as you can verify.

5. The log of the GM of a and b is (using the power rule and the product rule for logs from Section 9.3)

$$\log g = \log \sqrt{ab} = \log(ab)^{1/2} = \frac{1}{2}\log(ab) = \frac{1}{2}\big(\log a + \log b\big), \tag{10.77}$$

which is the AM of the logs of a and b, as desired. We didn't bother writing the base of the log in the above equation, because the result is true for any value of the base.

Remark: We noted in the text that the geometric mean g of a and b is halfway between a and b *multiplicatively*, as opposed to additively. The result of this exercise tells us that if we take the log of the numbers, we end up with an additive result: $\log g$ is halfway between $\log a$ and $\log b$ *additively* (it's their average). For example, 100 is the geometric mean of 10 and 1000, and $\log_{10} 100 = 2$ is halfway between $\log_{10} 10 = 1$ and $\log_{10} 1000 = 3$. (We could have chosen any base here, but 10 makes the numbers be nice.) This result is equivalent to the fact that on the horizontal axis in the log plot back in Fig. 9.14, the 2 is halfway between the 1 and the 3. ♣

6. Expanding the square gives

$$\big(\sqrt{a} - \sqrt{b}\big)^2 \geq 0 \quad \Longrightarrow \quad a - 2\sqrt{ab} + b \geq 0 \quad \Longrightarrow \quad \frac{a+b}{2} \geq \sqrt{ab}\,, \tag{10.78}$$

as desired.

If $a = b$, then the "$\geq$" sign in the first of the above equations is an "=" sign. This sign carries through, so it ends up in the last equation too. Therefore, if $a = b$ then equality holds in Eq. (10.15).

Similarly, if $a \neq b$, then the "$\geq$" sign in the first of the above equations is a strict inequality "$>$" sign. This sign carries through, so it ends up in the last equation too. Therefore, if $a \neq b$ then equality does *not* hold in Eq. (10.15).

Putting these results together, we see that equality holds in Eq. (10.15) *if and only if* $a = b$. In the other proofs below, we won't keep repeating this fact. But you should verify in each proof that the line of reasoning implies that equality holds if and only if $a = b$.

7. Expanding the square and then adding $4ab$ to both sides turns $(a-b)^2$ into $(a+b)^2$, which is helpful because we're trying to produce a result involving $a+b$:

$$\begin{aligned} (a-b)^2 \geq 0 \quad &\Longrightarrow \quad a^2 - 2ab + b^2 \geq 0 \\ \Longrightarrow \quad a^2 + 2ab + b^2 \geq 4ab \quad &\Longrightarrow \quad (a+b)^2 \geq 4ab. \end{aligned} \tag{10.79}$$

Taking the square root and then dividing by 2 gives Eq. (10.15), as desired. (We're assuming that both a and b are positive, which means that the product ab is positive. So we can take its square root.) This proof is very similar to the preceding one, so perhaps it shouldn't count as a separate proof.

If the steps in Eq. (10.79) look familiar, it's because they're the same (in reverse) as the steps in Eq. (8.139) in the solution to Exercise 8.21.

8. (a) The AM of $a = \mu - d$ and $b = \mu + d$ is μ, due to how we constructed these forms of a and b. (Or you can just calculate $(a+b)/2$.) And the GM is

$$\text{GM} = \sqrt{ab} = \sqrt{(\mu - d)(\mu + d)} = \sqrt{\mu^2 - d^2} \leq \sqrt{\mu^2} = \mu = \text{AM}. \tag{10.80}$$

So GM $\leq$ AM, as we wanted to show. This proof boils down to the fact that the difference-of-squares result in Eq. (3.23) means that the GM always involves subtracting the positive (or zero) quantity d^2 from μ^2, yielding a quantity that is less than (or equal to) μ^2.

(b) The GM of $a = g/r$ and $b = gr$ is g, due to how we constructed these forms of a and b. (Or you can just calculate $\sqrt{ab}$.) And the AM is

$$\text{AM} = \frac{a+b}{2} = \frac{g/r + gr}{2} = \frac{g}{2}\left(\frac{1}{r} + r\right) \geq \frac{g}{2}(2) = g = \text{GM}, \tag{10.81}$$

where we have used the result from Example 8.9 to say that $r + 1/r \geq 2$. So AM $\geq$ GM, as we wanted to show.

9. Fig. 10.13 shows the four helpful rectangles. The area of the overall square, which is $(a+b)^2$, is larger than the total area of the four rectangles, which is $4 \cdot ab$. We have therefore demonstrated that $(a+b)^2 \geq 4ab$, as desired.

If Fig. 10.13 looks familiar, that's because it's the same as Fig. 3.6 back in Exercise 3.4. In that exercise we showed that the areas of the five sub-regions add up properly to the area of the overall square. This fact isn't necessary for the present purposes. All we need to know here is that the unshaded square in the middle of Fig. 10.13 has an area that is at least zero.

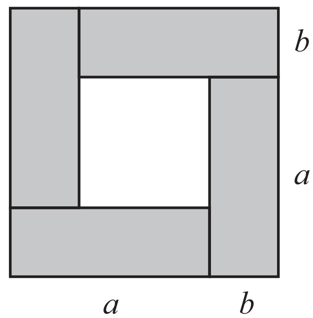

Figure 10.13: A geometric derivation of the AM-GM inequality. The area of the overall square is larger than the total area of the four rectangles.

10. The diameter of the semicircle is $a + b$, so the radius is $r = (a + b)/2$. This is therefore the length of the taller dashed line in Fig. 10.2.

    The shorter dashed line is the altitude h of the large right triangle. To find its length, consider the two smaller right triangles that the large one is divided into. These two right triangles are similar (from Exercise A.12), so we can equate the ratios of the short leg to the long leg:

$$\frac{\text{short leg}}{\text{long leg}}: \qquad \frac{h}{a} = \frac{b}{h} \quad \Longrightarrow \quad h^2 = ab \quad \Longrightarrow \quad h = \sqrt{ab}. \tag{10.82}$$

    We already derived this result in the third remark on page 586, where we invoked the lengths we found in Fig. A.52 (which were also obtained via similar-triangle reasoning).

    Since the taller dashed line in Fig. 10.2 is $r = (a+b)/2$ (which is the AM), and the shorter dashed line is $h = \sqrt{ab}$ (which is the GM), we conclude that $(a+b)/2 \geq \sqrt{ab}$, as desired. Note that the r line is in fact always at least as tall as the h line, because the vertical radius hits the semicircle at its highest point.

    REMARK: If you're wondering why the top angle in the large right triangle in Fig. 10.2 is a right angle, this can be demonstrated as follows. Due to the three radial lengths r shown in Fig. 10.14, triangles ADC and BDC are isosceles (see Section A.2.2), which means that they each have two equal angles. We have labeled the equal pairs as α and β, as shown.

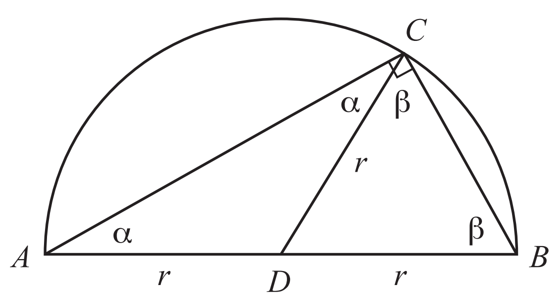

Figure 10.14: Proving that an angle inscribed in a semicircle equals 90°.

The three angles in triangle ABC are α, β, and $\alpha+\beta$. Therefore, since the sum of the angles in any triangle is $180°$ (see Section A.2.1), we have

$$(\alpha)+(\beta)+(\alpha+\beta)=180^\circ \implies 2\alpha+2\beta=180^\circ \implies \alpha+\beta=90^\circ. \tag{10.83}$$

And since $\alpha+\beta$ is the angle $\angle ACB$ in question, we see that this angle equals $90°$, as we wanted to show.

This $90°$ result is a special case of the more general fact that if an angle is inscribed in a circle, it equals half of the intercepted arc. (The proof is a variation of the one above.) In the case at hand, the intercepted arc is the other half of the circle (not shown), which spans an angle of $180°$. And half of this is $90°$. ♣

11. RMS $\geq$ AM: Squaring both sides of the RMS $\geq$ AM inequality gives

$$\begin{aligned} \sqrt{\frac{a^2+b^2}{2}} \geq \frac{a+b}{2} \quad &\implies \quad \frac{a^2+b^2}{2} \geq \frac{a^2+2ab+b^2}{4} \\ \implies \quad 2a^2+2b^2 \geq a^2+2ab+b^2 \quad &\implies \quad a^2+b^2 \geq 2ab \\ \implies \quad a^2-2ab+b^2 \geq 0 \quad &\implies \quad (a-b)^2 \geq 0, \end{aligned} \tag{10.84}$$

which is true. Since all of the above steps are reversible, the initial RMS $\geq$ AM inequality follows from the (true) final $(a-b)^2 \geq 0$ inequality.

GM $\geq$ HM: Rearranging the GM $\geq$ HM inequality gives

$$\sqrt{ab} \geq \frac{2ab}{a+b} \quad \implies \quad \frac{a+b}{2} \geq \sqrt{ab}, \tag{10.85}$$

which is simply the AM-GM inequality, which we know is true. (Or you can prove it again by multiplying though by 2 and squaring both sides, and then rearranging to obtain $(a-b)^2 \geq 0$.)

Alternatively, note that the product of the AM in Eq. (10.1) and the HM in Eq. (10.13) equals ab, which is the square of the GM. So

$$(\mathrm{GM})^2 = (\mathrm{AM})(\mathrm{HM}) \quad \implies \quad \mathrm{GM} = \sqrt{(\mathrm{AM})(\mathrm{HM})}. \tag{10.86}$$

That is, the GM is the geometric mean of the AM and HM. The GM must therefore lie between the AM and HM. And since we know that the GM is smaller than the AM, it must in turn be larger than the HM, as desired.

12. If we apply the AM-GM inequality to the quantities a^2 and b^2, and use the fact that $a^2+b^2=c^2$, we obtain

$$\frac{a^2+b^2}{2} \geq \sqrt{a^2b^2} \quad \implies \quad \frac{c^2}{2} \geq ab. \tag{10.87}$$

The desired maximum value of the product ab is therefore $c^2/2$. This is achieved (that is, equality holds in the inequality) when $a^2=b^2 \implies a=b$, in which case it follows that both a and b are equal to $c/\sqrt{2}$ (since $a^2+b^2=c^2$). This corresponds to a 45-45-90 right triangle (see Section A.7).

Since the area of a right triangle is $ab/2$ (see Section A.4.2), we see that the maximum area of a right triangle with a given hypotenuse c is $(c^2/2)/2 = c^2/4$, and this is achieved when $a=b=c/\sqrt{2}$. For further discussion of this, see the solution to Exercise 8.30.

