Math 3110: Number Theory

Exploration 1: Integer linear equations

January 11, 2016

1 Linear equations with integer coefficients

We have several intertwined goals in this exploration. First of all, we want to find solutions to equations of the form:

$$ax + by = n \tag{E}$$

This is the sort of thing you have done in high school algebra, but there is a twist: we are going to presume that a, b, and n are *integers* and we are only going to look for integer solutions x and y.

We will discover that this problem has a lot to do with prime numbers and prime factorization, which makes it the starting point for the whole subject of number theory.

Question 1. Which problem do you expect to be harder: finding integer solutions to (E) or finding real solutions? What about finding rational solutions?

The integers, rationals, and reals are all examples of *number systems*, what mathematicians call *commutative rings*. In the next few weeks we will meet many more number systems, in fact infinitely many more!

One of the best ways to get a handle on a new problem is to start thinking about examples. Here are a few suggestions:

Question 2. Try to find all integer solutions x and y to the following equations. (Suggestion: Find the easiest ones and solve those first. You don't have to solve all of these right now.)

(E1) $x + y = -7$

(E2) $4x + 6y = 8$

(E3) $4x + 6y = 2$

(E4) $6x + 10y = 8$

(E5) $6x + 9y = 11$

(E6) $8x + 16y = 24$

Can you find one solution? More than one? Can you find all the solutions? Can you prove that your list of solutions is complete?

Make up more examples of your own. Try to look for patterns, or maybe a procedure you can always use to find solutions.

Question 3. One very powerful method for solving linear equations over the real numbers is called the *substitution* method. We solve for one variable in terms of the others and then make subsitutions to reduce a system of many equations to just one equation. In this case we just have one equation, but we could still try solving for y in terms of x:

$$y = \frac{n - ax}{b}$$

Does this help us find integer solutions to Equation (E)?

Question 4. Do equations of the form (E) always have integer solutions? If so, can you prove that solutions always exist? If not, can you find a way to tell, based on the numbers a, b, and n, whether (E) has a solution? Suggestion: think about Equation (E5).

Question 5. When a, b, and n are selected so that (E) has at least one solution, how many solutions do you find? Is there any discernable pattern in the solutions?

Here we are starting to see that we can divide our study of equations (E) into two parts. Does a solution to (E) exist? And if a solution does exist, how many solutions are there?

We are going to analyze the existence question first.

2 An ideal solution

We are going to turn the question around a little. Instead of fixing a, b, and n and asking whether (E) can be solved, we are going to hold just a and b fixed and ask which integers n can be *represented* in the form (E).

Let's introduce some notation:

$$L(x, y) = ax + by$$

We say that an integer n is *representable* by L if there are some integers x and y such that $L(x, y) = n$. Our original question was about solving $L(x, y) = n$. In our new terminology, that would be asking if some particular n is representable by L. Of course, we also need to find *which* x and y represent n, but we'll save that problem for later.

> Given two integers a and b, let I be the set of all integers that are representable by L. Here it is in symbols:
>
> $$I = \{ax + by \ : \ x, y \in \mathbb{Z}\}$$

How can we tell if an integer n is in I?

Question 6. For each of the equations in Question 2, try to find a simple description of I. You don't have to do all of the equations, but try to find a complete description of I for at least one of them. Can you conjecture a general pattern about the relationship between the set I and the numbers a and b? (Hint: Try to find the smallest positive integer in I.)

Theorem 7. *Show that I has the following properties:*

(i) If n and m are in I then $n+m$ and $n-m$ are in I.

(ii) If z is an integer and n is in I then zn is in I.

Sets with these properties are known as ideals *and they are very important in number theory, algebraic geometry, commutative algebra, and many other fields of mathematics. We may see more of them later in the semester, and you can learn a lot more about them in an abstract algebra class.*

The proof of the next theorem uses an important tool called the division algorithm, which is a careful formulation of the result of long division from elementary school:

Theorem 8. *For any integers n and $d \neq 0$ there is a unique pair of integers q and r such that*

$$n = qd + r$$

and $0 \leq r < d$. We call q the quotient *of n by d and we call r the* remainder *of n when divided by d.*

Theorem 9. *Let d be the smallest positive integer in I.*

(i) Prove that d divides every element of I.

