states the set of natural numbers with no least element must be empty. A set \( S \) of natural numbers is empty if and only if \( S \cap \{n\} = \emptyset \) for all \( n \in \mathbb{N} \). Thus it suffices to prove that if \( S \) has no least element, then \( S \cap \{n\} = \emptyset \) for all \( n \in \mathbb{N} \). We prove the conclusion by induction on \( n \).

Since \( S \subseteq \mathbb{N} \) and \( S \) has no least element, \( 1 \notin S \), so \( S \cap \{1\} = \emptyset \). For the induction step, suppose that \( S \cap \{n\} = \emptyset \). Since \( S \) has no least element, we therefore have \( n + 1 \notin S \), since \( n + 1 \) is the least natural number among those not in \( S \). Now we have \( S \cap \{n+1\} = \emptyset \).

3.65. Employers and thieves. Each employer has one apprentice. When an apprentice is a thief, everyone knows except the thief's employer. The mayor declares: "At least one apprentice is a thief. Each thief is known to be a thief by everyone except his/her employer, and all employers reason perfectly. If during the \( i \)th day from now you are able to conclude that your apprentice is a thief, you must come to the village square at the next noon to denounce your apprentice." The villagers gather at noon every day thereafter to see what will happen. If in fact \( k \geq 1 \) of the apprentices are thieves, then their employers denounce them on the \( k \)th day.

The proof is by induction on \( k \). Basis step \( (k=1) \). When there is exactly one thief, the thief's employer knows of no thieves. Since the employer knows there is at least one thief, his apprentice must be a thief.

Induction step \( (k=n+1) \). The induction hypothesis states that when there are actually \( n \) thieves, they will be denounced on the \( n \)th day. When there are \( n+1 \) thieves, every employer knows of \( n+1 \) thieves or of \( n \) thieves. An employer who knows of \( n \) thieves knows that there must actually be \( n \) thieves or \( n+1 \) thieves, depending on whether his/her apprentice is a thief. If there were actually \( n \) thieves, then by the induction hypothesis they would be denounced on the \( n \)th day. Since this doesn't happen (there is no one who knows of fewer than \( n \) thieves), there can't be only \( n \) thieves. Hence there must be \( n+1 \) thieves. The employers who know of only \( n \) thieves conclude this after waiting past noon on the \( n \)th day, so they denounce their employees on the \( n+1 \)th day.

4. BIJECTIONS AND CARDINALITY

4.1. Summation of \((120102)_{3}\) and \((110222)_{3}\) in base 3, with check in base 10. When the sum of the entries in a column is at least 3, the number of 3s "carries" to the next column, as in decimal addition.

<table>
<thead>
<tr>
<th>base 3</th>
<th>conversion to base 10</th>
<th>base 10</th>
</tr>
</thead>
<tbody>
<tr>
<td>120102</td>
<td>1 \cdot 243 + 2 \cdot 81 + 0 \cdot 27 + 1 \cdot 9 + 0 \cdot 3 + 2 \cdot 1</td>
<td>416</td>
</tr>
<tr>
<td>110222</td>
<td>1 \cdot 243 + 1 \cdot 81 + 0 \cdot 27 + 2 \cdot 9 + 2 \cdot 3 + 2 \cdot 1</td>
<td>350</td>
</tr>
<tr>
<td>1001101</td>
<td>1 \cdot 729 + 0 \cdot 243 + 0 \cdot 81 + 1 \cdot 27 + 1 \cdot 9 + 0 \cdot 3 + 1 \cdot 1</td>
<td>766</td>
</tr>
</tbody>
</table>

4.2. \(333_{(12)}\) is larger than \(3333_{(5)}\). Let \(x = 333_{(12)} = 3 \cdot 111_{(12)}\) and \(y = 3333_{(5)} = 3 \cdot 1111_{(5)}\). It suffices to compare 144 + 12 + 1 and 125 + 25 + 5 + 1. The first is larger (by 1), so \(x > y\).

4.3. Squares in base 10. The square of the number obtained by appending 5 to the base 10 representation of \(n\) is \((10n + 5)^2 = 100n^2 + 100n + 25\). The last two digits are 25. The number obtained by appending 25 to the base 10 representation of \(n(n + 1)\) is \(100n(n + 1) + 25\). These are the same number.

4.4. Another temperature scale. If the conversion of Fahrenheit temperature \(x\) to \(T\) is \(ax + b\), then changes of fixed amount in \(x\) correspond to changes of fixed amount on the \(T\) scale. Thus the Fahrenheit temperature corresponding to \(T\) temperature 50 is the average of the Fahrenheit temperatures corresponding to \(T\) temperatures 20 and 80. If water freezes at \(T\) temperature 20 and boils at \(T\) temperature 80, then the Fahrenheit temperature corresponding to 50 is the average of the Fahrenheit temperatures 32 (freezing) and 212 (boiling). The answer is 90.

4.5. A finite set \(A\) has a nonidentity bijection to itself if and only if it has at least two elements. With one element, the only function is the identity. When \(A\) has at least two elements, we let \(x, y\) be distinct elements in \(A\). Let \(f(x) = y\), \(f(y) = x\), and \(f(a) = a\) for every \(a \in A\) other than \(x, y\). By construction, the image is all of \(A\), and no two elements of \(A\) are mapped to the same element of \(A\), so \(f\) is a bijection other than the identity.