13. With $a=x$ and $b=1-x$, the AM-GM inequality in Eq. (10.15) becomes

$$\frac{x+(1-x)}{2} \ge \sqrt{x(1-x)} \implies \frac{1}{2} \ge \sqrt{x(1-x)} \implies \frac{1}{4} \ge x(1-x). \qquad (10.88)$$

So $1/4$ is the desired maximum value. Equality is achieved when $a=b$, which implies $x=1-x \implies x=1/2$.

REMARK: The same reasoning quickly shows that the maximum value of $x(c-x)$ is $c^2/4$, and that this is achieved when $x=c/2$. The geometric interpretation of this result is that if half the perimeter of a rectangle is c (which means that the sides take the form of x and $c-x$, for some value of x), then the maximum value of the area $x(c-x)$ is $c^2/4$, and this is achieved when $x=c/2$, which means that the rectangle is a square. The $c^2/4$ result here is the same as the $P^2/16$ result in Exercise 8.28(a), because P was the entire perimeter, whereas c is half the perimeter. So $P=2c$. ♣

14. We'll use the general AM-GM inequality in Eq. (10.16), with four numbers a, b, c, and d. We'll choose $a=b=c=x/3$ and $d=1/x^3$. Eq. (10.16) then gives

$$\frac{1}{4}\left(\frac{x}{3}+\frac{x}{3}+\frac{x}{3}+\frac{1}{x^3}\right) \ge \sqrt[4]{\frac{x}{3}\cdot\frac{x}{3}\cdot\frac{x}{3}\cdot\frac{1}{x^3}} \implies x+\frac{1}{x^3} \ge 4\sqrt[4]{\frac{1}{27}} = \frac{4}{3^{3/4}}. \qquad (10.89)$$

So the minimum value of $x+1/x^3$ is $4/3^{3/4} = 1.75$. This is achieved when all four quantities in the inequality are equal, that is, when $x/3 = 1/x^3 \implies x^4 = 3 \implies x = 3^{1/4} = 1.32$. You can check that plugging this value of x into $x+1/x^3$ yields $4/3^{3/4}$.

15. With $a=1-x$ and $b=(1+2x)/2$ (it's fine to introduce the multiplicative factor of 2), the AM-GM inequality in Eq. (10.15) becomes

$$\frac{(1-x)+(1+2x)/2}{2} \ge \sqrt{(1-x)\cdot(1+2x)/2} \implies \frac{3}{4}\sqrt{2} \ge \sqrt{(1-x)(1+2x)}$$
$$\implies \frac{9}{8} \ge (1-x)(1+2x). \qquad (10.90)$$

So $9/8$ is the desired maximum value. Equality is achieved when $a=b$, which implies $1-x=(1+2x)/2 \implies x=1/4$. Note that the solution to this problem didn't involve three terms, as Example 10.3 did. We simply needed to divide one of the two given terms by 2. (Multiplying the other one by 2 would also work.)

16. We'll use the general AM-GM inequality in Eq. (10.16), with five numbers. We'll choose $a=b=c=x^2/3$ and $d=e=1/(2x^3)$. These five quantities add up to x^2+1/x^3, and

their product is a constant. Eq. (10.16) gives

$$\frac{1}{5}\left(\frac{x^2}{3}+\frac{x^2}{3}+\frac{x^2}{3}+\frac{1}{2x^3}+\frac{1}{2x^3}\right) \ge \sqrt[5]{\frac{x^2}{3}\cdot\frac{x^2}{3}\cdot\frac{x^2}{3}\cdot\frac{1}{2x^3}\cdot\frac{1}{2x^3}}$$
$$\Longrightarrow \quad x^2+\frac{1}{x^3} \ge 5\sqrt[5]{\frac{1}{27}\cdot\frac{1}{4}} = \frac{5}{3^{3/5}2^{2/5}} = 1.96. \qquad (10.91)$$

So the minimum value of x^2+1/x^3 is 1.96. This is achieved when all five quantities in the inequality are equal, that is, when $x^2/3 = 1/(2x^3) \implies x^5 = 3/2 \implies x = (3/2)^{1/5} = 1.08$.

17. In the a_n notation, we want to show that the relation $a_n = (a_{n-1}+a_{n+1})/2$ is true for any n (except $n = 1$, because a_1 doesn't have a neighbor on its left side). Using the expression for a_n in Eq. (10.24), the $a_n = (a_{n-1}+a_{n+1})/2$ relation becomes

$$\begin{aligned} a+(n-1)d &= \frac{1}{2}\Big(\big[a+\big((n-1)-1\big)d\big]+\big[a+\big((n+1)-1\big)d\big]\Big) \\ &= \frac{1}{2}\Big(2a+(2n-2)d\Big) \\ &= a+(n-1)d. \end{aligned} \qquad (10.92)$$

The two sides agree, as desired.

18. With $n = 7$, Eq. (10.24) tells us that $a_7 = a+(7-1)d$. Since $a = 5$ and $a_7 = 59$, this equation becomes $59 = 5+6d \implies 54 = 6d \implies d = 9$. The seven terms in the sequence are therefore 5, 14, 23, 32, 41, 50, 59.

You can also solve this problem in your head. Since 59 is 54 more than 5, you know that 6 times the spacing d (yes, just 6 and not 7, due to the fencepost theorem) must equal 54. The spacing d is therefore 9.

19. (a) Let the two spacings be d_1 and d_2. And let the common first term be a (this won't matter; it will cancel). Then Eq. (10.24) tells us that

$$a+(10-1)d_1 = a+(5-1)d_2 \quad\Longrightarrow\quad 9d_1 = 4d_2 \quad\Longrightarrow\quad \frac{d_1}{d_2}=\frac{4}{9}. \qquad (10.93)$$

So the ratio is $4/9$ (or $9/4$, depending on how you look at it). The first sequence has the smaller d.

The important point here is that the desired ratio of the d's is *not* equal to $5/10 = 1/2$. The fencepost theorem tells us that only 9 of the spacings are needed to get to the 10th term, and only 4 of them are needed to get to the 5th term. So the ratio is $4/9$.

(b) If $d_1 = 4$, then Eq. (10.93) tells us that $d_2 = 9$. So if the common first term is 2, the two sequences are:

$$\begin{aligned} \text{First}: &\;\; 2,\ 6,\ 10,\ 14,\ 18,\ 22,\ 26,\ 30,\ 34,\ 38,\ldots \\ \text{Second}: &\;\; 2,\ 11,\ 20,\ 29,\ 38,\ldots \end{aligned} \qquad (10.94)$$

The 10th and 5th terms are correctly equal, with 38 being their common value.

20. The given rule tells us that

$$\begin{aligned} a_2 = Aa_1^2 + Ba_1 &\implies 8 = A \cdot 4^2 + B \cdot 4 \implies 2 = 4A + B, \\ a_3 = Aa_2^2 + Ba_2 &\implies 12 = A \cdot 8^2 + B \cdot 8 \implies 3 = 16A + 2B. \end{aligned} \tag{10.95}$$

This is a system of two equations in the two unknowns A and B. Solving for B in the first equation and plugging the result into the second gives

$$3 = 16A + 2(2 - 4A) \implies -1 = 8A \implies A = -1/8. \tag{10.96}$$

And then either equation quickly gives $B = 5/2$. So the rule is

$$a_n = -\frac{1}{8}a_{n-1}^2 + \frac{5}{2}a_{n-1}. \tag{10.97}$$

You can double check that this does indeed turn 4 into 8, and then 8 into 12.

If you apply the rule to $a_3 = 12$, you will obtain $a_4 = (-1/8)12^2 + (5/2)12 = -18 + 30 = 12$. And then you're stuck at 12 for all future terms. So the sequence governed by our strange rule is 4, 8, 12, 12, 12, etc., whereas the standard arithmetic sequence would be 4, 8, 12, 16, 20, etc. But given only the first three terms, there's no way to know which sequence it is.

Furthermore, there is an infinite number of other possible rules, for example, $a_n = Aa_{n-1}^7 + B/a_{n-1}^5$. You could solve for the A and B that make this rule work by using the same strategy as in Eq. (10.95), although the numbers would be messier. However, with the exception of this exercise, we'll stick to the standard kinds of sequences in this chapter, namely arithmetic and geometric (and also a combination of these in the subsection on page 619).

Remark: This remark isn't important for our general discussion of sequences, but it's an excuse to introduce an interesting technique. We saw above that if you start with the number 4 and successively generate new terms via Eq. (10.97), you will get stuck at 12. You might be wondering if this is a coincidence. If you start with another number, will you also get stuck at 12? Or maybe some other value? You should try a few other starting numbers and see what happens when you repeatedly apply Eq. (10.97), before reading further.

To answer this, note that the righthand side of Eq. (10.97) can be expressed as the function $f(x) = -x^2/8 + 5x/2$. Type this function into Desmos, including the "$f(x) =$." Then type "$f(s)$" in a new box and enable the slider for s. (We've picked the letter s for "starting" value.) The numerical value of $f(s)$ will appear in the lower-right corner of the $f(s)$ box. Change the slider bounds (by clicking on either bound) to $-10 \le a \le 30$, with a step of 0.1.

Then type $f(f(s))$, and $f(f(f(s)))$, and so on in new boxes, up to, say, six f's. Composing the function f with itself like this will generate the successive terms in the sequence defied by Eq. (10.97), because that equation tells you to take the a term you just generated and plug it into the function on the righthand side of Eq. (10.97) to obtain the next a term. And that's exactly what composing a function with itself does.

As you vary the slider for s, you will find that for many different values of s (but not all), the terms in the sequence eventually get very close to 12. More precisely, you will find that:

- If $0 < s < 20$, the terms approach 12.
- If $s = 0$, the terms are all 0.
- If $s = 20$, the second term is 0, and then they stay at 0.
- If $s < 0$ or $s > 20$, the terms head off to $-\infty$.

So there is indeed something special about the number 12 (and also about the numbers 0 and 20). It turns out that there is a simple graphical way of seeing why. Fig. 10.15 shows what happens if you start with $s = 3$ and then generate successive terms in the sequence by repeatedly composing the function $f(x)$ with itself.

