# UMASS AMHERST MATH 300 SP '05, F. HAJIR

HOMEWORK 9: COMPLEX NUMBERS

## 1. In the beginning ...

In the beginning, there is the number 1. Then $1+1$ makes 2, $1+2$ makes 3, and the rest is history. We get all the positive whole numbers. After a while, we ponder the reverse of this "adding 1" process and discover the "take away 1" process, which gives us $3-1$ is 2, $2-1$ is 1, $1-1$ makes ... a new and wondrous number, namely 0. Moreover, $0-1$ makes $-1$, $-1$ minus 1 makes $-2$ and so on giving us the negative numbers. The new system of numbers,

$$\mathbb{Z} = \{\ldots, -2, -1, 0, 1, 2, \ldots\}$$

has many wonderful qualities and internal relationships. For instance, $-2$ is defined to be $-1-1$ but it is also $(-1)+(-1)$. There are quite a few popular books which discuss the history of 0 and negative numbers. According to an ancient tradition of mathematics, which lasted through the fifteenth and sixteenth centuries in European mathematical circles (!), negative solutions of simple equations were not accepted as "proper" well-behaved solutions and were discarded at the end of the solving process. These days, of course, any accountant knows that negative numbers cannot be simply discarded ... or do they? [Can you say "Enron?" How about "Worldcom?"]

Once one chooses to accept the system $\mathbb{Z}$ of integers as a legitimate collection of numbers, it is difficult to avoid noting its charms. There they are, the integers, marching off to infinity at perfectly regular discrete intervals in two directions, with a neutral point 0 smack in the middle perfectly balancing out the positive and negative integers. As we discussed earlier, there are two basic *simply fantasmagoric, luscious* binary operations defined on $\mathbb{Z}$, namely $+$ and $\times$. What this implies is that for any fixed integer $k \in \mathbb{Z}$ we can attach two *actions* $\tau_k$ and $\delta_k$ corresponding to addition by $k$ and multiplication by $k$. We use $\tau_k$ to stand for translation by $k$ and $\delta_k$ for dilation by $k$ in order to emphasize the geometric meaning of these operators: $\tau_k$ shifts all the integers $k$ places (to the right if $\text{sign}(k) = +1$ and to the left if $\text{sign}(k) = -1$), and $\delta_k$ dilates everything (streches $\mathbb{Z}$) by a factor of $k$ (if $\text{sign}(k) = -1$ the there is a stretch by the factor $|k|$ followed by a 180-degree rotation of the line). To give a precise algebraic definition, for each $k \in \mathbb{Z}$, we have two self-maps $\tau_k, \delta_k$ of $\mathbb{Z}$, namely

$$\begin{aligned} \tau_k : \mathbb{Z} \to \mathbb{Z} \text{ defined by} \qquad & \tau_k(n) = k + n \text{ for all } n \in \mathbb{Z} \\ \delta_k : \mathbb{Z} \to \mathbb{Z} \text{ defined by} \qquad & \delta_k(n) = kn \text{ for all } n \in \mathbb{Z}. \end{aligned}$$

Note that $\tau_k$ is a bijection for all $k \in \mathbb{Z}$. But $\delta_k$ is a bijection for only *very few* integers $k$. Which ones?! Well, the map $\delta_k$ is injective as long as $k \neq 0$ (prove it!). But the image of $\delta_k$ is what we have been calling $k\mathbb{Z}$, namely the set of all multiples of $k$. The latter is all of $\mathbb{Z}$ if and only if $k = \pm 1$. Note that $\pm 1$ are the only integers which have no prime divisors (they are "units"). They are also the only ones whose multiplicative reciprocal belongs to $\mathbb{Z}$.

Any algebraic manipulation with integers can be reinterpreted in terms of the "geometry" of the maps $\tau : \mathbb{Z} \to \mathrm{Maps}(\mathbb{Z}, \mathbb{Z})$ and $\delta : \mathbb{Z} \to \mathrm{Maps}(\mathbb{Z}, \mathbb{Z})$. For instance, why were we motivated to create the negative numbers in the first place? Because the translation map $\tau_1$ (or rather its restriction to $\mathbb{N}$[1]) naturally wants to have an inverse map. Now $\tau_1|_{\mathbb{N}}$ doesn't quite have an inverse because 1 is not in its image! Thus, we are led to creating a symbol 0 which serves as $\tau_1^{-1}(1)$. One then proves that $\tau_0$ is the identity map, i.e. $0 + n = 0$ for all $n$. By the same process, $-1 = \tau_1^{-1}(0)$, and now one proves that $\tau_{-1} = \tau_1^{-1}$!

Note that for any integer $k$, $\tau_k$ is the compositum of $k$ copies of $\tau_1$, resp. $\tau_{-1}$, if $k$ is positive, resp. negative.

Recalling how the lack of surjectivity of $\tau_k$ led to the creation of negative numbers, we recall that the maps $\delta_k : \mathbb{Z} \to \mathbb{Z}$ are injective for $k \neq 0$, but they turn out not to be surjective for $|k| > 1$. So, for $|k| > 1$, and $n \in \mathbb{Z}$, we want an element $\delta_k^{-1}(n)$, which is in general not in $\mathbb{Z}$. So, we "create" a new set $\mathbb{Q}$ which is the smallest set which fills in the "missing" numbers. Namely, we create a set $\mathbb{Q}$ by defining on the set $\mathbb{Z} \times \mathbb{Z}_0$[2] the equivalence relation $(a, b) \sim (c, d)$ if and only if $ad - bc = 0$, then we put $\mathbb{Q} = (\mathbb{Z} \times \mathbb{Z}_0)/\sim$ for the set of equivalence classes. We think of the equivalence class $\widetilde{(a, b)}$ as the fraction $a/b$. We have a natural injection $\mathbb{Z} \hookrightarrow \mathbb{Q}$ given by $n \mapsto \widetilde{(n, 1)}$. In this way, we "think of" $\mathbb{Z}$ as a subset of $\mathbb{Q}$. On $\mathbb{Q}$, we know how to define addition and multiplication by the "think-of-them-as-fractions" yoga, namely

$$\widetilde{(a, b)} + \widetilde{(c, d)} = \widetilde{(ad + bc, bd)}, \qquad \widetilde{(a, b)} \times \widetilde{(c, d)} = \widetilde{(ac, bd)}.$$

Now the maps $\tau_k, \delta_k : \mathbb{Q} \to \mathbb{Q}$ is easy to define, just as before as addition and multiplication by $k$ maps, with the advantage, now, that $\tau_k$ is a bijection of $\mathbb{Q}$ to itself for all $k$ and $\delta_k$ is a bijection of $\mathbb{Q}$ to itself for all $k \neq 0$. The fact that we cannot make $\tau_0$ into a bijection by extending its field of definition is due to its being so *horrifically* (is that a word?) non-injective!

