Solutions to March 2008 Problems

Problem 1. Find integers a, b, c, d, and n such that

$$\frac{1}{\sqrt{2}+\sqrt{3}+\sqrt{6}} = \frac{a+b\sqrt{2}+c\sqrt{3}+d\sqrt{6}}{n}.$$

Solution. Presumably we are not being sent on a wild goose chase, and there are such integers. So we will proceed on the assumption that there are, and on that assumption find out what they must be. Once we have found candidates, it is an easy matter to verify that they work. Rewrite our equation as

$$(\sqrt{2}+\sqrt{3}+\sqrt{6})(a+b\sqrt{2}+c\sqrt{3}+d\sqrt{6}) = n.$$

Expand, and gather like terms together. After some computation, the left-hand side becomes:

$$(2b + 3c + 6d) + (a + 3d + 3c)\sqrt{2} + (a + 2d + 2b)\sqrt{3} + (c + b + a)\sqrt{6}.$$

We would like the above expression to be a non-zero integer n. Actually, we don't really need n, a, b, c, and d to be integers, they can be non-zero rationals. This is because if we find rationals that work, we can always find integers that work by multiplying everything by an appropriate integer.

Let us look for rationals a, b, c, and d, and n, not all 0, that do the job. Look at the coefficients of $\sqrt{2}$, $\sqrt{3}$, and $\sqrt{6}$ in the above expression. If these can be made all equal to 0, everything will be fine. (In fact, it can be shown that these coefficients *must* be made equal to 0 if we are to get a rational, but we do not need to prove that.) To make the coefficients 0, we need to solve the following system of equations.

$$a+3c+3d=0; \quad a+2b+2d=0; \quad a+b+c=0.$$

We have 3 equations in 4 unknowns. Let's put $d=-1$ to see what happens. (If some non-zero d will do the job, then $d=-1$ also will.) So we get the equations $a+3c=3$, $a+2b=2$, $a+b+c=0$. This system of equations is easy to solve. After a while we get $a=-12$, $b=7$, $c=5$. Then $2b+3c+6d=23$. (We got "lucky," and got an integer solution straightaway.) We conclude that our desired representation holds, with $n=23$, $a=-12$, $b=7$, $c=5$, and $d=-1$. All the other solutions can in fact be obtained by multiplying the one we have given by an integer, so our solution is almost unique.

Another Way. Look at $\sqrt{2}+\sqrt{3}+\sqrt{6}$, and replace $\sqrt{3}$ by $-\sqrt{3}$. Do the replacing also at the "hidden" $\sqrt{3}$: note that $\sqrt{6}=\sqrt{2}\sqrt{3}$, so $\sqrt{6}$ should be replaced by $-\sqrt{6}$. So we obtain $\sqrt{2}-\sqrt{3}-\sqrt{6}$. Now calculate the product

$$(\sqrt{2}+\sqrt{3}+\sqrt{6})(\sqrt{2}-\sqrt{3}-\sqrt{6}).$$

After not much work, we arrive at the pleasantly simple expression

$$-7-6\sqrt{2}.$$

Multiply this by $-7+6\sqrt{2}$. We get -23. It follows that

$$(\sqrt{2}+\sqrt{3}+\sqrt{6})(\sqrt{2}-\sqrt{3}-\sqrt{6})(-7+6\sqrt{2})=-23.$$

Now we are almost finished. We get that

$$\frac{1}{\sqrt{2}+\sqrt{3}+\sqrt{6}}=\frac{(\sqrt{2}-\sqrt{3}-\sqrt{6})(-7+6\sqrt{2})}{-23}.$$

The "bottom" of the right-hand side is an integer, as desired, though we might like to multiply top and bottom by -1 to make the bottom positive. As to the top of the right-hand side, it doesn't yet look quite right, but if we multiply it out we get something of the right shape.

Problem 2. There are 20 people in a tango class. How many ways are there to divide them up into 10 dance couples? (This is 21st century Canada, there should be no assumptions based on sexual identity.)