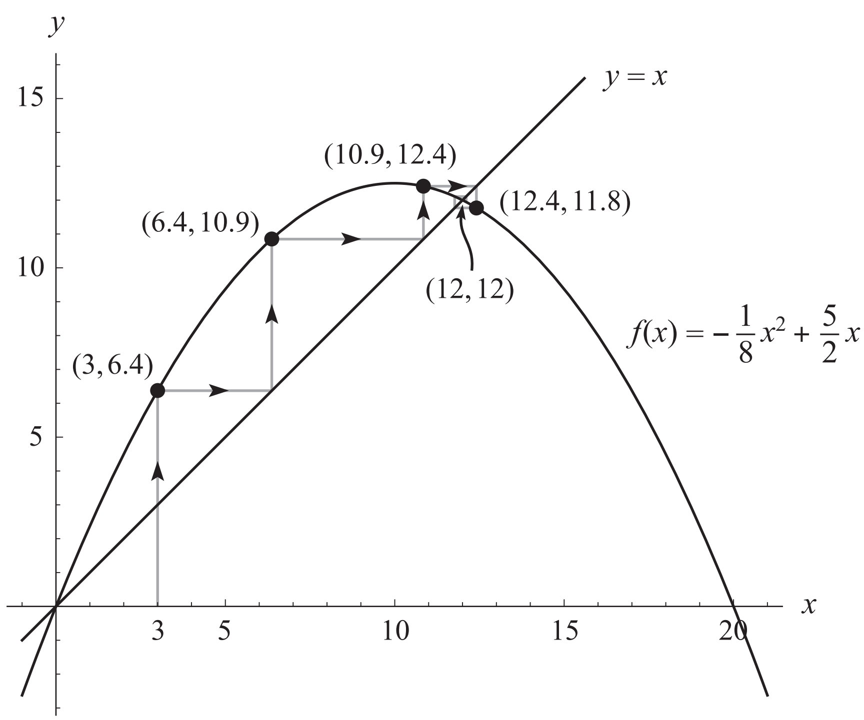

Figure 10.15: The graphical explanation of why the sequence described by Eq. (10.97) always approaches 12, if the starting value s is between 0 and 20.

The first segment of the gray path in the figure tells you that $f(3) = 6.4$. So 6.4 is the second term in the sequence. If you then move horizontally to the $y = x$ line, this makes your new x value be 6.4. Moving vertically then tells you that $f(f(3)) = f(6.4) = 10.9$. So 10.9 is the third term in the sequence. Moving horizontally to the $y = x$ line makes your new x value be 10.9, and then moving vertically tells you that $f(f(f(3))) = f(10.9) = 12.4$.

So 12.4 is the fourth term in the sequence. And so on. We see from the figure that the path "spirals" in to the point $(12, 12)$. Therefore, the terms in the sequence approach 12.

You can verify that for any starting value s between 0 and 20, the gray path will approach the point $(12, 12)$. So the terms in the sequence approach 12, as we stated in the first bullet point above. You can also verify that if $s < 0$ or $s > 20$, the path will head off down to the left, to $-\infty$. Lastly, if $s = 0$ or $s = 20$, the second term is 0, and then you stay there.

Hopefully you found this remark interesting. But as mentioned above, none of this is important for our discussion of sequences. ♣

21. When $a = 1$ and $d = 1$, Eq. (10.27) correctly gives

$$S_n = \frac{n\big(2 \cdot 1 + (n-1) \cdot 1\big)}{2} = \frac{n(n+1)}{2}. \tag{10.98}$$

And when $a_1 = 1$ and $a_n = n$, Eq. (10.28) also gives $S_n = n(1+n)/2$.

22. Plugging the eight numbers into your calculator hopefully yielded a sum of 168.

When using Eq. (10.27), the first term in the sequence is $a = 7$, the difference is $d = 4$, and the number of terms in $n = 8$. So Eq. (10.27) gives

$$S = \frac{n\big(2a + (n-1)d\big)}{2} = \frac{8\big(2 \cdot 7 + (8-1)4\big)}{2} = \frac{8(14+28)}{2} = 168. \tag{10.99}$$

When using Eq. (10.28), the first term is $a_1 = 7$ and the last term is $a_8 = 35$. So Eq. (10.28) gives $S = 8(7+35)/2 = 168$.

23. Eq. (10.24) tells us that the 9th term equals $a + (9-1)d$. Using the given information, this becomes $29 = 5 + 8d \implies d = 3$. Eq. (10.27) then gives the sum as

$$S = \frac{n\big(2a + (n-1)d\big)}{2} = \frac{23\big(2 \cdot 5 + (23-1)3\big)}{2} = 874. \tag{10.100}$$

Alternatively, Eq. (10.24) tells us that the last (23rd) term in the sequence is $5+(23-1)\cdot 3 = 71$. So Eq. (10.28) gives the sum as $23(5+71)/2 = 874$.

24. Plugging the six k values into the three sums gives

$$\begin{aligned}&4\big[(-2)^2 + (-1)^2 + 0^2 + 1^2 + 2^2 + 3^2\big] + 4\big[(-2) + (-1) + 0 + 1 + 2 + 3\big]\\ &\qquad + \big[1+1+1+1+1+1\big]\\ &= 4\big[19\big] + 4\big[3\big] + \big[6\big] = 94,\end{aligned} \tag{10.101}$$

as desired.

25. The sum can be written as

$$7 + 11 + 15 + 19 + 23 + 27 + 31 + 35 = \sum_{k=0}^{7} (7 + k \cdot 4), \tag{10.102}$$

as you can verify by plugging the successive values of k into the righthand side. If you want to break up the sum, you can write it as

$$\begin{aligned} &= \sum_{k=0}^{7} 7 + 4\sum_{k=0}^{7} k = (7+7+7+7+7+7+7+7) + 4(0+1+2+3+4+5+6+7) \\ &\qquad\qquad = 56 + 4\cdot 28 = 168, \end{aligned} \tag{10.103}$$

in agreement with Eq. (10.29).

You can also write the sum as $\sum_{k=1}^{8}\big(7+(k-1)\cdot 4\big)$, as you can verify. Note that when a $\sum$ is written like this in a line of the text (as opposed to in a displayed equation, like in Eq. (10.102)), the upper and lower limits of k are written to the right of the $\sum$, because there isn't enough space above and below it.

26. This exercise is the general version of the preceding one. In the $\sum$ notation, the series in Eq. (10.26) can be written as

$$\sum_{k=0}^{n-1}(a+kd). \tag{10.104}$$

There are also other ways to write the sum, for example,

$$\sum_{k=1}^{n}\big(a+(k-1)d\big) \qquad \text{or} \qquad \sum_{k=13}^{n+12}\big(a+(k-13)d\big). \tag{10.105}$$

In both of these sums, the coefficient of d correctly runs from 0 to $n-1$. The first sum is a reasonable and helpful way to write things, but the second one isn't. Nothing is gained (and clarity is lost) by writing the sum in this confusing way.

The sum in Eq. (10.104) can be broken up into the two sums,

$$\sum_{k=0}^{n-1} a + \sum_{k=0}^{n-1} kd = na + d\sum_{k=0}^{n-1} k = na + d\cdot\frac{(n-1)n}{2}, \tag{10.106}$$

in agreement with Eq. (10.30). The na here comes from the fact that we're adding up n a's (yes, there are n of them since k runs from 0 to $n-1$). And we have used Eq. (4.55), with n replaced by $n-1$, to evaluate the $\sum k$ sum. The 0 value of k doesn't contribute anything to this sum, but that's fine. In contrast, the 0 value of k *does* contribute to the $\sum a$ sum, because it yields an a just like every other value of k does.

REMARK: In Eq. (10.104) we made sure to use the full "$k=$" notation below the $\sum$, because the sum involves other letters (namely a and d). The "$k=$" isn't necessary if there are only k's involved, as in, say, $\sum_2^7(k^2+1/k)$. However, when there are other letters, we need to indicate which one we're summing over. The other letters are then assumed to be constants. For example, if we instead summed over a or d in Eq. (10.104), we would obtain these two sums:

$$\begin{aligned}\sum_{a=0}^{n-1}(a+kd) &= (0+kd)+(1+kd)+(2+kd)+\cdots+(n-1+kd)\\ &= \frac{(n-1)n}{2}+nkd,\\ \sum_{d=0}^{n-1}(a+kd) &= (a+0)+(a+k)+(a+2k)+\cdots+(a+(n-1)k)\\ &= na+\frac{(n-1)n}{2}\cdot k. \qquad (10.107)\end{aligned}$$

The righthand sides here illustrate the fact that if you're ever confused about what a sum means when written in the $\sum$ notation, you should plug in some values of the index (a or d here) and see what the terms look like.

The first of the above sums involves no a's in the end, because we summed over a. Likewise, the second sum involves no d's. (And the result in Eq. (10.106) involves no k's.) The second sum takes the same form as Eq. (10.106), with d replaced by k. ♣

27. The two ways are

$$\left(\sum_1^4 k\right)+1 \qquad \text{and} \qquad \sum_1^4 (k+1), \qquad (10.108)$$

which give numerical results of

$$(1+2+3+4)+1=11 \qquad \text{and} \qquad 2+3+4+5=14. \qquad (10.109)$$

The first of these is the one that is technically correct. If you want the 1 to be included in the sum, you should explicitly use parentheses, as we did in the second way. However, in practice, if all of the terms involve k (which isn't the case with the 1 above), as in, say, $\sum 2k^3+5k^2-3k$, then it is clear that the sum is intended to include all three terms. This is true because if k is being summed over in the first term (the $2k^3$), then k can't be a constant that sits outside the sum, like the above 1 could be.

28. In the a_n notation, we want to show that the relation $a_n=\sqrt{a_{n-1}a_{n+1}}$ is true for any n (except $n=1$, because a_1 doesn't have a neighbor on its left side). Using the expression for a_n in Eq. (10.37), the $a_n=\sqrt{a_{n-1}a_{n+1}}$ relation becomes

$$ar^{n-1}=\sqrt{ar^{(n-1)-1}\cdot ar^{(n+1)-1}}=\sqrt{a^2r^{2n-2}}=ar^{n-1}. \qquad (10.110)$$

The two sides agree, as desired.

29. With $n=7$, Eq. (10.37) tells us that $a_7=ar^6$. Plugging the given values into this equation yields

$$24=3r^6 \implies 8=r^6 \implies r=(2^3)^{1/6}=2^{1/2}=\sqrt{2}. \qquad (10.111)$$

The seven terms in the sequence are therefore 3, $3\sqrt{2}$, 6, $6\sqrt{2}$, 12, $12\sqrt{2}$, 24.

You can also solve this problem in your head. Since 24 is 8 times 3, you know that 6 powers of r (yes, just 6 and not 7, due to the fencepost theorem) must equal 8. The ratio r is therefore $8^{1/6} = 2^{1/2}$.

30. The corresponding legs of the (similar) light-shaded right triangle (associated with the squares with sides a and b) are likewise the side length a and half the difference $(b-a)/2$. The ratio of the short leg to the long leg is the same in both triangles, so we have

$$\frac{\text{short leg}}{\text{long leg}}: \quad \frac{(b-a)/2}{a} = \frac{(c-b)/2}{b} \quad \Longrightarrow \quad b(b-a) = a(c-b) \tag{10.112}$$
$$\Longrightarrow \quad b^2 - \cancel{ba} = ac - \cancel{ab} \quad \Longrightarrow \quad b^2 = ac \quad \Longrightarrow \quad b = \sqrt{ac}.$$

So b is the geometric mean of a and c, as desired. Equivalently, the $b^2 = ac$ relation in the second-to-last equation can be rewritten as $b/a = c/b$. This equality of ratios means that the three sides form a geometric sequence. As in Example 10.5, this result also holds for all other squares we could draw, extending to the right or the left in Fig. 10.8. So the sides all form a geometric sequence.