4.6. The function giving each day of the week the number of letters in its English name is not injective. Two days are mapped to the same integer: \(f(\text{Sunday}) = f(\text{Monday}) = 6\), but \(\text{Sunday} \not= \text{Monday}\).

4.7. Injectivity and surjectivity of functions from \(\mathbb{R}^2\) to \(\mathbb{R}\).

a) \(A(x, y) = x + y\). The addition function is surjective. For each \(b \in \mathbb{R}\), we have \(A(b, 0) = b\). It is not injective, since also \(A(b - 1, 1) = b\).

b) \(M(x, y) = xy\). The multiplication function is surjective. For each \(b \in \mathbb{R}\), we have \(M(b, 1) = b\). It is not injective, since also \(M(b/2, 2) = b\).

c) \(D(x, y) = x^2 + y^2\). This function is not surjective, since no negative number belongs to the image. It is not injective, since \(D(0, a) = D(a, 0)\) even though \((a, 0) \not= (a, 0)\) when \(a \not= 0\).

4.8. Examples of composition. If \(f(x) = x - 1\) and \(g(x) = x^2 - 1\), then \(f \circ g\) and \(g \circ f\) are defined by \((f \circ g)(x) = x^2 - 2\) and \((g \circ f)(x) = x^2 - 2x\).

4.9. If \(f\) and \(g\) are monotone functions from \(\mathbb{R}\) to \(\mathbb{R}\), then \(g \circ f\) is also monotone—TRUE. The composition is decreasing if one of \(f, g\) is decreasing and the other is decreasing. The composition is increasing if \(f\)
and \( g \) are both increasing or both decreasing. Given \( x < y \), application of the functions reverses the order for each of \( \{ f, g \} \) that is decreasing and preserves the order for each of \( \{ f, g \} \) that is increasing. Since \( f \) and \( g \) are monotone, this is independent of the choice of \( x \) and \( y \), so the claimed statements hold.

4.10. Linear functions and their composition. Let \( f(x) = ax + b \) and \( g(x) = cx + d \) for constants \( a, b, c, d \) with \( a \) and \( c \) not zero.

Both \( f \) and \( g \) are bijections. For each real number \( y \), the number \( (y - b)/a \) is defined and is the only choice of \( x \) such that \( f(x) = y \). Thus \( f \) is both surjective and injective. The same analysis applies to \( g \).

The function \( g \circ f - f \circ g \) is neither injective nor surjective. Note that \( (g \circ f)(x) = c(ax + b) + d \) and \( (f \circ g)(x) = a(cx + d) + b \). The difference \( h \) is defined by \( (h)(x) = cax + cb + d - acx - ad - b = cb - ad + d - b \). Thus \( h \) is a constant function. It maps all of \( \mathbb{R} \) to a single element of \( \mathbb{R} \), so it is neither injective nor surjective.

4.11. Multiplication by 2 defines a bijection from \( \mathbb{R} \) to \( \mathbb{R} \) but not from \( \mathbb{Z} \) to \( \mathbb{Z} \). Let \( f \) denote the doubling function. For \( y \in \mathbb{R} \), the number \( 2x/2 \) is the unique real number such that \( f(x) = y \). When \( y \in \mathbb{Z} \) and \( y \) is odd, \( y/2 \not\in \mathbb{Z} \). Hence odd numbers are not in the image of \( f: \mathbb{Z} \to \mathbb{Z} \).


a) Every decreasing function from \( \mathbb{R} \) to \( \mathbb{R} \) is surjective—FALSE. Let \( f(x) = \begin{cases} -x & \text{if } x \leq 0 \\ 0 & \text{if } x > 0 \end{cases} \).

b) Every nondecreasing function from \( \mathbb{R} \) to \( \mathbb{R} \) is injective—FALSE. The constant function \( f \) defined by \( f(x) = 0 \) is nondecreasing but not injective.

c) Every injective function from \( \mathbb{R} \) to \( \mathbb{R} \) is monotone—FALSE. The function \( f \) defined by \( f(0) = 0 \) and \( f(x) = 1/x \) for \( x \neq 0 \) is injective but not monotone. The function is decreasing on every interval not containing 0, but \( f(x) \) is positive when \( x \) is positive and negative when \( x \) is negative.

d) Every surjective function from \( \mathbb{R} \) to \( \mathbb{R} \) is unbounded—TRUE. When \( f \) is surjective to \( \mathbb{R} \), every real number appears in the image, which means that there is no bound on the absolute value of numbers in the image.

e) Every unbounded function from \( \mathbb{R} \) to \( \mathbb{R} \) is surjective—FALSE. Define \( f \) by \( f(x) = 0 \) for \( x \leq 0 \) and \( f(x) = x \) for \( x > 0 \). This function is unbounded but has no negative numbers in its image.

4.13. The difference between abc and cba, added to its own reverse, yields 1089 (given that \( a \neq c \)). We may assume that \( a > c \). The digits of \( abc - cba \) are \((a-c-1), 9, (10+c-a)\), so \( abc - cba = 100(a-c-1) + 90 + (10+c-a) \). The reverse of this is \( 100(10+c-a) + 90 + (a-c-1) \). Summing the two expressions yields \( 100(10 - 1) + 180 + (10 - 1) = 1089 \).