We discussed earlier in the semester that the Greeks were quite happy with their system of numbers ($\mathbb{Q}$) until they discovered that the equation $x^2 - 2 = 0$ does not have a solution in this number system. The Greeks, knew, however, that the quantity $\sqrt{2}$ can be approximated as well as one wishes by a sequence of rational numbers, e.g. $1, 1.4, 1.41, 1.414, \ldots$. What is amiss is that this sequence does not have a limit in the set $\mathbb{Q}$ itself. Thus, what is needed is to have a number system in which all sequences of rational numbers which "should" tend to an acutal number *do* have a limit. Thus, another extension of their number system was required. The eventual solution was to write rational numbers in decimal expansion, and then note that there are lots of decimal expansions that do not express rational numbers, because the decimal expansion of a rational number always ends in a repeating finite pattern. Thus, the set of all real numbers, $\mathbb{R}$ is defined to be the set of numbers expressible as a decimal expansion. One then shows that any sequence of real numbers which "should" have a limit does indeed have a limit. (The concept of "should have a limit" which I am leaving vague here can be made precise of course and goes under the name of "Cauchy sequence" for the fabulous mathematician Auguste Louis Cauchy. You will learn more about this in Math 523).

[1]Whenever $f : X \to Y$ is a map, and $S$ is a subset of $X$, we get a map $S \to Y$ in a simple way, namely by restricting the "domain" to $S$. In other words, we define the *restriction of $f$ to $S$*, denoted $f|_S$ by $f|_S : S \to Y$, $s \mapsto f(s)$ for all $s \in S$.

[2]Recall that $\mathbb{Z}_0 = \mathbb{Z} \setminus \{0\}$

Thus, the set of real numbers, $\mathbb{R}$ is "complete," it has all the properties that one would want. Translation by a real number is invertible. Multiplication by a real number other than 0 is invertible. The real numbers have no "wholes," meaning any sequence of reals that are getting closer and closer together actually converge to a real number. Moreover, Newton showed how to find real solutions of cubic, fifth degree, and higher odd degree polynomial equations by an "iteration" procedure. But the seeming "perfection" of the real number system is tarnished by "only" one minor defect, namely, the simple equation $x^2 + 1 = 0$ does not have any roots! Thi story goes back thousands of years: from ancient times, we have known that there is a simple "quadratic formula" for solving all quadratic equations, namely, if

$$ax^2 + bx + c = 0, \qquad a, b, c \in \mathbb{R}, a \neq 0,$$

then we put

$$\Delta = b^2 - 4ac, x_1 = \frac{-b + \sqrt{\Delta}}{2a}, \qquad x_2 = \frac{-b - \sqrt{\Delta}}{2a},$$

and then $ax^2 + bx + c = a(x - x_1)(x - x_2)$, so that $x_1, x_2$ are the roots of the equation.

Great. The only trouble is how to interpret the symbol $\sqrt{\Delta}$ when $\Delta$ is a negative real number. In other words, how does one solve the equation $y^2 - \Delta = 0$ when $\Delta < 0$? If $y$ is a real quantity, then $y^2 \geq 0$ so $y^2$ simply cannot equal $\Delta < 0$. The easy answer is to say, Dude, you just proved these equations don't have solutions, so just let it rest man! However, if you've followed the "arc" of the story so far, you will have noticed that almost every time that a certain "solution" did not seem to exist to a simple problem, there was a way to create a larger set within which a solution does exist and this new set is a bigger and better place to do mathematics. So the ambitious answer would be: Dude, you have shown there are no **real** numbers $y$ that satisfy the equation $y^2 = \Delta$ when $\Delta < 0$, so are there some **"non-real"** numbers that do satisfy the equation?!

Historically what happened is that people came to realize that if they sometimes allowed the use of *symbols* such as $\sqrt{-1}$ in their calculations, as long as these symbols were handled with care, then they could be manipulated in the usual way and (this next bit was very important historically) they could use these manipulations to find *real* solutions of other equations! This was a great advantage to those who braved the new world of "imaginary" numbers as they came to be called. There is a fascinating story here [for which I highly recommend the book "Imagining Numbers" by Professor Barry Mazur as summer reading for all of you], but let's move on to describing the set $\mathbb{C}$ of complex numbers, a place where all polynomial equations have solutions!

Where to start? We want to solve $y^2 = \Delta$ where $\Delta$ is negative, so how about we try $y^2 = -1$ for starters. We know that $y$ cannot be a real quantity, so we just invent a symbol (traditionally the symbol used is $i$ and we will stick with that) and stipulate that this symbol designates a fixed object with the property that $i^2 = -1$. We will then want this $i$ to interact with real numbers in reasonable ways, so for example, we should be able to add 5 to $i$ to get an object $i + 5$. Also we should be able to multiply by real numbers, say $5i$, $-i$ should be admitted to our system of numbers. We will want some standard commutative/associative rules, so that, for example, $i + i = 2 \cdot i = i \cdot 2$, $2 \cdot (3i) = 6i$ etc. In short, we want to define a new system of numbers to be all those that are "generated" by the real numbers $\mathbb{R}$ and this one new symbol $i$ but keeping the usual nice rules of addition and multiplication. So, we put

$$\mathbb{C} = \{a + bi \mid a, b \in \mathbb{R}\},$$

and call this the set of complex numbers. Note that we can inject $\mathbb{R}$ into $\mathbb{C}$ via $a \mapsto a + 0i$. In this way, we think of $\mathbb{R}$ as a subset of $\mathbb{C}$. So, a complex number $z$ is nothing other than a pair $(a, b)$ of real numbers. What's the big deal about that? Well, as a set, maybe, that's what $\mathbb{C}$ is but on this set we now define some cool operations. Namely, given $z, w \in \mathbb{C}$, we write $z = a + bi$ and $w = c + di$, then we define $z + w = (a + c) + (b + d)i$, i.e. we add two complex numbers coordinate-wise, that's easy. Multiplication is a much niftier operation, so let's be careful here. How should we define $zw = (a + bi)(c + di)$? Recalling that $i^2 = -1$ is its defining quality, its *raison d'être*, its *mantra*, its ... enough already, and that we want to maintain the usual algebraic rules, we compute

$$(a + bi)(c + di) = ac + adi + bic + bidi = ac + (ad + bc)i - bd = (ac - bd) + (ad + bc)i.$$

