20. The graduate studies committee has asked the graduate students to provide the faculty with a list of three nominees to represent the students on the committee. After much discussion, three nominees are put forward, and you are asked to rank them, with the rank representing preference (i.e., 1 is most preferred, 2 is second best, and 3 is the third-best choice). The nominees are Beta, a seventh-year student who recently defended his prospectus; Gamma, a second-year student who is very bright but tends to dominate seminar discussion; and Alpha, a fourth-year year student who is preparing for her exams and is widely viewed as level-headed and realistic. Provide your pairwise preference rankings of each candidate. Check to see whether your rankings are transitive. If you have been assigned this problem for class, bring your ordering to class so that the class can determine whether it is transitive at the aggregate level under pairwise majority rule.

21. Recall the first question of Section 8.2 in Chapter 1. There we asked you to pick ideal spending points for three parties, as well as a status quo and a bill, and conjecture about whether or not it would pass. Now we want you to go further. Write a utility function for each party that is largest at that party's ideal point. How does that function decrease with distance from the ideal point? Try to draw a curve around each ideal point that gives the same utility to the corresponding party for every point on the curve. These are called *indifferences curves*, as the party is indifferent between all points on the curve. Draw these for all parties and see whether you can answer your earlier conjecture.

22. Propose and justify a quadratic utility function as representing the preferences of some political actor over something.

Chapter Four

Limits and Continuity, Sequences and Series, and More on Sets

We have now covered most of the building blocks we need to introduce the more complex topics in the remainder of the book, but there remain a few more important ideas to address. Specifically, we must describe properties of functions and sets related to *limit behavior*, the behavior of these mathematical constructs as we approach some point in or out of their domains. Those readers with a stronger math background will likely have seen many of these ideas before, though they may have not seen them presented in this fashion, and so could benefit from skimming this chapter. We expect many readers will not have spent a great deal of time on these topics prior to this course, and thus will benefit from a more thorough reading.

The first section discusses sequences and series, which are necessary background for the rest, along with being useful in their own right in political science. The second introduces the notion of a limit. The third uses both limits and sequences to explore some additional properties of sets necessary for understanding calculus. Finally, the fourth treats continuity, a fundamental property of functions that affects our ability to maximize them, and thus find optimal choices.

4.1 SEQUENCES AND SERIES

4.1.1 Sequences

A **sequence** is an ordered list of numbers. For example, $\{1, 2, 3, 4\ldots\}$ is a sequence. So is $\{\frac{3}{10}, \frac{3}{100}, \frac{3}{1{,}000}, \frac{3}{10{,}000}, \ldots\}$. Sequences can be infinite (i.e., continue forever), as depicted above, or they can be finite: $\{2, 1, 3, 1, 4, 1, 5, 1\}$. All sequences are countable, in that one can assign a natural number to each element in the sequence. For this reason we typically identify an element in the sequence with the notation x_i, where i is an index corresponding to its place in the sequence. Sequences usually start at index 0 or 1. To avoid confusion on this point, we can write the whole sequence in shorthand as $\{x_i\}_{i=1}^{N}$. If N is finite, it is a finite sequence; if $N = \infty$, it is an infinite one.

Sequences are rarely just random numbers placed in a row, and usually are generated from some underlying process, particularly when the sequences are infinite. The elements in the first example above can be represented by $x_i = i$, and the elements in the second example by $x_i = \frac{3}{10^i}$. We can write these sequences as $\{i\}_{i=1}^{\infty}$ and $\{\frac{3}{10^i}\}_{i=1}^{\infty}$.

Subsequences are just parts of larger sequences. The sequence $\{2, 4, 6, \ldots\}$ is a **subsequence** of $\{1, 2, 3, 4 \ldots\}$, which one can represent as $x_i = 2i$.

4.1.2 Series

A **series** is the sum of a sequence. Put differently, a series is a sequence with addition operators between each of the elements (e.g., $\{1 + 2 + 3 + 4 + \ldots\}$ or $\{\frac{3}{10} + \frac{3}{100} + \frac{3}{1,000} + \frac{3}{10,000} + \ldots\}$. We can write this as $\sum_{i=1}^{N} x_i$, where $N = \infty$ for our two examples.

Finite series are easy to grasp because calculating the sum of a finite sequence is straightforward (we add the terms). One can also consider infinite series, and many times these are easy (e.g., the infinite sequence $\{1, 2, 3, 4, \ldots \infty\}$ sums to infinity). In other cases, though, they are more complex. We do not have a great need to sum many different series in political science, but there are some of importance, so it's worth discussing briefly how one can sum a series.

A common series that you will see time and again in game theory is $\sum_{t=0}^{\infty} \delta^t$, where $\delta < 1$. The story behind this series goes something like this. State A and state B are interacting in a repeated game (a game in which the actors have repeated chances to interact). In each period of the game, both states decide whether or not to cooperate (on defense policy, free-trade agreements, and the like) or to defect (by making private deals, raising trade barriers, and so on). After the states have decided simultaneously what to do, they receive some measure of utility, which we'll call their payoffs from the period. Since the game is repeated, this happens again and again, indefinitely. By definition, strategies must cover decisions by the states across all time. Choosing the optimal strategy requires that we know what the total payoff is across time, which means adding up all the individual periods' payoffs.

Before we can do so, we must ask ourselves how we should treat the payoffs for future periods. If they are as valuable to us as present payoffs are, then we'll just be adding something of roughly the same magnitude an infinite number of times, which would give an infinite total payoff. Is that reasonable? Game theorists generally answer no—future payoffs are worth less than present ones. Why? Well, the game might end at any time if, say, one state ceases to be or undergoes a radical shift in governance. If there is a chance that this might happen in any period, then we have to discount the payoffs for future periods that might not occur. Even if this is not the case, future payoffs may still be worth less. If I have \$5 now, I can put it in a bank (in theory) and receive some rate of return, getting back \$5(1+r) later. The \$5 I may get later is worth less than this \$5(1+r). Reversing things, compared to my \$5 now, the \$5 later is only worth $\$\frac{5}{1+r}$ or $\$5\delta$, where $\delta < 1$. This δ is known as a **discount rate** as it represents the rate at which the actors value future utility relative to present utility.

This can be repeated for every period onward, yielding the series $\sum_{t=0}^{\infty} \$5\delta^t$.

