A1. How many positive integers N satisfy all of the following three conditions?

(i) N is divisible by 2020.
(ii) N has at most 2020 decimal digits.
(iii) The decimal digits of N are a string of consecutive ones followed by a string of consecutive zeros.

Answer. $504 \cdot 1009 = 508536$.

Solution 1. A positive integer N satisfying (iii), with j ones followed by k zeros, has the form

$$N = \frac{10^j - 1}{9} \cdot 10^k$$

where $j \geq 1$, $k \geq 0$, and $j + k \leq 2020$. Note that $2020 = 20 \cdot 101$, so to satisfy (i) the integer N must be divisible by 101 and end in at least two zeros (so $k \geq 2$). If 101 divides N then 101 divides $M = 10^j - 1$. A quick check shows that $M \equiv 0, 9, 99, 90 \mod 101$ when $j \equiv 0, 1, 2, 3 \mod 4$. Consequently, 4 must divide j. (One can see directly that the conditions $k \geq 2, 4|j$ are necessary and sufficient by noting that 101 divides 1111 but not 1, 11 or 111.)

If $j = 4m$, then for N to satisfy (ii) also, we need $2 \leq k \leq 2020 - 4m$, for a total of $2019 - 4m$ possible values of k. The total number of integers N satisfying all the conditions is therefore

$$\sum_{m=1}^{504} (2019 - 4m) = 2019 \cdot 504 - 4 \cdot \frac{504 \cdot 505}{2} = 504 \cdot (2019 - 1010) = 504 \cdot 1009 = 508536.$$

Solution 2. As in the first solution, it is straightforward to show that the acceptable numbers N are those for which there are at most 2020 decimal digits, consisting of j ones with $4|j$ followed by k zeros with $k \geq 2$. By introducing additional "phantom" digits z at the beginning of the number, we can convert it to a string of length exactly 2020 of the form $zzz \cdots z111 \cdots 1000 \cdots 0$. We now show that the set of such strings is in bijective correspondence with a set of size $\binom{1009}{2} = 508536$. To see this, remove the final two zeros from the string, and group the remaining 2018 positions in the string into consecutive pairs. Then any choice of 2 of these 1009 pairs corresponds to a unique string of the desired form, as follows. If the two chosen pairs have an even number of pairs between them, put a z in each position before the first chosen pair, put 11 for each of the chosen pairs and all pairs in between, and put a 0 in each position after the second chosen pair, for example:

$$\underbrace{xx}\ \underbrace{xx}_{\text{Choose}}\ \underbrace{xx}\ \underbrace{xx}\ \underbrace{xx}_{\text{Choose}}\ \underbrace{xx} \cdots \mapsto zz \mid 11 \mid 11 \mid 11 \mid 11 \mid 00 \ \cdots .$$

If the two chosen pairs are separated by an odd number of pairs, do the same except for replacing the chosen pairs by $z1$ and 10, respectively, for example:

$$\underbrace{xx}\ \underbrace{xx}_{\text{Choose}}\ \underbrace{xx}\ \underbrace{xx}_{\text{Choose}}\ \underbrace{xx} \cdots \mapsto zz \mid z1 \mid 11 \mid 10 \mid 00 \ \cdots .$$

Note that in either case, the resulting number of ones is divisible by 4. Erasing the digits z and restoring the two zeros that were removed at the end of the string, we get every acceptable number N exactly once from some choice of 2 of the 1009 consecutive pairs.

Solution 3. As in the first solution, the positive integers N satisfying conditions (i) and (iii) have j ones followed by k zeros, with $4|j$, $j \geq 1$, and $k \geq 2$. Thus if we let b_m be the number of m-digit positive integers with these properties, we have the generating function

$$\sum_{m=0}^{\infty} b_m x^m = \frac{x^4}{1-x^4} \cdot \frac{x^2}{1-x}\ .$$

Hence the generating function for the number $B_m = \sum_{k \leq m} b_k$ of such integers with at most m digits is

$$\sum_{m=0}^{\infty} B_m x^m = \frac{1}{1-x} \cdot \sum_{m=0}^{\infty} b_m x^m = \frac{x^6}{(1-x)^2(1-x^4)} = \frac{x^6(1+x+x^2+x^3)^2}{(1-x^4)^3}.$$

Because

$$\frac{1}{(1-y)^3} = \sum_{k=0}^{\infty} \binom{k+2}{2} y^k$$

and $x^6(1+x+x^2+x^3)^2 = x^6 + 2x^7 + 3x^8 + 4x^9 + 3x^{10} + 2x^{11} + x^{12}$, we can read off the answer

$$B_{2020} = \binom{504}{2} + 3\binom{505}{2} = 126756 + 381780 = 508536.$$

A2. Let k be a nonnegative integer. Evaluate

$$\sum_{j=0}^{k} 2^{k-j} \binom{k+j}{j}.$$

Answer. $2^{2k} = 4^k$.

Solution 1. Consider lattice paths of length $2k+1$ that begin from the origin, and where each step is either of the form $(x,y) \mapsto (x+1,y)$ or $(x,y) \mapsto (x,y+1)$; thus these paths will be in the first quadrant. There are 2^{2k+1} such paths, and by reflective symmetry, exactly half of them reach the line $x = k+1$ (the other half reach the line $y = k+1$). So there are 2^{2k} such paths that have at least $k+1$ "right steps" among their $2k+1$ steps.

On the other hand, any of those paths first touches the line $x = k+1$ at some point $(k+1, j)$, which means that the previous step came from (k, j). There are exactly $\binom{k+j}{j}$ paths from $(0,0)$ to (k,j), and 2^{k-j} possible paths after $(k+1,j)$. Summing over the possible values for j, which range from 0 to k, gives the sum from the problem statement, and thus that sum equals 2^{2k}.

Solution 2. Let $S(k)$ denote the sum from the problem statement. Then using basic properties of binomial coefficients, one finds that for $k \geq 0$,

$$\begin{aligned}
S(k+1) &= \sum_{j=0}^{k+1} 2^{k+1-j} \binom{k+1+j}{j} \\
&= \sum_{j=0}^{k+1} 2^{k+1-j} \left(\binom{k+j}{j} + \binom{k+j}{j-1} \right) \\
&= 2 \sum_{j=0}^{k+1} 2^{k-j} \binom{k+j}{j} + \sum_{j=0}^{k} 2^{k-j} \binom{k+j+1}{j} \\
&= 2\, S(k) + \binom{2k+1}{k+1} + \frac{1}{2} \left(S(k+1) - \binom{2k+2}{k+1} \right) \\
&= 2\, S(k) + \frac{1}{2}\, S(k+1) + \binom{2k+1}{k+1} - \frac{1}{2} \binom{2k+2}{k+1} \\
&= 2\, S(k) + \frac{1}{2}\, S(k+1).
\end{aligned}$$

Therefore $S(k+1) = 4\, S(k)$, and since $S(0) = 1$, by induction we have $S(k) = 4^k$ for all k.

