PUMaC 2023 Power Round Solutions

Frank Lu

April 2023

1 Rings and Fields

1.1 Rings and Ideals

Problem 1.1.1. Here are some more examples, and a non-example, of rings:

1. Show that $2\mathbb{Z}$, the set of even integers, is not a ring. (Hint: which property does it fail? In general, for questions of this nature, it is helpful to go through the properties and figure out which ones are or are not satisfied).

2. Show that $\mathbb{C}[x]$, the set of polynomials in one variable x with complex coefficients, is a ring (under the standard addition and multiplication operations of polynomials)

3. Show that the subset of polynomials in $\mathbb{C}[x]$ whose x coefficient is 0 forms a ring (with the same addition and multiplication as for $\mathbb{C}[x]$).

Solution. **1.** We claim that the set of even integers does not have a multiplicative identity. Suppose for the sake of contradiction such an identity e existed. Then, we would require that $2e = 2$. But then this requires that $e = 1$, which is not an even integer. This yields the desired contradiction, and so $2\mathbb{Z}$ is not a ring under the typical addition and multiplication operations.

2. We show that the set of polynomials with complex coefficients form a ring, by checking each of the conditions. For the first condition, for any polynomials $f, g \in \mathbb{C}[x]$, we can write $f(x) = \sum\limits_{i=0}^{n} f_i x^i$ and $g(x) = \sum\limits_{i=0}^{m} g_i x^i$, where $f_i, g_i \in \mathbb{C}$. We may furthermore suppose that $n = m$, by adding more terms of the form $0x^i$ to these polynomials. Then, we observe that $f(x) + g(x) = \sum\limits_{i=0}^{n} (f_i + g_i)x^i$ and

$$f(x)\cdot g(x) = \sum_{i=0}^{n}\sum_{j=0}^{n} f_i g_j x^{i+j} = \left(\sum_{i=0}^{n}\sum_{j=0}^{i}(f_j g_{i-j})x^i\right) + \left(\sum_{i=n+1}^{2n}\sum_{j=i-n}^{n}(f_j g_{i-j})x^i\right) = \sum_{i=0}^{2n}\left(\sum_{\substack{0\le j\le n\\ 0\le i-j\le n}} f_j g_{i-j}\right) x^i$$

which are both polynomials with complex coefficients. We could have stopped at the expression $\sum\limits_{i=0}^{n}\sum\limits_{j=0}^{n} f_i g_j x^{i+j}$, but it is useful to have the explicit formula for the coefficients written out.

For associativity, suppose we are given three polynomials $f(x) = \sum\limits_{i=0}^{n} f_i x^i, g(x) = \sum\limits_{i=0}^{n} g_i x^i, h(x) = \sum\limits_{i=0}^{n} h_i x^i$ (again, by the above argument, we can suppose that the sums are over the same range by adding additional $0x^i$ terms). Then,

$$\begin{aligned}(f(x)+g(x))+h(x) &= \sum_{i=0}^{n}(f_i+g_i)x^i + \sum_{i=0}^{n} h_i x^i = \sum_{i=0}^{n}((f_i+g_i)+h_i)x^i \\ &= \sum_{i=0}^{n}(f_i+(g_i+h_i))x^i = \sum_{i=0}^{n} f_i x^i + \sum_{i=0}^{n}(g_i+h_i)x^i \\ &= f(x)+(g(x)+h(x)).\end{aligned}$$

Similarly, we can check that

$$\begin{aligned}(f(x)\cdot g(x))\cdot h(x) &= \left(\sum_{i=0}^{2n}\left(\sum_{\substack{0\le j\le n\\ 0\le i-j\le n}} f_j g_{i-j}\right)x^i\right)\cdot\sum_{i=0}^{n} h_i x^i \\ &= \sum_{i=0}^{3n}\left(\sum_{\substack{0\le j\le 2n\\ 0\le i-j\le n}} (\sum_{\substack{0\le k\le n\\ 0\le j-k\le n}} f_k g_{j-k})h_{i-j}\right)x^i \\ &= \sum_{i=0}^{3n}\left(\sum_{\substack{0\le j\le 2n\\ 0\le i-j\le n}} \sum_{\substack{0\le k\le n\\ 0\le j-k\le n}} f_k g_{j-k} h_{i-j}\right)x^i \\ &= \sum_{i=0}^{3n}\left(\sum_{\substack{0\le k\le n\\ 0\le i-j\le n\\ 0\le j-k\le n}} f_k g_{j-k} h_{i-j}\right)x^i,\end{aligned}$$