4.14. Finding the q-ary expansion of \( n + 1 \) from the q-ary expansion of \( n \). The idea is to add 1 in base \( q \). Let \( a_n, \ldots, a_0 \) be the q-ary expansion of \( n \). If \( a_0 = q - 1 \), let \( b_0 = a_0 + 1 \), and let \( b_i = a_i \) for \( i > 0 \). Otherwise, let \( j \) be the greatest index such that \( a_i = q - 1 \) for \( 0 \leq i \leq j \). Let \( b_0 = 0 \) for \( 0 \leq i \leq j \), let \( b_{j+1} = a_{j+1} + 1 \), and let \( b_i = a_i \) for \( i > j + 1 \).

By construction, \( 0 \leq b_i \leq q - 1 \) for all \( i \), so \( b \) is the q-ary expansion of some number. The contribution from indices greater than \( j + 1 \) is the same. By the geometric sum, the value of the expansion \( b \) is one more than the value of the expansion \( a \).

4.15. By induction on \( k \), the known weights \( \{1, 3, \ldots, 3^k-1\} \) suffice to measure the weights 1 through \((3^k-1)/2\) on a balance scale. Basis Step: For \( k = 1 \), the single known weight 1 balances 1. Induction Step: Suppose that the statement holds when the parameter is \( k \). When we add \( 3^k \) as the \( k + 1 \)th known weight, we can still weigh the numbers \( 1, \ldots, (3^k - 1)/2 \) as done previously, without using the new weight.

The new weight by itself can balance \( 3^k \). We can balance \( 3^k - 1, \ldots, 3^k - (3^k - 1)/2 \) by putting the new weight on the light side of earlier configurations. Since \( 3^k - (3^k - 1)/2 = (3^k + 1)/2 \), this fills the gap between the earlier configurations and \( 3^k \). We can balance weights \( 3^k + 1, \ldots, 3^k + (3^k - 1)/2 \) by putting the new weight \( 3^k \) on the heavy side of earlier configurations. Since \( 3^k + (3^k - 1)/2 = (3^{k+1} - 1)/2 \) and we left no gaps, we have balanced all the desired weights. Thus the claim holds also for \( k + 1 \).

Comment: The largest weight balanced by \( k \) weights occurs when all the known weights are on the same side. This value is \( \sum_{i=0}^{k-1} 3^i \), which by the geometric sum equals \((3^k - 1)/2\).

4.16. Using weights \( w_1 \leq \cdots \leq w_n \) on a two-pan balance, where \( S_j = \sum_{i=1}^{j} w_i \) every integer weight from 1 to \( S_n \), can be weighed if and only if \( w_1 = 1 \) and \( w_{j+1} - 1 \leq S_j + 1 \) for \( 1 \leq j < n \). For sufficiency, we use induction on \( n \). When \( n = 1 \), the condition forces \( w_1 = 1 \), and the weight 1 can be balanced. For the induction step, consider \( n > 1 \), and suppose that the condition is sufficient for \( n - 1 \) weights. For \( 1 \leq i \leq n \), the induction hypothesis implies that we can weigh \( i \) using \( \{w_1, \ldots, w_{i-1}\} \). With \( w_n \) also available, we can also weigh \( w_n - i \) and \( w_n + i \), so we can weigh every weight from \( w_n - S_{n-1} \) to \( w_n + S_{n-1} = S_n \) using \( \{w_1, \ldots, w_n\} \). Since \( w_n - S_{n-1} \leq S_{n-1} + 1 \) by hypothesis, we can weigh every weight up to \( S_n \).

For necessity, suppose we can balance all weights from 1 to \( S_n \). The second largest possibility is \( S_n - w_1 \), required to be \( S_n - 1 \), so \( w_1 = 1 \). If \( w_{j+1} > 2S_j + 1 \) for some \( j \), then we know \( W = S_n - 2S_j - 1 \); we claim that \( W \) cannot be weighed. The largest weight achievable without putting all of \( \{w_{j+1}, \ldots, w_n\} \) in one pan is \( S_n - w_{j+1} < W \), but the smallest weight achievable using all of \( \{w_{j+1}, \ldots, w_n\} \) in one pan is \( S_n - 2S_j \), which exceeds \( W \).
4.17. Winning positions in Nim. We prove by strong induction on the total number of coins that a position is winning (for the second player who leaves it) if and only if for all $j$, the number of pile-sizes whose binary representation has a 1 in the $j$th place is even. By $j$th place we mean contributions of $2^j$. Let $s_j$ be the number of pile-sizes whose binary representation has a 1 in the $j$th place. Let (6) denote the condition that each $s_j$ is even.

The condition (6) holds when the (starting) number of coins is 0. Since Player 1 cannot move, we view this as Player 2 having taken the last coin. This is the only position with 0 coins. For every $j$, we have $s_j = 0$. Thus the position is winning and satisfies (6).

When the (starting) number of coins is larger, suppose first that some of the $s_j$'s are odd. We show that some amount can be taken from some pile to leave them all even. By the induction hypothesis, Player 1 thus leaves a winning position, and therefore Player 2 loses. To find a winning move, let $J$ be the largest $j$ such that $s_j$ is odd, and let $S = \{ j : s_j \text{ is odd} \}$. Since $s_j$ is odd, some pile-size has a 1 in position $J$. We want to take coins from this pile $P$ change its binary representation $b$ in the positions indexed by $S$.