Interestingly, if instead of growing squares we have growing *rectangles* with the *same width* (so that the tops and bottoms of the rectangles form standard staircases), then the vertical sides of the rectangles form an *arithmetic* sequence, instead of the geometric sequence we found above for the growing squares. You can quickly verify this by drawing a picture. You will see that successive sides have a constant *difference* between them, as opposed to the constant *ratio* between them that we found above.

31. (a) From Eq. (10.37), the nth term in a geometric sequence takes the form of $a_n = ar^{n-1}$. So the relation $a_n = a_{n-1} + a_{n-2}$ becomes

$$ar^{n-1} = ar^{n-2} + ar^{n-3}. \tag{10.113}$$

Canceling the common factor of ar^{n-3} yields

$$r^2 = r + 1 \quad \Longrightarrow \quad r^2 - r - 1 = 0. \tag{10.114}$$

The quadratic formula then gives the two roots as

$$r_{\pm} = \frac{1 \pm \sqrt{5}}{2}. \tag{10.115}$$

You should square each of these r_+ and r_- roots and check explicitly that they satisfy $r^2 = r + 1$.

Note that the above two roots each make the relation $a_n = a_{n-1} + a_{n-2}$ be true for *any* n, because the n disappeared in going from Eq. (10.113) to Eq. (10.114). The results for r are therefore independent of n.

(b) We want to show that $a_n = Ar_+^{n-1} + Br_-^{n-1}$ satisfies the relation $a_n = a_{n-1} + a_{n-2}$ for all n, and for any values of A and B. That is, we want to show that

$$Ar_+^{n-1} + Br_-^{n-1} = \left(Ar_+^{n-2} + Br_-^{n-2}\right) + \left(Ar_+^{n-3} + Br_-^{n-3}\right). \tag{10.116}$$

We know that r_+ satisfies Eq. (10.113). So if we relabel $a \to A$ in that equation (since a can take on any arbitrary value), then the $Ar_+^{n-2} + Ar_+^{n-3}$ part of the sum on the righthand side of Eq. (10.116) can be rewritten as Ar_+^{n-1}, which agrees with what appears on the lefthand side. The same goes for the r_- terms (the ones with a B). So Eq. (10.116) is indeed true, for any values of n, A, and B.

(c) We must now find the values of A and B that make the first two terms in the sequence defined by $a_n = Ar_+^{n-1} + Br_-^{n-1}$ be equal to 1. Letting $n = 1$ and $n = 2$ (which means that the exponents are 0 and 1) in this expression for a_n, and demanding that the results equal 1, gives

$$
\begin{aligned}
a_1 = 1 \quad &\Longrightarrow \quad Ar_+^0 + Br_-^0 = 1 \quad \Longrightarrow \quad A + B = 1, \\
a_2 = 1 \quad &\Longrightarrow \quad Ar_+^1 + Br_-^1 = 1 \quad \Longrightarrow \quad A\left(\frac{1+\sqrt{5}}{2}\right) + B\left(\frac{1-\sqrt{5}}{2}\right) = 1.
\end{aligned}
\tag{10.117}
$$

There are various ways to solve this system of equations for A and B. We'll use the first equation to write $B = 1 - A$ and then plug this into the second equation. The result is

$$
A\left(\frac{1+\sqrt{5}}{2}\right) + (1-A)\left(\frac{1-\sqrt{5}}{2}\right) = 1. \tag{10.118}
$$

Collecting the A terms on the left, and putting the regular numbers on the right, gives

$$
\sqrt{5}A = \frac{1+\sqrt{5}}{2} \quad \Longrightarrow \quad A = \frac{1}{\sqrt{5}} \cdot \frac{1+\sqrt{5}}{2}, \tag{10.119}
$$

which can conveniently be written as $A = r_+/\sqrt{5}$. In a similar manner (or by using $B = 1 - A$), you can show that $B = -r_-/\sqrt{5}$. The general expression for the nth term, $a_n = Ar_+^{n-1} + Br_-^{n-1}$, therefore becomes

$$
\begin{aligned}
a_n = Ar_+^{n-1} + Br_-^{n-1} &= \frac{r_+}{\sqrt{5}} \cdot r_+^{n-1} - \frac{r_-}{\sqrt{5}} \cdot r_-^{n-1} = \frac{1}{\sqrt{5}}\left(r_+^n - r_-^n\right) \\
\Longrightarrow \quad &\boxed{a_n = \frac{1}{\sqrt{5}}\left[\left(\frac{1+\sqrt{5}}{2}\right)^n - \left(\frac{1-\sqrt{5}}{2}\right)^n\right]}
\end{aligned}
\tag{10.120}
$$

This is the general formula for the nth term in the Fibonacci sequence. You can check that when $n = 1$, 2, and 3 the formula correctly yields the first three terms in the sequence: 1, 1, 2. Of course, we explicitly constructed the a_n in Eq. (10.120) to have the property that it equals 1 when $n = 1$ and 2; that's what we did in Eq. (10.117). But it's still good to double check. You are encouraged to also check that $n = 4$ yields the fourth term, 3. The $n = 1$ check is quick, and for $n = 2$, 3, and 4 the algebra isn't too bad, because many of the terms cancel in the binomial expansions.

We'll now present a long list of remarks about the Fibonacci numbers. It isn't necessary to read these, and there are probably more details here than you care to know. But some of them are quite interesting, so dive in if you wish!

REMARKS:

(1) Before solving this problem, it isn't at all obvious that the terms in the Fibonacci sequence should all be able to be written in the form of $a_n = Ar_+^{n-1} + Br_-^{n-1}$, where simply increasing the power of the r's by 1 gives you the next term. But we showed that this expression does in fact produce all of the terms, assuming that A and B are chosen correctly.

(2) Once we know that the formula in Eq. (10.120) works for $n = 1$ and 2, it is guaranteed to work for all other n, because the formula was constructed so that the relation $a_n = a_{n-1} + a_{n-2}$ is satisfied. Given that this relation holds, the first two terms completely determine all the other terms in the sequence. That is, any sequence that satisfies $a_n = a_{n-1} + a_{n-2}$ and has its first two terms equal to 1 must be the sequence in Eq. (10.40). So there was technically no need to check the formula for $n = 3$; it's guaranteed to work. But it's still nice to see how the $a_3 = 2$ value pops out of the algebra.

(3) Eq. (10.120) also works for $n = 0$, because it yields $a_0 = 0$. And you can consider the Fibonacci sequence to start with the two numbers 0 and 1. The next term is 1, and then 2, and so on, as in Eq. (10.40).

(4) The minus sign between the two binomial expansions in Eq. (10.120) means that for any n, the 1st, 3rd, 5th, etc., terms in the binomial expansions will always cancel (because they are positive in both expansions); these terms all involve an even power of $\sqrt{5}$. Conversely, the 2nd, 4th, 6th, etc., terms in the binomial expansions will always add (because they are negative in the second expansion, and there is an overall minus sign between the expansions); these terms all involve an odd power of $\sqrt{5}$. Therefore, for all n, every term in the combined numerator in Eq. (10.120) will contain a leftover factor of $\sqrt{5}$. This will then cancel with the $\sqrt{5}$ in the denominator out front. So a_n will never contain a leftover $\sqrt{5}$, which is correct since the Fibonacci numbers are all simple integers.

(5) Since the exponents in the final result in Eq. (10.120) are n, you might think that it would have been better to define a_n as $Ar_+^n + Br_-^n$ instead of the $Ar_+^{n-1} + Br_-^{n-1}$ expression we used. That would certainly work, but the system of equations in Eq. (10.117) would be messier, so the calculation would take longer. The new results of $A = 1/\sqrt{5}$ and $B = -1/\sqrt{5}$ would be simpler, but the end result in Eq. (10.120) would be the same.

(6) The $(1+\sqrt{5})/2 = 1.618$ and $(1-\sqrt{5})/2 = -0.618$ numbers in Eq. (10.120) are the *golden ratio* and its inverse (or rather the negative of its inverse). We encountered the golden ratio back in Example 8.5. It shows up in all sorts of unexpected places in math, including the Fibonacci sequence! If we label the golden ratio as

$$\varphi \equiv \frac{1+\sqrt{5}}{2} \qquad \text{(Golden ratio)} \tag{10.121}$$

then $(1-\sqrt{5})/2$ equals both $-1/\varphi$ and $1-\varphi$, as you can check. So the result in Eq. (10.120) can be written in these two ways:

$$a_n = \frac{1}{\sqrt{5}}\big(\varphi^n - (-1/\varphi)^n\big) \qquad \text{or} \qquad a_n = \frac{1}{\sqrt{5}}\big(\varphi^n - (1-\varphi)^n\big). \tag{10.122}$$

(7) For large n, Eq. (10.120) allows us to easily produce an approximate expression for a_n. Since the $\big((1-\sqrt{5})/2\big)^n$ term in Eq. (10.120) equals $(-0.618)^n$, and since 0.618 is smaller than 1, its nth power is small if n is large. So for large n we can roughly ignore the second term in Eq. (10.120), leaving us with only the first term. For large n, the nth Fibonacci number a_n is therefore approximately equal to

$$\boxed{a_n \approx \frac{1}{\sqrt{5}}\left(\frac{1+\sqrt{5}}{2}\right)^n = \frac{\varphi^n}{\sqrt{5}}} \tag{10.123}$$

Furthermore, n doesn't need to be all that large in order for this to be a good approximation, because $(0.618)^n$ becomes small very quickly. For example, if $n = 8$ then $\big((1+\sqrt{5})/2\big)^n/\sqrt{5}$ equals 21.01. And if $n = 11$ it equals 88.998. These results are very close to 21 and 89, which are the 8th and 11th numbers in Eq. (10.40).

This means that you never need to bother with the second term in Eq. (10.120). You can simply use a calculator to evaluate the first term, and then pick the integer that the result is very close to. For example, $\big((1+\sqrt{5})/2\big)^{20}/\sqrt{5} = 6765.00003$, so the 20th Fibonacci number is 6765. If your goal is to find the nth Fibonacci number for just a single n, and if n is large, this method is much quicker than calculating the Fibonacci numbers in order by successively adding up the two most recent numbers to obtain the next.