(ii) Conclude that

$$I = d\mathbb{Z} = \{kd \,:\, k \in \mathbb{Z}\}$$

and that d is a divisor of both a and b.

(iii) Prove that if e divides both a and b then e divides d.

(iv) Deduce that d is the greatest common divisor *of a and b.*

(v) Prove that, for any integers a, b, and n, the equation (E) can be solved by integers x and y if and only if n is divisible by the greatest common divisor of a and b.

Question 10. Describe a process by which you can determine if (E) has an integer solution.

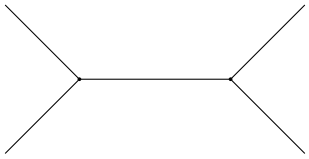
3 Visualizing ideals

This is great, but it's less than half way to a solution of our original problem. How are we supposed to find the solutions to the equation, once we know it has one? For that matter, how are we supposed to find the greatest common divisor of a and b so that we can test whether (E) has a solution in the first place?

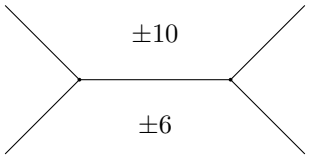
To help us understand the answers to these questions, we will take a visual approach to what we did in the last section. We will study the ideal:

$$I = \{10x + 6y \mid x, y \in \mathbb{Z}\}$$

Begin with a picture like this:



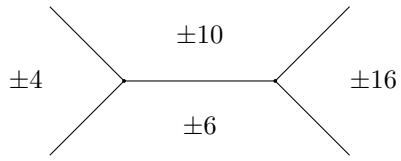
We start labelling the regions with numbers we know are in I. If $x = \pm 1$ and $y = 0$, we can get ± 10, so we put that in the upper region; if $x = 0$ and $y = \pm 1$ we can get ± 6, so we put that in the lower region:



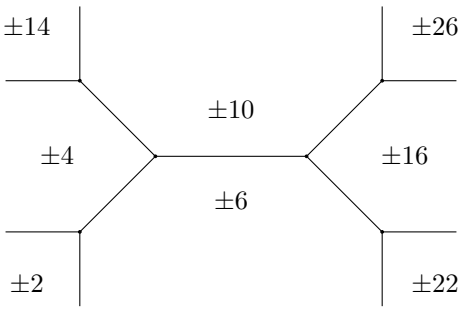
We know that if we can make find numbers n and m in the ideal then we can also find $n + m$ and $n - m$. Since we already have ± 10 and ± 6 there are 4 ways to combine these and get new elements of the ideal:

$$16 = 10 + 6 \qquad 4 = 10 - 6$$
$$-16 = -10 - 6 \qquad -4 = -10 + 6$$

These fit into two groups: ± 16 and ± 4. We add these to the picture:



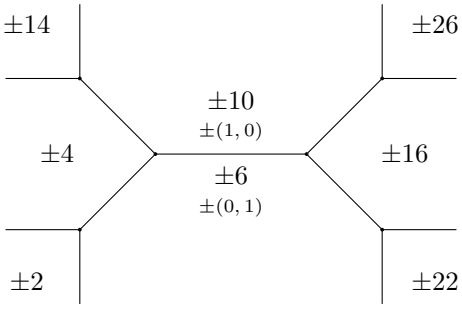
But now that we have ± 4 and ± 16, we can find some more elements of I. Combining ± 4 and ± 10 gives ± 6, which we already had, and ± 14. Likewise combining ± 4 and ± 6, we can get ± 2. We add these (and a few others) to the picture:



Now we've found ± 2, and we know (by inspection) that 2 is the greatest common divisor of 10 and 6, so $I = 2\mathbb{Z}$.

Question 11. Formulate a strategy based on the above picture to find the greatest common divisor of any two integers.

We still want to know how to find x and y such that $10x + 6y = 2$. We can figure this out by keeping track of the values of x and y that yield each entry.