For each position $j \in S$ where $b$ has a 1 in position $j$, we take $2^j$ coins from $P$. For each position $j \in S$ where $b$ has a 0 in position $j$, we add $2^j$ coins to $P$. Because $\sum_{j=1}^{J} 2^j$ is less than $2^J$, the total of these adjustments is a positive number less that the size of $P$, so we have obtained a legal move that achieves (6). As we have remarked, the induction hypothesis implies that Player 1 wins.

When each $s_j$ is even, every move changes the binary representation of one pile. Thus it changes the parity of some $s_j$, and therefore Player 1 cannot produce a smaller position that satisfies (6). By applying the method described above, Player 2 can now produce a position satisfying (6). By the induction hypothesis, such a position is winning, so the original position is a winning position to leave.

4.18. Exponentiation to a positive odd power is a strictly increasing function. We use induction on $k$ to prove this for the power $2k - 1$. Basis step ($k = 1$). Here exponentiation is the identity function: $x < y$ implies $x < y$.

Induction step. Suppose that exponentiation to the power $2k - 1$ is strictly increasing. Then $x^{2k-1} < y^{2k-1}$ when $x < y$. If $0 < x < y$, then $0 < x^2 < y^2$, and multiplying the two inequalities yields $x^{2k-1} < y^{2k-1}$. If $x < 0 \leq y$, then $x^{2k-1}$ is negative and $y^{2k-1}$ is nonnegative, so $x^{2k-1} < y^{2k-1}$. If $x < 0 \leq y$, then $0 \leq -y < -x$, and we have proved that $-y^{2k-1} < -x^{2k-1}$. Since an odd power of $-1$ is $-1$, this yields $-x^{2k-1} > y^{2k-1}$, and thus $x^{2k-1} < y^{2k-1}$.

Solutions to $x^n = y^n$. All pairs with $x = y$ are solutions. When $n$ is odd, the exponentiation is strictly increasing, and hence in this case there are no other solutions. When $n$ is even, the solutions are $x = \pm y$. To show that there are no other solutions, it suffices to show that exponentiation to the $n$th power is injective from the set of positive real numbers to itself. This follows by an induction like that above.

4.19. For $k \in \mathbb{N}$, the only solution to $\sum_{j=0}^{2k} x^{2k-j} = 0$ is $(x, y) = (0, 0)$. For $(x, y)$ satisfying the equation, multiplying both sides by $(x - y)$ yields $x^{2k+1} - y^{2k+1} = 0$. Since exponentiation to an odd power is injective, this requires $x = y$. Among solution pairs with $x = y$, the equation reduces to $(2k+1)x^{2k} = 0$. The only solution of this is $x = 0$, so the only solution of the original equation is $(x, y) = (0, 0)$, which indeed works.

4.20. Properties of the map $f: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ defined by $f(x, y) = (ax - by, bx + ay)$, where $a, b$ are fixed parameters with $a^2 + b^2 \neq 0$.

a) $f$ is a bijection. As proved in Example 4.12, the function $f: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ defined by $f(x, y) = (ax + by, cx + dy)$ is a bijection if and only if $ad - bc \neq 0$. In this problem, the values taken by $a, b, c, d$ are $-a, b, a, b$ respectively, and hence $ad - bc$ becomes $a^2 + b^2$, which by hypothesis is non-zero. Hence the function given here is a bijection, by Example 4.12.

To prove directly that $f$ is a bijection, we show directly that $f$ is both surjective and injective, meaning that for every element $(r, s)$ in the target there is exactly one element $(x, y)$ in the domain such that $f(x, y) = (r, s)$. To prove injectivity, suppose $f(x, y) = f(x', y')$. This requires $ax - by = ax' - by'$ and $bx + ay = bx' + ay'$. Subtracting $b$ times the first equation from the second yields $(a^2 + b^2)y = (a^2 + b^2)y'$, and hence $y = y'$, since $a^2 + b^2 \neq 0$. Similarly, adding $a$ times the first equation to $b$ times the second yields $(a^2 + b^2)x = (a^2 + b^2)x'$, so $x = x'$. We have proved that if element $(x, y), (x', y')$ of the domain have the same image, then they must be the same element (no collapsing).

To prove surjectivity, we show that every element $(r, s)$ in the target is the image of some element of the domain. A suitable element $(x, y)$ must satisfy $r = ax - by$ and $s = bx + ay$. Because $a^2 + b^2 \neq 0$, we can solve this system of equations to find such a pair $(x, y)$. The formula for $(x, y)$ appears in part (b).

b) Formula for $f^{-1}$. When $f$ is a bijection, the inverse function $f^{-1}$ gives for each element of the target the unique element of the domain that maps to it. Computing the inverse function may allow us to prove surjectivity and injectivity simultaneously. In this example, the inverse image of the element $(r, s)$ in the target is the set of solutions $(x, y)$ to the system $r = ax - by$ and $s = bx + ay$. Because $a^2 + b^2 \neq 0$, there is a unique solution (existence implies surjectivity of $f$, uniqueness implies injectivity of $f$). The unique solution of the system is $x = \frac{r + bs}{a^2 + b^2}$ and $y = \frac{-r + as}{a^2 + b^2}$. Hence the inverse function is $f^{-1}(r, s) = \left( \frac{r + bs}{a^2 + b^2}, \frac{-r + as}{a^2 + b^2} \right)$. 