We can pull the \$5 out, so we just need to sum $\sum_{t=0}^{\infty} \delta^t$. It turns out this sum is $\frac{1}{1-\delta}$. To see this, note that $\sum_{t=0}^{\infty} \delta^t = \frac{1}{1-\delta}$ implies that $(\sum_{t=0}^{\infty} \delta^t)(1-\delta) = 1$ or $(1+\delta+\delta^2+\ldots)(1-\delta) = 1$. Multiplying out the LHS gives $1-\delta+\delta-\delta^2-\delta^2+\ldots$. All the infinite terms in δ cancel, yielding a sum of 1.

Sometimes scholars impose a finite number of rounds on an iterated (repeated) game instead, if they believe this best captures the structure of the political interactions they are studying. We can derive the sum of the finite sequence $\sum_{t=0}^{N} \delta^t$ from the infinite one by breaking up the infinite one into parts. First, note that $\sum_{t=0}^{\infty} \delta^t = \sum_{t=0}^{N} \delta^t + \sum_{t=N+1}^{\infty} \delta^t$. The second term on the RHS, since it goes on forever, is the same as $\delta^{N+1} \sum_{t=0}^{\infty} \delta^t$. Since $\sum_{t=0}^{\infty} \delta^t = \frac{1}{1-\delta}$, we can replace the infinite sums on both sides with $\frac{1}{1-\delta}$ to get $\frac{1}{1-\delta} = \sum_{t=0}^{N} \delta^t + \frac{\delta^{N+1}}{1-\delta}$. Isolating the finite sum yields our answer: $\sum_{t=0}^{N} \delta^t = \frac{1-\delta^{N+1}}{1-\delta}$.

4.1.3 Why Should I Care?

Some readers might be thinking, "This is pretty abstract. Did we really have to go through this?" In addition to providing some useful experience practicing manipulation of algebra and indices, this material will be useful for students in formal theory courses. You've already seen the utility of series in the extended example above; it is this math which underlies the importance of the "shadow of the future." However, sequences are also common even in games that are not repeated. Extensive form games, in which players alternate taking actions in some order, model sequential political interactions. These include bargaining scenarios and the legislative process of bill proposal, voting, and veto. In these cases the structure of the game has a fixed sequence. If you study game theory, you will be introduced to the terms *history*, *terminal history*, *subhistory*, and *proper subhistory*, and these are all defined in terms of mathematical sequences. Some histories are finite and others are infinite, but all histories express a sequence of events in the game.

For example, a game's history might consist of the House Speaker's choice of the open or closed rule, followed by the choice of a congressional committee to invest in specialized knowledge, the subsequent bill it proposes, and finally the floor of Congress's response to that proposal (Gilligan and Krehbiel, 1989). Or a game's history could consist of a series of alternating offers in an attempt to reach a bargain, with each player getting to make a new offer if she rejects the previous one. If an offer is accepted the game's history ends, but a rejection extends the history by an additional offer (and decision over it). This alternating-offer, infinite-horizon (because the game may never end, in theory) model of bargaining was proposed and solved by Rubinstein (1982) and has become an influential model of bargaining in political science (e.g., Baron and Ferejohn, 1989).[1] In Rubinstein's model, the player making the first offer seeks

[1] These games are also called "divide-the-dollar" games because they often focus on how to distribute one unit's worth of goods among the players.

to provide just enough of an incentive via the offer to make the other actor indifferent between taking the offer immediately and rejecting it in favor of making her own offer later. Taking the offer can be attractive because the discount factor is less than one; i.e., future periods' payoffs are not worth as much as the present period's, implying that waiting to accept an offer is costly for both players. In equilibrium, the first player's offer is just enough to take advantage of this cost, and the second player accepts the offer immediately, ending the game in the first period. Note that even though the game ends immediately, the potential future history of the game (i.e., the expected future sequence of actions) matters as it determines the payoff each player could expect to get should she reject the other player's offer.

Other games of note in political science have histories that are not truncated and so are truly infinite. The infinitely repeated prisoner's dilemma is an example of this. The prisoner's dilemma is an interaction in which mutual cooperation is beneficial to both players, but each player also has an incentive to deviate from this cooperation, defecting against his partner. When the game is played once—a one-shot game—mutual defection is the rule. The same is true for any finite history. But in an infinite history, the future gains to cooperative play and the lack of an end to the game prove sufficient to allow for mutual cooperation. Attention to future payoffs and their dependence on present actions is often called the "shadow of the future" in political science. Here sequences describe not only the equilibrium path of cooperation but also the punishment strategies each player uses to enforce cooperation. We provide an example of such a strategy below.

Game theory utilizes other sequences as well. For example, as you go on, you will encounter what is known as a *sequential equilibrium*. That concept is defined as the limit of a sequence.[2]

Last, one cannot study difference equation models without understanding sequences and series (and their limits), as the concept of a solution to such equations appeals to sequences and series (see Blalock, 1969, chap. 5, and Huckfeldt, Kohfeld, and Likens, 1982, chaps. 1–3).

4.2 LIMITS

A **limit** describes the behavior of a function, sequence, or series of numbers as it approaches a given value. That perhaps sounds rather strange and may

[2]Game theory has explicit mathematical foundations. In fact, even though game theory has most strongly influenced the field of economics, the scholars who created game theory were mathematicians. John Nash, who won a Nobel Prize in Economics for his foundational work in game theory, is a mathematician, not an economist, as were John von Neumann and Oskar Morgenstern, who are widely recognized as the founders of game theory. Though we refer only to a generic equilibrium in the text, there are actually several different (though closely related) equilibrium concepts in game theory, of which sequential equilibrium is one. However, you are more likely to encounter Nash equilibrium and subgame perfect equilibrium early in a first course in game theory.

produce questions about why one would want to know the behavior of a function, sequence, or series as it approaches a particular value. It turns out that knowing the limit can be quite useful. More specifically, the limit of a function can help us determine what the rate of change of the function is, and knowing the rate of change is useful in both statistics and formal models. Further, as noted above, the limit of a sequence or a series is important for the study of game theory and other formal models, and it is also appealed to in some areas of statistics.

4.2.1 Limits and Sequences

Limits are connected to sequences in a fundamental way. Consider our two example infinite sequences from the previous section, $\{i\}_{i=1}^{\infty}$ and $\{\frac{3}{10^i}\}_{i=1}^{\infty}$. We might want to know what is the "endpoint" of these sequences, despite knowing that the sequences continue without end. Does this idea have any meaning?