The $a_n \approx \varphi^n/\sqrt{5}$ result in Eq. (10.123) tells us that for large n, the ratio of a_{n+1}/a_n approaches the golden ratio, $\varphi = (1+\sqrt{5})/2$ (as we observed in the remark in the solution to Exercise 8.18). Even if n isn't that large, this is still a good approximation. For example, $a_{11}/a_{10} = 1.61818$, which is very close to the golden ratio, $\varphi \approx 1.61803$. The Fibonacci sequence is therefore *approximately* a geometric sequence. And the approximation gets better and better as n increases.

(8) What is the sum S_n of the first n terms in the Fibonacci sequence? If you calculate the sum for a few values of n, you will quickly discover that it is given by $S_n = a_{n+2} - 1$. For example, the sum of the first nine terms (up to 34) is $S_9 = 88$, which is 1 less than $a_{11} = 89$.

The hard way of proving this $S_n = a_{n+2} - 1$ result is to use the finite-geometric-series result from Eq. (10.67) to sum up the terms in the two geometric series generated by the two terms in the expression for a_n in Eq. (10.120). This is more tractable if you use the first of the formulas in Eq. (10.122) and sum up the φ^n series and the $(-1/\varphi)^n$ series. You will need to use the defining property of φ (namely $\varphi^2 = \varphi + 1$ from Eq. (10.114)) to simplify your result. You will eventually obtain a sum equal to $a_{n+2} - 1$, although it takes a bit of messy algebra.

The quick way of proving the $S_n = a_{n+2} - 1$ result is to use mathematical induction, which you'll learn about in the next chapter. See Exercise 11.9.

The intermediate way of proving the result is to write down the Fibonacci sequence out to the nth term, and then below that write it out to the $(n+1)$th term, but shifted one place to the left, as shown in Table 10.2, where we have chosen $n = 7$. (Imagine that there are + signs between the numbers.) Adding up the two terms in each column produces the next term in the sequence, as shown, since that's how the Fibonacci sequence is defined. So the sum of the two rows equals $S_{n+2} - 1$, because we now have the sum up to 34 (the 9th term), but the first 1 is missing. (Stop reading here if you want to figure out on your own how the $S_n = a_{n+2} - 1$ result follows from Table 10.2.)

	1	1	2	3	5	8	13	=	S_n
1	1	2	3	5	8	13	21	=	S_{n+1}
1	2	3	5	8	13	21	34	=	$S_{n+2} - 1$

Table 10.2: Finding the S_n sum of the Fibonacci sequence, for $n = 7$.

The table tells us that $S_n + S_{n+1} = S_{n+2} - 1$. Solving for S_n gives

$$S_n = (S_{n+2} - S_{n+1}) - 1 \quad \Longrightarrow \quad S_n = a_{n+2} - 1, \tag{10.124}$$

as desired. We used the fact that $S_{n+2} = S_{n+1} + a_{n+2}$, since S_{n+2} contains one more term (namely a_{n+2}) than S_{n+1}.

(9) The procedure we used to solve this exercise can be generalized in a number of ways:

1. What if a sequence is governed by the same $a_n = a_{n-1} + a_{n-2}$ rule as the Fibonacci sequence, but the first two terms aren't 1's? The same procedure works, with a slight modification. We obtain the same r values in Eq. (10.115), but the righthand sides of the equations in Eq. (10.117) are modified, because the 1's get replaced with the new first two terms. (The lefthand sides of the equations are the same.) So we have to solve a different system of equations, which will give us different values of A and B (with the same values of r_+ and r_-).
2. What if a sequence is governed by a different rule, for example $a_n = 2a_{n-1} + 3a_{n-2}$? The same procedure works, except that Eq. (10.114) becomes $r^2 = 2r + 3$. The quadratic formula will then give new values of r_+ and r_-, so the system of equations in Eq. (10.117) will be different. But the general strategy is the same. See Exercise 10.32.
3. What if we have a more complicated rule like $a_n = a_{n-1} + 3a_{n-2} + 2a_{n-3}$, or even just $a_n = a_{n-1} + 2a_{n-3}$? The same procedure works (in theory), although we now end up with a *cubic* equation (as you can verify), as opposed to the quadratic equation in Eq. (10.114). Cubic equations are much harder to solve, so the problem is much messier. Additionally, Eq. (10.117) gets replaced with a system of *three* equations in three unknowns, because you need to specify the first three terms to get the sequence going. ♣

32. All of the explanations here are the same as in the solution to Exercise 10.31, but the equations are modified. Eq. (10.113) is now

$$ar^{n-1} = 2ar^{n-2} + 3ar^{n-3}, \tag{10.125}$$

so Eq. (10.114) becomes

$$r^2 = 2r + 3 \implies r^2 - 2r - 3 = 0. \tag{10.126}$$

This equation quickly factors into $(r-3)(r+1)$, so the roots are $r_+ = 3$ and $r_- = -1$. Or you can use the quadratic formula.

Using the general expression $a_n = Ar_+^{n-1} + Br_-^{n-1}$, and letting $n = 1$ and $n = 2$, the system of equations in Eq. (10.117) is now

$$\begin{aligned} a_1 = 1 \quad &\implies \quad Ar_+^0 + Br_-^0 = 1 \quad \implies \quad A + B = 1, \\ a_2 = 1 \quad &\implies \quad Ar_+^1 + Br_-^1 = 1 \quad \implies \quad A(3) + B(-1) = 1. \end{aligned} \tag{10.127}$$

Adding these two equations and dividing by 2 quickly gives $A = 1/2$. And then either equation gives $B = 1/2$. The general expression for the nth term, $a_n = Ar_+^{n-1} + Br_-^{n-1}$, is therefore

$$a_n = \frac{3^{n-1} + (-1)^{n-1}}{2}. \tag{10.128}$$

You can check that when $n = 1, 2, 3, 4, 5$, and 6, this formula correctly yields the first six terms in the sequence: 1, 1, 5, 13, 41, 121. Checking the values here is much cleaner than it was in Exercise 10.31 with all the $\sqrt{5}$'s.

33. Fig. 10.16 shows a way of seeing why the sum is 2. Every term that is added on brings the sum halfway closer to 2. For example, $1 + 1/2 + 1/4 + 1/8 = 15/8$. And then when you add on $1/16$, the sum increases halfway from $15/8$ (which can be written as $30/16$) to 2 (which can be written as $32/16$), so it becomes $31/16$.

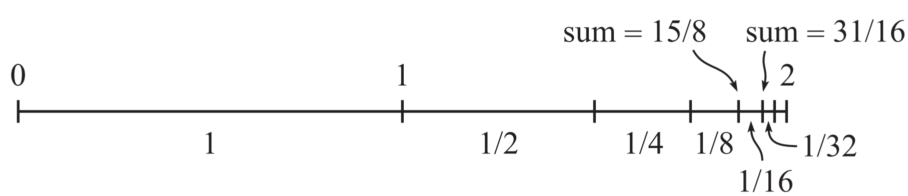

Figure 10.16: A geometric way of seeing why $1 + 1/2 + 1/4 + 1/8 + \cdots = 2$.

This "halfway" effect is clear if we write $1 + 1/2 + 1/4 + 1/8$ as $2 - 1/8$. This number is $1/8$ less than 2, so adding on $1/16$ cuts the $1/8$ shortfall in half. The shortfalls therefore always take the form of $1/2^n$. So when n gets very large, the shortfalls approach 0, which means that the sum approaches 2.

34. Let the second series have first term a and ratio r. Then the first series has first term $2a$ and ratio $r/2$. Equating the sums given by Eq. (10.46) yields

$$\frac{2a}{1-r/2} = \frac{a}{1-r} \implies 2-2r = 1-r/2 \implies 1 = 3r/2 \implies 2/3 = r. \tag{10.129}$$

So the second series has $r = 2/3$, and the first series has $r/2 = 1/3$. Note that a cancels in the above equation, so it doesn't matter what its value is. If we let $a = 1$ for simplicity, then the two series are

$$2 + \frac{2}{3} + \frac{2}{9} + \frac{2}{27} + \cdots \qquad \text{and} \qquad 1 + \frac{2}{3} + \frac{4}{9} + \frac{8}{27} + \cdots . \tag{10.130}$$

The first series starts off larger, but its terms decrease faster. These two effects cancel, and the sums are equal. From Eq. (10.46), the two sums are $2/(1-1/3) = 3$ and $1/(1-2/3) = 3$, which are correctly equal.

35. In the suggested grouping, note the following fact: If the denominator of a fraction is increased, then the fraction decreases. This means that if we replace the $1/3$ in the given sum by $1/4$, we decrease the sum. So the given sum is *larger* than the sum with $1/3$ replaced by $1/4$. Likewise, the given sum is larger than the sum with each of $1/5$, $1/6$, and $1/7$ replaced by $1/8$. The given sum is therefore larger than

$$1 + \frac{1}{2} + \left(\frac{1}{4} + \frac{1}{4}\right) + \left(\frac{1}{8} + \frac{1}{8} + \frac{1}{8} + \frac{1}{8}\right) + \cdots = 1 + \frac{1}{2} + \frac{1}{2} + \frac{1}{2} + \cdots = \infty. \tag{10.131}$$

This sum is infinite because the $1/2$'s go on forever. There is an infinite number of sets of parentheses, each with a value of $1/2$. And infinity times $1/2$ equals infinity. Therefore, the original sum, which is larger than the sum in Eq. (10.131), is also infinite.

There is nothing special about powers of 2 in the above strategy. Grouping the terms in Eq. (10.54) according to powers of 3 (that is, $1/2$ and $1/3$, then $1/4$ through $1/9$, then $1/10$ through $1/27$, etc.) works just as well, as you can verify. Likewise for powers of any other integer.

36. In Exercise 10.35, we made the denominators of some of the fractions *larger*, which produced a sum that was *smaller*. The original sum was therefore *larger* than the new sum, which we showed was infinite. The original sum was therefore also infinite.

In the present case, we'll make the denominators of some of the fractions *smaller*, which will produce a sum that is *larger*. The original sum is therefore *smaller* than the new sum. So if we can show that the new sum is finite, then the original sum must also be finite.

Here is the grouping (note that the groups *start* with powers of 2, instead of *ending* with them as they did in Exercise 10.35):

$$1 + \left(\frac{1}{2^2} + \frac{1}{3^2}\right) + \left(\frac{1}{4^2} + \frac{1}{5^2} + \frac{1}{6^2} + \frac{1}{7^2}\right) + \cdots \tag{10.132}$$

And here is the new sum with smaller denominators:

$$\begin{aligned}
&1+\left(\frac{1}{2^2}+\frac{1}{2^2}\right)+\left(\frac{1}{4^2}+\frac{1}{4^2}+\frac{1}{4^2}+\frac{1}{4^2}\right)+\cdots\\
&=1+2\cdot\frac{1}{2^2}+4\cdot\frac{1}{4^2}+8\cdot\frac{1}{8^2}+\cdots\\
&=1+\frac{1}{2}+\frac{1}{4}+\frac{1}{8}+\cdots=\frac{1}{1-1/2}=2, \qquad (10.133)
\end{aligned}$$

where we have used Eq. (10.44). Since this result of 2 is finite, and since the given sum in Eq. (10.132) is smaller than the sum in Eq. (10.133), the given sum is also finite (less than 2), as we wanted to show.