c) A geometric interpretation of $f$ when $a^2 + b^2 = 1$. This uses the distance from the origin to a point $(x, y)$ in $\mathbb{R}^2$, defined to be $\sqrt{x^2 + y^2}$. The distance from the origin to the image point $(ax - by, bx + ay)$ is $\sqrt{(a^2 + b^2)(x^2 + y^2)}$, which equals the distance from the origin to $(x, y)$ if $a^2 + b^2 = 1$. Hence the effect of $f$ on the vector $(x, y)$ is to rotate it around the origin. Every vector is rotated through the same angle; in particular, when $a = 0$ and $b = 1$, the function rotates everything by 90 degrees counterclockwise. Proving that every vector is rotated by the same amount relies on knowing that the angle between two vectors is determined by their dot product divided by the product of their lengths. Considering the old vector $(x, y)$ and the new vector $(ax - by, bx + ay)$, their dot product is $a^2 x^2 - b^2 xy + bx^2 + ay^2 = a(x^2 + y^2)$, and the product of their lengths is $x^2 + y^2$. The ratio is $a$, independent of the element $(x, y)$, so every point in the plane is rotated by the same amount.

4.21. The number of subsets of $[n]$ with odd size equals the number of subsets of $[n]$ with even size, where $n \in \mathbb{N}$, bijectively.

Let $A$ be the collection of even subsets of $[n]$, and let $B$ be the collection of odd subsets. For each $x \in A$, define $f(x)$ as follows:

$$f(x) = \begin{cases} f(x) - \{n\} & \text{if } n \in x \\ f(x) \cup \{n\} & \text{if } n \notin x \end{cases}$$

By this definition, $|A|$ and $|f(x)|$ differ by one, so $f(x)$ is a set of odd size, and $f$ maps $A$ to $B$.

We claim that $f$ is a bijection. Consider distinct $x, y \in A$. If both contain or both omit $n$, then $f(x)$ and $f(y)$ agree on whether they contain $n$ but differ otherwise. If exactly one of $x, y$ contains $n$, then exactly one of $\{f(x), f(y)\}$ contains $n$. Thus $x \neq y$ implies $f(x) \neq f(y)$, and $f$ is injective. If $x \subseteq B$, then flipping whether $n$ is present in $x$ yields a subset $x$ such that $f(x) = x$, so $f$ is also surjective. Thus $f$ is a bijection.

When $n = 0$, there is one even subset and no odd subset. The bijection fails because $\emptyset = \emptyset$ and there is no element $n$ to change.

Alternatively, one can define a function $g: B \to A$ by the same rule used to define $f$ (switching the domain and target), and observe that $g \circ f$ is the identity function on $A$ and $f \circ g$ is the identity function on $B$. This implies that $g$ is the inverse of $f$ and thus that $f$ is a bijection and $|A| = |B|$. Without knowing $|A| = |B|$, it does not suffice to show that only one of the compositions is the identity.

4.22. The formula $f(x) = \frac{2x - 1}{2(x-1)}$ defines a bijection from $(0, 1)$ to $\mathbb{R}$.

$f$ is injective. Suppose that $f(x) = f(y)$. From $\frac{2x - 1}{2(x-1)} = \frac{2y - 1}{2(y-1)}$, we obtain $(2x - 1)(1 - y) = (2y - 1)(1 - x)$, which simplifies to $2x^2 - 2x - 4xy^2 = 2x^2 - 2x - 4x^2y$ and then $2(y^2 - x^2) - 2(y - x) = 4xy(y - x)$. If $y \neq x$, then we can divide by $2(y - x)$ to obtain $y + x - 1 = 2xy$. Rewriting this as $-xy = (x - 1)(y - 1)$ makes it clear that there is no solution when $x, y \in (0, 1)$, since the left side is negative and the right side is positive.

$f$ is surjective. Suppose that $f(x) = b$; we solve for $x$ to obtain $x \in (0, 1)$ such that $f(x) = b$. Observe that $b = 0$ is achieved by $x = 1/2$, so we may assume that $b \neq 0$. Clearing fractions leads to $xb - x^2b = x - 1/2$, or $bx^2 + (1 - b)x - 1/2 = 0$. The quadratic formula yields

$$x = \frac{b - 1 \pm \sqrt{b^2 + 1}}{2b}.$$

The magnitude of the square root is larger than $|b|$. Therefore, choosing the negative sign in the numerator yields a negative $x$, which is not in the domain of $f$. We therefore choose the positive sign.

If $b > 0$, then the square root is less than $b + 1$, and we obtain $x < \frac{b - 1 \pm \sqrt{b^2 + 1}}{2b} = 1$. Also the square root is bigger than $1$, so $x > 0$. If $b < 0$, then let $b' = -b$. The formula for $x$ becomes $x = \frac{b + 1 - \sqrt{b^2 + 1}}{2b}$, where $b' > 0$. The square root is strictly between 1 and $b' + 1$, so $x$ is strictly between 1/2 and 0. In each case, we have found $x$ in the domain of $f$, such that $f(x) = b$.

4.23. Functions from $\mathbb{R}$ to $\mathbb{R}$.

a) $f(x) = x^3 - x + 1$. This function is surjective, like all cubic polynomials, but it is not injective, since $f(1) = f(-1) = 1$. The formula defines a bijection from $S$ to $S$, where $S = [1, \infty)$.

b) $f(x) = \cos(\pi x/2)$. This function is not surjective, since the value of cosine is always between $-1$ and 1. Also it is not injective; the value at every odd integer is 0. Nevertheless, when the domain and target are restricted to the interval $[0, 1]$, $f$ is a bijection.