Solution 3. Note that the desired sum

$$\sum_{j=0}^{k} 2^{k-j} \binom{k+j}{j} = 2^k \sum_{j=0}^{k} 2^{-j} \binom{k+j}{k}$$

is the coefficient of x^k in the polynomial

$$\begin{aligned}
P_k(x) &= 2^k \sum_{j=0}^{k} 2^{-j} (1+x)^{k+j} \\
&= 2^k (1+x)^k \sum_{j=0}^{k} \left(\frac{1+x}{2} \right)^j \\
&= 2^k (1+x)^k \, \frac{1 - \left(\frac{1+x}{2}\right)^{k+1}}{1 - \frac{1+x}{2}} \\
&= 2^{k+1} (1+x)^k \, \frac{1 - \left(\frac{1+x}{2}\right)^{k+1}}{1-x} \\
&= \left[2^{k+1} (1+x)^k - (1+x)^{2k+1} \right] \frac{1}{1-x} \\
&= \left[2^{k+1} (1+x)^k - (1+x)^{2k+1} \right] (1 + x + x^2 + \ldots).
\end{aligned}$$

But this coefficient can also be expressed as

$$2^{k+1} \sum_{j=0}^{k} \binom{k}{j} - \sum_{j=0}^{k} \binom{2k+1}{j} = 2^{k+1} \cdot 2^k - \frac{1}{2} \cdot 2^{2k+1} = 2^{2k} = 4^k,$$

as claimed.

A3. Let $a_0 = \pi/2$, and for $n \geq 1$, let $a_n = \sin(a_{n-1})$. Determine whether

$$\sum_{n=1}^{\infty} a_n^2$$

converges.

Answer. The series diverges.

Solution 1. Note that $a_1 = 1$; we now show by induction on n that for all $n \geq 1$, $a_n \geq 1/\sqrt{n}$. Note that on the interval $(0, \pi/2)$, $\sin(x)$ is increasing and $\sin x > x - x^3/6$ by Taylor's theorem with remainder (because the fifth derivative, $\cos x$, is positive on the interval). In particular, from the induction hypothesis,

$$a_{n+1} = \sin a_n \geq \sin\left(\frac{1}{\sqrt{n}}\right) > \frac{1}{\sqrt{n}} - \frac{1}{6}\left(\frac{1}{\sqrt{n}}\right)^3.$$

On the other hand,

$$\begin{aligned}\frac{1}{\sqrt{n}} - \frac{1}{\sqrt{n+1}} &= \frac{\sqrt{n+1}-\sqrt{n}}{\sqrt{n}\sqrt{n+1}} \\ &= \frac{1}{(\sqrt{n+1}+\sqrt{n})\sqrt{n}\sqrt{n+1}} \\ &> \frac{1}{(3\sqrt{n})\sqrt{n}(2\sqrt{n})} = \frac{1}{6}\left(\frac{1}{\sqrt{n}}\right)^3,\end{aligned}$$

so

$$a_{n+1} > \frac{1}{\sqrt{n+1}},$$

completing the induction. But then $\sum_{n=1}^{\infty} a_n^2$ diverges because it is greater than the harmonic series $\sum_{n=1}^{\infty} \frac{1}{n}$.

Solution 2. Because we have $0 < \sin x < x$ for $x \in (0, \pi/2]$, the sequence (a_n) is monotonically decreasing to a limit $L \in [0, 1]$. By the continuity of the sine function we must have $L = \sin(L)$, so $L = 0$. Now it follows from L'Hôpital's rule that

$$\begin{aligned}\lim_{n\to\infty}\left(\frac{1}{a_{n+1}^2} - \frac{1}{a_n^2}\right) &= \lim_{n\to\infty}\frac{a_n^2 - \sin^2(a_n)}{a_n^2\sin^2(a_n)} = \lim_{x\to 0}\frac{x^2 - \sin^2 x}{x^2\sin^2 x} \\ &= \lim_{x\to 0}\frac{x - \sin x}{x^3}\cdot\frac{x+\sin x}{x}\cdot\left(\frac{x}{\sin x}\right)^2 = \frac{1}{6}\cdot 2\cdot 1^2 = \frac{1}{3}.\end{aligned}$$

We can then apply the Stolz-Cesàro theorem to get

$$\lim_{n\to\infty}\frac{1/a_n^2}{n} = \lim_{n\to\infty}\frac{1/a_{n+1}^2 - 1/a_n^2}{(n+1)-n} = \frac{1}{3}.$$

Hence $a_n^2 \sim 3/n$ as $n \to \infty$ and the given series diverges.

A4. Consider a horizontal strip of $N + 2$ squares in which the first and the last square are black and the remaining N squares are all white. Choose a white square uniformly at random, choose one of its two neighbors with equal probability, and color this neighboring square black if it is not already black. Repeat this process until all the remaining white squares have only black neighbors. Let $w(N)$ be the expected number of white squares remaining.

Find

$$\lim_{N\to\infty} \frac{w(N)}{N}.$$

Answer. $\dfrac{1}{e}$.

Solution. Note that $w(0) = 0$, $w(1) = 1$, and $w(2) = 1$ (as eventually one of the two white squares will turn black). Let $N \geq 3$, and number the original white squares from 1 to N. For the first step of the process, there are $2N$ equally likely possible "outcomes", consisting of a choice of a white square along with a choice of one of its neighbors to be colored. Two of these outcomes (when white square 1 or N is chosen along with its neighbor that is already black) result in no change and can therefore be disregarded. The other $2N-2$ outcomes all result in a white square being colored black. If that is the white square numbered k, then the configuration that results is equivalent to a pair of strips like the original one, but with N replaced by $k-1$ for one of the strips and by $N-k$ for the other, so that the expected number of white squares at the end of the process, given a particular value of k, will be $w(k-1) + w(N-k)$. Of the $2N-2$ outcomes that lead to a white square being colored black, one has $k = 1$, one has $k = N$, and the other values of k, with $2 \leq k \leq N-1$, occur twice each. (The white squares at the ends are neighbors of a white square in just one way, the other white squares are adjacent to white squares on either side.) Because these $2N-2$ outcomes are equally likely, we can conclude that

$$\begin{aligned} w(N) &= \frac{1}{2N-2}\big(w(0) + w(N-1) + 2(w(1) + w(N-2)) + \\ &\qquad\qquad \cdots + 2(w(N-2) + w(1)) + w(N-1) + w(0)\big) \\ &= \frac{1}{N-1}\Big(w(0) + w(N-1) + 2\sum_{n=1}^{N-2} w(n)\Big). \end{aligned}$$

Thus $(N-1)w(N) = w(N-1) + 2\sum_{n=1}^{N-2} w(n)$. Subtracting this equation from the same one in which N is replaced by $N+1$, we get $Nw(N+1) - (N-1)w(N) = w(N) + w(N-1)$, that is,

$$w(N+1) = w(N) + \frac{1}{N}w(N-1). \qquad (*)$$

The solutions of this second-order linear homogeneous recurrence relation are known to be linear combinations of any two independent solutions. By inspection, one solution is $w(N) = N + 1$; we will use reduction of order to find the general solution.

Substituting $w(N) = (N+1)q(N)$ into $(*)$ yields $(N+2)q(N+1) = (N+1)q(N) + q(N-1)$, which can be rewritten as $(N+2)[q(N+1) - q(N)] = -[q(N) - q(N-1)]$. Thus if we let $\Delta(N) = q(N) - q(N-1)$, we have

$$\Delta(N+1) = -\frac{1}{N+2}\,\Delta(N)\,,$$

which implies that

$$\Delta(N) = C\,\frac{(-1)^N}{(N+1)!}$$

for some constant C. Therefore,

$$q(N) = q(0) + \sum_{n=1}^{N} \Delta(n) = q(0) + C\left[\sum_{n=1}^{N} \frac{(-1)^n}{(n+1)!}\right],$$

from which we see that the general solution to $(*)$ is given by

$$w(N) = a \cdot (N+1) + C \cdot (N+1)\left[\sum_{n=1}^{N} \frac{(-1)^n}{(n+1)!}\right].$$

From the initial conditions we have $w(0) = 0$, so $a = 0$, and then $w(1) = 1$, so $C = -1$ and

$$w(N) = (N+1)\left[\sum_{j=2}^{N+1} \frac{(-1)^j}{j!}\right] = (N+1)\left[\sum_{j=0}^{N+1} \frac{(-1)^j}{j!}\right].$$

Finally,

$$\lim_{N\to\infty} \frac{w(N)}{N} = \lim_{N\to\infty}\left[\sum_{j=0}^{N+1} \frac{(-1)^j}{j!}\right] = \frac{1}{e}.$$

A5. Let a_n be the number of sets S of positive integers for which

$$\sum_{k\in S} F_k = n,$$

where the Fibonacci sequence $(F_k)_{k\geq 1}$ satisfies $F_{k+2} = F_{k+1} + F_k$ and begins $F_1 = 1, F_2 = 1, F_3 = 2, F_4 = 3$.
Find the largest integer n such that $a_n = 2020$.