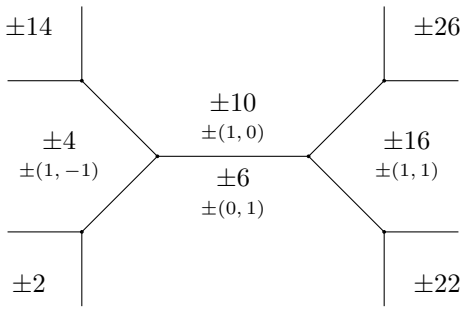
Remember that if

$$ax_1 + by_1 = n_1$$
$$ax_2 + by_2 = n_2$$

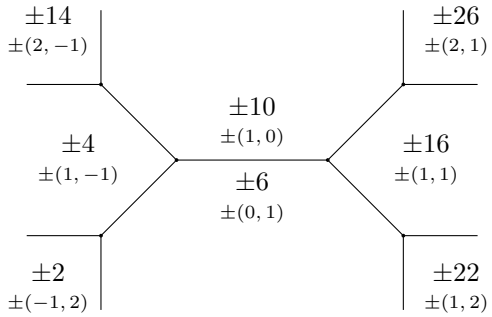
then

$$a(x_1 + x_2) + b(y_1 + y_2) = n_1 + n_2$$
$$a(x_1 - x_2) + b(y_1 - y_2) = n_1 - n_2.$$

Thus we can figure out how to label ± 16 and ± 4 from the labels on ± 10 and ± 6: Since $\pm 16 = \pm(10+6)$ and we get 10 from $(x, y) = (1, 0)$ and we get 6 from the label $(x, y) = (0, 1)$, we get 16 from the label $(x, y) = (1, 1)$. Similarly we have $4 = 10 - 6$, so we should subtract the labels, $(1, 0)$ and $(0, 1)$ and label 4 by $\pm(1, -1)$.



Now we can use the same process to figure out the label on $pm2$ and the other remaining entries. We got ± 2 as $6-4$ and we got 6 from $(0,1)$ and we got 4 from $(1,-1)$, so we get 2 from $(0,1)-(1,-1)=(-1,2)$. So we should label ± 2 with $\pm(-1,2)$.



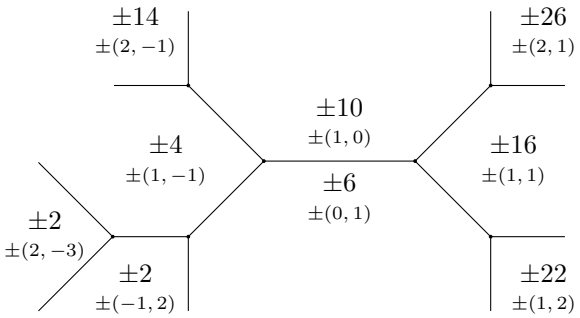
There are a few things to be careful about when you are executing this algorithm. We have always been careful to keep the signs in the right order. When we label an entry by $\pm c$ we make sure that we put the coefficients $\pm(x,y)$ below it in such a way that $c=ax+by$ and not $-c=ax+by$. You may even want to omit the $\pm$ symbols when you do this by hand to help keep straight what is going on.

Question 12. Can you find solutions to these equations?

(E7) $21x+27y=3$

(E8) $481x+128y=1$

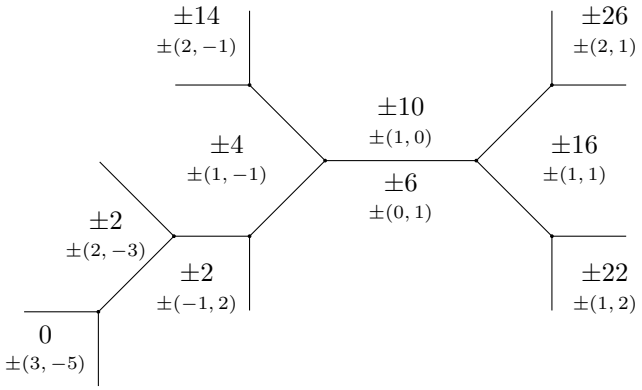
Let's see what happens if we keep going a few more steps in the topograph.