To answer this, we start with some definitions. A **limit of a sequence** $\{x_i\}$ is a number L such that $\lim_{i\to\infty} x_i = L$. The expression is read as "the limit of x_i as i approaches infinity is L." It means that as you traverse the sequence further and further (i.e., the index i gets bigger and bigger) the elements in the sequence get closer and closer to L. They may never equal L as long as i is finite, but they get arbitrarily close.[3]

Of course, this only makes sense if the elements *do* get closer to something. If they just oscillate between -1 and 1 forever, for instance, then there is no one value they approach. We say a sequence (or series, or function) **converges** if it has a finite limit, and **diverges** if it either has no limit or has a limit of $\pm\infty$.

Let's now return to our examples. The sequence $\{i\}_{i=1}^{\infty}$ gets larger and larger forever and approaches infinity, so it diverges. In contrast, the sequence $\{\frac{3}{10^i}\}_{i=1}^{\infty}$ gets smaller and smaller, approaching zero as $i \to \infty$. Thus, its limit is zero. Some useful limits to remember are $\lim_{i\to\infty} \delta^i = 0$ if $|\delta| < 1$,[4] and $\lim_{i\to\infty} \frac{1}{i^z} = 0$ if $z > 0$.

It is important to note that limits differ in general from the **extreme values** of a sequence. The extreme values of a set are the minimum and maximum values in that set. Some students mistakenly assume that the limit of a sequence is one or both of the extreme value(s) of that sequence. This may be true, but it may not. For example, in the sequence $\{1, 0, \frac{1}{2}, \frac{1}{2}, \ldots\}$ the limit is $\frac{1}{2}$, but the maximum is 1 and the minimum is 0.

4.2.1.1 Why Should I Care?

Though limits of series are more often used in formal theory than are limits of sequences, understanding the latter helps us to understand the strategic logic of repeated behavior. Because payoffs are discounted with a discount factor

[3]More formally, one can choose any tiny number $\epsilon > 0$ and find an N such that, for all $i > N$ (i.e., for all elements of the sequence further along than is N), $|L - x_i| < \epsilon$.

[4]To see this it might help to plug in a fraction for δ, say $\delta = \frac{1}{2}$. Then $\delta^2 = \frac{1}{4}$, $\delta^3 = \frac{1}{8}$, etc., with a limit of 0.

less than one, from the standpoint of the present time, payoffs accrued in the future get smaller and smaller the longer one must wait to receive them. In the limit, the size of the payoffs goes to zero, which is a necessary condition for an individual to want to trade off present gains for future losses. If the limit were *not* zero, total payoffs in the limit would effectively exceed payoffs in any finite time period, and no one would ever trade off present gains for permanent future losses. This would take away a key strategic insight of repeated behavior.

4.2.2 Limits and Series

The **limit of a series** is much like that of a sequence, except that one is looking for the sum of all elements in an infinite sequence rather than the "endpoint." If the series is $\sum_{i=1}^{N} x_i$, then the limit is $\lim_{N\to\infty}\sum_{i=1}^{N} x_i = S$. You've already seen the limits of two series. The limit $\lim_{N\to\infty}\sum_{i=1}^{N} i = \infty$, and so the series is divergent. The limit $\lim_{N\to\infty}\sum_{t=0}^{N}\delta^t = \frac{1}{1-\delta}$, so the series converges to $\frac{1}{1-\delta}$.

We can use this second example to connect sequences with series. Let $S_N = \sum_{t=0}^{N}\delta^t$ and $S_\infty = \lim_{N\to\infty}\sum_{t=0}^{N}\delta^t$. Construct the sequence S, where $S = \{S_i\}_{i=0}^{\infty}$; this is a sequence of partial sums. We calculated that $S_N = \frac{1-\delta^{N+1}}{1-\delta}$ in the previous section. Since $\delta < 1$, as $N \to \infty$, $\delta^{N+1} \to 0$, so $S_N \to \frac{1}{1-\delta} = S_\infty$. Thus the sequence of partial sums converges to S_∞.

What about our other example? That infinite series is $\lim_{N\to\infty}\sum_{i=1}^{N}\frac{3}{10^i}$. This is not immediately obvious, but as with many series we can write out some terms to try and see a pattern. In this case, using decimals makes it more clear. The sum is $.3 + .03 + .003 + .0003 + \ldots = .\bar{3} = \frac{1}{3}$.[5] Thus the series converges to $\frac{1}{3}$.

As a further illustration, consider the paradox of Zeno's runner. Zeno was a Greek philosopher who produced a number of apparent paradoxes. One of them involves a sequence and is thus of some interest for our purposes. Zeno asks us to consider a runner who is to complete a course from point A to point B. Call the distance between the points one unit (perhaps a mile). Imagine that the runner completes half the distance from A to B, and then completes half the remaining distance, and again half the remaining distance, and so on. If we think about the runner proceeding in this manner, does it not follow that the runner will never reach the endpoint, B?[6]

To answer this, start by noting that the sequence of distances the runner traverses is $\{\frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \frac{1}{16}\ldots\}$. This can be represented as the element $x_i = \frac{1}{2^i}$. To see whether the runner reaches the goal, we want to calculate the sum of all these distances traveled, which is a series with this element. This is $S_N = \sum_{i=1}^{N}\frac{1}{2^i}$, where we have again used S_N to refer to a partial sum. We can write down the first few elements of a sequence of these partial sums to get a pattern:

[5]The bar over the .3 indicates that this number repeats in the decimal indefinitely.

[6]Zeno's paradox of the runner relies on the observation that an infinite number of points exists between any two points, and thus regardless of how many times the runner travels half the remaining distance, he will never reach point B. You will find a silly cartoon about it at `http://xkcd.com/1153/`.

$\{\frac{1}{2}, \frac{3}{4}, \frac{7}{8}, \ldots\}$. Each element gets closer and closer to 1. From this, we conjecture that the series converges to 1 in the limit.[7] Thus, in the limit, the runner reaches point B.[8]

4.2.2.1 Why Should I Care?

Let's return once more to the infinitely repeated prisoner's dilemma. To determine when it is beneficial to cooperate and when one would prefer to defect, one must consider the total payoff that one receives for cooperating and compare that to the total payoff one would receive if one were to defect and suffer the punishment. To compute these payoffs, one must add a sequence of infinite payoffs, which is to say, compute an infinite series.

Let's see how this works. Assume the payoff if both players cooperate in any period is 4. What total payoff does one get if both players cooperate forever? Well, in each period one gets the payoff 4, discounted by one more multiple of δ for every period that has passed. In other words, the total payoff is $4 + 4\delta + 4\delta^2 + \ldots$. Rewriting, this is the infinite series we discuss above, multiplied by 4: $\lim_{N\to\infty}\sum_{t=0}^{N} 4\delta^t = \frac{4}{1-\delta}$.

