Mathematics 414, Solutions to Problem Set #5

Problem 1. Let n be a positive integer. Using school level methods, show that $1 + 1/2^2 + 1/3^2 + \cdots + 1/n^2 < 1.7$. (Euler showed that the infinite sum $1 + 1/2^2 + 1/3^2 + \cdots$ is equal to $\pi^2/6$, but I know of no high school level proof.) Hint: Approximate the "tail" (the stuff from $1/(k+1)^2$ on, for suitable k), by something easily summed.

Solution. The terms are positive, so it is enough to show that even if n is large, our sum is less than 1.7. So let n be large, and suppose that $k < n$. Our sum is equal to

$$1 + \frac{1}{2^2} + \cdots + \frac{1}{k^2} + \frac{1}{(k+1)^2} + \frac{1}{(k+2)^2} + \cdots + \frac{1}{n^2}.$$

The "tail," the sum from $1/(k+1)^2$ on, is less than

$$\frac{1}{k(k+1)} + \frac{1}{(k+1)(k+2)} + \cdots + \frac{1}{(n-1)(n}.$$

But in general $\frac{1}{i(i+1)} = \frac{1}{i} - \frac{1}{i+1}$. Thus the tail is less than

$$\left(\frac{1}{k} - \frac{1}{k+1}\right) + \left(\frac{1}{k+1} - \frac{1}{k+2}\right) + \cdots + \left(\frac{1}{n-1} - \frac{1}{n}\right).$$

This last sum collapses to $\frac{1}{k} - \frac{1}{n-1}$, which is less than $1/k$.

We conclude that the tail is less than $1/k$. It follows that if $n \geq 3$, then our full sum is less than

$$\left(1 + \frac{1}{2^2} + \frac{1}{3^2}\right) + \frac{1}{3}$$

(the $1/3$ term is our (over)estimate of the tail). The above sum is $1.69\overline{4}$, which is less than 1.7.

Comment. There is also an interesting induction proof. From the fact that the sum up to the $1/n^2$ term is less than 1.7, we obviously cannot conclude in any simple way that the sum up to the $1/(n+1)^2$ term is less than 1.7, since the second sum is obviously *larger* than the first. But we can push our way through by *strengthening* the induction hypothesis. Here are the details. Let $S(n)$ be the sum of the terms up to $1/n^2$.

We will prove by induction that for any positive integer $n \geq 3$, $S(n) < 1.7 - 1/n$. The result is true (barely) when $n = 3$, by a straightforward calculation. Assume that the result is true when $n = k$, where $k \geq 3$. We show that the result is true when $n = k + 1$.

Note that $S(k+1) = S(k) + 1/(k+1)^2$. Thus by the induction hypothesis,

$$S(k+1) = S(k) + \frac{1}{(k+1)^2} < 1.7 - \frac{1}{k} + \frac{1}{(k+1)^2} = 1.7 - \left(\frac{1}{k} - \frac{1}{(k+1)^2}\right).$$

However,

$$\frac{1}{k} - \frac{1}{(k+1)^2} = \frac{k^2+k+1}{k(k+1)^2} = \frac{1}{k+1}\frac{k^2+k+1}{k^2+k} > \frac{1}{k+1}.$$

It follows that $S(k+1) < 1.7 - 1/(k+1)$.

This shows that $S(n) < 1.7$ for all n, since it is obvious that $S(n) < 1.7$ when $n = 1$ and $n = 2$, and for $n \geq 3$ we have just proved the stronger result $S(n) < 1.7 - 1/n$.

Note that the above argument owes a lot to our first strategy for estimating $S(n)$. If, with no previous experimentation, we looked for a function $G(n)$ for which we could prove by an easy induction that $S(n) < 1.7 - G(n)$, the search might be frustrating.

Problem 2. What is the sum of the *real* roots of $x^4 - 2x^3 + x^2 - 2x + 1$? Hint: Divide by x^2 and let $u = x + 1/x$.

Solution. Since $x = 0$ is not a root, our equation is equivalent to the equation

$$x^2 - 2x + 1 - \frac{2}{x} + \frac{1}{x^2} = 0 \quad \text{that is} \quad \left(x^2 + \frac{1}{x^2}\right) - 2\left(x + \frac{1}{x}\right) + 1 = 0.$$

Let $u = x + 1/x$. Then $u^2 = x^2 + 2 + 1/x^2$, and therefore $x^2 + 1/x^2 = u^2 - 2$. So x is a solution of our equation if and only if

$$(u^2 - 2) - 2u + 1 = 0.$$

We are looking at the equation $u^2 - 2u - 1 = 0$. The roots are $1 \pm \sqrt{2}$.

We have seen in earlier problems that if x is real and positive, then $x + 1/x \geq 2$. Thus if x is negative, then $x + 1/x \leq -2$. But $|1 - \sqrt{2}| < 2$, so if x is real, we cannot have $u = 1 - \sqrt{2}$. (We could find this out later, by trying to solve for x, and finding that the roots are not real. But why bother with a fruitless quest?

So we are looking at the equation $x + 1/x = 1 + \sqrt{2}$, or equivalently $x^2 - (1 + \sqrt{2})x + 1 = 0$. There is no need to find the roots, since we only want their sum, which is the negative of the coefficient of x, that is, $1 + \sqrt{2}$.