Remarks:

1. Since the given sum is finite, it must take on a definite value. Is there a way to determine this value? Yes, but it involves some higher-level math. So you'll just have to accept here that the sum in Eq. (10.56) has the rather fascinating value of $\pi^2/6 \approx 1.64$ (which is indeed less than 2). It is by no means obvious that the sum of the reciprocals of the squares of the positive integers should involve π, which is a number that is generally associated with circles. But it does!

   If you apply the (slightly shifted) grouping in Exercise 10.35 (where the groups end with powers of 2) to the present sum, you can show that this sum is *larger* than $3/2$. Its value must therefore lie somewhere between $3/2$ and 2. The true value of $\pi^2/6 \approx 1.64$ does indeed lie between these bounds.

2. In Exercise 10.35 we showed that $\sum_1^\infty 1/k$ is infinite. And in the present exercise we showed that $\sum_1^\infty 1/k^2$ is finite. What about powers of k that are between 1 and 2? For example, is $\sum_1^\infty 1/k^{1.5}$ finite or infinite?

   The strategy we used in Eq. (10.133) works just as well when the power is 1.5 instead of 2. We'll again decrease the denominators to obtain a larger new sum, so we again know that the given sum is smaller than the new sum, which is

$$\begin{aligned}
&1+\left(\frac{1}{2^{1.5}}+\frac{1}{2^{1.5}}\right)+\left(\frac{1}{4^{1.5}}+\frac{1}{4^{1.5}}+\frac{1}{4^{1.5}}+\frac{1}{4^{1.5}}\right)+\cdots\\
&=1+2\cdot\frac{1}{2^{1.5}}+4\cdot\frac{1}{4^{1.5}}+8\cdot\frac{1}{8^{1.5}}+\cdots\\
&=1+\frac{1}{2^{0.5}}+\frac{1}{4^{0.5}}+\frac{1}{8^{0.5}}+\cdots\\
&=1+\frac{1}{2^{0.5}}+\frac{1}{(2^{0.5})^2}+\frac{1}{(2^{0.5})^3}+\cdots\\
&=\frac{1}{1-1/2^{0.5}}=\frac{1}{1-0.707}=3.41. \qquad (10.134)
\end{aligned}$$

   The given sum is therefore less than 3.41, which means that it is finite.

   More generally, if we let the power be $1+m$ instead of 1.5, you can quickly show that the $1/(1-1/2^{0.5})$ result in Eq. (10.134) becomes $1/(1-1/2^m)$. For example, if the power is 1.1 (that is, $m = 0.1$), we obtain $1/(1 - 1/2^{0.1}) = 14.9$. And since we again know that the given sum is less than this new sum of 14.9, the given sum is finite. Likewise, if the power is 1.01 (that is, $m = 0.01$), we obtain $1/(1 - 1/2^{0.01}) = 144.8$. And since the given sum is less than this, it is finite. No matter how small m is, as long as it is positive, the $1/(1 - 1/2^m)$ result will be a finite positive number, which means that the given sum is finite. This $1/(1 - 1/2^m)$ result holds for larger m too. For example, if the power is 5, then $m = 4$, so the sum is less than $1/(1 - 1/2^4) = 1.067$.

However, the $1/(1 - 1/2^m)$ result isn't valid for $1 + m$ powers that are less than or equal to 1, because then m is less than or equal to zero. This means that $1/2^m$ is greater than or equal to 1, which in turn means that $1/(1 - 1/2^m)$ is negative (or undefined if $m = 0$). This is clearly an incorrect value of the sum, since the sum is positive. The reason for this incorrect answer is that the series in the fourth line of Eq. (10.134) is now an increasing geometric series (with $r \geq 1$), which makes Eq. (10.44) invalid.

The answer to the question we posed at the beginning of this remark is therefore: The cutoff between an infinite sum and a finite sum occurs when the power equals 1. That is, $\sum_1^\infty 1/k$ is infinite, but $\sum_1^\infty 1/k^p$ is finite for any $p > 1$ (even, say, $p = 1.000001$). Note that for $p < 1$, the denominators are smaller than those in Eq. (10.54), which means the fractions are larger. So the sum is certainly also infinite when $p < 1$. ♣

37. The straight line AB in Fig. 10.9 connects the upper-right corners of all the squares. We'll show below that this line is indeed straight. But accepting this fact for the moment, we can use similar triangles to find the desired sum $S = 1 + r + r^2 + r^3 + \cdots$.

Note that the partial square drawn on the left side of Fig. 10.9 has side length $1/r$. This is true because we can extend the squares in either direction, yielding a geometric sequence of side lengths of the form, $\dots, 1/r^2, 1/r, 1, r, r^2, \dots$.

The overall right triangle ABC is similar to all of the little right triangles on top of the squares, because their hypotenuses all make the same angle with the horizontal (assuming correctly that AB is a straight line). We can therefore equate the ratios of corresponding sides in the similar triangles. Consider the overall triangle ABC, along with the second little triangle from the left (the one with height $1 - r$), although any of the other little triangles would work just as well. Equating the ratios of the long leg to the short leg in these two similar triangles gives:

$$\frac{\text{long leg}}{\text{short leg}}: \quad \frac{CB}{AC} = \frac{r}{1-r} \implies \frac{S}{1/r} = \frac{r}{1-r} \implies Sr = \frac{r}{1-r} \implies S = \frac{1}{1-r}, \tag{10.135}$$

in agreement with Eq. (10.44), as desired.

Let's now show that AB is in fact a straight line. This is true because the slopes of all the little right triangles on top of the squares are the same. The leftmost triangle has a slope of rise/run $= -(1/r - 1)/1$, which can be rewritten as $-(1 - r)/r$. The second triangle also has a rise-over-run slope of $-(1 - r)/r$. The third triangle's slope is $-(r - r^2)/r^2$, which also equals $-(1 - r)/r$. In general, the slopes all take the form of $-(r^n - r^{n+1})/r^{n+1}$, which reduces to $-(1-r)/r$. Therefore, since the slopes are all the same, the path AB is indeed a straight line.

This result is similar to the one in Exercise 10.30, but in reverse. In that exercise, we were given the straight lines that formed the angle, and then we showed that the sides of the squares formed a geometric sequence. In the present exercise, we were given the geometric sequence of side lengths, and then we showed that AB is a straight line.

In solving this exercise, it is tempting to just assume that AB is a straight line, since that's how it's drawn. But technically this needs to be demonstrated. If we had drawn other shapes instead of squares (for example, rectangles with a common width, and heights that form a geometric sequence), then the path connecting the upper-right corners would *not* be straight. You can verify this with a drawing, and/or by calculating the slopes of the individual hypotenuses.

38. (a) The areas of the squares are 1^2, $(1/2)^2$, $(1/4)^2$, etc. Adding these up yields a geometric series with a first term of 1 and a ratio of $1/4$. So Eq. (10.44) gives the total area of the squares as

$$1+\frac{1}{4}+\frac{1}{16}+\frac{1}{64}+\cdots=\frac{1}{1-1/4}=\frac{1}{3/4}=\frac{4}{3}. \tag{10.136}$$

As in the preceding exercise, the line AB is in fact straight.

(b) The overall right triangle ABC is a 45-45-90 triangle (see Section A.7 in Appendix A) with leg length 2. The 45° angles follow from the fact that triangle ABC is similar to triangle EFG, and the latter is a 45-45-90 triangle because its two legs are both equal to $1/2$ (since $EG = 1 - 1/2 = 1/2$). The leg lengths of 2 for triangle ABC follow from the fact that $AD = 1$ (since AED is a 45-45-90 triangle and $DE = 1$). Alternatively, you can add up the sides of all the squares along the bottom edge. This gives $1 + 1/2 + 1/4 + \cdots$, which equals $1/(1 - 1/2) = 2$ from Eq. (10.44).

The area of the overall right triangle ABC is therefore $bh/2 = (2)(2)/2 = 2$. So the first question mark in Eq. (10.57) equals 2, because A_{sq} and A_{tri} add up to the area of the overall triangle.

The second question mark in Eq. (10.57) also equals 2, since each square has twice the area of the sub-triangle above it, because cutting a square along a diagonal yields two copies of the triangle above it.

The two equations in Eq. (10.57) are therefore

$$A_{\text{sq}} + A_{\text{tri}} = 2 \qquad \text{and} \qquad A_{\text{sq}} = 2A_{\text{tri}}. \tag{10.137}$$

Substituting the second equation into the first gives $2A_{\text{tri}} + A_{\text{tri}} = 2 \implies A_{\text{tri}} = 2/3$. We need to double this to obtain A_{sq}, so $A_{\text{sq}} = 4/3$, in agreement with the result in part (a).

39. (a) In the square in Fig. 10.11(a), the largest shaded square has $1/4$ of the area of the overall square. The next shaded square is $1/4$ as large as the first one, so it has $1/4^2$ of the area of the overall square. And so on, with the shaded fractions equalling $1/4^3$,

$1/4^4$, etc. The series in Eq. (10.58) therefore represents the fraction of the overall area that is shaded. But this fraction is simply $1/3$, because for every size of the shaded squares, there are two identical unshaded squares (above, and to the right). This means that the unshaded area is twice the shaded area. The unshaded and shaded fractions must therefore be $2/3$ and $1/3$, respectively, because these are the two numbers that add up to 1 and have a ratio of 2. (The two fractions must add up to 1, since every part of the area is either unshaded or shaded.) To derive this formally: If x is the shaded fraction, then $2x + x = 1 \implies 3x = 1 \implies x = 1/3$.

The same reasoning applies to the triangle in Fig. 10.11(a). The largest shaded triangle has $1/4$ of the area of the overall triangle. (This is true because the overall triangle can be built up from four sub-triangles – the "upside down" largest shaded triangle in the middle, plus three other "right side up" triangles of that size.) The next shaded triangle is $1/4$ as large as the first one (by the same reasoning), so it has $1/4^2$ of the area of the overall triangle. And so on, with the shaded fractions equalling $1/4^3$, $1/4^4$, etc. And just as with the squares, there are two identical unshaded triangles (to the right, and to the left) for every size of the shaded triangles. So the shaded fraction of the overall triangle is again $1/3$.

(b) The reasoning in part (a) isn't restricted to nice geometrical shapes like squares and triangles. The "boring" rectangle in Fig. 10.11(b) works just as well. Or actually even better, because it works for a general integer N. We've chosen $N = 5$ for concreteness in the figure, but the following reasoning works for any integer N.