What if one defects? The answer depends on the punishment the other player will enact, which is represented by an infinite sequence of actions. There are many possible punishments, but we consider the simplest and most robust one: grim trigger. The grim trigger punishment (aka Nash reversion strategy) requires that one defect against one's opponent for all time after the opponent defects once. In other words, one defection leads to mutual defection forever, since there is no reason for the original defector to cooperate in the future given her opponent's permanent defection. The total payoff in this case is the payoff for the one-period defection in the presence of her opponent's cooperation (often called the temptation payoff), plus the infinite series of payoffs arising from mutual defection. Let the temptation payoff be 6, and the mutual defection payoff be 2. Then the total payoff is $6 + \lim_{N\to\infty}\sum_{t=1}^{N} 2\delta^t$. Note that the infinite sum here begins with $t = 1$, since the zeroth period is the one in which the first deviation from cooperation happens, giving a payoff of 6. But since the

[7]There are other ways to see this. You can draw a number line and mark a (generously sized) interval from 0 to 1 on it. Sector it by following the sequence of Zeno's runner: mark the halfway point, then mark the halfway point of the remainder, and so on. Do so until you are bored with it (or run out of room, as the case may be). What number do you get closer to each time you identify the next value? Would the number you are getting close to change if you continued the process beyond the number of times you did it? Nope. You are moving closer and closer to 1 no matter how many times you repeat the procedure, just as you are moving closer and closer to infinity as you add the positive integers or move closer and closer to $\frac{1}{3}$ as you add 3 divided by ever higher powers of 10.

[8]The reader might notice that this doesn't really answer the question, since who runs an infinite number of distances anyway? The trick, recognized by Aristotle, is that the time it takes to run the ever-smaller distances also decreases. This, however, leads to a discussion of infinitesimal change, a topic we cover in Part II.

sum is infinite, we can shift the series index by one if we pull out a δ, like so: $\lim_{N\to\infty}\sum_{t=1}^{N} 2\delta^t = \lim_{N\to\infty} 2\delta \sum_{t=1}^{N} \delta^{t-1} = \lim_{N\to\infty} 2\delta \sum_{t=0}^{N} \delta^t = \frac{2\delta}{1-\delta}$.[9]

Thus, we need to compare the cooperation payoff, $\frac{4}{1-\delta}$, to the payoff for defecting, $6 + \frac{2\delta}{1-\delta}$. If the first exceeds the second, then cooperation is possible. This happens when $\frac{4}{1-\delta} \geq 6 + \frac{2\delta}{1-\delta}$ or $4 - 2\delta \geq 6 - 6\delta$, which implies $\delta \geq \frac{1}{2}$. In other words, whenever the discount factor is sufficiently high, implying that people value future payoffs sufficiently much, cooperation can be maintained by the threat of a complete breakdown of cooperation.

4.2.3 Limits and Functions

Recall that a function is a mapping of the values in one set to the values in another set (or, equivalently, a graph of that mapping, or, equivalently again, an equation that describes that graph). If we want to know the value of a function for a given x, we can plug that x into the function and calculate its value. For example, if $f(x) = x^2$, then the value of $f(x)$ for $x = 2$ is 4 (since $2^2 = 4$). However, even though we can always find the value of the function at a point, as long as that point is in the function's domain, it's not always obvious how the values of the function *near* that point relate to the value at that point.

Why do we care about values close to some point $x = c$? There are several reasons, but the one of most importance to us relates to the roles of the limit in calculus, the topic of Part II of the book. Let's consider just one example. We are often interested in political science in the rate of change of a function at a point. We briefly mentioned this above when discussing the linear equation. If $y = \alpha + \beta x$, then the rate of change of y with x is given by β, and is constant for all x. This β is the important output of a statistical analysis, as it tells us the degree to which the concept represented by x affects the concept affected by y. Discerning such relationships, of course, is central to research in political science.

This still doesn't explain why we need limits, though. After all, the slope of the linear equation is constant. So let's try instead $y = \beta x^2$. It turns out, as you'll see in Part II, that the rate of change of y with x now is given by $2\beta x$, and so varies with the value of x. In this case, y is changing increasingly fast with x as x increases.

The way we figure out the value of the slope at any point is to look at the values of the function on both sides of a point $x = c$, but very close to that point. This difference tells us how much y changes for a small change in x. As we make this change smaller and smaller (i.e., take the limit as it goes to zero), we get a better and better approximation of the slope at that point, which is exact at the limit. Thus, limits of functions allow us to figure out the rate at which independent variables affect dependent variables, *even when the dependence of one on the other is complex.*

[9] The key here is that the sum from $t = 1$ to infinity of δ^{t-1} is the same as the sum from $t = 0$ to infinity of δ^t. To see this, plug in the first few values of t and observe that the sum in each case is the same. Any difference in the sum with the shifted index would manifest at the end of the series, but it has no end, so there is no difference.

Table 4.1: Limit of $f(x) = x^2$ as $x \to 2$

x	$f(x)$
1.9	3.61
1.95	3.8025
1.98	3.9204
1.99	3.9601
2.0	4
2.01	4.0401
2.02	4.0804
2.05	4.2025
2.1	4.41

So, what is the **limit of a function**? For the function $y = f(x)$, a limit is the value of y that the function tends toward given arbitrarily small movements toward a specific value of x, say $x = c$. The limit either exists or does not exist for a given value of x, and if it does exist we can calculate it. Formally, much as for a sequence we can write $\lim_{x\to c} f(x) = L$.[10] One reads that notation as "the limit of f of x as x approaches c is L." Unlike a sequence, however, it's possible to approach a point (other than $\pm\infty$) from two different directions. If x approaches c from above (i.e., x decreases toward c), then we write $\lim_{x\to c^+} f(x) = L^+$. If x approaches c from below, we write $\lim_{x\to c^-} f(x) = L^-$. If the limits from above and below are equal, so that $L = L^+ = L^-$, then the function has a unique limit at c.[11]

Let's consider some examples, starting with a straightforward one: the limit of $f(x) = x^2$ as $x \to 2$. At $x = 2$, the function, as we have noted, takes value 4. What about near 2? When we are evaluating the limit of a function we want to know about *the function's behavior* at and near a given value, not the function's value. We want to ask: As we move a small distance away from $x = 2$ in either direction, does the function return a value that is only a small distance from $y = 4$? Let's look at a table to try to get a fix on this. Try the following x values: 1.9, 1.95, 1.98, 1.99 and 2.1, 2.05, 2.02, and 2.01. Table 4.1 reports both the x and y values.