So that's what we do: on the set $\mathbb{C}$ we define

$$(a + bi) + (c + di) = (a + c) + (b + d)i, \qquad (a + bi)(c + di) = (ac - bd) + (ad + bc)i.$$

One then checks that all the usual algebraic rules do in fact hold!

Now, in this new wonderful set $\mathbb{C}$, what can we then say about the equation $y^2 = -1$? Well, naturally $i$ is a solution of it. But even more than that, $-i$ is also a solution! Just plug it in and see:

$$(-i)^2 = (-i)(-i) = (-1)(-1)(i)(i) = i^2 = 1.$$

So we have $(y - i)(y + i) = y^2 - i^2 = y^2 + 1$, i.e. $y^2 + 1$ now factors completely and has exacly two roots in $\mathbb{C}$, namely $i, -i$. How do we know? Because if $zw = 0$ (where $z, w \in \mathbb{C}$), then either $z = 0$ or $w = 0$. We say that $\mathbb{C}$ has no *zero-divisors*.

The observant reader might now be thinking "Dude, like we had to do, like, *all this* just to solve, um, $y^2 = -1$?! And now we have to go through all these hoops again to solve, like, $y^2 = -2$, and $y^2 = -3$ etc.?" No, Dude, we don't. Watch: if $i^2 = -1$, and $\Delta = -r$ for some real number $r > 0$, then we know we have a positive real number $\sqrt{r}$, so we observe that $\pm i\sqrt{r}$ are two solutions of $y^2 = \Delta$. (As before, these are the only solutions in $\mathbb{C}$).

Consequently, we have the following theorem.

**Theorem 1.1.** *If $a, b, c \in \mathbb{R}$ with $\Delta = b^2 - 4ac = -r < 0$, then the equation $az^2 + bz + c = 0$ has exaclty two solutions in $\mathbb{C}$, namely*

$$z_\pm = \frac{-b \pm i\sqrt{r}}{2a}.$$

Much more is in fact true. It turns that if we consider any polynomial equation with coefficients of any degree $n \geq 1$ in $\mathbb{C}$, then it will always factor completely into linear factors over $\mathbb{C}$. This is known as the fundamental theorem of algebra (although it is an analytic fact!), and the first proof of it was given by Gauss in 1799. Keep in mind Gauss was born in 1777. Moroever, the field $\mathbb{C}$ does not have any "holes" meaning it is complete, meaning any sequence in $\mathbb{C}$ which ought to converge does in fact do so. Thus, the set of complex numbers is some kind of heavenly place to work with numbers.

**Theorem 1.2** (Fundamental Theorem of Algebra)**.** *Suppose $f(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0 x$ where $a_k \in \mathbb{C}$ for $k = 1, \ldots, n$, and $a_n \neq 0$. Then there exist integers $r, s \geq 0$ with $r + 2s = n$, real numbers $\alpha_1, \ldots, \alpha_r$ and non-real complex numbers $\beta_1, \ldots, \beta_s$ such that*

$$f(x) = a_n(x - \alpha_1) \cdots (x - \alpha_r) \cdot (x - \beta_1) \cdots (x - \beta_s) \cdot (x - \overline{\beta_1}) \cdots (x - \overline{\beta_s}).$$

*The numbers $r, s$ as well as the unordered sequences $\alpha_1, \ldots, \alpha_r$ and $\beta_1, \cdots, \beta_s$ are uniquely determined by $f$.*

You will probably study a proof of this wonderful theorem in Math 421 using the concept of differentiablity for a function $f : \mathbb{C} \to \mathbb{C}$. Just to give a very simple example, suppose $w \in \mathbb{C}$ and $w \neq 0$. Then the equation $z^2 - w$ has two roots. Namely, if we write $w = re^{i\theta}$, with $r > 0$, then with $z_\pm = \pm\sqrt{r}e^{i\theta/2}$, we have $z^2 - w = (z - z_+)(z - z_-)$.

## 2. A constructive existence theorem

Some of you might feel lulled into a sense of complacency by the previous section, culminating in Theorem 1.1. Others might be feeling somewhat unsatisfied or unconvinced on the point of the "reality" of the solution there proposed. Oh Oh, here comes another one of those `$&%(*^&$%` plays!

==========

The reality of imaginary numbers

*Javier:* Hey, professor Dude. So you're saying you can solve an equation like $z^2 = -1$ by "**inventing**" a symbol $i$ and calling that the solution. What kind of a cop-out is that? I thought this course was supposed to be about proving/justifying everthing that we do logically. How can you just say "Oh happy me, let me just *invent* a solution, la di da!" and expect us to accept that?! What would you do if I just "invented" answers and gave them names to whatever equation you gave me on an exam?

*Kristen:* Way harsh.

*Farshid:* No, no, Javier has a valid point here. The point is: fine, maybe wonderful things will happen if I just invent a solution to some equation, but until I *prove* that a solution does exist, it is a valid question to ask how I know that the equation $z^2 = -1$ has a solution. This issue was actually the subject of bitter controversy in the 17th and 18th centuries. Gauss, for instance, was very critical of how his predecessors just posited (philosophers love that word) the existence of solutions and went merrily along playing with these posited (invented out of thin air) solutions. Gauss was the first person to put the complex numbers on firm ground and prove what we now call The Fundamental Theorem of Algebra (more on that later).

*Matt T.:* So you're saying that complex numbers don't really exist?

*Farshid:* Oh not at all, I'm just saying I haven't yet proved that they exist. I will do so now.