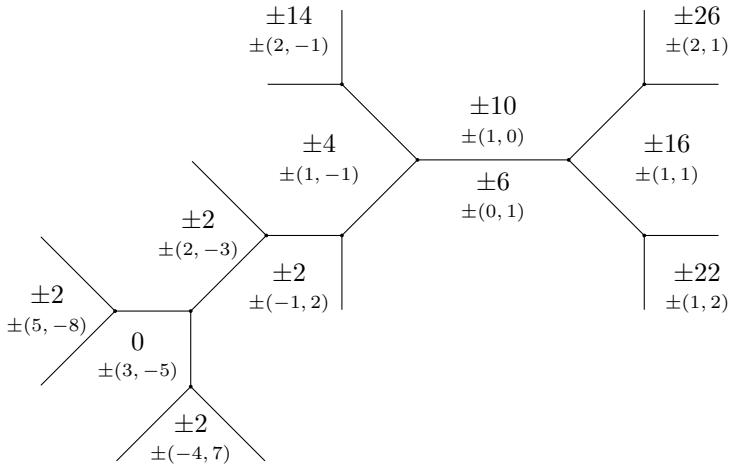
This gives us another way to build the greatest common divisor!

$$2 = 10(-1) + 6(2) = 10(2) + 6(-3)$$

Let's go even further:



Now we have a 0 in the diagram. Adding or subtracting zero from something doesn't change anything so it's pretty easy to predict what the next entries in the topograph will be:



This gives us even more ways of producing ± 2. In fact, it gives us an infinite collection of solutions to the equation $10x + 6y = 2$: whenever

$$(x, y) = (-1, 2) + k(3, -5)$$

for some $k \in \mathbb{Z}$, we will have

$$10x + 6y = 2.$$

Question 13. Are these all of the solutions to $10x + 6y = 2$? Or are there more?

Question 14. Can you find a way to make this process more efficient? Try this process out on the following ideal:

$$I = \{61x + 19y \mid x, y \in \mathbb{Z}\}$$

Do you find yourself doing several steps that could be done at once? Can you find a way to perform this algorithm without using the topograph?

4 The topography of the rational numbers (part 1)

5 Prime factorization

Definition 15. An integer u is called a *unit* if there is some other integer v such that $uv = 1$. An integer p is said to be

(i) *irreducible* if p is not a unit and in every factorization $p = ab$, either a or b is a unit;

(ii) *prime* if p is not a unit and whenever a and b are any two integers such that p divides ab, either p divides a or p divides b.

Theorem 16. *An integer p is prime if and only if it is zero or it is irreducible.*

Theorem 17. *Any nonzero integer can be written in as a product of irreducible integers, and the irreducible integers involved are unique up to reordering and change of sign.*

6 Changing variables

One of the most powerful techniques we have to solve equations is to *change variables*. You already know of a lot of examples that fit in this mold: u-substitution and trigonometric substitution for computing antiderivatives; completing the square for solving quadratic equations; Gaussian elimination for solving linear equations. Our technique from the last section can also be understood as a sequence of changes of variables.

Question 18. Let's introduce two new variables:

$$x = x_1 - y_1$$
$$y = y_1$$

Make this subsitution into Equation (E4) of Question 2. Is this equation any easier to solve than the original?

Now introduce some more variables:

$$x_1 = x_2$$
$$y_1 = y_2 - x_2$$

Substitute these into the result of your previous subsitution. Is the equation looking simpler? What are the solutions?

Let's make another subsitution:

$$x_2 = x_3 - y_3$$
$$y_2 = y_3$$

And one last one:

$$x_3 = x_4 - y_4$$
$$y_3 = y_4$$

By now the equation should be simple enough to find all of the integer solutions, and there shouldn't be any question about whether you have a complete list of solutions or not. But now you have solved Equation (E) in terms of x_4 and y_4. Figure out what the answers are in terms of the original variables x and y.

Question 19. What kinds of changes of variables can you make without affecting whether or not the equation has a solution? How do the solutions change when you make a change of variables?

Question 20. Do you think you could solve any equation of the form (E)? Describe a process—an *algorithm*—by which you could solve any such equation.