The overall rectangle in Fig. 10.11(b) is divided into five sub-rectangles (let's call these "Level 1" rectangles). So the dark-shaded rectangle on the left has $1/5$ of the area of the overall rectangle. The second-to-left Level 1 rectangle is further subdivided into five Level 2 rectangles, and the leftmost one of these is shaded with a medium shade. (We've chosen different shadings so that you can tell the rectangles apart and they don't all blend together. We didn't need to worry about this with the squares and triangles in part (a).) This shaded Level 2 rectangle has $1/5$ of $1/5$ (which is $1/5^2$) of the overall area.

Continuing on, the second-to-left Level 2 rectangle is further subdivided into five Level 3 rectangles, and the leftmost one of these is shaded with a light shade. This shaded Level 3 rectangle has $1/5$ of $1/5^2$ (which is $1/5^3$) of the overall area. And so on. The fraction of the overall area that is eventually shaded is $1/5+1/5^2+1/5^3+\cdots$.

From the figure, we see that for every shaded rectangle at a given level, there are *three* identical unshaded rectangles at that same level (the three rightmost ones). This means that the unshaded area is 3 times the shaded area. The unshaded and shaded fractions must therefore be $3/4$ and $1/4$, respectively. (Formally, if x is the shaded fraction, then $3x + x = 1 \implies 4x = 1 \implies x = 1/4$.) So the sum of the $1/5+1/5^2+1/5^3+\cdots$ series is $1/4$.

For a general N, the above 5 gets replaced with N, and the 4 gets replaced with $N-1$. So the sum of the series in Eq. (10.59) is $1/(N-1)$, as desired. In short, if there are N rectangles at each level, then one of them is shaded, another one is further subdivided, and $N-2$ of them are unshaded. The unshaded area (at a given level) is therefore $N-2$ times the shaded area. So the unshaded fraction is $(N-2)/(N-1)$, and the shaded fraction is $1/(N-1)$, because these are the two numbers that add up to 1 and have a ratio of $N-2$. (Formally, if x is the shaded fraction, then $(N-2)x + x = 1 \implies (N-1)x = 1 \implies x = 1/(N-1)$.)

40. (a) The setup is shown in Fig. 10.17. The cyclists starts at the dots labeled 0, which are a distance d apart. The bird starts at the left cyclist. When it reaches the right cyclist, the two cyclists' positions are labeled 1.

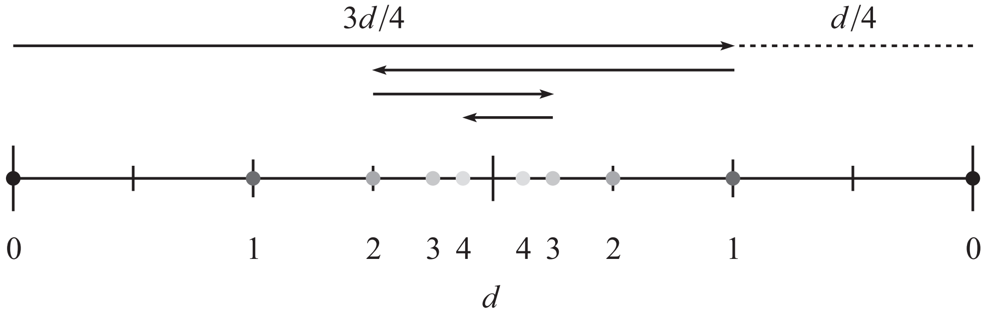

Figure 10.17: Successive positions of the cyclists, along with the distances the bird travels.

As indicated in the figure, the bird travels a distance of $3d/4$ during the first stage, because it travels 3 times as far as the right cyclist (since their speeds are 30 and 10), which means that we need to find two distances that are in the ratio of 3 to 1, and that add up to d. These distances are quickly seen to be $3d/4$ and $d/4$. Or you can just solve the equation $3x + x = d \implies x = d/4$, which then gives the bird's distance as $3x = 3d/4$.

From Fig. 10.17, we see that the two dots labeled 1 are a distance $d - d/4 - d/4 = d/2$ apart, because the cyclists have each moved $d/4$ closer to the center. The second stage of the bird's flight is therefore just a repeat of the first stage, with the only change being that the initial separation is $d/2$ instead of d (and with "right" replaced with "left," but that doesn't matter for the distances). From the same reasoning as above (since the speeds are still 30 and 10, so they are still in a 3-to-1 ratio), the bird will travel $3/4$ of the $d/2$ distance to the left cyclist, which means they will meet at the (left) dot labeled 2. So during the second stage the bird travels a distance of $(3/4)(d/2)$.

At the positions labeled 2, the cyclists are now $d/4$ (half of $d/2$) apart. So the same reasoning works again with an initial separation of $d/4$. The bird therefore travels a distance of $(3/4)(d/4)$ in the third stage. And then $(3/4)(d/8)$ in the fourth stage. And so on. The cyclists' separation at the start of each stage is successively cut in half, so the distance the bird travels in each stage (which is always $3/4$ of the initial separation) is likewise cut in half.

There is technically an infinite number of dots in Fig. 10.17, but we have drawn them only up to the end of the fourth stage. (Eventually, of course, the distances become smaller than the bird itself. And also the bird needs some time to turn around. But we won't worry abut these things in this idealized math problem!) We have shaded the pairs of dots differently, to make it easier to tell which dot goes with which.

We have shown that the stage distances form a geometric sequence with first term $3d/4$ and ratio $1/2$. So Eq. (10.46) gives the sum of the geometric series as

$$\frac{3d}{4} + \frac{3d}{4}\cdot\left(\frac{1}{2}\right) + \frac{3d}{4}\cdot\left(\frac{1}{2}\right)^2 + \frac{3d}{4}\cdot\left(\frac{1}{2}\right)^3 + \cdots = \frac{3d/4}{1-1/2} = \frac{3d}{2}. \tag{10.138}$$

REMARK: Aside from the given distance d, the only other thing that matters in the above solution is the ratio of the two speeds. If the speeds were 60 and 20, instead of 30 and 10, we would end up with the same answer of $3d/2$. The entire cycling/flying process would take half as much time, but all of the distances would be the same.

What is the answer to the problem (that is, what is the total distance the bird travels) if the ratio of the speeds takes the general value of n? (So we solved the $n = 3$ case above.) You should stop reading here if you want to solve this extension on your own.

The setup involving a general ratio n of speeds is shown in Fig. 10.18, where we have drawn only the first stage. The bird travels a distance of $nd/(n+1)$ during the first stage, because it travels n times as far as the right cyclist, and we quickly see that $nd/(n+1)$ and $d/(n+1)$ are the two distances that have a ratio of n and add up to d. Or you can just solve the equation $nx + x = d \implies x = d/(n+1)$, which then gives the bird's distance as $nx = nd/(n+1)$.

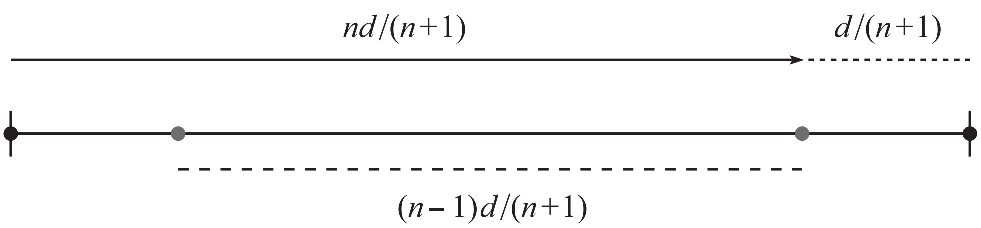

Figure 10.18: Redrawing Fig. 10.17 for a general ratio n of speeds, with only the first stage shown.

From Fig. 10.18, we see that at the end of the first stage, the cyclists are a distance $d - d/(n+1) - d/(n+1) = (n-1)d/(n+1)$ apart, because they have each moved $d/(n+1)$ closer to the center. The second stage of the bird's flight is therefore just a repeat of the first stage, with the only change being that the initial separation is $(n-1)d/(n+1)$ instead of d. The initial separation is therefore decreased by the factor $(n-1)/(n+1)$, which means that the bird's flight distance is also. This applies to every successive stage, so the bird's distances form a geometric sequence with first term $nd/(n+1)$ and ratio $(n-1)/(n+1)$. From Eq. (10.46) the total distance the bird travels is then

$$\frac{\dfrac{nd}{n+1}}{1-\dfrac{n-1}{n+1}} = \frac{\dfrac{nd}{n+1}}{\dfrac{2}{n+1}} = \frac{nd}{2}. \tag{10.139}$$

This agrees with the result in Eq. (10.138) when $n = 3$. ♣

(b) Each cyclist travels a distance $d/2$ by the time they meet in the middle. Since the bird travels 3 times as fast (30 mph vs. 10 mph), and since it travels for the same amount of time (this fact is important), it must travel 3 times as far. Its total distance is therefore $3 \cdot d/2$, as we found above.

This solution is *much* quicker than the one in part (a)! And in the more general case where the ratio of speeds is n, this method quickly reproduces the result of $n \cdot d/2$ in Eq. (10.139).

REMARKS:

1. There is a famous story about someone giving this problem, or a slight variation, to John von Neumann, who was a brilliant mathematician known for (among many other talents) the ability to do complex calculations in his head. He solved this problem very quickly, leading the other person to remark (these aren't exact quotes, but they capture the idea), "Oh, so you know the trick!" He responded, "What trick? I just summed the infinite geometric series!"
2. As a bonus, not only is this second solution quicker, it is more general. The solution in part (a) required that everything happen along a straight line. But what if the bike path is curved (as it is on the cover of this book!), and what if the bird flies in random zig-zaggy directions and even takes a detour over to a picnic to have some cake? As long as the bird always travels at n times the speed of the cyclists, the second solution shows that by the time the cyclists have traveled a distance $d/2$, the bird has traveled a distance $n \cdot d/2$. It can even end up in a different location from the cyclists.
3. The second solution also works in the case where the cyclists have different speeds. (The first solution is rather messy in this case.) Let's say the cyclists' speeds are 10 and 20 (in mph), and the bird's speed is 40. Then the cyclists close the initial d separation at a relative speed of $10 + 20 = 30$. So the time it takes to meet (that is, to decrease the separation to zero) is $d/30$. The bird travels at a speed of 40 for this time, so it travels a distance of $40 \cdot d/30 = 4d/3$.
   In general, if the cyclists' speeds are c_1 and c_2, and if the bird's speed is v, then the time is $d/(c_1 + c_2)$, so the bird travels a distance of $v \cdot d/(c_1 + c_2) = d \cdot v/(c_1 + c_2)$. For the original problem, this expression correctly gives $d \cdot 30/(10 + 10) = 3d/2$. ♣

41. (a) Since A and B lie on the $y = x^2$ parabola, their heights (y coordinates) are $(x - a)^2$ and $(x + a)^2$. Expanding the squares, and then adding them and dividing by 2, we find that the average of the two heights is $x^2 + a^2$. This is the height of D.