Solution. Imagine lining up the people in order of age, or beauty, or (forgot to mention these are UBC students) by Student Number. Take the first person in the list, and *choose* one of the remaining people for her/him to be partnered with. There are 19 "remaining" people, so the choosing can be done in 19 ways. Once the choosing is done, there are 18 people left. Take the first person in the list who is as yet unpartnered, and choose one of the 17 others for her/him to be partnered with. For every one of the 19 ways of doing the first partnering, there are 17 ways of doing the second partnering, for a total so far of $19\cdot 17$. Now there are 16 people left. Take the first person in the list left unpartnered, and choose a partner for her/him from the remaining 15 people. We can see that there are $19\cdot 17\cdot 15$ ways of doing the first three partnerings. Continue. The total number of ways to do the partnering is

$$19\cdot 17\cdot 15\cdot\cdots\cdot 5\cdot 3\cdot 1.$$

(The 1 at the end is superfluous, but looks nice.)

Comment 1. We describe a somewhat more formal way of doing exactly the same thing. Let $P(2k)$ be the number of ways of pairing off a group of $2k$ people. Now imagine $2n$ UBC people ready to be paired off. Line them up in some order, say by Student Number, from smallest to largest. Take the person with lowest student number. There are $2n-1$ ways to choose a partner for her/him. For every one of these ways, we have $2n-2$ people left, so by definition there are $P(2n-2)$ ways to pair them. We have just proved that

$$P(2n) = (2n-1)P(2n-2).$$

Using this recurrence relation repeatedly, we quickly find our previous expression for $P(20)$.

Comment 2. We have 20 UBC students, and want to form 4 basketball teams (define a team as 5 people). Line the students up by Student Number. Take the leftmost person. There are $\binom{19}{4}$ ways to choose the 4 people who are to join her/him on the team. Now there are 15 people left. Take the leftmost person not yet chosen, and choose 4 people to join her/him. This can be done in $\binom{14}{4}$ ways. Continue. The total number of ways is $\binom{19}{4}\binom{14}{4}\binom{9}{4}\binom{4}{4}$. It is very easy to get the wrong answer by arguments that sound plausible, and that all overcount.

Another Way. First we choose 10 people from the 20 to take on the traditional male role ("leading"), or, if you prefer, choose 10 people to wear the traditional female clothes. There are $\binom{20}{10}$ ways of doing this. Line up the leaders in order of Student Number. The first leader can be paired off with one of the remaining 10 people in 10 ways. For everyone of these ways, the second leader can be paired off with one of the remaining 9 people in 9 ways, and so on. We conclude that the number of ways of making 10 tango pairs *with designated leaders* is

$$\binom{20}{10} \cdot 10 \cdot 9 \cdot \cdots \cdot 2 \cdot 1.$$

Call this number $L(20)$. Is this equal to $P(20)$? No! Take any particular way of pairing off people. Line up the pairs, as $\pi_1, \pi_2, \ldots, \pi_n$. In each pair, the 2 people in the pair argue about who is to lead. For each pair, there are 2 choices for who is to lead. So once the pairing has been done, there are 2^{10} patterns of which one of the pair is to lead. It follows that $L(20) = 2^{10}P(20)$. From our earlier expression for $L(20)$, we conclude that

$$P(20) = \frac{\binom{20}{10} \cdot 10 \cdot 9 \cdot \cdots \cdot 2 \cdot 1}{2^{10}}.$$

The "top" is $\frac{20!}{(10!)(10!)}10!$, that is, $20!/10!$. But

$$20! = (19 \cdot 17 \cdot \cdots \cdot 1)(20 \cdot 18 \cdot \cdots \cdot 2) = (19 \cdot 17 \cdot \cdots \cdot 1)(2^{10})(10 \cdot 9 \cdot \cdots \cdot 1).$$

Now it is easy to see that $P(20)$ as calculated here gives the same answer as in the first solution.