the last equality coming from the fact that $0 \le k, j-k \le n$ implies that $0 \le j \le 2n$. But then we can rewrite this as

$$\sum_{i=0}^{3n}\left(\sum_{\substack{0\le k\le n\\ 0\le i-k\le 2n}} \sum_{\substack{0\le i-j\le n\\ 0\le j-k\le n}} f_k g_{j-k} h_{i-j}\right)x^i = \sum_{i=0}^{3n}\left(\sum_{\substack{0\le k\le n\\ 0\le i-k\le 2n}} f_k(\sum_{\substack{0\le i-j\le n\\ 0\le j-k\le n}} g_{j-k} h_{i-j})\right)x^i$$

$$= \sum_{i=0}^{n} f_i x^i \cdot \left(\sum_{i=0}^{2n}(\sum_{\substack{0\le j\le n\\ 0\le i-j\le n}} g_{i-j} h_j)x^i\right) = f(x)\cdot(g(x)\cdot h(x)).$$

This gives us associativity.

Next, for commutativity, we can quickly check that, given any polynomials f, g, which we can write in the form above, that

$$f(x) + g(x) = \sum_{i=0}^{n}(f_i + g_i)x^i = \sum_{i=0}^{n}(g_i + f_i)x^i = g(x) + f(x)$$

and

$$f(x) \cdot g(x) = \left(\sum_{i=0}^{2n} \left(\sum_{\substack{0 \le j \le n \\ 0 \le i-j \le n}} f_j g_{i-j} \right) x^i \right) = \left(\sum_{i=0}^{2n} \left(\sum_{\substack{0 \le j \le n \\ 0 \le i-j \le n}} g_j f_{i-j} \right) x^i \right) = g(x) \cdot f(x),$$

where the second-to-last equality arises from replacing the index j with $i - j$.

Using these formulas, we can check that the polynomial 0 is the additive identity, as

$$0 + \sum_{i=0}^{n} f_i x^i = \sum_{i=0}^{n}(0 + f_i)x^i = \sum_{i=0}^{n} f_i x^i,$$

and 1 is the multiplicative identity, as $1 \cdot f(x) = \left(\sum_{i=0}^{2n} \left(\sum_{\substack{0 \le j \le n \\ 0 \le i-j \le n}} a_j f_{i-j} \right) x^i \right)$, where a_j is 1 for $j = 0$ and 0 otherwise. But then this is just $\sum_{i=0}^{n} f_i x^i$.

For additive inverse, we note that given $f(x) = \sum_{i=0}^{n} f_i x^i$, the polynomial $\sum_{i=0}^{n} (-f_i)x^i$ is the additive inverse, as adding $f(x)$ to this yields $\sum_{i=0}^{n}(f_i - f_i)x^i = 0$. Finally, for the distributive law, we note that

$$\begin{aligned} f(x) \cdot (g(x) + h(x)) = f(x) \cdot \sum_{i=0}^{n}(g_i + h_i)x^i &= \left(\sum_{i=0}^{2n} \left(\sum_{\substack{0 \le j \le n \\ 0 \le i-j \le n}} f_j(g_{i-j} + h_{i-j}) \right) x^i \right) \\ &= \left(\sum_{i=0}^{2n} \left(\sum_{\substack{0 \le j \le n \\ 0 \le i-j \le n}} f_j g_{i-j}) \right) x^i \right) + \left(\sum_{i=0}^{2n} \left(\sum_{\substack{0 \le j \le n \\ 0 \le i-j \le n}} f_j h_{i-j}) \right) x^i \right) \\ &= f(x) \cdot g(x) + f(x) \cdot h(x). \end{aligned}$$

Having verified all of the conditions for the set of polynomials with complex coefficients, we thus see that this forms a ring with the typical addition and multiplication operations.

3. We note that we have done a lot of the work already: we just need to check that conditions 1, 4, and 5 hold for this subring, since the associativity, commutativity, and distributivity come from the work we did in part 2. Furthermore, we can quickly check 4, since $0, 1$ both are polynomials with x coefficient 0.

For 5, the additive inverse of a polynomial $\sum\limits_{i=0}^{n} f_i x^i$ is $\sum\limits_{i=0}^{n}(-f_i)x^i$. But the former lies in the set of polynomials whose x coefficient is 0 if and only if $f_i = 0$, which holds if and only if $-f_i = 0$. Thus, any polynomial f in our set has its additive inverse in the set. We thus just need to check that condition 1 is satisfied.