Answer. $n = F_{4040} - 1$.

Solution 1. We will show that for any integer $m \geq 2$, the largest n such that $a_n = m$ is $n = F_{2m} - 1$; therefore, the answer is $F_{4040} - 1$.

Note that for every positive integer n, the inequalities $F_{2m-2} \leq n \leq F_{2m} - 1$ are satisfied for a unique $m \geq 2$. Thus it is enough to show the two following facts, each of which will be proved by induction on m:

a) For $n = F_{2m} - 1$, we have $a_n = m$.
b) Whenever $F_{2m-2} \leq n \leq F_{2m} - 1$, we have $a_n \geq m$.

Proof of a): For the base case $m = 2$, we have $n = F_4 - 1 = 2$ and the sets S with $\sum_{k\in S} F_k = n$ are $\{1, 2\}$ and $\{3\}$. Now let $n = F_{2m+2} - 1$. To get a decomposition $n = \sum_{k\in S} F_k$ of the desired form, we can add F_{2m+1} to any such decomposition of $F_{2m} - 1$; by the induction hypothesis, there are exactly m such. If a decomposition of n does *not* include F_{2m+1}, it must include all earlier F_k, because $n = F_{2m} + F_{2m-1} + \cdots + F_1$; this identity yields one additional decomposition, for a total of $m + 1$, completing the induction.

Proof of b): We'll use $m = 2$ and $m = 3$ as base cases, checking that $a_1 = 2$ (the sets are $\{1\}$ and $\{2\}$) and $a_2 = 2$ for $m = 2$, and that $a_3 = 3, a_4 = 3, a_5 = 3, a_6 = 4, a_7 = 3$ for $m = 3$. Now suppose that $F_{2m} \leq n \leq F_{2m+2} - 1$ and $m \geq 3$. Then we can write $n = F_q + \ell$, where $q = 2m$ or $q = 2m + 1$ and $0 \leq \ell \leq F_{q-1} - 1$. We distinguish three cases.

(1) If $0 \le \ell < F_{q-3}$, we can get a decomposition of n by adding F_q to any decomposition of ℓ (of which there are at least 2, unless $\ell = 0$, in which case we only have $n = F_q$) as well as by adding F_{q-1} to any decomposition of $F_{q-2} + \ell$ (which is a number less than F_{q-1}, so it cannot create repetition). By the induction hypothesis, there are at least m of the latter, so there are at least $m+1$ different decompositions of n, as desired.

(2) If $F_{q-3} \le \ell < F_{q-3} + F_{q-4} = F_{q-2}$, then we can get a decomposition of n by adding F_q to any decomposition of ℓ, and we can also get one by adding $F_{q-1} + F_{q-2}$ to any decomposition of ℓ. Because $F_{2m-4} \le \ell \le F_{2m} - 1$, by induction hypothesis there are at least $m-1$ decompositions of ℓ, so there are at least $2(m-1)$ decompositions of n. Because $m \ge 3$, we have $2(m-1) \ge m+1$, as desired.

(3) The final case occurs when $F_{q-2} \le \ell \le F_{q-1} - 1$. In this case, we can get a decomposition of n by adding F_q to any decomposition of ℓ; by the induction hypothesis, there are at least m such decompositions. We can also find at least one other decomposition by starting with $F_{q-1} + F_{q-2}$ and continuing to add the largest possible distinct Fibonacci number to keep the sum $\le n$, that is, by using the "greedy" algorithm. (This works because $F_q + \ell \le F_q + F_{q-1} - 1 = F_{q+1} - 1 = F_{q-1} + F_{q-2} + \cdots + F_1$, and if any Fibonacci number is not used, the sum of the remaining ones is greater than or equal to the one skipped, so the algorithm can continue until the sum reaches n.) So once more there are at least $m+1$ different decompositions of n, and the proof is complete.

Solution 2. We start by showing that every positive integer n can be written *uniquely* as a sum $n = \sum_{k \in S} F_k$ for which the set S contains no two consecutive integers and contains only integers that are at least 2. (This way of writing n will be referred to as the base-Fibonacci representation of n; the first few are $1 = F_2$, $2 = F_3$, $3 = F_4$, $4 = F_2 + F_4$.) To show that such a representation of n is possible, choose the largest m for which $F_m \le n$, and note that then $n < F_{m+1}$, so $n - F_m < F_{m-1}$. By induction, $n - F_m$ has a base-Fibonacci representation, for which the set of subscripts cannot include $m-1$, and adding F_m gives a base-Fibonacci representation for n. To show uniqueness of the representation, note that if we have such a representation $n = \sum_{k \in S} F_k$ and j is the largest element in S, then

$$
\begin{aligned}
n &\le F_j + F_{j-2} + F_{j-4} + \cdots \\
&= (F_{j+1} - F_{j-1}) + (F_{j-1} - F_{j-3}) + \cdots \\
&= F_{j+1} - 1\,,
\end{aligned}
$$

so j must be the largest integer with $F_j \le n$; then $n - F_j = \sum_{k \in S - \{j\}} F_k$ is a base-Fibonacci representation, which is unique by induction.

Next, we identify any nonempty finite set S of positive integers with the finite sequence $(s_k)_{k \ge 1}$ of zeros and ones, ending with a 1, such that $s_k = 1 \Leftrightarrow k \in S$. In particular, we can then define

$$f(S) = \sum s_k F_k = \sum_{k \in S} F_k\,,$$

and we are looking for the largest n such that there are exactly 2020 sets S with $f(S) = n$. We will show, more generally, that the largest n for which there are exactly $m \ge 2$ sets with $f(S) = n$ is $n = F_{2m} - 1$, so that the answer is $F_{4040} - 1$.

To begin, note that for any positive integer n, one set S_0 with $f(S_0) = n$ is given by the base-Fibonacci representation of n, and any other set S with $f(S) = n$ can be transformed into S_0 as follows. Replace any occurrence of consecutive terms $(1, 1, 0)$ in the sequence by $(0, 0, 1)$; if there is no occurrence of $(1, 1, 0)$ but the sequence starts with $(1, 0)$, replace that start by $(0, 1)$. Repeat these "moves" until no further move is possible, at which point we must have arrived at S_0, because there are no longer consecutive 1's in the sequence and the sequence starts with 0. Therefore, we can count the number of sets with $f(S) = n$ by counting the number of sets that can be transformed into S_0 by a sequence of such moves, or equivalently the number of sets that can be obtained from S_0 (including S_0 itself) by reversing these moves. These reverse moves send $(0, 0, 1)$ anywhere in the sequence to $(1, 1, 0)$ (we'll refer to this as an A move) or $(0, 1)$ at the beginning of the sequence to $(1, 0)$ (a B move).