Regardless of whether we begin at the top of the table or the bottom of the

[10] The similarities to the limit of a sequence do not end there. On the one hand, one can view a sequence as a function with a domain of the natural numbers; its limit is then just the limit of the function at infinity. On the other hand, one can construct a sequence with elements equal to $f(c + \frac{1}{n})$ or $f(c - \frac{1}{n})$. Its limit is the limit of the function from above or below, respectively.

[11] More formally, one can choose any tiny number $\epsilon > 0$ and find a $\delta > 0$ such that $|L - f(x)| < \epsilon$ whenever $0 < x - c < \delta$ for the limit from above, and whenever $0 < c - x < \delta$ for the limit from below.

table,[12] we can see that the closer we get to $x = 2$, the closer we get to $f(x) = 4$. In other words, the behavior of $f(x)$ around $x = 2$ is smooth: the function does not produce any surprising values in its sequence. If we graphed the function, it would not produce any surprising jumps in its graph around $x = 2$ (we do this in the exercises below). So we can conclude that the function $f(x) = x^2$ has a limit of 4 at $x = 2$.[13]

What often strikes students as odd about an example like this is that we can simply calculate that $f(x) = 4$ at $x = 2$, so what is all this business about saying that the limit of $f(x)$ at $x = 2$ is 4? Isn't it obvious that it is 4? Well, yes, the value of the function at $x = 2$ is indeed 4, and this is the same as the limit in this case, but information about the function at a point does not alone tell us whether $f(x)$ has a limit at $x = 2$. Put another way, the existence of a limit depends upon information about *the behavior of the function near a specific value of the variable(s) defined in the function,* and that is not the same as the value of the function at a specific value of the variable(s). Instead, we need to evaluate the behavior of the function near that point, as we did in Table 4.1.

Still don't believe us? Try this example:

$$f(x) = \begin{cases} x^2 & : x < 2, \\ (x-2)^2 & : x \geq 2. \end{cases}$$

The limit from below is still 4 at $x = 2$, since the function is the same as before. But now the limit from above is $0 \neq 4$. So this function has no limit at all. Is this a realistic scenario? Well, that depends on what the function is intended to represent, but certainly piecewise functions[14] see use in political science. For example, payoffs could decrease after some bureaucratic deadline is reached or after brinkmanship goes too far and leads to war rather than continued bargaining. Or a sudden change such as a shift in the party in power might lead to reduced circumstances for the party now out of power.

We don't need to invoke piecewise functions to find those that lack limits, though. Consider $f(x) = \frac{1}{x}$. As x approaches 0 from below, $f(x)$ gets ever more negative without limit. However, as x approaches 0 from above, $f(x)$ gets ever more positive without limit.[15] Thus, not only does this function have no finite limit at 0, it has no definable limit at all!

Before moving on, we consider some properties of limits. For any $f(x), g(x)$ that *both* have a well-defined limit at $x = c$, we have that (the last one as long

[12] If we think about this with respect to a number line, then we would say "regardless of whether we approach 2 from the left or the right."

[13] More formally, for any $\delta > 0$, $f(2-\delta) \to f(2)$ and $f(2+\delta) \to f(2)$ as $\delta \to 0$.

[14] A piecewise function is one that behaves differently depending on the value of x, and thus can be written in pieces. We introduce these in Chapter 3.

[15] If this is not clear, replicate the type of analysis in Table 4.1.

as $\lim_{x\to c} g(x) \neq 0$):

$$\begin{aligned} \lim_{x\to c}(f(x) + g(x)) &= \lim_{x\to c} f(x) + \lim_{x\to c} g(x), \\ \lim_{x\to c}(f(x) - g(x)) &= \lim_{x\to c} f(x) - \lim_{x\to c} g(x), \\ \lim_{x\to c}(f(x)g(x)) &= (\lim_{x\to c} f(x))(\lim_{x\to c} g(x)), \\ \lim_{x\to c}(f(x)/g(x)) &= (\lim_{x\to c} f(x))/(\lim_{x\to c} g(x)). \end{aligned}$$

These are helpful, but don't cover all cases. The most notable of these is the situation in which the value of the function is undefined at c. For example, let $f = \frac{x^2-4}{x-2}$. This function is undefined at $x = 2$. However, it has a limit of 4 at $x = 2$. How did we get this? One way is to factor the numerator into $(x-2)(x+2)$, and then divide the numerator and denominator by $(x-2)$. This is not allowed at $x = 2$, because of division by zero, but it is just fine at all points other than $x = 2$, which are the points we need to calculate to figure out the limit. Canceling the $(x-2)$ leaves $x+2$, which approaches 4 from above and below as $x \to 2$.

Even when the numerator and denominator of a rational function do not cancel, one can still use the same approach when the limit is $x \to \infty$ and the numerator and denominator of the function both go to ∞ as well. In this case, just pick the terms on both bottom and top that are the biggest for any finite value of x, and then cancel common terms as above.[16] For example, the limit of $f = \frac{x^3+2x^2}{3x^3+2x-1}$ as $x \to \infty$ is $\frac{1}{3}$, which is what you get when you consider only the terms of highest order (the biggest ones for all finite x) in both the numerator and denominator, and then cancel the x^3.

4.2.4 Why Should I Care?

We have already discussed the utility of the limit of a function in determining the rate at which independent variables affect dependent variables. Further, by discussing the importance of summing infinite series of payoffs in repeated games in game theory, we implicitly illustrated the benefits on limits of series, since an infinite series is just the limit of a finite series as the number of terms goes to infinity. There are many other examples like these. For instance, it is possible to develop a number of ideas relevant to the study of time series (e.g., permanent shocks with respect to unit roots, or impulse response functions), though these concepts are frequently presented without reference to the limit of a sequence.

Limits are also a building block in several other concepts that are frequently used in political science. The first is the maximum or minimum of a function. More specifically, the limit is a stepping stone to understanding derivatives, and derivatives are used to find maximum and minimum values of a function.