*Alby:* No way, dude, there is no way you can do it.

*Farshid:* Let me try and the class, including you, can be the judge of that. Before I start, let's agree on what it is I'm setting out to do. We all agree that the equation $z^2 = -1$ has no solution with $z \in \mathbb{R}$. What I want to construct is a set, let me call it $\mathbf{C}$ for now, which contains $\mathbb{R}$ but also contains a special element $i$ such that $i^2 = -1$. Actually what I will do is slightly different. I will construct a set $\mathbf{C}$ which does not contain $\mathbb{R}$ strictly speaking, but an exact copy $\mathbf{R}$ of the set of real numbers written in a slightly unusual way. To explain what I mean, recall when we constructed the rational numbers as $\mathbb{Z} \times \mathbb{Z}_0/\sim$ i.e. as equivalence classes $\widetilde{(a,b)}$ of pairs $(a, b)$ of integers with $b \neq 0$ with $(a, b) \sim (c, d)$ if and only if $ad = bc$?

Well the set $\mathbb{Z}$ is not strictly speaking a subset of $\mathbb{Q} = \mathbb{Z} \times \mathbb{Z}_0/\sim$ now, is it? But we have a natural injection $\mathbb{Z} \hookrightarrow \mathbb{Q}$ by sending $a \mapsto \widetilde{(a,1)}$. In the same way, we we will have a very natural injective map $\mathbb{R} \hookrightarrow \mathbf{C}$ whose image $\mathbf{R}$ we will think of as a copy of the real numbers.

*Brent:* I think, with that explanation, you've confused the heck out of half the class and put the other half to sleep.

*Farshid:* You're right, sorry, let me just spit it out. Alright, here we go. You remember matrices, right? Let's look at the set $\mathbf{M}$ of all $2 \times 2$ matrices with real entries:

$$\mathbf{M} = \{\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \mid a, b, c, d \in \mathbb{R}\}.$$

I want to look at those matrices that satisfy $a = d$ and $b = -c$.

*Matt L.:* Why? That seems like a whacko idea.

*Farshid:* You'll see. So let me define $\mathbf{C}$ to be the matrices that have repeating diagonal entries and whose "anti-diagonal" entries are negatives of each other, i.e.

$$\begin{aligned} \mathbf{C} &= \{\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathbf{M} \mid c = -b, d = a\} \\ &= \{\left(\begin{smallmatrix} a & b \\ -b & a \end{smallmatrix}\right) \mid a, b \in \mathbb{R}\}. \end{aligned}$$

If we define a map $m : \mathbb{R}^2 \to \mathbf{M}$ by $(a, b) \mapsto m_{a,b}$ where

$$m_{a,b} = \left(\begin{smallmatrix} a & b \\ -b & a \end{smallmatrix}\right),$$

then $\mathbf{C}$ is simply the image of the map $m$. You will admit, I hope, that the existence of the set $\mathbf{C}$ is not in question.

*Class:* (Grudingly) Granted. Get to the point, Dude.

*Farshid:* Thank you. Next, I want to show you that the real numbers live very happily inside $\mathbf{C}$, can you guess how, Kate?

*Kate:* I guess the real numbers correspond to matrices with zeros in the "$b$" and "$-b$" position.

*Farshid:* Brilliant, yes! How did you come up with that guess?

*Kate:* Well, I went with the coincidence-of-notation theory. In other words, I ...

*Farshid:* (interrupting) Wait, what was that? What's the "coincidence- of-notation theory?"

*Kate:* What I mean is that the set $\mathbf{C}$ you have defined is supposed to be like the complex numbers, right?

*Farshid:* Yeah, so?

*Kate:* And you said the elements of $\mathbb{C}$ (what you talked about before) are $a + bi$ with real numbers $a, b$. But now the elements of $\mathbf{C}$ are matrices $\left(\begin{smallmatrix} a & b \\ -b & a \end{smallmatrix}\right)$. Notice how the same letters are used? Coincidence? I personally don't think so. I think in some twisted way $a + bi$ the complex number is associated with $\left(\begin{smallmatrix} a & b \\ -b & a \end{smallmatrix}\right)$ the matrix. So, to answer your question about which of these matrices corresponds to a real number, I went back to ask the same about the $a + bi$ and it's pretty obvious which of those are real numbers, the ones with $b = 0$. So then going back the matrices, I guess the "real ones" are the ones with $b = 0$ i.e. with zeros on the anti-diagonals.

*Farshid:* Wow, that is great detective work and perfectly correct in its findings, although I suspect the "coincidence-of-notation" theory can lead you down the road of confusion and false turns too. But in this case, I gotta hand it to you, you nailed it. We do define $\mathbf{R}$ to be

the set of matrices

$$\begin{aligned}\mathbf{R} &= \{\left(\begin{smallmatrix} a & b \\ -b & a \end{smallmatrix}\right) \in \mathbf{C} \mid b = 0\} \\ &= \{\left(\begin{smallmatrix} a & 0 \\ 0 & a \end{smallmatrix}\right) \mid a \in \mathbb{R}\}.\end{aligned}$$

In other words, $\mathbf{R}$ is the image of the (clearly injective) map $\mathbb{R} \hookrightarrow \mathbf{C}$ defined by $a \mapsto \left(\begin{smallmatrix} a & 0 \\ 0 & a \end{smallmatrix}\right)$.

*Julie:* Will this discussion ever come back to the equation $z^2 = -1$?

*Farshid:* Yes, very soon. Good point. So how do we express the usual number $-1$ in this weird matrixy world, Julie?

*Julie:* I guess as $\left(\begin{smallmatrix} -1 & 0 \\ 0 & -1 \end{smallmatrix}\right)$?

*Farshid:* Yes, so we look for a matrix $Z$ (of the right type, i.e. with negative antidiagonals and equal entries on the diagonal) such that when you multiply it (matrix multiplication) by itself you get the matrix Julie just said. Can anyone see such a matrix?

*Class:* Silence.

*Farshid:* Okay, if I gave you ten minutes, even five, even three, to work together, and you really went at it, I'm sure you'd find it by experimentation, but that would be too obvious and boring a way to get the answer. Any other ideas for how to proceed?

*Jonah:* Oh come on man. For Pete's sake, you know the matrix, just tell us what it is!