Comment 3. In the solution above, we used a strategy of controlled multiple counting. But we know by what factor we overcounted, so by dividing by the right thing, in this case 2^{10}, we get the right number.

Another Way. The next solution also works by multiple counting. This time our first attempt will be deliberately wrong, but fixable.

The first pair can be chosen in $\binom{20}{2}$ ways. For every such choice, there are $\binom{18}{2}$ ways of choosing the second pair, for a total of $\binom{20}{2}\binom{18}{2}$ ways of choosing the first two pairs. Continue. We get $\binom{20}{2}\binom{18}{2} \cdot \binom{4}{2}\binom{2}{2}$ ways of choosing all the pairs.

But this isn't right! For one thing, it can be quickly checked that it gives an answer which is different from the one of the first two solutions. Let's look at the expression we got for the number of ways to pick the "first" two pairs. Think of 4 people, A, B, C, and D. The pair {A, B} was among the pairs counted by $\binom{20}{2}$, as was the pair {C, D}. Once we had chosen {A, B} as the "first" pair, the pair {C, D} is among the $\binom{18}{2}$ "second" pairs. However, we can reverse the roles of {A, B} and {C, D}, picking the second first and the first second. However, that gives the same *pairings*: we are multiply counting.

We can fix this by realizing that what we have just counted is the number of ways of picking pairs *and* lining up the pairs in a row. Let $R(20)$ be the number of ways of doing that: we have just counted $R(20)$.

Let us count $R(20)$ another way: pick the pairs (there are $P(20)$ ways of doing this), and *then* line them up (as pairs). There are 10! ways of doing this. We conclude that $R(20) = 10!P(20)$, and we conclude that

$$P(20) = \frac{\binom{20}{2}\binom{18}{2}\cdots\binom{4}{2}\binom{2}{2}}{10!}.$$

If we simplify, we will get exactly the same answer as we got in the first two solutions.

Another Way. We sketch an approach that involves more extreme overcounting. Pick a person, then pick a partner for that person, then pick another person, pick a partner, and so on. Basically we are lining up the 20 people in a row, and saying the first two people are partners, the next two people are partners, and so on. There are 20! ways to do this.

The number 20! overcounts the truth in two different ways. Think of the ordering ABCDEF...ST. Possibly interchanging A and B, also (independently) possibly interchanging C and D, also E and F, and so on would produce the same pairings. The number of ways to interchange is 2^{10}. So we must divide by this. But once we have done that, we have obtained the pairs *lined up* in a certain way. The same pairings could have been lined up in 10! different ways. As in the previous solution,correct for this by dividing by 10!. We obtain the answer $20!/(10!2^{10})$.

Problem 3. A collection of 6 different integers is chosen at random from the set {1, 2, 3, ..., 48, 49}, with all choices equally likely. What is the probability that no 2 of these 6 numbers are consecutive?

Solution. There are many possible approaches. We describe a particularly efficient approach, in which the numbers 49 and 6 play a very minor role.

Imagine picking 6 integers in the interval from 1 to 49, in such a way that no two consecutive integers are picked. Let the integers be, in increasing order, a_1, a_2, a_3, a_4, a_5, and a_6.

Let $x_1 = a_1 - 0$, let $x_2 = a_2 - a_1 - 1$, $x_3 = a_3 - a_2 - 1$, $x_4 = a_4 - a_3 - 1$, $x_5 = a_5 - a_4 - 1$, $x_6 = a_6 - a_5 - 1$, and $x_7 = 50 - a_6$. Note that for example x_2 is the number of integers in the "gap" between a_2 and a_1. The numbers x_3, x_4, x_5, and x_6 have similar descriptions.

The numbers x_1 and x_7 are a little special. A good way to visualize them is to expand our collection of numbers to the integers from 0 to 50. Then x_1 is the number of integers from 0 (inclusive) to the first number picked, that is, in the interval $[0, a_1)$, while x_7 is the number of integers in the interval $(a_6, 50]$.