[16] A more general way to deal with limits like these (and one that makes this procedure make quite a bit more sense) invokes l'Hôpital's rule, which says that $\lim_{x\to c} \frac{f(x)}{g(x)} = \lim_{x\to c} \frac{f'(x)}{g'(x)}$ if $\lim_{x\to c} f(x) = \lim_{x\to c} g(x) = 0$ or $= \pm\infty$ and $\lim_{x\to c} \frac{f'(x)}{g'(x)}$ exists. Here $f'(x)$ and $g'(x)$ are first derivatives, which we'll cover in Part II.

When might one want to know the minimum or maximum of a function? Some theories assume that an actor wants to maximize something (e.g., their power, utility, time in office, etc.). If one writes down a function to describe the individual's power, utility, time in office, etc., then one will be interested in limits (and derivatives) to help one determine whether a maximum exists and what it is. In Chapter 8 of this book we put these tools to work to do precisely that.

In addition, statistical analysis often leads one to have interest in minima and maxima of functions. For instance, one might want to minimize the predicted error in a regression or maximize the likelihood of matching the data to the statistical model that the researcher specified. Again, the limit (and the derivative) are useful for that purpose.

Limits also arise in the definition of other important concepts in addition to those in calculus. We discuss two below: open and closed sets, and continuity of functions.

4.3 OPEN, CLOSED, COMPACT, AND CONVEX SETS

In Chapter 1 we discussed several properties of sets, but did not have the background to be as complete as we would have liked. Now we do and we introduce four new types of sets: **open**, **closed**, **compact**, and **convex**. All four will prove central to later developments in the book. We address them in the order stated.

Often one sees an open set defined as the opposite of a closed set, or vice versa. Since we are not overly concerned with formalism and want to highlight intuition, we will define both independently. An **open set** is one in which there is some distance (which may be arbitrarily small) that you may move in any direction within the set and stay in the set. We'll deal mostly with the concept of openness in reference to spaces that have a defined distance metric. In these cases we can be more intuitive. Picture a perfectly spherical ball of the same dimension as the space. So, if we're in normal three-dimensional space ($\mathbb{R}^3$), you should be picturing a sphere, while in two dimensions you should be imagining a circle. If you can find a small enough ball around any point such that the entirety of the ball remains in the space, then the space is open.[17]

The modal example for an open set is $(0, 1)$.[18] The point 1 is not in the set, but everything less than 1 (but greater than 0) is. Similarly, 0 is not in the set, but everything greater than 0 (but less than 1) is. So no matter which point less than 1 and greater than 0 you choose there is always some small enough distance that allows you to go a little to the left or a little to the right and

[17]Formally, define the distance between two points $x, y \in \mathbb{R}^n$ as $d(x, y) \geq 0$. Then a set $A \subset \mathbb{R}^n$ is open if for all $x \in A$, there exists an $\epsilon > 0$ such that all points $y \in A$ with $d(x, y) < \epsilon$ are in A too.

[18]Recall from Chapter 1 that this notation means the set of all real numbers between zero and one, exclusive, and that the use of curved brackets signifies openness.

stay in the set. The union of any number of open sets remains open, while the intersection of a finite number of open sets is also open.

Now consider the set $[0, 1]$. This is very similar, except for one important property: 0 and 1 are in the set. That means that one can't fit a ball around the points 0 or 1 and stay in the set, since *nothing* less than 0 or greater than 1 is in the set. So $[0, 1]$ isn't open. Instead, it is closed, though these terms are not opposites. For example, the empty set and the universal set are both open and closed, while the sets $(0, 1]$ and $[0, 1)$ are neither open nor closed. So how do we know $[0, 1]$ is closed? One definition of a closed set is that it is the complement of an open set, but this may not be that intuitive at this point. Instead we turn to limits.[19]

A **closed set** is one that contains all its limit points. What does that mean? It means that if you make a sequence with all its elements contained in a set A, then for a closed set the limit of this sequence must also be in A. More intuitively, start by choosing points within a set, such that the sequence of these points gets closer and closer to some other point. If a set is closed, then the point that this sequence approaches is also in the set. In other words, one can't leave a closed set by following a sequence that is otherwise within it to its limit. A set that is not closed, in contrast, can have the point the sequence approaches—the set's **limit point**—outside the set, even if all prior points in the sequence are in the set. In an open set, one can get closer and closer to a point, all the while staying in the set, but leave the set in the limit of the sequence.

Let's try this out with $[0, 1]$. We can choose a sequence with elements $x_i = \frac{1}{i}$ or $x_i = 1 - \frac{1}{i}$, depending on which limit we want to show. In both cases, for any finite value of i the corresponding element of the sequence is within A for either of $(0, 1)$ or $[0, 1]$. However, the limit of the first sequence is 0 and the limit of the second is 1, and neither of these points is in $(0, 1)$. They are both in $[0, 1]$, though. And any other sequence whose elements are all in $[0, 1]$ will also have a limit in $[0, 1]$. So $[0, 1]$ is closed. Now we can also make more sense of the other definition. The complement of a closed set is everything outside the closed set. In this case, the complement of $[0, 1]$ is $(-\infty, 0) \cup (1, \infty)$, which is the union of two open sets, which is open. The important thing is that the complement does not contain the points 0 and 1, since they are in $[0, 1]$. One can come ever closer to the boundary of a closed set from outside it without reaching it.

Any intersection of closed sets is closed, while the union of a finite number of closed sets is closed. One can get from an open (or other) set to a closed set by adding all the original set's limit points into the set. The closed set you get by doing this is called the closure of the original set. So $[0, 1]$ is the closure of $(0, 1)$.

A subset of $\mathbb{R}^n$ that is both closed and bounded—that is, the set contains all its limit points and can itself be contained within some finite boundary—is

[19]And thereby justify this section's inclusion in this chapter.

called **compact**.[20] Compact sets are in some sense self-contained: they don't go on forever, either by having no bound or by containing sequences that don't stay in the set in the limit. These properties help ensure that continuous functions defined over compact sets have nice properties, notably maxima and minima that are present in the set.