4.24. If $f$ and $g$ are surjective functions from $\mathbb{Z}$ to $\mathbb{Z}$, then the pointwise product of $f$ and $g$ need not be surjective. If $f$ and $g$ are defined by $f(x) = x$ and $g(x) = x$, then $f$ and $g$ are surjective, but $fg(x) = x^2$, and $fg$ does not map onto any negative integer. (Many other examples can be given.)

4.25. Formulas defining surjections from $\mathbb{N} \times \mathbb{N}$ to $\mathbb{N}$.

a) $f(a, b) = a + b$—NO. When $a, b \in \mathbb{N}$, $a + b \geq 1$, so the image does not contain 1.

b) $f(a, b) = ab$—YES. For $n \in N$, $f(n, 1) = n$, so $n$ is in the image.

c) $f(a, b) = ab(b + 1)/2$—YES. For $n \in \mathbb{N}$, $f(n, 1) = n$, so $n$ is in the image.

d) $f(a, b) = (a+1)(b+1)/2$—NO. When $a, b \in \mathbb{N}$, $(a+1)(b+1)/2 \geq 2$, so the image does not contain 1.
4.26. If there are positive constants \( c, \alpha \) such that, for all \( x, y \in \mathbb{R} \), 
\[ |f(x) - f(y)| \geq c|x - y|^\alpha, \] 
then \( f \) is injective. If \( f \) is not injective, then there are distinct numbers \( x, y \) such that \( f(x) = f(y) \). Since \( |x - y|^\alpha > 0 \), this contradicts the hypothesized condition.

4.27. Surjectivity and injectivity of polynomials. Consider an arbitrary quadratic polynomial, \( ax^2 + bx + c \), with \( a \neq 0 \). As in the derivation of the quadratic formula, we write \( ax^2 + bx + c = a(x + b/(2a))^2 - b^2/(4a) \). Since \( (x + b/(2a))^2 \geq 0 \), the value of the polynomial cannot be smaller than \( c - b^2/(4a) \) if \( a > 0 \), and it cannot be larger than \( c - b^2/(4a) \) if \( a < 0 \). Hence the function is not surjective. (Comment: Since equality holds when \( x = -b/(2a) \), this is where the extreme value of the quadratic occurs, and the extreme value equals \( c - b^2/(4a) \); this is consistent with problem 1 of homework 1.)

The polynomial \( x^3 - x + 1 \) is not injective, since it has the value 1 at more than one place (at \( x = 0 \) or \( x = \pm 1 \)). Until Chapter 4, we can only sketch a proof that this function is surjective. Note that \( x(x^2 - 1) + 1 \) is increasing when \( x > 1 \), because \( y > x > 1 \) implies \( y^2 - 1 > x^2 - 1 \), and then \( y(y^2 - 1) + 1 > x(x^2 - 1) + 1 \). Similarly, it is increasing when \( x < -1 \). If we believe in continuity and in the values getting arbitrarily far from 0, then the function is surjective.

4.28. The cubic polynomial defined by \( ax^3 + bx^2 + cx + d \) is injective if and only if \( b^2 - 3ac < 0 \).

The formula for the value of the general cubic polynomial at \( x \) is \( f(x) = ax^3 + bx^2 + cx + d \); these coefficients are known. Since multiplying the function by \( -1 \) doesn’t affect injectivity and quadratics are not injective, we may assume that \( a > 0 \).

We use a change of variables to reduce the problem to polynomials \( h \) of the form \( h(y) = y^3 + ry + d' \). We determine constants \( s, t \) so that substituting \( x = s(y + t) \) expresses \( ax^3 + bx^2 + cx + d \) as \( y^3 + ry + d' \), where \( r, d' \) are constants. That is,
\[ as^3(y + t)^3 + bs^2(y + t)^2 + cs(y + t) + d = y^3 + ry + d'. \]

Since polynomials are equal when their coefficients are equal, we set \( as^3 = 1 \) for the coefficient of \( y^3 \) and \( 3as^3t + bs^2 = 0 \) for the coefficient of \( y^2 \). This yields \( s = 1/(a)^{1/3} \) and \( t = -b/(3as) \). The resulting coefficient \( r \) for \( y^1 \) is \( 3as^3y^2 + 2bs^2y + cs \), which can be computed using the formulas for \( s \) and \( t \).

Let \( g(y) = s(y + t) \). When \( a \neq 0 \) and \( s, t \) are defined above, \( g \) is a bijection from \( \mathbb{R} \) to \( \mathbb{R} \) and \( h = f \circ g \). Thus \( f = h \circ g^{-1} \), and \( f \) will be injective if and only if \( h \) is injective.

The constant \( d' \) in the formula for \( h \) does not affect injectivity. Replacing it by 0 merely shifts the images. It suffices to consider \( y^3 + ry \). If \( y^3 + ry = z^3 + rz \) for some distinct \( y, z \), then dividing by \( y - z \) yields \( y^2 + yz + z^2 = z \).