Notice that given polynomials $\sum\limits_{i=0}^{n} f_i x^i$ and $\sum\limits_{i=0}^{n} g_i x^i$ in our set, we have $f_1 = g_1 = 0$, and their sum equals

$$\sum_{i=0}^{n}(f_i + g_i)x^i,$$

where the coefficient at $i = 1$ is $f_1 + g_1 = 0 + 0 = 0$, so this lies in the set of polynomials with x coefficient 0. Similarly, their product is $\sum\limits_{i=0}^{2n}\left(\sum\limits_{\substack{0\le j\le n\\ 0\le i-j\le n}} f_j g_{i-j}\right)x^i$, and the coefficient at $i = 1$ is $f_0 g_1 + f_1 g_0 = 0 + 0 = 0$, so this set is also closed under multiplication.

We have therefore shown that the set of polynomials whose x coefficient is 0 forms a ring, which is what we wanted to show.

Problem 1.1.2. (points) Let $\mathbb{Z}/n\mathbb{Z}$ be the set of remainders of integers upon division by n, where addition and multiplication are defined modulo n. For instance, when $n = 6$, we have that $4+5 = 3$, and $4 \cdot 5 = 2$. Prove that this is a ring.

Solution. We check through the properties of a ring. For property 1, we note that we can add and multiply remainders to get another remainder, by the definition of these operations. Properties 2 and 3 follow from the associativity and commutativity of addition in $\mathbb{Z}$. For instance, for associativity, $(r_1 + r_2) + r_3$ is equal to $(r_1 + r_2) + r_3 \pmod{n}$, which is equal to $r_1 + (r_2 + r_3) \pmod{n}$.

Property 4 follows by using 0 and 1 for the additive and multiplicative identities, respectively. Indeed, adding a multiple of n to an integer does not change its remainder upon divison by n, and given an integer m with remainder r (so $m = qn + r$ for some integer q), if x is remainder 1, so $x = q'n + 1$ for some integer q', thhen $mx = (qn + r)(q'n + 1) = r + n(rq' + q + qq'n)$, which has remainder r.

Property 5 follows by considering $n - r$ for each remainder r (unless $r = 0$, where we can take 0), since $n - r + r$ is 0 modulo n.

Finally, property 6 follows from the distributivity law on $\mathbb{Z}$, similarly to properties 2 and 3. Indeed, given remainders r_1, r_2, r_3, we have $r_1(r_2 + r_3) \equiv r_1r_2 + r_1r_3 \pmod{n}$. Therefore, it follows that $\mathbb{Z}/nZ$ is a ring, with addition and multiplication defined modulo n.

Problem 1.1.3. Given a ring R, show that there exists an element $x \in R$ such that for all $r \in R$, $r + xr = 0$. What element is this?

Solution. We claim that this element x is the additive inverse of the additive identity $1 \in R$. Such an element exists by condition 5 in the definition of a ring.

By conditions 4 and 6, we thus have that for all $r \in R$, $r + xr = 1 \cdot r + x \cdot r = (1 + x) \cdot r = 0 \cdot r$. But by the above proposition, $0 \cdot r = 0$, which is what we wanted to show.

Problem 1.1.4. (points) Show that the set of odd integers, as a subset of $\mathbb{Z}$, is not an ideal.

Solution. We notice that the set of odd integers is not closed under addition, since $1+1=2$, with 1 odd, but 2 is even (so not odd). Hence, the set of odd integers is not an ideal.

Problem 1.1.5. Determine, with proof, all the prime ideals of $\mathbb{C}[x]$. You may use, without proof, the Fundamental Theorem of Algebra (namely, that polynomials in $\mathbb{C}[x]$ can be written as a product of linear factors, and this product is unique up to the order of the linear factors).

Solution. We claim that the set of prime ideals in $\mathbb{C}[x]$ is precisely the set of ideals of the form $\langle x-a\rangle$, where $a \in \mathbb{C}$, and the set $\{0\}$, with $\{0\}$ the only nonmaximal prime ideal.

First, we show that these are all prime. To show that $\{0\}$ is prime, we observe that if f, g are nonzero polynomials, this means that their leading coefficient are both nonzero, so the leading coefficient of fg is the leading coefficient of f times that of g, which is nonzero (and therefore $fg \neq 0$). In other words, if $fg = 0$, then one of f, g is zero.