The length of CD equals the difference in the heights of D and C. And since the height of C is simply x^2 (because it lies on the parabola), the length of CD is $(x^2 + a^2) - x^2 = a^2$.

If you tilt your head sideways, triangle CDA has base $CD = a^2$. And its height is a, because the distance from A to the base CD is the difference in the x coordinates of A and C (or D), which equals a. The $bh/2$ area of triangle CDA is therefore

$a^2 \cdot a/2 = a^3/2$. The same reasoning applies to triangle CDB, so its area is also $a^3/2$. The total area of the dark-shaded triangle ABC is therefore $2(a^3/2) = a^3$, as we wanted to show.

REMARKS:

1. Remember that this a^3 result holds only if C is halfway between A and B (with regard to the x coordinates). However, the points A and B can be anywhere. They can even lie on the same side of the origin (both x values positive, or both negative).
2. It is interesting that no matter what the value of x is, as long as the x coordinates of A and B are $x - a$ and $x + a$, the length of CD is always a^2 (and so the area of triangle ABC is always a^3). If x is large (and a isn't too large), so that all three points A, B, C are far off to the right (and high up) on the parabola, then triangle ABC will be very tall and thin. So it is by no means obvious that the length of CD is still a^2. But this is in fact always the case.
   In the special case where $x = 0$, it is clear that CD has length a^2, because A, D, and B all lie on a horizontal line, and the y coordinate of points A and B on the parabola is $(\pm a)^2 = a^2$.
3. If you want some practice with algebra, you can derive the above a^3 result in the following lengthier manner. Drop vertical lines from A and B down to the x-axis, and consider the trapezoid formed by these two lines, along with line AB and the x-axis. And then form the analogous trapezoids associated with the lines AC and CB. The area of triangle ABC equals the area of the AB trapezoid minus the areas of the AC and CB trapezoids. You can show that this difference leads to an area of a^3. (The area of each of these trapezoids equals their base times the average of their side heights.) ♣

(b) The reasoning in part (a) can be repeated with triangle AEC. The only difference is that the a in part (a) is now $a/2$, because the horizontal span of triangle AEC is a, instead of the $2a$ for triangle ABC. Note that although the x value of E (which is $x - a/2$) is different from the x value of C (which is x), this is irrelevant because the a^3 answer in part (a) is independent of x.

The area of triangle AEC is therefore obtained by replacing a with $a/2$ in the above a^3 result for the area. So the area of triangle AEC is $(a/2)^3 = a^3/8$. The same result applies to triangle CFB, so its area is also $a^3/8$. The total area of the two light-shaded triangles is therefore $2(a^3/8) = a^3/4$.

Similarly, at the next stage we have four triangles, where the a in the a^3 result for the area is replaced with $a/4$. So their total area is $4(a/4)^3 = a^3/16$. And then in the next stage we have eight triangles with a total area of $8(a/8)^3 = a^3/64$. And so on. The total area of all of the triangles in all of the stages (that is, the area of the parabolic section) is therefore

$$a^3 + \frac{a^3}{4} + \frac{a^3}{16} + \frac{a^3}{64} + \cdots = a^3\left(1 + \frac{1}{4} + \frac{1}{4^2} + \frac{1}{4^3} + \cdots\right). \tag{10.140}$$

From Eq. (10.44) the geometric series in the parentheses has a sum of $1/(1-1/4) = 4/3$. The area of the parabolic section is therefore $4/3$ times the a^3 area of triangle ABC, as we wanted to show.

42. Multiplying the series in Eq. (10.60) by $1-r$ gives

$$\begin{aligned}(1-r)&(1+2r+3r^2+4r^3+5r^4+\cdots)\\ &= (1+2r+3r^2+4r^3+5r^4+\cdots) - (r+2r^2+3r^3+4r^4+5r^5+\cdots)\\ &= 1+r+r^2+r^3+r^4+\cdots\,. \end{aligned} \tag{10.141}$$

But we know that the sum of this geometric series is $1/(1-r)$. So if S is the desired sum in Eq. (10.60), we see that $(1-r)S = 1/(1-r)$. Therefore, $S = 1/(1-r)^2$, as we wanted to show.

43. Multiplying the given sum by $1-r$ yields

$$\begin{aligned}(1-r)&(1+3r+6r^2+10r^3+15r^4+\cdots)\\ &= (1+3r+6r^2+10r^3+15r^4+\cdots) - (r+3r^2+6r^3+10r^4+15r^5+\cdots)\\ &= 1+2r+3r^2+4r^3+5r^4+\cdots\,. \end{aligned} \tag{10.142}$$

From Eq. (10.63) (or the preceding exercise), we know that the sum of this series is $1/(1-r)^2$. So if S is the desired sum, we see that $(1-r)S = 1/(1-r)^2$. Therefore, $S = 1/(1-r)^3$.

If you continue in this manner, you can show that the analogous series whose sum equals $1/(1-r)^4$ is $1+4r+10r^2+20r^3+35r^4+\cdots$. You can think about why the various coefficients we've encountered in this exercise and the preceding one form diagonal strings of numbers in Pascal's triangle in Table 3.2.

44. From Eq. (10.46), the sum of an infinite series starting with a is $a/(1-r)$, and the sum of an infinite series starting with ar^{n+1} is $ar^{n+1}/(1-r)$. (The sum of an infinite series is always the first term divided by $1-r$.) As observed in the statement of this exercise, the difference of these two series is the finite series on the lefthand side of Eq. (10.67). Its sum is therefore the difference of the two sums we just found, which gives

$$\frac{a}{1-r} - \frac{ar^{n+1}}{1-r} = \frac{a-ar^{n+1}}{1-r}\,, \tag{10.143}$$

in agreement with Eq. (10.67).

45. With $a = r^n$ and a ratio of $1/r$, Eq. (10.67) gives the sum as

$$\frac{r^n - r^n(1/r)^{n+1}}{1-1/r} = \frac{r^n - 1/r}{1-1/r}\cdot\frac{r}{r} = \frac{r^{n+1}-1}{r-1} = \frac{1-r^{n+1}}{1-r}\,, \tag{10.144}$$

as desired.

46. Eq. (10.66) gives $1 + r + r^2 = (1 - r^3)/(1 - r)$. So the first expression for S equals

$$S = (1 + r^3)\frac{1 - r^3}{1 - r} = \frac{1 - (r^3)^2}{1 - r} = \frac{1 - r^6}{1 - r}, \tag{10.145}$$

in agreement with Eq. (10.65). We used the difference-of-squares result from Eq. (3.23) in the numerator here.

With r replaced by r^2, Eq. (10.66) gives $1+r^2+r^4 = \big(1-(r^2)^3\big)/(1-r^2) = (1-r^6)/(1-r^2)$. So the second expression for S equals

$$S = (1 + r)\frac{1 - r^6}{1 - r^2} = (1 + r)\frac{1 - r^6}{(1 + r)(1 - r)} = \frac{1 - r^6}{1 - r}, \tag{10.146}$$

as desired. You can also think of the binomials in Eq. (10.70) as (short) geometric series, and apply Eq. (10.66) to those too. You will again obtain the result in Eq. (10.65).

47. (a) Multiplying the $1 + r + r^2 + r^3$ factor by the 1 in the second factor, and then by the r^4, and then by the r^8, and so on, gives

$$(1 + r + r^2 + r^3) + (r^4 + r^5 + r^6 + r^7) + (r^8 + r^9 + r^{10} + r^{11}) + \cdots . \tag{10.147}$$

This is simply an infinite geometric series with first term 1 and ratio r, so Eq. (10.44) gives the sum as $1/(1 - r)$.

(b) Applying Eq. (10.66) to the finite series in the first factor gives a sum of $(1-r^4)/(1-r)$. Applying Eq. (10.44) to the infinite series in the second factor (with a ratio of r^4) gives a sum of $1/(1 - r^4)$. The product of these two sums is

$$\frac{1 - r^4}{1 - r} \cdot \frac{1}{1 - r^4} = \frac{1}{1 - r}, \tag{10.148}$$

as desired. You can check that the above method also works more generally for this product:

$$(1 + r + r^2 + \cdots + r^{n-1})(1 + r^n + r^{2n} + r^{3n} + \cdots). \tag{10.149}$$

48. In place of Eq. (10.61) (which contained an infinite number of rows, each of which was an infinite geometric series) we now have just n rows, each of which is a finite geometric series:

$$\begin{aligned}1 + 2r + 3r^2 + 4r^3 + \cdots + nr^{n-1} = 1 + r + r^2 + r^3 + \cdots + r^{n-2} + r^{n-1}& \\ + r + r^2 + r^3 + \cdots + r^{n-2} + r^{n-1}& \\ + r^2 + r^3 + \cdots + r^{n-2} + r^{n-1}& \\ \vdots\qquad& \\ + r^{n-2} + r^{n-1}& \\ + r^{n-1}&\end{aligned} \tag{10.150}$$

Applying Eq. (10.67) to each row, we obtain (instead of Eq. (10.62)):

$$\frac{1-r^n}{1-r} + \frac{r-r^n}{1-r} + \frac{r^2-r^n}{1-r} + \cdots + \frac{r^{n-1}-r^n}{1-r}\,. \tag{10.151}$$

The first of the two terms in the numerators form a finite geometric series, so Eq. (10.66) (with n replaced by $n-1$) gives a sum of $(1-r^n)/(1-r)$. The second of the two terms in the numerators are all r^n's, so those simply add. Factoring out the overall $1/(1-r)$, we see that the sum equals

$$\frac{1}{1-r}\left(\frac{1-r^n}{1-r} - nr^n\right) = \frac{1-r^n-nr^n(1-r)}{(1-r)^2} = \frac{1-(n+1)r^n+nr^{n+1}}{(1-r)^2}\,. \tag{10.152}$$

This isn't the prettiest answer, but you can check that it works for a few values of r and n. For example, if $r=2$ and $n=4$, the formula says that

$$\begin{aligned} 1+2\cdot 2+3\cdot 2^2+4\cdot 2^3 &= \frac{1-5\cdot 2^4+4\cdot 2^5}{1} \\ \Longrightarrow \quad 1+4+12+32 &= 1-80+128, \end{aligned} \tag{10.153}$$

which is indeed true, since both sides equal 49. In the case where $r=2$ and we keep n general, the series is $1+2\cdot 2+3\cdot 2^2+4\cdot 2^3+\cdots+n\cdot 2^{n-1}$, and the result in Eq. (10.152) becomes

$$\frac{1-(n+1)2^n+n\cdot 2^{n+1}}{1} = 1-(n+1)2^n+2n\cdot 2^n = (n-1)2^n+1, \tag{10.154}$$

which is a fairly clean result. When $n=4$, it correctly equals $3\cdot 2^4+1=49$.