Suppose that $n = F_{2m} - 1$. Then the base-Fibonacci representation of n is

$$n = F_3 + F_5 + \cdots + F_{2m-3} + F_{2m-1}\,,$$

corresponding to the sequence $0, 0, 1, 0, 1, 0, 1, \ldots, 0, 1$, and when we start reversing the moves we see that at every step there is only one choice, which is an A move; after k steps we will have

$$n = F_1 + F_2 + \cdots + F_{2k} + F_{2k+3} + F_{2k+5} + \cdots + F_{2m-3} + F_{2m-1}$$

and we get such representations for $k = 0, 1, \ldots, m - 1$, so there are exactly m sets S with $f(S) = n$.

To finish the proof, we now show by induction that if $n \geq F_{2m}$, there are more than m sets S with $f(S) = n$. Note that because $n \geq F_{2m}$, the base-Fibonacci representation of n corresponds to a sequence which has a 1 in, or to the right of, the $2m$th position. Thus it is enough to prove that for any base-Fibonacci sequence S_0 with its final 1 in either the $2m$th position or the $(2m + 1)$st position, there are at least $m + 1$ different outcomes (counting S_0 itself) of the A and B moves described above. We will do so by induction, using the rightmost string of at least two successive zeros that occurs in S_0 before the final 1. If there is no such string of zeros, then S_0 must be precisely of the form $0, 1, 0, 1, 0, 1, \ldots, 0, 1$ with its final 1 in the $2m$th position; in this case we can start with a B move and then make up to $m - 1$ A moves, so there are in fact exactly $m + 1$ different outcomes. If the rightmost string of at least two successive zeros actually comes at the very beginning of the sequence, say that the sequence starts with exactly z zeros ($z \geq 2$), so it consists of those z zeros followed by $1, 0, 1, 0, \ldots, 1, 0, 1$, say $(r+1)$ ones and r zeros. Then we can start with $\lfloor z/2 \rfloor$ A moves at the beginning of the sequence; if z is odd, we can follow those up with a B move, so, whether z is even or odd, we have a total of $\lfloor (z+1)/2 \rfloor$ moves available at the beginning of the sequence. After that we have an additional r A moves, as each 1 that has not been moved yet can be "pushed to the left" (using an A move) in its turn. In all, we have at least $\lfloor (z+1)/2 \rfloor + r + 1$ different outcomes (counting S_0 itself); meanwhile, the length of the sequence S_0 is $z+2r+1$. Whether this equals $2m$ or $2m + 1$, we have $\lfloor (z + 1)/2 \rfloor + r + 1 = m + 1$, so we are done in this case. We are left with the case that the rightmost string of at least two zeros in S_0 does not come at the beginning of the sequence; say it comes after a 1 in the ath position and consists of exactly z zeros, followed by $(r+1)$ ones and r zeros, alternating as in the previous case. By the induction hypothesis, there are at least $\lfloor a/2 \rfloor + 1$ different outcomes available (including the starting "state") for just the first a positions of the sequence. For each of these "partial" outcomes, the rest of the sequence, starting with the z consecutive zeros, can be treated as in the previous case, except that if z is odd, we do not have the follow-up B move available after the initial $\lfloor z/2 \rfloor$ A moves. Thus we have at least $\lfloor z/2 \rfloor + r + 1$ different possible outcomes for the part of the sequence after the first a positions, so overall we have at least

$$\begin{aligned}(\lfloor a/2\rfloor+1)(\lfloor z/2\rfloor+r+1) &\geq \lfloor a/2\rfloor+\lfloor z/2\rfloor+r+1+\lfloor a/2\rfloor\lfloor z/2\rfloor\\ &\geq \lfloor a/2\rfloor+\lfloor z/2\rfloor+r+2\end{aligned}$$

outcomes, because $a\geq 2$ and $z\geq 2$. The length of the sequence is $a+z+2r+1$. If this equals $2m$, then one of a and z is even (and the other is odd), so

$$\lfloor a/2\rfloor+\lfloor z/2\rfloor+r+2=a/2+z/2-1/2+r+2=(a+z+2r+3)/2=m+1\,;$$

if the length equals $2m+1$, then a and z have the same parity and

$$\lfloor a/2\rfloor+\lfloor z/2\rfloor+r+2\geq a/2+z/2-1+r+2=(a+z+2r+2)/2=m+1\,.$$

This estimate concludes the proof.

A6. For a positive integer N, define the function

$$f_N(x)=\sum_{n=0}^{N}\frac{N+1/2-n}{(N+1)(2n+1)}\sin\big((2n+1)x\big)\,.$$

Determine the smallest constant M such that $f_N(x)\leq M$ for all N and all real x.

Answer. $M=\dfrac{\pi}{4}$.

Solution. Note that $f_N(x)$ is an odd function with period 2π. The following computation allows us to write its derivative in closed form:

$$\begin{aligned}
f_N'(x) &= \sum_{n=0}^{N}\frac{2N+1-2n}{2(N+1)}\cos((2n+1)x)=\sum_{n=0}^{N}\frac{2N+1-2n}{2(N+1)}\,\mathrm{Re}\left(e^{(2n+1)ix}\right)\\
&=\frac{1}{2(N+1)}\,\mathrm{Re}\left(e^{2(N+1)ix}\sum_{n=0}^{N}(2N+1-2n)e^{(2n-1-2N)ix}\right)\\
&=\frac{1}{2(N+1)}\,\mathrm{Re}\left(ie^{2(N+1)ix}\,\frac{d}{dx}\sum_{n=0}^{N}e^{(2n-1-2N)ix}\right)\\
&=\frac{1}{2(N+1)}\,\mathrm{Re}\left(ie^{2(N+1)ix}\,\frac{d}{dx}\left(e^{(-1-2N)ix}\,\frac{1-e^{2(N+1)ix}}{1-e^{2ix}}\right)\right)\\
&=\frac{1}{2(N+1)}\,\mathrm{Re}\left(ie^{2(N+1)ix}\,\frac{d}{dx}\,\frac{1-e^{-2(N+1)ix}}{e^{ix}-e^{-ix}}\right)\\
&=\frac{1}{4(N+1)}\,\mathrm{Re}\left(e^{2(N+1)ix}\,\frac{d}{dx}\,\frac{1-e^{-2(N+1)ix}}{\sin x}\right)\\
&=\frac{1}{4(N+1)}\,\mathrm{Re}\left(\frac{2(N+1)i}{\sin x}+\frac{(1-e^{2(N+1)ix})\cos x}{\sin^2 x}\right)\\
&=\frac{[1-\cos(2(N+1)x)]\cos x}{4(N+1)\sin^2 x}=\frac{\sin^2((N+1)x)}{2(N+1)\sin^2 x}\cos x\,.
\end{aligned}$$

In particular, the derivative has the same sign as $\cos x$. (This is still true where $\sin x=0$, because then $x=k\pi$ for some integer k, and for all $0\leq n\leq N$, $\cos((2n+1)k\pi)=\cos k\pi$, so that the first expression for $f_N'(x)$ above is a positive multiple of $\cos x$. Alternatively, one can use continuity of the derivative and l'Hôpital's rule.) It follows that $f_N(x)$ has its maximum value for $x = \pi/2$. That value is

$$
\begin{aligned}
f_N\left(\frac{\pi}{2}\right) &= \sum_{n=0}^{N} \frac{2N+1-2n}{(2N+2)(2n+1)} \cdot (-1)^n \\
&= \sum_{n=0}^{N} \left(\frac{1}{2n+1} - \frac{1}{2N+2}\right) \cdot (-1)^n \\
&= \begin{cases} \left(\sum\limits_{n=0}^{2M} \dfrac{(-1)^n}{2n+1}\right) - \dfrac{1}{4M+2} & \text{when } N = 2M \text{ is even} \quad \text{and} \\[2ex] \left(\sum\limits_{n=0}^{2M} \dfrac{(-1)^n}{2n+1}\right) - \dfrac{1}{4M+3} & \text{when } N = 2M+1 \text{ is odd.} \end{cases}
\end{aligned}
$$