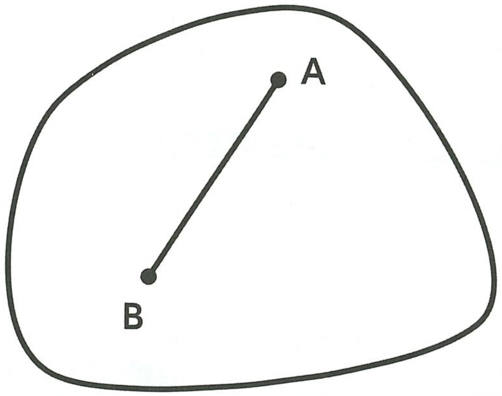

Figure 4.1: Convex Sets

Finally, a subset of $\mathbb{R}^n$ for which every pair of points in the set is joined by a straight line that is also in the set is called a **convex set**. Formally, a set is convex if for all points x and y in it and all $\lambda \in [0,1]$,[21] the point $(1-\lambda)x + \lambda y$ is also in the set. The **convex hull** of a set A is the set A plus all the points needed to make A convex. See Figures 4.1 and 4.2 for pictures of convex and nonconvex sets. Convex sets are useful when one is considering linear combinations of variables and wants to make sure that these combinations remain in the set. For example, a political party in a proportional representation system may get utility both from the coalition government policy and from the share of the cabinet it controls in the coalition government. If all linear combinations of policy and cabinet share are possible—i.e., if the set of possible government outcomes is convex—then the strategy the party might take in trying to form a government may be very different than if all combinations are not possible

[20] Recall the definition of bounded from Chapter 1. This is one of those points where we sacrifice formalism and generality for simplicity. Something bounded and self-contained seems compact in common language. The more general definition of compact—every open cover has a finite subcover—requires additional definitions and adds little intuition, in our opinion. One should be aware that the definition of compact as closed and bounded is a consequence of the Heine-Borel theorem and applies specifically to subsets of $\mathbb{R}^n$. However, these are by far the most common sets we will encounter in political science.

[21] Parameters that vary between zero and one in this way and multiply other variables are often called weights.

because the set of possible outcomes is not convex. A party might, for instance, seek out a much larger cabinet share if a larger share is the only way to secure a beneficial policy than if a preferred policy is available with a smaller share.

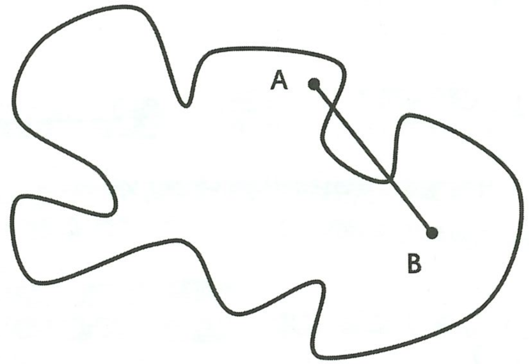

Figure 4.2: Nonconvex Sets

4.3.1 Why Should I Care?

This section may have seemed fairly abstract, but these set concepts will prove fairly central to the rest of this book and political science in general. Compactness in particular will appear time and again for the simple reason that it permits us to maximize (and minimize) functions over a domain, and maximizing (and minimizing) functions is, as we have noted, central to analysis in political science. We address this more in Part II, but an example can help us to see how this works. Consider the sets $A = (0,1)$ and $B = [0,1]$. B is compact; A is not. We ask: What's the maximum of the linear function $f(x) = x$ on each of these sets?

For B the answer is 1. One is the largest value in B, and so the biggest value that f can take on the domain B. But there is no corresponding biggest value of f on A because A *has no biggest value.* That's right, even though it is bounded by 1, there is no single largest value in it because one can always go just a bit closer to 1. Each movement closer to 1 increases f, but since the limit of a sequence approaching 1 is outside the set A, we can never increase f as much as possible, since we're limited to values within the domain A. So f has no maximum on the domain A. Isn't math fun?

4.4 CONTINUOUS FUNCTIONS

Intuitively, a **continuous function** is a function without sudden breaks in it. When you draw the graph of a continuous function, you never need to lift your pencil from the page. Put differently, the graph of a continuous function forms an unbroken curve, whereas the graph of a **discontinuous function** has at least one break in it. This break is usually called a jump, since the function jumps up or down quite a bit for a small change in x. We can define continuity in a couple of ways.

More formally, a continuous function is one for which an arbitrarily small change in x causes an arbitrarily small change in y for *all* values of x. That's why you never need to lift your pencil when drawing the graph of a continuous function: each time you move a little bit to the right, further along the x-axis, you need only move a little bit up or down along the y-axis as well.[22]

This is a fine definition (particularly the one in the footnote), but limits provide perhaps a more intuitive one. A function is continuous at a point if the limit of the function at that point exists and is equal to the value of the function at that point. In the language of math: $f(x)$ is continuous at $x = c$ if and only if $\lim_{x \to c} f(x) = f(c)$. A function that is continuous at all points c in its domain is called continuous. Why is this more intuitive (according to us)? Consider what this means. That the limit exists means that if I traverse the function toward c from above and from below, I get closer and closer to a point, and this point is the same on both sides. That means the two pieces of the function *should* connect at c. The only problem is that the function at c need not equal the limit at that point, as we have noted. So, the function is continuous only when the value of the function at c is equal to this limit; i.e., there is no gap in the function right at that point to force you to move your pencil. In a way, this concept is very similar to that of a closed set. A closed set contains the limit points of all sequences that lie within it. In a continuous function, the range of the function, loosely speaking, contains the limit of the function.

Let's look at some examples. Most of the functions described in Chapter 3 are continuous. So, affine (and linear) functions are continuous, quadratic and higher-order polynomials are continuous, logarithms are continuous over the region on which they are defined, and so on. In fact, nearly every function you've had occasion to consider has likely been continuous. But there are lots of functions that are not continuous. For instance, the function $f(x) = \frac{1}{x}$ is not continuous over the domain $\mathbb{R}$, i.e., the real number line, since no limit exists at zero. Put another way, there is no way you can get from $-\infty$ to $+\infty$ while crossing $x = 0$ without picking up your pencil!

The piecewise function described at the end of the section on limits is also

[22] More formally still: a function is continuous at some point $x = c$ if for every $\epsilon > 0$, there exists a $\delta > 0$ such that whenever $|x - c| < \delta$, $|f(x) - f(c)| < \epsilon$. In other words, points very near to c map to points very near to $f(c)$, and by going closer to c one can go as close as one wants to $f(c)$. A function that is continuous at all points c in its domain is called continuous.

not continuous. Why not? Consider the function again:

$$f(x) = \begin{cases} x^2 & : x < 2 \\ (x-2)^2 & : x \geq 2. \end{cases}$$

Its limit at $x = 2$ is 4 coming from the left and 0 coming from the right, so it's not continuous at $x = 2$. Could we change it to make it continuous? Well, yes. Since it's continuous at all points other than $x = 2$, all we need to do is either to shift the left piece down by 4, to $x^2 - 4$, or the right piece up by 4, to $(x-2)^2 + 4$. Then the value of each piece at $x = 2$ would be the same, as would be the limits from above and below. Of course, this would change the meaning of our function and so we might not want to do this.