Comment. If we have a polynomial equation where the coefficients read the same forwards as backwards (a *palindromic equation*, the same substitution strategy can be used to simplify the equation. The key fact is that $x^n + 1/x^n$ is a polynomial in the variable $u = x + 1/x$. We illustrate with $x^3 + 1/x^3$. Note that

$$\left(x + \frac{1}{x}\right)^3 = x^3 + 3x + \frac{3}{x} + \frac{1}{x^3},$$

and therefore $x^3 + 1/x^3 = u^3 - 3u$.

Problem 3. There are 6678 powers of 2 between 1 and 10^{2010}. Roughly how many have a decimal representation that begins with 1? Do not attempt to find an exact answer: informed speculation is good enough. (One might guess around 6678/9, but I think such a guess is unreasonable, and indeed it is wildly wrong.)

Solution. It is useful to calculate for a while, to get an idea of what is going on. Look at 2^0, 2^1, 2^2, 2^3, and so on. We could record the leftmost digit of 2^n, or more simply to write "Y" if 2^n begins with 1, and "N" if it does not. The first few entries are YNNNYNNYNNYNNNYNNYNNYNNNY. It looks as if there are never more than 3 consecutive N separating consecutive Y. And fairly often there are only 2 consecutive N separating consecutive Y. On the fairly sparse evidence that we have so far, 2 consecutive N occur more often than 3 consecutive N. Indeed the calculation we have made supports the conjecture that we have the pattern 3, 2, 2, 3, 2, 2, 3, 2, 2, and so on. Further calculation would tend to confirm that conjecture, though in fact it is false.

The small calculation we have done seems to show that a *lot more* than one-ninth of the powers of 2 begin with a 1. How many, roughly? If the conjectured pattern continues to hold (which actually is not quite true), there would be 3 Y's in the first 10 powers of 2, 3 Y's in the next 10, and so on. So the total number of powers of 2 from 1 to 10^{2010} would be about $(3/10)(6678)$, or roughly 2003.

All this is reasonably plausible, but we have only computed the first digit of 2^n for $0 \le n \le 24$, and 25 is a pretty small number. If we were doing a statistical study, such a tiny nonrandom sample would be considered laughably inadequate. We need to prove *something* that will enable us to deal with larger n.

Suppose that 2^n has first digit equal to 1, meaning that we wrote Y for 2^n. . We will say that n is of Type 1 if $1 \times 10^k \le 2^n < 1.25 \times 10^k$ for some integer k. Otherwise, that is, if $1.25 \times 10^k < 2^n < 2 \times 10^k$ for some k, we say that n is of Type 2.

If n is of Type 1, we have $2 \times 10^k \le 2^{n+1} < 2.5 \times 10^k$, $4 \times 10^k \le 2^{n+2} < 5 \times 10^k$, $8 \times 10^k \le 2^{n+3} < 10 \times 10^k$, and $1.6 \times 10^{k+1} \le 2^{n+4} < 2 \times 10^{k+1}$. This means that there are exactly 3 N's until the next Y.

If n is of Type 2, we have $2.5 \times 10^k < 2^{n+1} < 4 \times 10^k$, $5 \times 10^k < 2^{n+2} < 8 \times 10^k$, $1 \times 10^{k+1} < 2^{n+3} < 1.6 \times 10^{k+1}$. This means that there are exactly 2 N's until the next Y.

Thus some of our empirical observations on the first 25 powers of 2 have now been proved: the "gap" is always 3 or 2. It follows that at least $1/4$ of the powers of 2 between 1 and 10^{2010} begin with a 1.

We can get some extra information almost for free. Suppose we are looking at a type 1 Y. Then the Y that we get after the 3 N's is of Type 2, since it lies between $1.6 \times 10^{k+1}$ and $2 \times 10^{k+1}$. So the next Y comes up after 2 more N's. And this Y lies between $1.28 \times 10^{k+2}$ and $1.6 \times 10^{k+2}$, so it is of Type 2. We conclude that if we are at a Type 1 number, and we count forward until we have seen 3 more Y, we will go forward by exactly 10 steps.

If we are looking at a Y of Type 2, we can't be as definite. The Y we get next might be of Type 1 or of Type 2. But if it is of Type 1, then for sure the next Y is of Type 2. So if we are at a Type 2 number, and we count forward until we have seen 3 more Y, we will go forward at most 10 steps. (It need not be exactly 10. With a bit of effort one can find three consecutive Y of Type 2. They are rare.)

So if we start with a Y, and go until the *third* Y after that, we will have incremented the exponent of 2 by no more than 10 (sometimes it will be 9). It follows that at least $3/10$ of the powers of 2 between 1 and 10^{2010} begin with a 1.

Comment. It turns out that the long run fraction of the powers of 2 that begin with 1 is $\log_{10} 2$, about 0.301. Our fairly simple minded calculations have given a result (the fraction is at least $3/10$) which is a remarkably good approximation of the truth! A full proof of the $\log_{10} 2$ result is accessible to a Math undergraduate, but not easy.

In this problem, there is almost nothing special about 2. Let a be a positive integer which is not a power of 10. The long run fraction of the powers of a that begin with 1 is about $\log_{10} 2$. The game can be played with other digits. If d is a digit, in the long run the fraction of the powers of a that start with the digit d is $\log_{10}(1 + 1/d)$.

The following seems to be often empirically more or less true. Take a very diverse (whatever that means) collection of numerical physical data. Then roughly $\log_{10} 2$ of the numbers start with 1, roughly $\log_{10}(3/2)$ of the numbers start with 2, and so on. The phenomenon is not universal, and is not completely understood, though some partial explanations have been given. The empirical observation about the distribution of initial digits is usually called *Benford's Law*.