From here it is straightforward to check that $f_{2M}(\pi/2) - f_{2M-1}(\pi/2) = \frac{1}{(4M+1)(4M+2)}$ and $f_{2M+1}(\pi/2) - f_{2M}(\pi/2) = \frac{1}{(4M+2)(4M+3)}$, so the maximum value $f_N(\pi/2)$ is an increasing function of N. Thus the least upper bound on $f_N(x)$ that is valid for all N and x is

$$
\lim_{N\to\infty} f_N\left(\frac{\pi}{2}\right) = \sum_{n=0}^{\infty} \frac{(-1)^n}{2n+1} = \arctan(1) = \frac{\pi}{4}\,.
$$

Variant. An alternate computation of the derivative $f_N'(x)$ uses the trigonometric sums given by the
Lemma:

$$
\sum_{n=0}^{N} \cos[(2n+1)x] = \frac{\sin[(N+1)x]\cos[(N+1)x]}{\sin x}\,, \qquad \sum_{n=0}^{N} \sin[(2n+1)x] = \frac{\sin^2[(N+1)x]}{\sin x}\,.
$$

Proof: Let $z = e^{ix}$, and note that

$$
\begin{aligned}
\sum_{n=0}^{N} e^{i(2n+1)x} &= \sum_{n=0}^{N} z^{2n+1} = \frac{z - z^{2N+3}}{1-z^2} \\
&= \frac{z^{2N+2}-1}{z-1/z} = z^{N+1}\left(\frac{z^{N+1}-z^{-(N+1)}}{z-1/z}\right) \\
&= (\cos[(N+1)x] + i\sin[(N+1)x])\,\frac{\sin[(N+1)x]}{\sin x}.
\end{aligned}
$$

Taking real and imaginary parts, we get the desired sums, proving the lemma.

We now use the identity

$$
\frac{N+1/2-n}{(N+1)(2n+1)} = \frac{1}{2n+1} - \frac{1}{2(N+1)}
$$

to split $f_N(x)$ into two pieces:

$$
\begin{aligned}
f_N(x) &= \sum_{n=0}^{N} \frac{\sin[(2n+1)x]}{2n+1} - \frac{1}{2(N+1)} \sum_{n=0}^{N} \sin[(2n+1)x] \\
&= \sum_{n=0}^{N} \frac{\sin[(2n+1)x]}{2n+1} - \frac{1}{2(N+1)} \frac{\sin^2[(N+1)x]}{\sin x} \,,
\end{aligned}
$$

using the second part of the lemma. The derivative is therefore

$$
\begin{aligned}
f_N'(x) &= \sum_{n=0}^{N} \cos[(2n+1)x] - \frac{1}{2(N+1)} \left(\frac{2(N+1)\sin[(N+1)x]\cos[(N+1)x]}{\sin x} - \frac{\sin^2[(N+1)x]}{\sin^2 x} \cdot \cos x \right) \\
&= \frac{\sin^2((N+1)x)}{2(N+1)\sin^2 x} \cos x \,,
\end{aligned}
$$

because the first two summands on the right cancel by the first part of the lemma.

B1. For a positive integer n, define $d(n)$ as the sum of the digits of n when written in binary (for example, $d(13) = 1+1+0+1 = 3$). Let

$$S = \sum_{k=1}^{2020} (-1)^{d(k)} k^3.$$

Determine S modulo 2020.

Answer. 1990.

Solution 1. We will show that if we let $d(0) = 0$, start the sum at $k = 0$ (which does not change its value), and group the terms from that start in groups of sixteen each, then each complete group contributes 0 to the sum. Therefore, S is equal to the sum starting at $2016 = 16 \cdot 126$, that is,

$$S = \sum_{k=2016}^{2020} (-1)^{d(k)} k^3.$$

From the binary expansion $11,111,100,000$ of 2016, we observe that
$d(2016) = 6,\ d(2017) = 7,\ d(2018) = 7,\ d(2019) = 8,\ d(2020) = 7$ and so

$$\begin{aligned} S &= (2016)^3 - (2017)^3 - (2018)^3 + (2019)^3 - (2020)^3 \\ &\equiv (-4)^3 - (-3)^3 - (-2)^3 + (-1)^3 = -64 + 27 + 8 - 1 = -30 \equiv 1990 \pmod{2020}. \end{aligned}$$

It remains to prove the desired cancellation, which results from the polynomial $P(x) = x^3$ having degree 3 and from $2^{3+1} = 16$. In fact, we can apply the following lemma.

Lemma: Suppose $P(x)$ is a polynomial of degree $n \geq 0$, let $N = 2^{n+1} - 1$, and let $m \geq 0$ be any integer. Then

$$Q_m = \sum_{j=0}^{N} (-1)^{d\left(m2^{n+1}+j\right)}\ P(m2^{n+1} + j) = 0.$$

Proof: Note that for $0 \leq j \leq N$, there are no carries in binary in the addition $m2^{n+1} + j$, so $d\left(m2^{n+1} + j\right) = d\left(m2^{n+1}\right) + d(j)$ and we have $Q_m = (-1)^{d\left(m2^{n+1}\right)} q_m(0)$, where

$$q_m(x) = \sum_{j=0}^{N} (-1)^{d(j)} p_m(x+j)\,,\ p_m(x) = P(m2^{n+1} + x).$$

Now $p_m(x)$, a translate of $P(x)$, is again a polynomial of degree n, and so it is enough to show that if $p(x)$ is any polynomial of degree n and $q(x) = \sum_{j=0}^{N} (-1)^{d(j)} p(x+j)$, then $q(0) = 0$. In fact, we will see that $q(x)$ is identically zero. Define the difference operator Δ_k on polynomials by $\Delta_k R(x) = R(x+k) - R(x)$. We then have

$$\begin{aligned} q(x) &= \sum_{j=0}^{2^n-1} (-1)^{d(j)} p(x+j) + \sum_{j=2^n}^{N} (-1)^{d(j)} p(x+j) \\ &= \sum_{j=0}^{2^n-1} \left[(-1)^{d(j)} p(x+j) + (-1)^{d(j+2^n)} p(x+j+2^n)\right] \\ &= -\Delta_{2^n} \sum_{j=0}^{2^n-1} (-1)^{d(j)} p(x+j) \end{aligned}$$

because $d(j+2^n) = d(j)+1$ for $0 \le j \le 2^n - 1$. By continuing to halve the interval for j in this way we end up with

$$q(x) = (-1)^{n+1}\Delta_{2^n} \cdots \Delta_4\Delta_2\Delta_1 p(x).$$

However, each application of a difference operator Δ_k lowers the degree of a nonconstant polynomial by 1, so $\Delta_{2^{n-1}} \cdots \Delta_4\Delta_2\Delta_1 p(x)$ is constant and $q(x)$ is identically zero, completing the proof of the lemma.