Note that the x_i are all positive, and that once we know the a_i, we know the x_i, and vice-versa. Note also that

$$x_1 + x_2 + x_3 + x_4 + x_5 + x_6 + x_7 = 49 + 2 - 6.$$

Any choice of 6 numbers of which no two are consecutive yields in this way a sequence $x_1, x_2, \ldots, x_7$ of positive integers whose sum is 45. Conversely, any sequence $x_1, x_2, \ldots, x_7$ of positive integers whose sum is 45 uniquely identifies a collection of 6 numbers chosen from the interval from 1 to 49, with no two numbers consecutive.

Thus in order to count the number of ways of choosing 6 numbers, with no two consecutive, all we need to do is to count the number of solutions in positive integers of the equation $x_1 + \cdots + x_7 = 45$.

This problem is fairly easy to solve. Imagine putting 45 coins in a row, separated from each other by a little gap. Then there are 44 gaps. Now choose 6 of these gaps, and put a divider into each chosen gap. Let x_1 be the number of coins up to the first divider, x_2 the number of coins from the first divider to the second divider, and so on. Any choice of 6 gaps determines in this way a solution of $x_1 + \cdots + x_7 = 45$ in positive integers, and any solution determines where the dividers should go. So there are as many solutions as there are ways of placing dividers, namely $\binom{44}{6}$.

But there are $\binom{49}{6}$ ways of choosing 6 numbers in the interval from 1 to 49. It follows that the probability that no two of these numbers are consecutive is

$$\frac{\binom{44}{6}}{\binom{49}{6}}.$$

Now in principle we are finished. But if we want a numerical answer, we can note that $\binom{44}{6} = (44)(43)(42)(41)(40)(39)/6!$, with a similar expression for $\binom{49}{6}$. So our probability is

$$\frac{(44)(43)(42)(41)(40)(39)}{(49)(48)(47)(46)(45)(44)}.$$

To a few decimal places, this is 0.5048. Note that it follows that the probability that 2 or more of the numbers are adjacent is about 0.4952, almost 1/2. The probability of 2 or more adjacent numbers is much higher than most people think.

Problem 4. Calculate $\displaystyle\sum_{k=1}^{\infty} \frac{k}{k^4+k^2+1}$.

Solution. Note that for all k,

$$k^4+k^2+1=(k^2+k+1)(k^2-k+1). \tag{1}$$

Identity 1, like all algebraic identities, is easy to verify. In this case, just multiply out the right-hand side and simplify.

Let's see how we might *discover* Identity 1. The identity $x^3-1=(x-1)(x^2+x+1)$ is probably familiar. It follows, by putting $x=k^2$, that

$$k^6-1=(k^2-1)(k^4+k^2+1). \tag{2}$$

But

$$k^6-1=(k^3-1)(k^3+1)=(k-1)(k^2+k+1)(k+1)(k^2-k+1) \tag{3}$$
$$=(k^2-1)(k^2+k+1)(k^2-k+1). \tag{4}$$

Now a comparison of Equations 2 and 3 yields Identity 1.

It is easy to verify that

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

Let $f(x)=x^2-x+1$. It is easy to verify that $f(x+1)=x^2+x+1$. Now let $G(x)=(1/2)/(x^2-x+1)$. We conclude that for any integer k,

$$\frac{k}{k^4+k^2+1}=G(k)-G(k+1).$$

Define the function $S(n)$ by

$$S(n)=\sum_{k=1}^{n}\frac{k}{k^4+k^2+1}$$

By our previous calculations, we have

$$S(n)=(G(1)-G(2))+(G(2)-G(3))+(G(3)-G(4))+\cdots+(G(n)-G(n+1)).$$

Add up, and note the wholesale cancellation. We find that

$$S(n)=G(1)-G(n+1)=\frac{1}{2}-\frac{1}{2(n^2+n+1)}. \tag{5}$$

Now to find the sum asked for in the problem, look at Equation 5, and let n get very large. As n gets very large, the sum approaches $1/2$, so our sum is $1/2$.