Now, we show that $\langle x-a\rangle$ are prime, ideals. Suppose that fg lies in $\langle x-a\rangle$. Then, we note that the polynomial fg evaluates to 0 at $x = a$. In other words, $f(a)g(a) = 0$, meaning that one of $f(a), g(a)$ is zero. But that, in turn, implies that one of f, g lies in $\langle x-a\rangle$, which is what we wantedot show.

We show that all prime ideals must take one of the above forms. Suppose that I is a prime ideal that contains at least one nonzero polynomial f. By the fundamental theorem of algebra, we may write $f(x) = (x-a_1)(x-a_2)\cdots(x-a_n)$, where n is the degree of f and the a_i lie in $\mathbb{C}$. However, by the definition of prime, one of the $x-a_i$ lies in the ideal I. But then $\langle x-a_i\rangle \subset I$, since by definition of an ideal, $x-a_i \in I$ implies $o(x)(x-a_i) \in I$ for all $p(x) \in \mathbb{C}[x]$. However, if I is not equal to $\langle x-a_i\rangle$, then there is some polynomial $g \in I$ where $g(a_i) \neq 0$. In particular, as $g(x)-g(a_i)$ lies in $\langle x-a_i\rangle$, it follows that $g(a_i)$ lies in I, or that $1 \in I$. However, this means that $I = \mathbb{C}[x]$, which is a contradiction. Therefore, $I = \langle x-a_i\rangle$, which is what we wanted to show.

Problem 1.1.6. For each of the functions below, state whether they are injective, surjective, both, or neither.

1. The function $f(x) = |x|$ from the set of negative real numbers to the set of positive real numbers.

2. The function $f(x) = e^x$ from $\mathbb{R}$ to $\mathbb{R}$.

3. The function $f(x) = \sin x$ from $[0, 2\pi]$ to $[-1, 1]$.

Solution. **1.** The function $f(x) = |x|$ is both injective and surjective. For any positive real number r, $|-r| = r$, and $-r < 0$ is negative. Furthermore, if $|x| = |y|$, then $x = \pm y$. But x, y must be negative, so $x = y$.

2. This is injective, but not surjective. Note that $e^x > 0$ for all real x; in particular, this means that $e^x = -1$ has no solution, so this is not a surjective function. For injectivity: $e^x = e^y$ implies that, taking the natural logarithm, $x = y$.

3. This is surjective, but not injective. For surjectivity, note that for all $y \in [-1, 1]$, $\sin\sin^{-1}(y) = y$. For where it fails injectivity, note that $\sin 0 = \sin\pi = 0$.

1.2 A Family of Rings

Problem 1.2.1. Show that the set of real numbers $a + b\sqrt{2}$, where $a, b \in \mathbb{Q}$, forms a field, under the normal rules of addition and multiplication in $\mathbb{R}$. This set is sometimes notated as $\mathbb{Q}(\sqrt{2})$.

Solution. We first verify that this is a ring, and then show that every nonzero element has a multiplicative inverse. As this is a subset of $\mathbb{R}$, conditions 2, 3, and 6 are given to us automatically. Condition 4 follows since the identities 0 and 1 lie in $\mathbb{Q}$, and thus $\mathbb{Q}(\sqrt{2})$. Condition 5 follows from the fact that, for all $a+b\sqrt{2} \in \mathbb{Q}(\sqrt{2})$, $(-a)+(-b)\sqrt{2}$ is the additive inverse, and this also lies in $\mathbb{Q}(\sqrt{2})$.

To show that this is closed: for all elements $a+b\sqrt{2}$ and $a'+b'\sqrt{2}$, where $a, a', b, b' \in \mathbb{Q}$, we have $a+b\sqrt{2}+a'+b'\sqrt{2} = (a+a')+(b+b')\sqrt{2}$, and $(a+b\sqrt{2})(a'+b'\sqrt{2}) = (aa'+2bb')+(a'b+ab')\sqrt{2}$, both of which lie in $\mathbb{Q}(\sqrt{2})$ as $aa'+2bb', a'b+ab', a+a', b+b'$ all are rational numbers (as the rationals are closed under addition and multiplication).