Question 21. Here are a few more equations. Try out your techniques on them:

(E9) $1028x + 48y = 8$

(E10) $1028x + 48y = 16$

7 Change of basis

Now we are going to reinterpret the changes of variables from the last section in linear algebraic terms. We are going to think about $L(x, y) = ax + by$ as a function from the set $V = \mathbb{Z}^2$ to $\mathbb{Z}$:

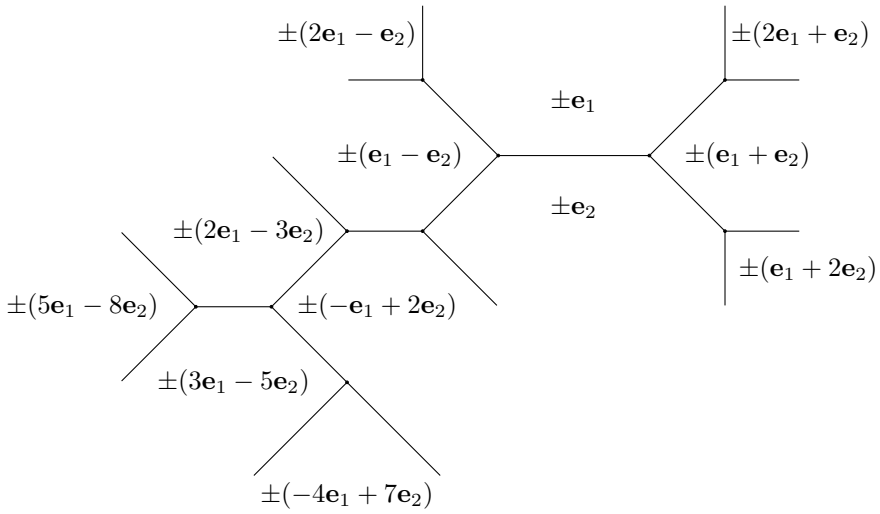
$$L : V \to \mathbb{Z}$$

There are two elements, $\mathbf{e}_1 = (1, 0)$ and $\mathbf{e}_2 = (0, 1)$ in V such that any element of $\mathbb{Z}^2$ can be written as a combination of $\mathbf{e}_1$ and $\mathbf{e}_2$ in a unique way:

$$(x, y) = x\mathbf{e}_1 + y\mathbf{e}_2$$

But $\mathbf{e}_1$ and $\mathbf{e}_2$ aren't the only vectors in V with this property. For example, the pairs $\mathbf{e}_1$ and $\mathbf{e}_1 + \mathbf{e}_2$ also have it. Any pair of vectors $\mathbf{v}$ and $\mathbf{w}$ such that every element of V has a unique expression as a linear combination of $\mathbf{v}$ and $\mathbf{w}$ is called a *basis* of V.

Here is a picture of the topograph, labelled using vector notation:



Theorem 22. *The two vectors on opposite sides of any edge in the topograph form a basis for V.*

What this means is that if $\mathbf{u}$ and $\mathbf{v}$ are on opposite sides of an edge, trying to solve the equation

$$L(x\mathbf{e}_1 + y\mathbf{e}_2) = n$$

for x and y is equivalent to trying to solve the equation

$$L(x'\mathbf{u} + y'\mathbf{v}) = n$$

for x' and y'. Of course, it may not be entirely clear *how* the problems are equivalent, but certainly if one problem has a solution then so does the other, and there is a one-to-one correspondence between the solutions to the two equations. Indeed, the reason for this is that both equations are ways of trying to solve the *vector* equation $L(\mathbf{w}) = n$, *in different bases.*

To get a feel for how this actually works, let's study the function below:

$$L(x\mathbf{e}_1 + y\mathbf{e}_2) = 10x + 6y$$

Question 23. Find formulas for:

$$\begin{aligned} L(x\mathbf{e}_1 + y\mathbf{e}_2) &= \\ L(x(\mathbf{e}_1 - \mathbf{e}_2) + y\mathbf{e}_2) &= \\ L(x(\mathbf{e}_1 - \mathbf{e}_2) + y(-\mathbf{e}_1 + 2\mathbf{e}_2)) &= \\ L(x(2\mathbf{e}_1 - 3\mathbf{e}_2) + y(-\mathbf{e}_1 + 2\mathbf{e}_2)) &= \\ L(x(3\mathbf{e}_1 - 5\mathbf{e}_2) + y(-\mathbf{e}_1 + 2\mathbf{e}_2)) &= \end{aligned}$$