*Farshid:* Okay, you're right, it's $\left(\begin{smallmatrix} -2342+\sqrt{5} & 19497+\sqrt{5} \\ -19497-\sqrt{5} & -2342+\sqrt{5} \end{smallmatrix}\right)$. Would you mind just squaring that matrix and verifying that it gives $\left(\begin{smallmatrix} -1 & 0 \\ 0 & -1 \end{smallmatrix}\right)$?

*Jonah:* (Annoyed, gets to work, nonetheless).

*Farshid:* While Jonah is working on that ...

*Shaohan:* I have an idea, which, in principle will work. To ease the notation, recall that the matrix $\left(\begin{smallmatrix} a & b \\ -b & a \end{smallmatrix}\right)$ has the name $m_{a,b}$. Then we want to find $a, b$ such that $m_{a,b}^2 = m_{-1,0}$.

*Farshid:* (Hiding his glee) Yes, go on.

*Shaohan:* The equation $m_{a,b}^2 = m_{-1,0}$ will give us 4 equations, one for each of the entries. We can try to solve those four equations to find the $a$ and the $b$.

*Farshid:* Excellent, you and Ashley and Jing work on that and see what you get. Jonah, how's it going?

*Jonah:* Still working on it, thanks to you jabbering away.

*Farshid:* (Not sorry at all) Sorry, keep working.

*Mike:* Jen and I think we've figured out the answer.

*Farshid:* Did you just experiment, trial-and-errorifically? That would be too dull.

*Mike:* Well, not exactly. I really liked Kate's theory, so we went with that to try to guess the anwer; it's more like "steal-and-check" as opposed to "trial-and-error." Anyway, remember how Kate said she thinks $m_{a,b}$ is secretly the number $a + bi$?

*Farshid:* I believe she used the phrase "in a twisted way."

*Mike:* Yes, well, that's not the point is it? Anyhoo, if $m_{a,b}$ is $a + bi$, and we're looking for what matrix represents $i$ right? Then we should

*Jen C.:* Look at the matrix $m_{0,1}$ because $i = a + bi$ with $a = 0$ and $b = 1$.

*Ashley:* Okay, we followed Shaohan's method and we got the matrix $\left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right)$, which is the matrix $m_{0,-1}$, as one solution.

*Farshid:* Jonah, any luck?

*Jonah:* Dude, I multiplied the whole darn thing out three times and I got the same answer three times and it doesn't give what you said!!

*Farshid:* (Feigning sorrow) Oh, sorry my man, my bad. But don't worry, we seem to be getting the right answer some other way. Let's note that $m_{0,1} = -m_{0,-1}$ so if one of them squared is $m_{-1,0}$ then so is the other one. Alright, let's check $m_{0,1}$ then:

$$m_{0,1}^2 = \left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right)\left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right) = \left(\begin{smallmatrix} -1 & 0 \\ 0 & -1 \end{smallmatrix}\right),$$

just as we wanted. Hurray. Notice that in the set $\mathbf{C}$, we have two solutions of $z^2 = -1$, namely $m_{0,1}$ and $m_{0,-1}$. So let us define $I = m_{0,1}$. This is very nice because then $m_{a,b} = a+bI$ with the usual operations of matrix addition and multiplying a matrix by a scalar. We have therefore settled the existence of a set $\mathbf{C}$ which contains a copy of the real numbers and in which the equation $z^2 = -1$ has a solution. Moreover, the usual algbraic operations on $\mathbb{R}$ extends to $\mathbf{C}$. In practice, we could continue to work with these matrices, but having now rigorously shown the existence of complex numbers (in a rather concrete way, in fact), we can now return with an easy conscience to the paradise of $\mathbb{C} = \{a + bi \mid a, b \in \mathbb{R}\}$; should we ever have any doubts, we can travel back to $\mathbf{C}$ to make sure our intuition is based on a firm foundation.

==========

## 3. The geometry of $\mathbb{C}$

We have seen that, geometrically speaking, the set $\mathbb{C}$ is just the plane $\mathbb{R}^2 = \{(a, b) \mid a, b \in \mathbb{R}\}$. Namely we have a very nice bijective map $\mathbb{R}^2 \to \mathbb{C}$ given by $(a, b) \mapsto a + bi$. Moreover on this set, there is a "vector space" rule for adding elements (geometrically it is the familiar "parallelogram rule") which is just gotten by adding coordinate-wise which is the usual way adding numbers in $\mathbb{R}^2$. On the set $\mathbb{C}$, however, we have defined a very cool *multiplication* rule which one probably would not have thought of if the point $(0, 1)$ had not been given the interpretation as the quantity $i = \sqrt{-1}$. This multiplication rule is commutative, associative and distributes over addition, i.e. $zw = wz$, $(zv)w = z(vw)$, and $z(v + w) = zv + zw$ for all $z, v, w \in \mathbb{C}$.

The map $\mathbb{R}^2 \to \mathbb{C}$ allows us to think of complex numbers as vectors in the plane. Continuing with the vector metaphor, If $z = a + bi$ with $a, b \in \mathbb{R}$, the real part of $z$ is defined to be $\Re(z) = a$ and the imaginary part of $z$ is defined to be $\Im(z) = b$. In other words, these are, respectively the $(1, 0)$ and $(0, 1)$ components of $(a, b)$. We always have $z = \Re(z) + \Im(z)i$. We define the *modulus* (or *length*) of $z = a + bi$ to be

$$|z| = \sqrt{\Re(z)^2 + \Im(z)^2} = \sqrt{a^2 + b^2}.$$

By definition, the modulus of $z$ is a non-negative real number, measuring its distance (in the usual sense) from the origin. The complex numbers with fixed real part lie on a vertical line, those with fixed imaginary part lie on a horizontal line, and those with fixed modulus lie on a circle centred at the origin. We have already encountered a symmetry of the complex numbers, the one where $i$ and $-i$ are interchanged. We define, for $z = a + bi$, its *complex conjugate* to be $\overline{z} = a - bi$. We have

$$\Re(z) = \frac{z + \overline{z}}{2}, \qquad \Im(z) = \frac{z - \overline{z}}{2}.$$

Note that $|z|^2 = z\overline{z}$.