Finally, say $a+b\sqrt{2} \neq 0$ is a nonzero element. Then, notice that $\frac{a}{a^2-2b^2} - \frac{b\sqrt{2}}{a^2-2b^2}$ is a multiplicative inverse if $a^2-2b^2 \neq 0$, as multiplying this with $a+b\sqrt{2}$ yields $\frac{(a-b\sqrt{2})(a+b\sqrt{2})}{a^2-2b^2} = \frac{a^2-2b^2}{a^2-2b^2} = 1$. Notice finally that $a^2-2b^2 = 0$ implies that $a^2 = 2b^2$, so either $b = 0$ (and so $a = 0$), or $\frac{a}{b}$ is $\sqrt{2}$, which we know is not possible. Thus, all nonzero elements in $\mathbb{Q}(\sqrt{2})$, $a+b\sqrt{2}$, are so $a^2-2b^2 \neq 0$, and thus have multiplicative inverses.

Therefore, it follows that $\mathbb{Q}(\sqrt{2})$ is a field, as desired.

Problem 1.2.2. Show that a ring R is a field if and only if it has exactly two ideals. Which two ideals are these?

Solution. Suppose that R is a field. Suppose that I is an ideal of R. Then, either $I = \{0\}$, which is an ideal, or I contains a nonzero element f. However, as the multiplicative inverse of f, f^{-1}, exists, we have that for all $r \in R$ that $rf^{-1}f = r \in I$, or that $I = R$. Thus, R has two ideals.

Now, suppose that R has two ideals. Then, in particular as $\{0\}$ and R are both ideals (and they are distinct since R contains a nonzero element), it follows these are all of the ideals. But then for all nonzero f, we have that $\langle f \rangle = R$. In particular, $1 \in \langle f \rangle$, so there exists an $r \in R$ so $rf = 1$, or that r is a multiplicative inverse of f. Therefore, R is a field, which is what we wanted to show.

Problem 1.2.3. Show that $\mathbb{Z}$ is a PID. To do this, given any ideal I of $\mathbb{Z}$, consider the smallest positive element in I, say i. Show that every element in the ideal has to be divisible by i.

Solution. We show that $\mathbb{Z}$ is a PID. Let I be an ideal of $\mathbb{Z}$. First, if $I = \{0\}$, then I is generated by the element 0. Otherwise, there exists a nonzero element $i \in I$; either $i > 0$, so there is some positive element in I, or $i < 0$. But if $i < 0$, then $(-1) \cdot i > 0$, and $(-1) \cdot i = -i \in I$.

Now, let r be the smallest positive element in I. We show that $\forall d \in I$, d is divisible by r. This will then show that $I = \langle r \rangle$.

Suppose for the sake of contradiction that an element $d \in I$ exists such that d is not divisible by r. Dividing d by r, we can then use the division algorithm to get $d = qr + r'$, where $0 < r' < r$ ($r' > 0$ by the fact that d is not divisible by r). However, $d, r \in I$, so $(-q) \cdot r$ and thus $d - qr$ both lie in I. But then $r' = d - qr \in I$, and $0 < r' < r$, contradicting the fact that r is the smallest positive element in I.

Thus, all the elements in I are divisible by r, and thus I is generated by r, meaning that $\mathbb{Z}$ is a PID, as desired.

Problem 1.2.4. Show that for any integral domain R, every prime element is irreducible.

Solution. Suppose that p is a prime element. Suppose that $p = ab$, where a, b are elements. Then, $\langle p \rangle$ is a prime ideal, meaning that, as $ab \in \langle p \rangle$, we have that either a or b lies in this ideal; without loss of generality, say this is a. Then, $a = pr$ for some $r \in R$, meaning that $brp = p$, or that

$(br-1)p=0$. As R is an integral domain and $p \neq 0$, we thus have that $br=1$; in other words, b has multiplicative inverse r, so b is a unit. We also note that p is not a unit, as otherwise $\langle p \rangle = R$, which is not prime.

Therefore, p must be irreducible (it cannot be written as a product of two nonunit elements).

Problem 1.2.5. Show that the set of elements $\mathbb{Z}[\sqrt{-13}]$, of the form $a+b\sqrt{-13}$, for $a,b \in \mathbb{Z}$, while an integral domain, is not a UFD, and therefore not a PID.

Solution. We first verify that this is a ring; with the operations inherited from $\mathbb{C}$, we just need to show that it contains the multiplicative and additive identities (which is clear, as $0,1 \in \mathbb{Z} \subset \mathbb{Z}[\sqrt{-13}]$), and is closed under addition, multiplication, and additive inverse.