Solution 2. Let $\overline{S}_{n,q} = \sum\limits_{k=0}^{n}(-1)^{d(k)}k^q \bmod 2020$, so we are looking for $\overline{S}_{2020,3}$. Note that $d(2k) = d(k)$ and $d(2k+1) = d(k)+1$, so splitting the sum into even and odd terms we get

$$\begin{aligned}
\overline{S}_{2020,3} &\equiv \sum_{k=0}^{1009}(-1)^{d(2k)}(2k)^3 + \sum_{k=0}^{1009}(-1)^{d(2k+1)}(2k+1)^3 \\
&= \sum_{k=0}^{1009}(-1)^{d(k)}\big(8k^3 - 8k^3 - 12k^2 - 6k - 1\big) \\
&\equiv -12\overline{S}_{1009,2} - 6\overline{S}_{1009,1} - \overline{S}_{1009,0}\,,
\end{aligned}$$

where the congruences are modulo 2020. Similarly, we have

$$\begin{aligned}
\overline{S}_{1009,2} &\equiv \sum_{k=0}^{504}(-1)^{d(2k)}(2k)^2 + \sum_{k=0}^{504}(-1)^{d(2k+1)}(2k+1)^2 \\
&= \sum_{k=0}^{504}(-1)^{d(k)}\big(4k^2 - 4k^2 - 4k - 1\big) \\
&\equiv -4\overline{S}_{504,1} - \overline{S}_{504,0}
\end{aligned}$$

and

$$\begin{aligned}
\overline{S}_{1009,1} &\equiv \sum_{k=0}^{504}(-1)^{d(2k)}(2k) + \sum_{k=0}^{504}(-1)^{d(2k+1)}(2k+1) \\
&= \sum_{k=0}^{504}(-1)^{d(k)}(-1) \\
&\equiv -\overline{S}_{504,0}\,.
\end{aligned}$$

Combining the results so far, we have

$$\begin{aligned}
\overline{S}_{2020,3} &\equiv -12(-4\overline{S}_{504,1} - \overline{S}_{504,0}) + 6\overline{S}_{504,0} - \overline{S}_{1009,0} \\
&\equiv 48\overline{S}_{504,1} + 18\overline{S}_{504,0} - \overline{S}_{1009,0}\,.
\end{aligned}$$

Note that

$$\begin{aligned}
\overline{S}_{504,1} &\equiv \sum_{k=0}^{252}(-1)^{d(2k)}(2k) + \sum_{k=0}^{251}(-1)^{d(2k+1)}(2k+1) \\
&= (-1)^{d(504)}(504) + \sum_{k=0}^{251}(-1)^{d(k)}(-1) \\
&\equiv (-1)^{d(504)}(504) - \overline{S}_{251,0} \\
&= 504 - \overline{S}_{251,0}\,,
\end{aligned}$$

because the binary expansion of 504 is 111111000, so that $d(504) = 6$ is even. Finally, $\overline{S}_{n,0} = 0$ whenever n is odd (using the same splitting into even and odd terms), so $\overline{S}_{503,0} = 0$ and $\overline{S}_{504,0} = (-1)^{d(504)} = 1$. It follows that

$$\begin{aligned}
\overline{S}_{2020,3} &\equiv 48(504 - 0) + 18 \cdot 1 - 0 \\
&= 24210 \equiv 1990.
\end{aligned}$$

B2. Let k and n be integers with $1 \le k < n$. Alice and Bob play a game with k pegs in a line of n holes. At the beginning of the game, the pegs occupy the k leftmost holes. A legal move consists of moving a single peg to any vacant hole that is further to the right. The players alternate moves, with Alice playing first. The game ends when the pegs are in the k rightmost holes, so whoever is next to play cannot move, and therefore loses. For what values of n and k does Alice have a winning strategy?

Answer. Alice has a winning strategy if and only if at least one of k and n is odd.

Solution. Number the holes, from left to right, $1, 2, \ldots, n$. We first show that when k and n are both even, Bob has a winning strategy. In this case we can divide the holes into disjoint adjacent pairs $P_i = \{2i - 1, 2i\}$ with $1 \le i \le n/2$. At the beginning of the game the pegs completely occupy the holes in the leftmost $k/2$ pairs, and all the holes in the remaining pairs are vacant. Thus Alice's first move must take a peg from an occupied pair of holes and place it in one of a vacant pair of holes. A winning strategy for Bob is to always take the other peg of the pair that Alice moved from and place it in the remaining hole of the pair that Alice moved to. Thus after each of Bob's moves, each of the pairs P_i either has pegs in both holes or in neither, whereas after each of Alice's moves, there are two of the pairs P_i with one peg each. In particular, Alice can never reach the ending position, and the game will end after one of Bob's moves.

If k and n are not both even, Alice always has a first move available which will leave Bob either with no moves at all, or with a position equivalent to the starting position of our game with even integers k_1 and n_1, $1 \le k_1 < n_1$. Thus by the case discussed in the previous paragraph, Alice (as the second player from that position) has a winning strategy. Specifically, if k and n are both odd, Alice can move the peg in hole k to hole n, leaving $k_1 = k - 1$ pegs at the beginning of a line of $n_1 = n - 1$ remaining holes. (If $k = 1$, the game is then over.) If k is odd and n is even, Alice can move the peg in hole 1 to hole n, winning the game immediately if $k = 1$ and otherwise leaving $k_1 = k - 1$ pegs at the beginning of a line of $n_1 = n - 2$ remaining holes. Finally, if k is even and n is odd, Alice can move the peg in hole 1 to hole $k + 1$, winning the game immediately if $n = k + 1$ and otherwise leaving $k_1 = k$ pegs at the beginning of a line of $n_1 = n - 1$ remaining holes. In each of these three cases, after making the indicated first move, Alice can use Bob's strategy from the previous paragraph to win.

Comment. The game with k pegs and n holes is equivalent to the game with $n-k$ pegs and n holes (moving the k pegs to the right is equivalent to moving the $n-k$ vacant spaces to the left). This symmetry can be used to reduce the three cases considered in the second paragraph to just two.

B3.

Let $x_0 = 1$, and let δ be some constant satisfying $0 < \delta < 1$. Iteratively, for $n = 0, 1, 2, \ldots$, a point x_{n+1} is chosen uniformly from the interval $[0, x_n]$. Let Z be the smallest value of n for which $x_n < \delta$. Find the expected value of Z, as a function of δ.

Answer. The expected value is $1 + \ln(1/\delta)$.

Solution 1. Let $\rho_n(x)$ be the probability density for the location of x_n. Note that $0 \le x_n \le 1$ for all n, so these density functions all have support $[0, 1]$. They can be found recursively from $\rho_1(x) = 1$ and

$$\rho_{n+1}(x) = \int_{y=x}^{1} \rho_n(y)\, \frac{dy}{y}\,.$$

This yields

$$\rho_2(x) = \int_{y=x}^{1} \frac{dy}{y} = -\ln(x), \quad \rho_3(x) = \int_{y=x}^{1} (-\ln y)\, \frac{dy}{y} = \frac{[-\ln(x)]^2}{2}\,,$$

which suggests that in general

$$\rho_n(x) = \frac{[-\ln(x)]^{n-1}}{(n-1)!}\,;$$

this is straightforward to check by induction.