You might recall that an alternate method for locating points in the plane is called *polar coordinates*. It specifies a point $z$ in the plane by giving a pair $(r, \theta)$ where $r, \theta$ are real numbers. Here, $|r|$ is the modulus of $z$ (we allow $r$ to be negative, in which case we move backwards along the ray indicated by $\theta$) and $\theta$ is an angle, measured in radians, with $\theta = 0$ being the angle subtended by the $x$-axis and positive angles indicating a counterclockwise motion. In some text, when we want to indicate a complex number $z$ whose polar coordinates are $r$ and $\theta$, one writes $z = r\text{cis}\theta$. A little recollection of trigonometry suffices for noting that $r\text{cis}\theta = r(\cos\theta + i\sin\theta)$. If $z = |z|(\cos\theta + i\sin\theta)$, then the angle $\theta$ is called the *argument* of $z$, sometimes denoted $\arg z$. Note that if $z \neq 0$, $\arg z$ is determined only up to an integer multiple of $2\pi$, and $\arg 0$ is not well-defined at all! For instance, Brian might say that "the" argument of $i$ is $\pi/2$, and Rick might say that $i$ has argument $-3\pi/2$. They would both be right.

Here is a wonderful and useful theorem.

**Theorem 3.1.** *For a complex number $z = r(\cos\theta + i\sin\theta)$ with $r, \theta \in \mathbb{R}$, we have $z = re^{i\theta}$.*

*Euler's Proof.* For this proof, I will assume that you are familiar with Taylor series from calculus. Recall the following lusciously and everywhere converging power series representations:

$$\begin{aligned} e^z &= \sum_{k=0}^{\infty} \frac{z^k}{k!} \\ \cos z &= \sum_{j=0}^{\infty} \frac{(-1)^j z^{2j}}{(2j)!} \\ \sin z &= \sum_{m=0}^{\infty} \frac{(-1)^k z^{2m+1}}{(2m+1)!}. \end{aligned}$$

It turns out that these power series converge and are valid even if the argument is a complex number. Before we plug in, let's take a peek ahead and notice that the expression $i^k$ will appear in the formulas; this is a periodic function of period four, giving $1, i, -1, -i, 1, i, -1, -i, 1, \ldots$, so we will split our sum into the even $k$'s giving alternating $1, -1$ and the odd $k$'s giving alternating $i, -i$.

So, we plug in $z = i\theta$ into the first one and compute:

$$\begin{aligned} e^{i\theta} &= \sum_{k=0}^{\infty} \frac{i^k\theta^k}{k!} \\ &= \sum_{k\text{even}} \frac{i^k\theta^k}{k!} + \sum_{k\text{odd}} \frac{i^k\theta^k}{k!} \\ &= \sum_{j=0}^{\infty} \frac{i^{2j}\theta^{2j}}{(2j)!} + \sum_{m=0}^{\infty} \frac{i^{2m+1}\theta^{2m+1}}{(2m+1)!} \\ &= \sum_{j=0}^{\infty} \frac{(-1)^j\theta^{2j}}{(2j)!} + \sum_{m=0}^{\infty} \frac{(-1)^m i\theta^{2m+1}}{(2m+1)!} \\ &= \cos(\theta) + i\sin(\theta). \end{aligned}$$

Presto. □

Now let's pull a few rabbits out of our hat with this fabulous result. First of all, let's say $z_1, z_2$ are points on the unit circle, say $z = e^{i\theta_1}$ and $w = e^{i\theta_2}$. Remembering rules for how to multiply exponentials, we have $zw = e^{i(\theta_1+\theta_2)}$. In other words, multiplying two points on the unit circle keeps them on the unit circle but adds their arguments.

If $n \in \mathbb{Z}$, clearly $z^n = (e^{i\theta})^n = e^{in\theta}$, or $z^n = \cos n\theta + i \sin n\theta$, which seems to indicate that if you multiply $z$ by itself $n$ times, then that just makes it go $n$ times further along the circle.

Note that if $t \in \mathbb{R}$, then $|e^{it}| = \sqrt{\cos^2 t + \sin^2 t} = 1$. Let us solve the equation $z^n = 1$ where $n \in \mathbb{N}$. Rewriting $z = re^{i\theta}$ with $r \geq 0$, we have $r^n e^{in\theta} = 1$ giving $r^n = 1$. Since $r \geq 0$, we conclude that $r = 1$. Moreover, $\arg(1) = \arg(e^{in\theta})$, so we have $n\theta = 0 + 2\pi k$ with $k \in \mathbb{Z}$. In other words, $\theta$ can take $n$ values $\{2\pi k/n \mid k = 0, 1, 2, \ldots, n-1\}$ which are distinct modulo $2\pi\mathbb{Z}$. These give $n$ equally-spaced points on the unit circle, 1 being one of them.

To see what multiplication of complex numbers means geometrically, for each $w \in \mathbb{C}$, let's define $\delta_w : \mathbb{C} \to \mathbb{C}$ by $z \mapsto wz$, i.e. $\delta_w(z) = wz$. This is just an extension of the maps $\delta_k$ we defined for integers $k$ to all of $\mathbb{C}$. What does the map $\delta_5$ do? Let's see, if $z = a + bi$, then $5z = 5a + 5bi$, so it multiplies the imaginary and real parts by 5. The point $z$ will move along the line from 0 to $z$ and travel to a point 5 times as far from 0 as it was. So $\delta_5$ is a radial "expansion-by-5" map. You can see that for all $s \in \mathbb{R}_{>0}$, that $\delta_s$ is an expansion by $s$ map. (If $0 < s < 1$, then it's more of a dilation, isn't it?) You should verify that $\delta_{-1}$ is rotation by 180 degrees around 0. More generally, if $s \in \mathbb{R}$ is negative, then the map $\delta_s$ is expansion by $|s|$ followed by a 180 degree rotation around 0. Of course, $\delta_0$ is the killer map that sends everybody and her cousin to 0.