But given $a+b\sqrt{-13}, c+d\sqrt{-5} \in \mathbb{Z}[\sqrt{-13}]$ their sum is $(a+c)+(b+d)\sqrt{-13}$, their product is $(ab-13bd)+(ad+bc)\sqrt{-13}$, and the additive inverse of $a+b\sqrt{-5}$ is $(-a)+(-b)\sqrt{-13}$, which for $a,b,c,d \in \mathbb{Z}$ lie in $\mathbb{Z}[\sqrt{-13}]$ since $\mathbb{Z}$ is a ring. Hence, $\mathbb{Z}[\sqrt{-13}]$ is a ring.

Furthermore, it is an integral domain. Indeed, if $(a+b\sqrt{-13})(c+d\sqrt{-13})=0$, then we have that $(ac-13bd)+(ad+bc)\sqrt{-13}=0$, or that $13bd-ac=(ad+bc)\sqrt{-13}$. But the right-hand side is not an integer unless $ad+bc=0$, and so $ac-13bd=0$ too. Now, suppose that $a+b\sqrt{-13} \neq 0$, so we have one of a,b is nonzero. If it is a, then notice that $ad+bc=0$ implies that $d=\frac{-bc}{a}$, and so $ac-13bd=ac+\frac{13b^2c}{a}=0$, or that $a^2c+13b^2c=0$, or $(a^2+13b^2)c=0$. If $a^2+5b^2=0$, then as squares of integers are positive, we would need $a,b=0$, contradiction. Thus, $c=0$, and therefore $d=0$. Similarly, if $b \neq 0$, we have that $c=\frac{-ad}{b}$, and so $ac-13bd=\frac{-a^2d}{b}-13bd=0$, or that $a^2d+13b^2d=0$, where the same argument lets us conclude that $c=d=0$. Hence, R is an integral domain.

To show it is not a UFD, consider the factorizations $14=2\cdot 7=(1+\sqrt{-13})(1-\sqrt{-13})$. On the one hand, notice all the elements involved are irreducible. Indeed, if $2=(a+b\sqrt{-13})(c+d\sqrt{-13})$, then $ac-13bd=2$ and $ad+bc=0$. But then either a or b is nonzero. But notice then that $(a-b\sqrt{-13})(c-d\sqrt{-13})=2$ as well, meaning that multiplying these implies that $(a^2+13b^2)(c^2+13d^2)=4$. But notice that a^2+13b^2 cannot equal 2. Hence, either $a^2+13b^2=1$ or $c^2+13d^2=1$, meaning that one of $a+b\sqrt{-13}, c+d\sqrt{-13}$ is ± 1, and so 2 is irreducible. By the same logic, we have that 7 is irreducible.

As for $1 \pm \sqrt{-13}$, notice that if $(1+\sqrt{-13})=(a+b\sqrt{-13})(c+d\sqrt{-13})$, then $ac-13bd=1, ad+bc=1$, and so similarly we find that $(1-\sqrt{-13})=(a-b\sqrt{-13})(c-d\sqrt{-13})$, so multiplying these together yields $(a^2+13b^2)(c^2+13d^2)=14$. But a^2+5b^2, c^2+5d^2 are positive, and cannot equal 2 or 7, so one of these is 1, and so one of $a+b\sqrt{-13}, c+d\sqrt{-13}$ is ± 1, so $1 \pm \sqrt{-13}$ is irreducible.

However, notice that $1 \pm \sqrt{-13}$ are not divisible by 2, since for any $a+b\sqrt{-13}$, we have $2(a+b\sqrt{-13})=2a+2b\sqrt{-13}$, where a,b are integers, and so we would need $1 \pm \sqrt{-13}=2a+2b\sqrt{-13}$, or $(2a-1)=(\pm 1-2b)\sqrt{-13}$. But this cannot happen as $\sqrt{-13}$ is not rational.

In particular, we have two factorizations where one of them is not simply a rearrangement of the other, or has different unit multiples. Thus, $\mathbb{Z}[\sqrt{-13}]$ is not a UFD, ergo not a PID.