Let q_n be the probability that $x_n < \delta$ but $x_{n-1} \ge \delta$, that is, the probability that $Z = n$. Then $q_1 = \delta$, and for $n \ge 2$ we have

$$\begin{aligned}
q_n &= \int_0^{\delta} \rho_n(x) - \rho_{n-1}(x)\, dx \\
&= \int_0^{\delta} \frac{[-\ln(x)]^{n-1}}{(n-1)!} - \frac{[-\ln(x)]^{n-2}}{(n-2)!}\, dx \\
&= \left.\frac{x[-\ln(x)]^{n-1}}{(n-1)!}\right|_0^{\delta} \\
&= \frac{\delta[-\ln(\delta)]^{n-1}}{(n-1)!}\,.
\end{aligned}$$

Finally, the expected value of Z is

$$\begin{aligned}
E(Z) &= \sum_{n=1}^{\infty} n q_n \\
&= \delta + \sum_{n=2}^{\infty} n \frac{\delta[-\ln(\delta)]^{n-1}}{(n-1)!} \\
&= \sum_{m=0}^{\infty} (m+1) \frac{\delta[-\ln(\delta)]^{m}}{m!} \\
&= \sum_{m=0}^{\infty} \frac{\delta[-\ln(\delta)]^{m}}{m!} + \sum_{m=1}^{\infty} \frac{\delta[-\ln(\delta)]^{m}}{(m-1)!} \\
&= \delta \exp[-\ln(\delta)] - \ln(\delta) \cdot \delta \exp[-\ln(\delta)] \\
&= 1 + \ln(1/\delta).
\end{aligned}$$

Solution 2. A short calculation shows that if X is a uniform random variable on $[0,1]$, then $U = -\ln X$ is an exponential random variable with expected value $\lambda = 1$, and probability density function $p_U(t) = e^{-t}$ for $t \geq 0$. Note that this applies to each of $X_1 = x_1/x_0$, $X_2 = x_2/x_1$, $\ldots$, $X_n = x_n/x_{n-1}$, and that the product $X_1 X_2 \cdots X_n$ equals x_n. Thus if $U_1, U_2, \ldots, U_n, \ldots$ are the corresponding exponential random variables, the problem is equivalent to finding the expected number $Z = Z(D)$ of i.i.d. samples that must be taken to get $U_1 + \cdots + U_Z > D$, where $D = -\ln \delta = \ln(1/\delta)$. By "First-Step Analysis" (considering what the situation is after the first sample) we see that

$$\begin{aligned}
Z(D) &= 1 + \int_0^D Z(D-t) p_U(t)\, dt = 1 + \int_0^D Z(D-t) e^{-t}\, dt \\
&= 1 + e^{-D} \int_0^D Z(u) e^{u}\, du\,.
\end{aligned}$$

Differentiating with respect to D gives

$$\begin{aligned}
Z'(D) &= -e^{-D} \int_0^D Z(u) e^{u}\, du + e^{-D} Z(D) e^{D} \\
&= -(Z(D) - 1) + Z(D) = 1,
\end{aligned}$$

where the second line follows by substituting for the integral using the First-Step Analysis equation. Thus $Z(D) = Z(0) + D = 1 + D = 1 + \ln(1/\delta)$.

Solution 3. Note that given $x_{k-1} \geq \delta$, the probability that x_k is not smaller than δ is $(x_{k-1} - \delta)/x_{k-1} \leq 1 - \delta$. Hence the probability that $Z = n$ is bounded above by $(1-\delta)^{n-1}$. Thus the expected value of Z is bounded by the sum of the convergent series $\sum_{n=1}^{\infty} n(1-\delta)^{n-1}$, and thus is finite.

Let $f(\delta)$ be this expected value, as a function of δ. Note that this function is monotone decreasing. If after one step of the iteration we are at $x_1 \geq \delta$, then rescaling by a factor $1/x_1$, we see that we have essentially returned to the original problem but with δ replaced by δ/x_1. Thus

$$f(\delta) = 1 + \int_\delta^1 f(\delta/x)\, dx.$$

Letting $g(t) = f(1/t)$ and making the substitution $u = tx$, this becomes

$$g(t) = 1 + \frac{1}{t}\int_1^t g(u)\,du.$$

Since f is monotone decreasing, g is monotone increasing and hence integrable. Thus it follows from this functional equation that g is continuous for $t > 0$. Hence the integral in the functional equation is a differentiable function of t, and it follows that g is differentiable.

Multiplying both sides of the functional equation by t and then taking the derivative of both sides leads to

$$g(t) + tg'(t) = 1 + g(t), \quad \text{so} \quad tg'(t) = 1.$$

Integrating and using the initial condition $g(1) = 1$, we get $g(t) = 1 + \ln t$ and hence $f(\delta) = 1 + \ln(1/\delta)$.

B4. Let n be a positive integer, and let V_n be the set of integer $(2n+1)$-tuples $\mathbf{v} = (s_0, s_1, \cdots, s_{2n-1}, s_{2n})$ for which $s_0 = s_{2n} = 0$ and $|s_j - s_{j-1}| = 1$ for $j = 1, 2, \cdots, 2n$. Define

$$q(\mathbf{v}) = 1 + \sum_{j=1}^{2n-1} 3^{s_j},$$

and let $M(n)$ be the average of $\dfrac{1}{q(\mathbf{v})}$ over all $\mathbf{v} \in V_n$.

Evaluate $M(2020)$.

Answer. $\dfrac{1}{4040}$.

Solution. We will show that $M(n) = \dfrac{1}{2n}$ for all n, by partitioning V_n into subsets such that the average of $\dfrac{1}{q(\mathbf{v})}$ over each subset is $\dfrac{1}{2n}$. First note that giving an element $\mathbf{v} \in V_n$ is equivalent to giving a sequence of length $2n$ consisting of symbols U (for "up") and D (for "down") so that each symbol occurs n times in the sequence; the symbol in position i is U or D according to whether $s_i - s_{i-1}$ is 1 or -1. With this representation of elements of V_n, there is a natural "cyclic rearrangement" map $\sigma : V_n \to V_n$ which moves each of the symbols one position back cyclically, that is, the symbol in position 1 moves to position $2n$, and for every $j > 1$ the symbol in position j moves to position $j-1$. In terms of the $(2n+1)$-tuples $\mathbf{v} = (s_0, s_1, \cdots, s_{2n-1}, s_{2n})$, this works out to

$$\sigma(\mathbf{v}) = (t_0, t_1, \cdots, t_{2n-1}, t_{2n}) \quad \text{where} \quad t_j = s_{j+1} - s_1\,,$$

with the understanding that subscripts are taken modulo $2n$. (Note that $t_0 = t_{2n} = 0$ and that $|t_j - t_{j-1}| = |s_{j+1} - s_j| = 1$.)

From the representation using the symbols U and D, we see that $\sigma^{2n}(\mathbf{v}) = \mathbf{v}$. In particular, for any $\mathbf{v} \in V_n$, the list of elements $\mathbf{v}, \sigma(\mathbf{v}), \sigma^2(\mathbf{v}), \ldots, \sigma^{2n-1}(\mathbf{v})$ runs through the orbit under σ of $\mathbf{v}$ a whole number of times. So the average of $\dfrac{1}{q(\mathbf{w})}$ for $\mathbf{w}$ on that list of elements is the same as the average over the orbit of $\mathbf{v}$; because the orbits partition V_n, it is enough to show that this average is $\dfrac{1}{2n}$ for any $\mathbf{v}$.

Now note that

$$
\begin{aligned}
\frac{1}{q(\sigma(\mathbf{v}))} &= \frac{1}{1+\sum\limits_{j=1}^{2n-1} 3^{s_{j+1}-s_1}} \\
&= \frac{3^{s_1}}{3^{s_1}+\sum\limits_{j=1}^{2n-1} 3^{s_{j+1}}} = \frac{3^{s_1}}{q(\mathbf{v})}
\end{aligned}
$$

because $3^{s_{2n}} = 1$. Applying this with $\mathbf{v}$ replaced by $\sigma(\mathbf{v})$ yields

$$\frac{1}{q(\sigma^2(\mathbf{v}))} = \frac{3^{s_2-s_1}}{q(\sigma(\mathbf{v}))} = \frac{3^{s_2}}{q(\mathbf{v})}$$

and similarly, by induction on j,

$$\frac{1}{q(\sigma^j(\mathbf{v}))} = \frac{3^{s_j}}{q(\mathbf{v})}.$$

To average $\dfrac{1}{q(\mathbf{w})}$ over the list $\mathbf{v}, \sigma(\mathbf{v}), \sigma^2(\mathbf{v}), \ldots, \sigma^{2n-1}(\mathbf{v})$, we add these answers for $j = 0, 1, \ldots, 2n-1$ and divide by $2n$. But the sum of these answers is $\dfrac{q(\mathbf{v})}{q(\mathbf{v})} = 1$, so we are done.