Another Way. Let's proceed more experimentally, by calculating the "sum" of the first 1 term(s), the sum of the first 2 terms, the sum of the first 3 terms, and so on for a while.

The first term is $1/3$, the second term is $2/21$, the third term is $3/91$, the fourth term is $4/273$, and the fifth term is $5/651$.

So the sum of the first 1 term(s) is $1/3$. The sum of the first 2 terms is $(1/3) + (2/21)$, that is, $9/21$, which simplifes to $3/7$. The sum of the first 3 terms is $(3/7) + (3/91)$, which simplifies to $6/13$. The sum of the first 4 terms is $6/13 + 4/273$, which simplifies to $10/21$. And the sum of the first 5 terms simplifies to $15/31$. There are certainly strong hints of a pattern. The denominators are 3, $3+4$, $3+4+6$, $3+4+6+8$, and $3+4+6+8+10$. The numerators are (almost) $1/2$ of the denominator, with the fit getting closer as we go on. It is natural to guess that the "infinite" sum we are looking for will turn out to be $1/2$. So far we only have guesses, based on rather skimpy numerical evidence.

We put flesh on the guesses, by hoping that if we sum to the k-th term, our denominator will be $3+4+6+\cdots+2n$, or more attractively $1+2(1+2+\cdots+n)$. Summing the arithmetic progression, we get a conjectured denominator of n^2+n+1. We also conjecture that the numerator is $(n^2+n)/2$.

Define the function $T(k)$ by

$$T(k) = \frac{(k^2+k)/2}{k^2+k+1}.$$

We would like to show that for any n, the sum of the first n terms is $T(n)$.

For $k \geq 2$, calculate $T(k) - T(k-1)$. If our formula is right, then the difference should be the k-th term. After a little work, we find that indeed $T(k) - T(k-1) = k/(k^4+k^2+1)$. The work is made easier if we make the preliminatry observation that $T(k) = (1/2)(1 - 1/(k^2+k+1))$. Note that the first term of our sum happens to be $T(1) - T(0)$. It follows that

$$\sum_{k=1}^{n} \frac{k}{k^4+k^2+1} = (T(1) - T(0)) + (T(2) - T(1)) + \cdots + (T(n) - T(n-1)).$$

Note the wholesale cancellations (telescoping). Our sum turns out to be $T(n) - T(0)$, or more simply, since $T(0) = 0$, our sum is $T(n)$.

That is what we wanted to show. And now that we have a formula for the sum of the first n terms, we can find the "infinite" sum as in the first solution.

Problem 5. Some non-negative integers can be expressed as the sum of two perfect squares. For example, $0 = 0^2 + 0^2$, $1 = 0^2 + 1^2$, $2 = 1^2 + 1^2$, $4 = 0^2 + 2^2$, and $5 = 1^2 + 2^2$. Some, like 25, 50, 65, and others, can even be so expressed in more than one way. But for example 3, 6, and 7 cannot be expressed as a sum of two squares. Among the integers from 0 to 999999, which ones are more common, the ones that can be expressed as the sum of two perfect squares, or the ones that cannot be so expressed?

Solution. We could experiment, by finding out, for various fairly small integers N, how many integers from 0 to N can be represented as the sum of two squares. From 0 to 9, there are 7; from 0 to 19 there are 12; from 0 to 29 there are 16; from 0 to 39 there are 20. It looks like a pretty close race, though the sums of two squares got a big head start in the interval from 0 to 9, and now seem to be losing ground. It seems very difficult to carry out the calculations by hand all the way up to 999999, but it would not be too hard to write a computer program to do the job. But we will take another approach.

If n can be expressed as a sum of two perfect squares, then there exist integers a, b, where without loss of generality we may take $a \le b$, such that $n = a^2 + b^2$. Note that $b \le 999$. We will count the number of pairs (a, b) where $0 \le a \le b \le 999$. There are $\binom{1000}{2}$ pairs that have $a < b$, and 1000 pairs that have $a = b$, for a total of 500500 pairs. So no more than 500500 of our numbers can be expressed as a sum of two squares.