1.3 Product Rings, Quotient Rings and More Examples

Problem 1.3.1. Prove that the operations are well-defined. That is, if $r_1'+I=r_1+I$ and $r_2'+I=r_2+I$, then

$$(r_1+I)+(r_2+I)=(r_1'+I)+(r_2'+I)$$

and

$$(r_1 + I) \cdot (r_2 + I) = (r_1' + I) \cdot (r_2' + I).$$

Solution. Suppose that $r_1' + I = r_1 + I$ and $r_2' + I = r_2 + I$, where $r_1', r_1, r_2', r_2 \in R$. Then, we know that for each $r \in (r_1 + r_2) + I$, by definition $r = r_1 + r_2 + i$ for some $i \in I$. However, $r_1 \in r_1' + I$, meaning that $r_1 = r_1' + i'$, and similarly $r_2 = r_2 + i''$, where i', i'' lie in I. But then $r = r_1' + r_2' + (i + i' + i'') \in r_1' + r_2' + I$. Therefore, $(r_1 + I) + (r_2 + I) \subset (r_1' + I) + (r_2' + I)$. By symmetry, we can run the same argument with r_i, r_i' swapped, meaning that these two are equal.

In particular, the value of $(r_1 + I) + (r_2 + I)$ is independent of the choices of r_1, r_2 to represent the set $r_1 + I$, and so is really a function of the set.

Similarly, $r \in (r_1 + I) \cdot (r_2 + I) = r_1 r_2 + I$ means that $r = r_1 r_2 + i$ for some $i \in I$. But again we then have that $r = r_1' r_2' + r_2' i' + r_1' i'' + i' i'' + i$. Since I is an ideal, $i, i', i'' \in I$ means that $r_2' i' + r_1' i'' + i' i'' + i \in I$, and so $r \in r_1' r_2' + I$. Therefore, $r_1 r_2 + I \subset r_1' r_2'$. Again, by symmetry, we may swap the roles of the r_i and r_i' to get that these two sets are equal. Again, it follows that $(r_1 + I) \cdot (r_2 + I)$, the operation we defined above, is independent of the specific choice of r_1, r_2 (and only depends on the set). This shows that our operations are well-defined, which is what we wanted to show.

Problem 1.3.2. Prove that R/I is a ring, equipped with the operations we defined above.

Solution. We run through the conditions again. The fact that R is closed under addition and multiplication is clear, since $(r_1 + I) + (r_2 + I) = (r_1 + r_2) + I$ and $(r_1 + I) \cdot (r_2 + I) = r_1 r_2 + I$ are both elements in R/I. Associativity, commutativity, and distributivity follow from the conditions for R. Indeed, for associativity we have that for all $r_1 + I, r_2 + I, r_3 + I \in R/I$ that

$$((r_1 + I) + (r_2 + I)) + (r_3 + I) = ((r_1 + r_2) + I) + (r_3 + I) = ((r_1 + r_2) + r_3) + I$$

$$= (r_1 + (r_2 + r_3)) + I = (r_1 + I) + ((r_2 + I) + (r_3 + I))$$

and

$$((r_1 + I) \cdot (r_2 + I)) \cdot (r_3 + I) = ((r_1 r_2) + I) \cdot (r_3 + I) = ((r_1 r_2) r_3) + I$$
$$= (r_1 (r_2 r_3)) + I = (r_1 + I) \cdot ((r_2 + I) \cdot (r_3 + I)).$$

For commutativity, we check that

$$(r_1 + I) + (r_2 + I) = (r_1 + r_2) + I = (r_2 + r_1) + I = (r_2 + I) + (r_1 + I)$$

and

$$(r_1 + I) \cdot (r_2 + I) = (r_1 r_2) + I = (r_2 r_1) + I = (r_2 + I) \cdot (r_1 + I),$$

and for distributivity we have

$$(r_1 + I) \cdot ((r_2 + I) + (r_3 + I)) = (r_1 + I) \cdot ((r_2 + r_3) + I) = (r_1 (r_2 + r_3) + I = (r_1 r_2 + r_1 r_3) + I$$

$$= (r_1 r_2 + I) + (r_1 r_3 + I) = (r_1 + I) \cdot (r_2 + I) + (r_1 + I) \cdot (r_3 + I).$$

For the identity, we verify that $0 + I$ is the additive identity and $1 + I$ is the multiplicative one, since for all $r \in R$, $(0 + I) + (r + I) = (0 + r) + I = r + I$ and $(1 + I) \cdot (r + I) = 1 \cdot r + I = r + I$. For the additive inverse, given $r + I \in R/I$, $(-r) + I \in R/I$ is the inverse, since $(r + I) + ((-r) + I) = (r + (-r)) + I = 0 + I$, the additive identity.