Now find all solutions to the equation:

$$10x + 6y = 2$$

Now let's think about our process for solving equations like (E). We are really making a sequence of changes of basis to make our equation simpler and simpler. All we really need to keep track of are the coefficients of the equation in our new basis, and how to express the new basis in terms of the original basis. Here is the solution to $10x + 6y = 2$ in these terms:

Step 1.	$\mathbf{u} = \mathbf{e}_1$ $\mathbf{v} = \mathbf{e}_2$	$L(x\mathbf{v} + y\mathbf{w}) = 10x + 6y$
Step 2.	$\mathbf{u} = \mathbf{e}_1 - \mathbf{e}_2$ $\mathbf{v} = \mathbf{e}_2$	$L(x\mathbf{v} + y\mathbf{w}) = 4x + 6y$
Step 3.	$\mathbf{u} = \mathbf{e}_1 - \mathbf{e}_2$ $\mathbf{v} = -\mathbf{e}_1 + 2\mathbf{e}_2$	$L(x\mathbf{v} + y\mathbf{w}) = 4x + 2y$
Step 4.	$\mathbf{u} = 2\mathbf{e}_1 - 3\mathbf{e}_2$ $\mathbf{v} = -\mathbf{e}_1 + 2\mathbf{e}_2$	$L(x\mathbf{v} + y\mathbf{w}) = 2x + 2y$
Step 5.	$\mathbf{u} = 3\mathbf{e}_1 - 5\mathbf{e}_2$ $\mathbf{v} = -\mathbf{e}_1 + 2\mathbf{e}_2$	$L(x\mathbf{v} + y\mathbf{w}) = 2y$

This last equation is easy to solve: $y = 1$ and x can be any integer. This means that the solutions to the equation $L(\mathbf{w}) = 2$ are

$$\mathbf{w} = \mathbf{v} + k\mathbf{u} = (-\mathbf{e}_1 + 2\mathbf{e}_2) + k(3\mathbf{e}_1 - 5\mathbf{e}_2) = (-1, 2) + k(3, -5)$$

where $k \in \mathbb{Z}$. That's the same solution we found before. In fact, now we know these are all solutions because $\mathbf{u}$ and $\mathbf{v}$ form a basis for V.

We can make this process more amenable to calculation by hiding the topograph and just keeping track of the table. In fact, we only need to keep track of the coefficients:

Step 1. $\left(\begin{array}{cc} 10 & 6 \\ \hline 1 & 0 \\ 0 & 1 \end{array}\right)$

Step 2. $\left(\begin{array}{cc} 4 & 6 \\ \hline 1 & 0 \\ -1 & 1 \end{array}\right)$

Step 3. $\left(\begin{array}{cc} 4 & 2 \\ \hline 1 & -1 \\ -1 & 2 \end{array}\right)$

Step 4. $\left(\begin{array}{cc} 2 & 2 \\ \hline 2 & -1 \\ -3 & 2 \end{array}\right)$

Step 5. $\left(\begin{array}{cc} 0 & 2 \\ \hline 3 & -1 \\ -5 & 2 \end{array}\right)$

8 The topography of the rational numbers (part 2)

Now we can convince ourselves that the same entry does not appear more than once in the topograph. Indeed, let's suppose that we have any two adjacent entries $\pm\mathbf{u}$ and $\pm\mathbf{v}$ of the topograph. Since $\mathbf{u}$ and $\mathbf{v}$ form a basis for V, there is a function

$$L(x, y) = ax + by$$

such that $L(\mathbf{u}) = 0$ and $L(\mathbf{v}) = 1$.

Label each edge of the topograph so that it points in the direction of increase of the absolute value of L.

Theorem 24. *Every edge of the topograph that is not adjacent to* $\mathbf{u}$ *points away from* $\mathbf{u}$.