If \( r \) is negative, then \( (y, z) = (0, \sqrt[3]{-r}) \) is a solution, and the function is not injective. If \( r = 0 \), then there is no solution with \( y \neq z \) (since cubing is injective). If \( r \) is positive, then again there is no solution, because \( y^2 + yz + z^2 \) is never negative, which follows from \( y^2 + z^2 \geq 2|y||z| \) (AGM Inequality).

Thus \( h \) is injective if and only if \( r \geq 0 \), and this determines whether \( f \) is injective. Since we have assumed that \( a > 0 \), also \( s > 0 \). Canceling \( s \) from the formula for \( r \) yields \( 3as^2t^2 + 2bs^2t + c = 0 \). It suffices to consider the sign of this. From \( 3as^2t^2 + 2bs^2t = 0 \), we obtain \( t = -b/(3as) \). Thus we are interested in the sign of \( b^2/(3as) - 2b^2/(3as) + c \). This is positive if and only if \( b^2 - 3ac < 0 \).

Comment. The methods of calculus in Part IV would enable us to observe that a differentiable function from \( \mathbb{R} \) to \( \mathbb{R} \) is injective if and only if its derivative is never 0. The derivative of \( ax^3 + bx^2 + cx + d = 3ax^2 + 2bx + c \). This is never 0 if and only if \( 3ax^2 + 2bx + c = 0 \) has no solution. By the quadratic formula, the condition for this is \( 4b^2 - 12ac < 0 \), which is the same answer obtained above. This argument is shorter because it relies on the work of defining and studying the derivative.

4.29. Properties of three functions \( f, g, h \) mapping \( \mathbb{R} \) to \( \mathbb{R} \).

\[ f(x) = x/(1 + x^2), \quad g(x) = x^2/(1 + x^2), \quad h(x) = x^3/(1 + x^2). \]

\( a \) The functions \( f \) and \( g \) are not injective, but \( h \) is injective. Since \( g(x) = g(-x) \) for all \( x, g \) is not injective. For \( f \), this is less obvious. If we do not see immediately something like \( f(0) = f(1/2) \), then we try to prove that \( f \) is injective. Setting \( f(x) = f(y) \) and assuming \( x \neq y \) yields \( x + yx^2 = y + yx^2 \), which simplifies to \( x - y = xy(x - y) \) and reduces to \( 1 = xy \). When \( x \neq 0 \), we have \( f(x) = f(y) \) if and only if \( xy = 1 \).

For \( h \), again we set \( h(x) = h(y) \) and assume that \( x \neq y \). We obtain \( x^3 + x^3y^3 = y^3 + y^3x^3 \), which reduces to \( x + xy + y = -x^3 \) after we rearrange and divide by \( x - y \). Rewriting this as \( x^2(1 + y^3) + xy + y^3 = 0 \) yields a quadratic equation for \( x \) in terms of \( y \). Since \( b^2 - 4ac = y^3 - 2y^4(1 + y^3) < 0 \), there is no solution for \( x \). Hence there are no distinct \( x, y \) with \( h(x) = h(y) \).

\( b \) The functions \( f \) and \( g \) are not surjective. For all \( x, g(x) > 0 \). Further, \( \left( 1/3 \right)^{1/3} = 1/(3 \sqrt[3]{3}) < 1 \) for \( x \neq 0, \) so always \( 0 \leq g(x) < 1 \).

Also \( f(x) < 1 \) always. If \( x < 0 \), then \( f(x) < 0 \). If \( 0 \leq x < 1 \), then \( x/(1 + x^2) < x < 1 \). If \( x \geq 1 \), then \( x/(1 + x^2) = 1/(1 + x^2) < 1 \).
c) The graphs. Note that \( f(-x) = -f(x) \), \( g(-x) = g(x) \), \( h(-x) = -h(x) \). All are 0 at 0. For large \( x \), they are asymptotic to 0, 1, \( x \), respectively.

**4.30.** If \( a, b, c, d \) are given real numbers, and \( f: \mathbb{R}^2 \to \mathbb{R}^2 \) is defined by \( f(x, y) = (ax + by, cx + dy) \), then \( f \) is injective if and only if \( f \) is surjective. If \( ad - bc \neq 0 \), then the system \( ax + by = r \) and \( cx + dy = s \) has a unique solution pair \((x, y)\) for each choice of \((r, s)\). This implies that \( f \) is a bijection. Thus, when \( ad - bc \neq 0 \), \( f \) is both injective and surjective.

In the remaining case, we have \( ad - bc = 0 \). Given \( f(x, y) = (r, s) \), we can multiply the first equation by \( c \) and the second by \( a \) to obtain \( acx + bcy = cr \) and \( acx + ady = as \). Because \( ad = bc \), the left sides of these two equations are equal. Hence \((r, s)\) belongs to the image if and only if \( cr = as \). This does not include all of \( \mathbb{R}^2 \), so \( f \) is not surjective. Also, \( ad - bc = 0 \) implies that increasing \( x \) by \( b \) and decreasing \( y \) by \( a \) does not change \( ax + by \) or \( cx + dy \). Hence for each \((r, s)\) in the image, there are infinitely many choices of \((x, y)\) such that \( f(x, y) = (r, s) \).

By considering the two cases, we have that \( f \) is surjective if and only if \( ad - bc \neq 0 \), and that \( ad - bc \neq 0 \) if and only if \( f \) is injective.