This is an obvious *overestimate,* for two reasons: (i) Some integers, like 25, 50, 65, and many others, can be represented as a sum of squares in two or more fundamentally different ways. For example, 50 comes from the pair $(5, 5)$, but also from the pair $(1, 7)$. This means that counting the ordered pairs (a, b) with $0 \le a \le b \le 999$ overcounts the number of representable numbers, since sometimes two pairs represent the same number. (ii) Some of our 500500 pairs, like $(900, 950)$, give a sum of squares which is bigger than 999999, so again counting ordered pairs (a, b) with $0 \le a \le b \le 999$ overcounts the representable numbers.

We will be able to answer our question if we can whittle down our crude estimate of no more than 500500 by 501 or more. That can be done by taking advantage of observation (i) above, or of observation (ii), or a combination. We first use (i) alone. Note that $25 = 0^2 + 5^2 = 3^2 + 4^2$. But then for every positive integer k, $25k^2$ has at least two distinct representations, namely $25k^2 = 0^2 + (5k)^2$ and $25k^2 = (3k)^2 + (4k)^2$. How many of our numbers up to 999999 are of the type $25k^2$? This is easy to do with a calculator, or without. Calculate 999999/25, take the square root of the result. We get a number between 199 and 200, so there are 199 possibilities for k.

Now we use the fact that $50 = 5^2 + 5^2 = 1^2 + 7^2$. Like with 25, this means that there are at least two representations for any positive integer of the form $50k^2$, where k is a positive integer. To see how many such there are in our interval, calculate 999999/50, and calculate the square root (this time the calculator is handy). We find that there are 141 values of k that work.

In a similar way, from the doubly representable number 65, we get 124 at least doubly represented numbers in our interval, and from the doubly represented number 85 we get 108.

It is easy to verify that the at least doubly represented numbers we have obtained from 25, 50, 65, and 85 are all *different.* That gives us already $199 + 141 + 124 + 108$ "duplicates." So our crude estimate of no more than 500500 is too big by at least 572. We conclude that the integers in our interval which are *not* representable are in the majorit

The problem is now solved, but let's see what observation (ii) would have given us. Note that $708^2 + 708^2 > 999999$. So all ordered pairs (a, b) with $708 \le a \le b \le 999$ could not possibly give us a representation of a number in our interval. For $a = 708$, that gives 292 pairs that cannot work. For $a = 709$, it gives 291 such pairs, for $a = 710$ it gives 290 such pairs, all the way down to 1 pair for $a = 999$. That's 42778 pairs that were counted in our 500500 but can't work! The representable numbers are definitely in the minority.

We could push this weeding out further. For example, look at a ranging from 600 to 707. For each of these, $b = 800$ or more is too big. That gets rid of another 21,600 pairs.

We sketch an approach for getting rid of pairs (a, b) that are "too big" more systematically. The pairs (a, b) that aren't too big are pairs (a, b), with $0 \le a \le b$, and $a^2 + b^2 < 1000000$. Consider the part of the first quadrant that is inside the circle with center the origin and radius 1000, and that lies between the line $y = x$ and the y-axis. This is one-eighth of the circle, and its area is, with an error that is relatively small, equal to the number of pairs we are interested in. So the number of pairs has order of magnitude $1000000\pi/8$, about 400000.

Comment 4. Let $T(n)$ be the number of integers less than n which are the sum of two squares. From the work of Laudau, Ramanujan, Hardy, and many others, it is known that for n large,

$$T(n) \approx K \frac{n}{\sqrt{\ln n}},$$

where K is a constant known to be about 0.7642. This formula predicts that about 300000 integers less than 1000000 are the sum of two squares. It would be interesting, but tedious, to check *exactly* how many there are! Note that for large n, $T(n)$ is substantially smaller than what is obtained by weeding out pairs (a, b) that are "too large." So the work is much more subtle, and depends on getting very firm control over the average number of representations of numbers as a sum of two squares.