Now, what about $\delta_i$? We calculate that $i(a + bi) = ai - b = -b + ai$ so multiplying by $i$ sends the Euclidean point $(a, b)$ to the point $(-b, a)$. This represents a vector which is perpendicular to the vector $(a, b)$ and has moved counterclockwise. In other words, $a + bi$ has moved 90 degrees counterclockwise, with unchanged modulus. Aha, so $\delta_i$ has a nice geometric meaning, it is rotation by $\pi/2$ radians counterclockwise. In retrospect, we could have seen that coming, because $i = e^{i\pi/2}$ so if $z = re^{i\theta}$, then $iz = re^{i(\theta+\pi/2)}$. In other words, $|\delta_i z| = |z|$ and $\arg(\delta_i z) = \arg(z) + \pi/2$. If that's not rotation counterclockwise by $\pi/2$ then I'm Mr. Noodle.

What $\delta_w$ is geometrically for an arbitrary $w$ should now be clear, for instance by thinking of as a composite map.

## 4. Problems

1. (There are no zero-divisors in $\mathbb{C}$). Show that if $z, w \in \mathbb{C}$, and $zw = 0$ then either $z = 0$ or $w = 0$. (you may use the fact that this is true for $z, w \in \mathbb{R}$).

2. (a) (Every non-zero complex number is invertible). Show that for each $z \in \mathbb{C}$ such that $z \neq 0$, there exists a unique $w \in \mathbb{C}$ such that $wz = 1$, so it's okay to write $w = z^{-1} = 1/z$.
(b) Use (a) to give another proof of the statement in Problem 1.
(c) For $z = 3 + 4i$, determine $1/z$ and write it in the form $a + bi$ with real numbers $a, b$.

3. (a) Show that for $z \in \mathbb{C}$, $z = 0$ if and only if $|z| = 0$.
(b) Prove that $|zw| = |z||w|$.
(c) Prove using induction that for all $n \in \mathbb{Z}$, $|z^n| = |z|^n$.

4. (a) Show that for $z, w \in \mathbb{C}$, $|z - w|$ is the usual distance from $z$ to $w$.

(b) (Triangle Inequality) Give an algebraic proof of the fact that for $z, w \in \mathbb{C}$, $|z - w| \leq |z|+|w|$ and interpret this fact geometrically. Hint: First prove that if $u \in \mathbb{C}$, then $\Re(u) \leq |u|$. Next, argue that it suffices to show that $|z - w|^2 \leq (|z| + |w|)^2$. Now justify each step in the following:

$$|z - w|^2 = (z - w)(\overline{z} - \overline{w}) = |z|^2 + |w|^2 + 2\Re(-z\overline{w}) \leq |z|^2 + |w|^2 + 2|z\overline{w}| = (|z| + |w|)^2.$$

(c) Shade in the region $\{z \in \mathbb{C} \mid 1 \leq |z - i| \leq 2\}$. It is called an "annulus." Hint: $|z - i|$ is the distance from $z$ to $i$.

5. (a) Find four solutions in $\mathbb{C}$ of the equation $z^4 = 1$.

(b) Using your vast knowledge of trigonometry, evaluate $\zeta = \cos(\theta) + i\sin(\theta)$ where $\theta = 2\pi/6$.

(c) Verify that $1, \zeta, \zeta^2, \zeta^3, \zeta^4, \zeta^5$ are six distinct solutions of $z^6 = 1$. They are called the sixth roots of unity in $\mathbb{C}$.

(d) Draw a fairly accurate picture of the unit circle showing that the roots of $z^4 = 1$ and $z^6 = 1$ all lie on it. (Label the solutions). Use red for the 4 solutions of one equation and Blue for the six solutions of the other.

6. (Autour le théorème de De Moivre) For $z = r(\cos(\theta) + i\sin(\theta)) \in \mathbb{C}$, prove using induction on $n$ that for all $n \in \mathbb{Z}$, $z^n = r^n(\cos(n\theta) + i\sin(n\theta))$.

7. (Cardano's[3] Bag of Tricks, first published in his *Ars Magna*: he learned it from Tartaglia, who learned it from Fior who learned it from del Ferro) Suppose we want to solve the cubic equation $x^3 + px - q = 0$ where $p, q$ are real (or even complex!) numbers.

(a) Make the substitution $x = w - p/(3w)$ and clear all $w$'s from denominators to get a sixth degree equation in $w$.

(b) Letting $v = w^3$, now magically convert the sextic equation of (a) to a quadratic one in $v$!

(c) Use the quadratic formula to solve for $v$, getting the two solutions $v_\pm = Q \pm \sqrt{Q^2 + P^3}$ where $P = p/3$ and $Q = q/2$.

(d) Now $v = w^3$, so taking cube roots of $v_\pm$, we get six values for $w$. How many different values of $x = w - p/(3w)$ do you think we'll have?

(e) Use this procedure to find three roots of the equation $x^3 - 508x - 4368 = 0$.

(f) Use this procedure to find three roots of $x^3 - 8x - 32 = 0$.

(g) Use this procedure to find three roots of $x^3 - 125 = 0$.

(h) Use this procedure to find three roots of $x^3 + 8i = 0$.

8. Using the interpretation of $\mathbb{C}$ as matrices (i.e. as elements of $\mathbf{C}$), show that for $z, w \in \mathbb{C}$, $z = w$ if and only if $\Re(z) = \Re(w)$ and $\Im(z) = \Im(w)$.

9. Let $\mathbb{G} = \{a + bi \mid a, b \in \mathbb{Z}\}$ be the set of *Gaussian Integers*.

(a) Show that if $z \in \mathbb{G}$ and $1/z \in \mathbb{G}$, then $z \in \{\pm 1, \pm i\}$.

(b) Draw a picture of $\mathbb{G}$ as points of a "lattice" in the complex plane. Include all the points $a + bi$ with $|a| \leq 6$ and $|b| \leq 6$.

[3] www.lib.virginia.edu/science/parshall/cardano.htm

(c) Let us call a Gaussian integer $a + bi$ a *Gaussian prime* if $|z|^2 = a^2 + b^2$ is a prime number. Thus, $1 + i$ and $2 + i$ are Gaussian primes, but $3 + i$ is not. Indicate by a special color or symbol all the Gaussian primes in your picture of part (b). Do you see a four-fold symmetry in your picture? Try to explain this symmetry; you might find part (a) relevant.

(d) [optional] Do you think there are infinitely many Gaussian primes? Record any patterns you observe about them. Can you prove any of these observations?