These examples of discontinuity all relate to the lack of a limit at some point, but that's not necessary. Consider the function $y = \frac{x^2}{x}$. If you plot it over the range -5 to 5 it appears to be a straight line, and its limit is well defined at every point. However, division by zero is undefined, so the function is undefined at $x = 0$, as indicated by the open circle at $x = 0$ on the plot in Figure 4.3. This discontinuity is the easiest to remove; all we have to do is to define the function piecewise: $y = 0$ at $x = 0$ and $y = \frac{x^2}{x}$ everywhere else. This fixes the discontinuity and most likely does not change the meaning of the function either.

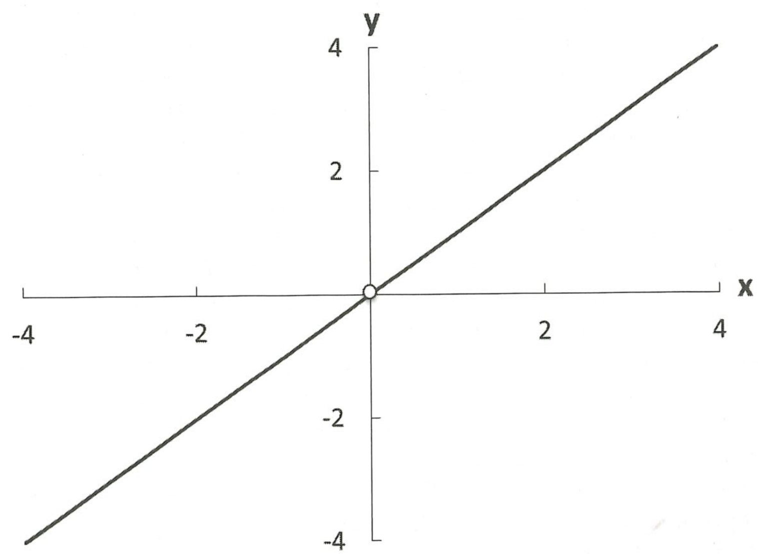

Figure 4.3: Graph of $y = \frac{x^2}{x}$, $x \in [-5, 5]$

4.4.1 Why Should I Care?

Continuity is perhaps one of the most central concepts you will encounter, for much the same reason that compactness proves useful: it helps us to maximize (or minimize) functions, and maximizing (or minimizing) functions is important in both game theory and statistics. To see how this works, consider one last time the piecewise function from above, but now let's say that the domain of x is only between 0 and 3. In other words, $x \in [0, 3]$, a compact subset of the real numbers. The expression $(x-2)^2$ never gets higher than 1 in this domain, so the maximum has to come from x^2. But the function only takes these values for $x < 2$. We can keep increasing the function by moving closer and closer to $x = 2$, but we never hit a maximum, just as when maximizing some functions on open sets.

We often assume continuous functions when representing preferences (i.e., a utility function) or the outcome of interactions between actors for this reason. This is particularly true when scholars work with functions that are only implicitly defined and are not given explicit functional forms. These implicitly defined functions are usually assumed continuous. Why might they make these assumptions? Well, the reason why they do not write down an explicit functional form is probably that they want their theories to be as general as possible. If they were to write down a specific functional form, then the theory would apply to that function, but not necessarily to others. So keeping it general expands the domain of their theories. Yet there are some functional forms that might produce an internal inconsistency in one's theory if they are not ruled out. For example, if one's theory relies on a result where one takes a derivative, then one will want to limit the theory to only those functions where it is possible to take a derivative. As we will learn in the next part of this book, the derivative is a type of limit. Thus, if a function is not continuous, there are points where it does not have a limit, and that means that the function is not differentiable (i.e., has no derivative) at those points. If one were not to limit one's theory to continuous functions, then one would be asserting that the theory is relevant to functions that have points over which the derivative is undefined, and then appealing to the derivative to substantiate one's result (or hypothesis). Doing so would be contradictory. Thus, to avoid contradiction, one begins with the assumption that the functions to which one's theory applies are continuous.

Though continuous functions are thus common, some solutions of problems require analysis of discontinuous functions. For example, in Chapter 3 we referred to best response correspondences in game theory, which describe what one player's optimal strategy (or strategies) is (are) given a particular choice in strategy by the other player. These are often discontinuous: one strategy may be optimal for some subset of the other player's strategy space, but a quite different one might be optimal for a different subset. For example, a particular distribution of the budget might be optimal if each party bargains in good faith, but if one of the parties asks for too much, the optimal response might abruptly drop to an offer of nothing. When one's best response to another actor's actions is discontinuous in those actions, one often needs an alternative technique to figure out the equilibrium of the interaction. One technique is monotone comparative statics, which allows you to relax assumptions on the continuity of best response functions (McCarty and Meirowitz, 2007; Ashworth and Bueno de Mesquita, 2005). So it is important to understand continuity and to be able to determine whether a function is continuous over all possible values of x or whether it is continuous only for some values of x.

4.5 EXERCISES

1. Draw a graph to show that the sequence $\{1, -1, 1, -1, 1, -1 \ldots\}$ is divergent.
2. Find the sum of the infinite series $\sum_{t=0}^{\infty} (\delta^t)^2$.
3. Show whether $f(x) = x + x^3$ has a limit at $x = 3$ and, if so, the value of the limit.
4. Show whether $f(x) = (x-3)(x+5)$ has a limit at $x = 4$ and, if so, the value of the limit.
5. Show whether $f(x) = \frac{3x^2-12}{x-2}$ has a limit at $x = 2$ and, if so, the value of the limit.
6. Show whether $f(x) = \frac{x^3-4}{x-2}$ has a limit at $x = 2$ and, if so, the value of the limit.
7. For each of the following sets, state whether they are (a) open, closed, both, or neither; (b) bounded; (c) compact; (d) convex:
   a) $[1, 3]$
   b) $(2, 5)$
   c) $[0, 6] \bigcup [10, 12]$
   d) $(2, 4) \bigcap [3, 4]$
   e) $[0, \infty)$
8. Is the function $f(x) = \frac{\ln(x)}{x}$ continuous for $x \in [2, \infty)$?
9. Is the function
$$f(x) = \begin{cases} x^3 - 3x + 4 & : x \leq 3, \\ x^2 & : x > 3, \end{cases}$$
continuous? If so, why? If not, what changes would make it continuous?