**4.31.** If \( f: A \to B \) is an increasing function, then \( f^{-1} \) is an increasing function. The contrapositive of the statement \( x < y \Rightarrow f(x) < f(y) \) is the statement \( f(x) \geq f(y) \Rightarrow x \geq y \). Writing \( u = f(x) \) and \( v = f(y) \) converts this to \( u \geq v \Rightarrow f^{-1}(u) \geq f^{-1}(v) \).

**4.32.** When \( F \) is a field, negation \((f)\) defines a bijection from \( F \) to itself; and reciprocal \((g)\) defines a bijection from \( \{0\} \) to itself. The field axioms imply that every element of \( F \) has a unique additive inverse, and every nonzero element of \( F \) has a unique multiplicative inverse. Given \( y \) in the target, these inverses are the unique elements \( x' \) and \( x \) such that \( f(x') = -x' = y \) and \( g(x) = x^{-1} = y \) (the latter applies only for \( y \neq 0 \)).

**4.33.** Composition of injections and surjections. Let \( f: A \to B \) and \( g: B \to C \), so \((g \circ f)(x) = g(f(x)) \) for all \( x \in A \).

a) The composition of two injections is an injection. Assume that \( f \) and \( g \) are injective. Suppose that \( (g \circ f)(x) = (g \circ f)(x') \), i.e., \( g(f(x)) = g(f(x')) \). Since \( g \) is injective, this implies \( f(x) = f(x') \). Since \( f \) is injective, this in turn implies \( x = x' \). Hence \((g \circ f)(x) = (g \circ f)(x')\) implies \( x = x' \), and \( g \circ f \) is injective.

Alternatively, consider the contrapositive. For \( x, x' \in A \) with \( x \neq x' \), we have \( f(x) \neq f(x') \) because \( f \) is injective, and then \( g(f(x)) \neq g(f(x')) \) because \( g \) is injective. Thus \( x \neq x' \) implies \((g \circ f)(x) \neq (g \circ f)(x')\), so \( g \circ f \) is injective.

b) The composition of two surjections is a surjection. Assume that \( f \) and \( g \) are surjective. Let \( z \) be an arbitrary element of \( C \). Since \( g \) is surjective, there is an element \( y \) in \( B \) such that \( g(y) = z \). Since \( f \) is surjective, there is an element \( x \in A \) such that \( f(x) = y \). Hence we have found an element of \( A \), namely \( x \), such that \((g \circ f)(x) = z \), and \( g \circ f \) satisfies the definition of a surjective function.

c) The composition of two bijections is a bijection. By (a) and (b), \( g \circ f \) is both injective and surjective and hence is a bijection, by definition.

d) If \( f: A \to B \) and \( g: B \to C \) are bijections, then \((g \circ f)^{-1} = f^{-1} \circ g^{-1} \). By part (c), \( g \circ f \) is a bijection from \( A \) to \( C \). Thus \( g \circ f \) is invertible, and the inverse is defined to be the function that yields the identity function on \( A \) when composed with \( g \circ f \). Let \( I_A \) and \( I_B \) denote the identity functions on \( A \) and \( B \). Letting \( h = f^{-1} \circ g^{-1} \), we use the associativity of composition to obtain \( h \circ (g \circ f) = f^{-1} \circ (g^{-1} \circ g) \circ f = f^{-1} \circ I_B \circ f = f^{-1} \circ I_B \circ f = I_A \). Thus \( h \) is the inverse of \( g \circ f \).

One can also argue more explicitly that \((g \circ f)^{-1} \) and \( f^{-1} \circ g^{-1} \) have the same domain and target and have the same value at each element of the domain, so they are the same function.

**4.34.** Composition of functions. Suppose that \( f: A \to B \), \( g: B \to C \), and \( h = g \circ f \).

a) If \( h \) is injective, then \( f \) is injective—TRUE. If \( f \) is not injective, then there exist two distinct elements \( x, y \in A \) such that \( f(x) = f(y) \). Since \( g \) is a function, this implies that \( g(f(x)) = g(f(y)) \). Since \( h = g \circ f \), we have obtained distinct elements \( x, y \in A \) such that \( h(x) = h(y) \), and hence \( h \) is not injective. We have proved the contrapositive, so the implication is true.

b) If \( h \) is injective, then \( g \) is injective—FALSE. Let \( A = \{1\} \), \( B = \{a, b\} \), and \( C = \{a\} \). Define \( f(1) = a \) and \( g(a) = g(b) = a \). Both \( f \) and \( h \) are injective, but \( g \) is not injective.

c) If \( h \) is surjective, then \( f \) is surjective—FALSE. Let \( A = \{1, 2\} \), \( B = \{a, b\} \), and \( C = \{a\} \). Define \( f(1) = f(2) = a \) and \( g(a) = g(b) = a \). Then \( h(1) = h(2) = a \), and \( h \) is surjective, but \( f \) is not surjective.

d) If \( h \) is surjective, then \( g \) is surjective—TRUE. If \( z = h(x) \), then \( z = g(f(x)) \). Thus the image of \( g \) contains the image of \( h \), which the hypothesis says is all of \( C \).

**4.35.** Composition of functions. Suppose \( f: A \to B \) and \( g: B \to A \).

a) If \( f(g(y)) = y \) for all \( y \in B \), then \( f \) need not be a bijection. For each \( y \in B \), \( y \) is the image under \( f \) of some element of \( A \), namely \( g(y) \). This