B5. For $j \in \{1, 2, 3, 4\}$, let z_j be a complex number with $|z_j| = 1$ and $z_j \neq 1$. Prove that $3 - z_1 - z_2 - z_3 - z_4 + z_1 z_2 z_3 z_4 \neq 0$.

Solution 1. Let $e_k(Z)$ denote the kth elementary symmetric function of $Z := (z_1, z_2, z_3, z_4)$, so that we want to show $3 - e_1(Z) + e_4(Z) \neq 0$. We will transform the variables first by $z_j = 1 - y_j$ and then by $y_j = 1/w_j$. The condition $z_j \neq 1$ becomes $y_j \neq 0$, so that w_j is indeed defined, while the condition $|z_j| = 1$ implies $\mathrm{Re}(w_j) = 1/2$. Meanwhile, using similar notation for the elementary symmetric functions of the y's and the w's, we find that

$$3 - e_1(Z) + e_4(Z) = e_2(Y) - e_3(Y) + e_4(Y) = \frac{1 - e_1(W) + e_2(W)}{e_4(W)}.$$

Now let

$$w_j = \frac{1}{2} + i\, v_j\,.$$

Then

$$w_j w_k = \frac{1}{4} - v_j v_k + \frac{i}{2}(v_j + v_k),$$

so for the symmetric functions we have

$$e_1(W) = \sum w_j = 2 + i\, e_1(V), \;\; e_2(W) = \sum_{j<k} w_j w_k = \frac{3}{2} - e_2(V) + \frac{3i}{2} e_1(V)$$

and it is enough to show that

$$1 - e_1(W) + e_2(W) = \frac{1}{2} + \frac{i}{2} e_1(V) - e_2(V)$$

is never zero for real $V = (v_1, v_2, v_3, v_4)$.

If this quantity were zero, taking real and imaginary parts we would have $e_1(V) = 0$, $e_2(V) = 1/2$. However, because V is real we have

$$e_1(V)^2 = \left(\sum v_j\right)^2 = \sum v_j^2 + 2\sum_{j<k} v_j v_k \geq 2e_2(V),$$

so those values for $e_1(V)$ and $e_2(V)$ are impossible.

Solution 2. We use a bilinear (linear fractional) transformation to map the circle $|z| = 1$ to the real line. Because $z_j \neq 1$, it seems natural to map 1 to the point at infinity; a transformation that will do these things is given by

$$w = i\,\frac{1+z}{1-z} \qquad \Leftrightarrow \qquad z = \frac{w-i}{w+i}\,.$$

To check, $|z| = 1$ implies $|w - i| = |w + i|$, from which it follows that w is real.

By (a significant amount of) direct computation we find that

$$\begin{aligned}&3 - z_1 - z_2 - z_3 - z_4 + z_1z_2z_3z_4 \\ &\qquad = \frac{8 - 4i(w_1 + w_2 + w_3 + w_4) - 4(w_1w_2 + w_1w_3 + w_1w_4 + w_2w_3 + w_2w_4 + w_3w_4)}{(w_1 + i)(w_2 + i)(w_3 + i)(w_4 + i)}.\end{aligned}$$

If this were zero for real numbers w_i, taking real and imaginary parts of the numerator we would get

$8-4(w_1w_2+w_1w_3+w_1w_4+w_2w_3+w_2w_4+w_3w_4) = 0$ and $w_1+w_2+w_3+w_4 = 0$, respectively.

However,

$$\begin{aligned}&8 - 4(w_1w_2 + w_1w_3 + w_1w_4 + w_2w_3 + w_2w_4 + w_3w_4) \\ &\qquad\qquad = 8 + 2\left[w_1^2 + w_2^2 + w_3^2 + w_4^2 - (w_1 + w_2 + w_3 + w_4)^2\right]\end{aligned}$$

is always positive when $w_1 + w_2 + w_3 + w_4 = 0$, completing the proof.

B6. Let n be a positive integer. Prove that

$$\sum_{k=1}^{n} (-1)^{\lfloor k(\sqrt{2}-1)\rfloor} \geq 0\,.$$

(As usual, $\lfloor x\rfloor$ denotes the greatest integer less than or equal to x.)

Solution. Let $\alpha = \sqrt{2} - 1$; note that this irrational number has the crucial property

$$\frac{1}{\alpha} = \sqrt{2} + 1 = 2 + \alpha.$$

Also, let $a_k = (-1)^{\lfloor k\alpha\rfloor}$. The sequence (a_k) is formed by the terms of the series whose partial sums we are looking at, and it consists of "signs" ±1. Because $1/3 < \alpha < 1/2$, the signs come in runs of two or three equal signs, starting with a run of two 1s because $\lfloor\alpha\rfloor = \lfloor 2\alpha\rfloor = 0$ and $\lfloor 3\alpha\rfloor = 1$. Now suppose we omit two of the signs from each run, so the runs of two equal signs are deleted altogether and each run of three equal signs is replaced by a single sign. Denote the new sequence of signs by (b_k); that is, b_k is the value taken by the k-th run of length 3 in the sequence (a_k).

We will show below that $a_k = b_k$. Assuming this for now, we can prove the desired result by a reduction argument, as follows. Suppose that the result is false, and let N be the least value of n for which $\sum_{k=1}^{n} a_k < 0$. Then a_{N-2}, a_{N-1}, a_N must be a run of three -1s, because every run of fewer -1s is preceded by a run of 1s of at least equal length, so that omitting both those runs can only decrease the partial sum. Suppose that up to and including this point, there are m runs of length 3. If we pass from the sequence (a_k) to the sequence (b_k), we delete two entries from each run, starting with two 1s and ending with two -1s, so the sum of all the terms will be unchanged. That is,

$$\sum_{k=1}^{m} b_k = \sum_{k=1}^{N} a_k < 0\,,$$

and because $a_k = b_k$ we have $\sum_{k=1}^{m} a_k < 0$. But $m \leq N/3$, contradicting the minimality of N.

It remains only to prove that $a_k = b_k$. Suppose that the k-th run of length 3 is the $(t+1)$-st run overall, that is, the run for which the floor is equal to t. Then $b_k = (-1)^t$. There are $2t+(k-1)$ terms before this run, so it consists of the terms with subscripts $2t+k, 2t+k+1$, and $2t+k+2$, and we have the inequalities

$$t < (2t+k)\alpha\,,\ (2t+k+2)\alpha < t+1\,.$$

Isolating k in each of these, we get

$$k > t(\frac{1}{\alpha} - 2) = t\alpha\,,\ k < (t+1)(\frac{1}{\alpha} - 2) = (t+1)\alpha\,,$$

so $t\alpha < k < (t+1)\alpha$ and thus $\lfloor k/\alpha \rfloor = t$. But $k/\alpha = k(2+\alpha) = 2k + k\alpha$, so $\lfloor k\alpha \rfloor = t - 2k$, $a_k = (-1)^{\lfloor k\alpha \rfloor} = (-1)^{t-2k} = (-1)^t = b_k$, and we are done.