9 More variables

Question 25 (Silverman, Exercise 6.4 (a))**.** Find all integers x, y, and z that satisfy the following equation:

$$6x + 15y + 20z = 1$$

Homework 26. Describe a procedure to find all integer solutions x, y, and z to equations of the form

$$ax + by + cz = n,$$

where all of a, b, c, and n are integers. Can you generalize your procedure to find integer solutions $x_1, \ldots, x_n$ to equations like (F), where all a_i and n are integers?

$$a_1x_1 + a_2x_2 + \cdots + a_\ell x_\ell = n \tag{F}$$

10 Simultaneous equations

Question 27. Find all solutions to the system of simultaneous equations:

$$\begin{aligned} 6x + 15y + 20z &= 1 \\ 11x + 35y - 30z &= 2 \end{aligned}$$

Hint: try making changes of variables and combining your equations in ways to make the problem easier.

Question 28. Write the equations from the last question as a matrix:

$$\left(\begin{array}{ccc|c} 6 & 15 & 20 & 1 \\ 11 & 35 & -30 & 2 \end{array}\right) \tag{M}$$

What is the effect of making the subsitution $x = x_1 - y_1$, $y = y_1$, and $z = z_1$? What is the effect of replacing the second equation with the sum of the first and second equations?

Question 29. Suppose that we have represented a system of linear equations with integer coefficients as a matrix, as in Equation (M). What operations can we perform on M that

(i) don't change the solutions of the equations?

(ii) don't change whether or not the equations have solutions?

(iii) don't change how many solutions there are?

Question 30. Consider a system of simultaneous equations:

$$\begin{aligned} a_{1,1}x_1 + a_{1,2}x_2 + \cdots + a_{1,\ell}x_\ell &= n_1 \\ a_{2,1}x_1 + a_{2,2}x_2 + \cdots + a_{2,\ell}x_\ell &= n_2 \\ &\vdots \\ a_{k,1}x_1 + a_{k,2}x_2 + \cdots + a_{k,\ell}x_\ell &= n_k \end{aligned} \tag{G}$$

for integers $x_1, \ldots, x_\ell$ whenever $a_{i,j}$ and n_i are all specified integers. Can you describe a strategy for solving equations of this form?

Definition 31. We will say that a $m \times n$ matrix $M = (a_{i,j})$ with integer entries is in *Smith normal form* if all of the following properties hold:

(i) $a_{i,j} = 0$ for whenever $i \neq j$;

(ii) $a_{i,i} \geq 0$ for all i; and

(iii) $a_{i,i} | a_{i+1,i+1}$ for all i.

The entries $a_{1,1}, a_{2,2}, \ldots$ are called the *elementary divisors* of M.

Definition 32. A *row operation* on a matrix is one of the following operations:

(i) adding an integer multiple of one row to another row;

(ii) multiplying a row by ± 1;

(iii) exchanging two rows.

For example, the following 3×5 matrix is in Smith normal form if $a_{1,1} | a_{2,2}$:

$$\begin{pmatrix} a_{1,1} & 0 & 0 & 0 & 0 \\ 0 & a_{2,2} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Theorem 33 (Elementary divisors theorem)**.** *Let M be a matrix with integer entries. Then one can perform row and column operations to bring M into Smith normal form.*

In fact, the Smith normal form of a matrix is *unique*. This is because the entries of the Smith normal form of a matrix have a useful characterization. For any integer matrix A, let $I_k(A)$ be the greatest common divisor of the $k \times k$ submatrices of A (obtained by crossing out an appropriate number of rows and columns).

Theorem 34. *(i) Show that if P is an invertible matrix then $I_k(PA) = I_k(A)$ and $I_k(AP) = I_k(A)$.*

(ii) Show that if A is in Smith normal form, with diagonal entries $a_{1,1}, a_{2,2}, \ldots$ then $I_k(A) = a_{1,1} \cdots a_{k,k}$.

(iii) Conclude that every integer matrix has a unique Smith normal form.

If you don't already know about groups, feel free to skip the next theorem. If you do already know about groups then you might enjoy trying to use the elementary divisors theorem to prove the next one.

Theorem 35 (structure of finitely generated abelian groups)**.** *Every finitely generated abelian group is a product of cyclic groups.*