10. Show that for $w \in \mathbb{C} = se^{i\alpha}$, the map $\delta_w : \mathbb{C} \to \mathbb{C}$ given by $\delta_w(z) = wz$ represents an expansion by the factor $|w|$ followed by a rotation by the angle $\alpha$.

11. (a) On the set $X = \mathbb{R}$, let us define the following equivalence relation: given $x, y \in \mathbb{R}$, we write $x \sim y$ if and only if $x - y \in \mathbb{Z}$. Check that $\sim$ is an equivalence relation.

Let $\widetilde{X} = \mathbb{R}/\sim$ be the associated quotient set. In parts (b) and (c) we will give two different "models" for understanding the set $\widetilde{X}$. First, recall that if $x \in \mathbb{R}$, then $[x]$ is the greatest integer less than or equal to $x$ and $\langle x \rangle = x - [x]$ is the fractional part of $x$. We have $x = n + \alpha$ with $n \in \mathbb{Z}$ and $\alpha \in [0, 1)$ if and only if $n = [x]$ and $\alpha = \langle x \rangle$. In other words, $[x]$ is "the stuff to the left of the decimal point" and $\langle x \rangle$ is "the decimal point and the stuff to the right of it."

(b) Show that the map $f : \mathbb{R} \to [0, 1)$ given by $f(x) = \langle x \rangle$ is surjective and its fibers are exactly the $\sim$-equivalence classes of $\mathbb{R}$ where $\sim$ is the equivalence relation defined above, i.e. $x \sim y$ if and only if $x - y \in \mathbb{Z}$. Use this fact to construct a natural bijective map $[0, 1) \to \widetilde{X}$.

(c) Let $S^1 = \{z \in \mathbb{C} \mid |z| = 1\}$ be the unit circle in the complex plane. We have seen that $\theta \mapsto e^{i\theta}$ gives a nice map from $\mathbb{R}$ onto $S^1$. Let us give a slight renormalization of this map, namely, we define $g : \mathbb{R} \to S^1$ by $g(x) = e^{2\pi i x}$. Show that $f$ is surjective and that the fibers of this map are exactly the $\sim$-equivalence classes of $\mathbb{R}$. Use this fact to construct a natural bijective map $\widetilde{X} \to S^1$.

## 5. Extra Credit

1. Prove that the points $z_1, z_2, z_3$ in the complex plane are vertices of an equilateral triangle if and only if

$$z_1^2 + z_2^2 + z_3^2 = z_1 z_2 + z_1 z_3 + z_2 z_3.$$

2. Let $\zeta = e^{2\pi i/5}$ so that $1, \zeta, \zeta^2, \zeta^3, \zeta^4$ are the vertices of a regular pentagon. The diagonals of this pentagon meet at the vertices of a smaller regular pentagon. Determine them.

3. (a) Show that for $A \neq 0$, the set of all points $(x, y)$ in $\mathbb{R}^2$ satisfying $Ax^2 + Ay^2 + Bx + Cy + D = 0$ is either empty or a circle. Determine the center and the radius. What happens when $A = 0$?

(b) Suppose $z_1, z_2 \in \mathbb{C}$ are distinct fixed points in $\mathbb{C}$ and $K$ is a fixed positive real number, $K \neq 1$. Show that the set of all $z \in \mathbb{C}$ satisfying

$$\frac{|z - z_1|}{|z - z_2|} = K$$

is a circle. Where is its center? What is its radius? How are $z_1, z_2$ positioned vis à vis this circle? If we keep $K$ fixed and move $z_1$ along a straight line toward $z_2$, what happens to the

center and radius of the circle? What happens when we move $z_1$ along the same straight line away from $z_2$? If we keep $z_1, z2$ fixed and move $K$ toward 0 or toward $\infty$, what happens to the circle? What happens when $K = 1$?

4. (a) Let $S$ be a set of size $n \geq 1$ and suppose $r$ is an integer in the range $0 \leq r \leq n$. Let

$$\mathcal{P}_r(S) = \{T \subseteq S \mid |T| = r\}$$

be the set of all subsets of $S$ of cardinality $r$. Use the multiplication counting principle to deduce that

$$|\mathcal{P}_r(S)| = \frac{n!}{r!(n-r)!}.$$

This number is often denoted by $\binom{n}{r}$.

(b) With the above notations for $n$ and $r$ and for variables $x$ and $y$, derive the binomial formula

$$(x+y)^n = \sum_{r=0}^{n} \binom{n}{r} x^r y^{n-r}.$$

5. (a) Use the well-ordering principle to prove the Principle of Double Induction: Suppose for each pair $(a, b) \in \mathbb{N} \times \mathbb{N}$, we have a statement $P(a, b)$. Suppose i) $P(1, 1)$ is true, and ii) Whenever $P(k, l)$ true for some $(k, l) \in \mathbb{N} \times \mathbb{N}$, then $P(k+1, l)$ and $P(k, l+1)$ are also true. Then $P(a, b)$ is true for all $(a, b) \in \mathbb{N}$.

(b) Now prove a slight modification: Suppose for all integers $n, r \geq 1$ with $r \leq n$, we have a statement $P(n, r)$. Suppose i) $P(1, 1)$ is true and ii) Whenever $P(k, l)$ is true for some $(k, l) \in \mathbb{N} \times \mathbb{N}$ with $l \leq k$, then $P(k+1, l)$ and $P(k+1, l+1)$ are true. Then $P(a, b)$ is true for all $(n, r) \in \mathbb{N}$ with $r \leq n$.

6. For a positive integer $n$, we let $I_n = \{k \in \mathbb{Z} \mid 1 \leq k \leq n\}$ be the set of integers from 1 to $n$. If $T$ is a subset of $I_n$, let $m_T$ be the least element of $T$. For $1 \leq r \leq n$, let $f(n, r)$ be the average, over all subsets $T$ of $I_n$ of cardinality $r$, of $m_T$. Recalling from problem 4 above that there are $\binom{n}{r}$ subsets of cardinality $r$ in $I_n$, we have, therefore,

$$f(n, r) := \frac{1}{\binom{n}{r}} \sum_{T \subseteq I_n, |T| = r} m_T.$$

Prove that

$$f(n, r) = \frac{n+1}{r+1}.$$