Having verified all of the conditions, it follows that R/I is a ring with the addition and multiplication operations we defined above.

Problem 1.3.3. Let R be a ring, and let I_1, I_2 be two ideals of R, such that $I_1 + I_2 = \{i_1 + i_2 | i_1 \in I_1, i_2 \in I_2\} = R$.

1. Show that $I_1 \cap I_2$ is an ideal.
2. Consider the homomorphism from $R/(I_1 \cap I_2)$ to $(R/I_1) \times (R/I_2)$ that sends $r + I_1 \cap I_2$ to $(r + I_1, r + I_2)$. Show that this map is well-defined and indeed a homomorphism.
3. Prove that the above map is injective.
4. Prove that the above map is surjective. As a suggestion on where to start, try considering any pair $(r_1 + I_1, r_2 + I_2)$, and the fact that $1 \in R = I_1 + I_2$.

Solution. **1.** First, suppose that $i, i' \in I_1 \cap I_2$. Then, $i, i' \in I_1$, so $i + i' \in I_1$. Similarly, $i + i' \in I_2$ since both i, i' lie in I_2. Therefore, $i + i' \in I_1 \cap I_2$.

Similarly, for all $i \in I_1 \cap I_2$ and $r \in R$, we have that $ri \in I_1$ and $ri \in I_2$, and so $ri \in I_1 \cap I_2$, meaning that $I_1 \cap I_2$ is an ideal, as desired.

2. We first argue that this map is well-defined. Indeed, suppose that $r + I_1 \cap I_2 = r' + I_1 \cap I_2$. Then, $r' \in r + I_1 \cap I_2$. Therefore, in particular, this means that $r' = r + i$, where $i \in I_1 \cap I_2$. In particular, this means that $r' \in r + I_1$ and $r' \in r + I_2$. Therefore, it follows that $r' + I_1 \subset r + I_1$, since each element of $r' + I_1$ can be written as $r' + i'$, where $i' \in I_1$, and so equals $r + i + i' \in r + I_1$. Similarly, we have that $r' + I_2 \subset r + I_2$. By symmetry, replacing the roles of r and r' yields that $r + I_1 \subset r' + I_1$ and $r + I_2 \subset r' + I_2$, meaning that $r + I_1 = r' + I_1$ and $r' + I_2 = r + I_2$. Therefore, this is a well-defined map $R/(I_1 \cap I_2)$ to $R/I_1 \times R/I_2$, the output value only depending on the set $r + I_1 \cap I_2$ and not the choice of representative.

To verify it is a ring homomorphism, we observe that given $r_1, r_2 \in R$, if ϕ is this map, we note that

$$\begin{aligned}\phi(r_1 + I_1 \cap I_2) + \phi(r_2 + I_1 \cap I_2) &= (r_1 + I_1, r_1 + I_2) + (r_2 + I_1, r_2 + I_2)\\ &= ((r_1 + r_2) + I_1, (r_1 + r_2) + I_2) = \phi((r_1 + I_1 \cap I_2) + (r_2 + I_1 \cap I_2))\end{aligned}$$

and

$$\begin{aligned}\phi(r_1 + I_1 \cap I_2) \cdot \phi(r_2 + I_1 \cap I_2) &= (r_1 + I_1, r_1 + I_2) \cdot (r_2 + I_1, r_2 + I_2)\\ &= ((r_1 r_2) + I_1, (r_1 r_2) + I_2) = \phi((r_1 + I_1 \cap I_2) \cdot (r_2 + I_1 \cap I_2)),\end{aligned}$$

and furthermore that $1 + I_1 \cap I_2$ is sent to $(1 + I_1, 1 + I_2)$. Noting that $1 + I_1$ is the multiplicative identity of R/I_1, $1 + I_2$ the multiplicative identity of R/I_2, we see that $(1 + I_1, 1 + I_2)$ is the multiplicative identity of the product ring. Thus, ϕ is a ring homomorphism, which is what we wanted.

3. Suppose that $\phi(r + I_1 \cap I_2) = \phi(r' + I_1 \cap I_2)$; by subtracting, this holds if and only if $\phi((r - r') + I_1 \cap I_2)$ is the additive identity in $R/I_1 \times R/I_2$. Then, it follows that, in particular, $(r - r') + I_1 = 0 + I_1$ and $(r - r') + I_2 = 0 + I_2$. But this implies that $r - r' \in I_1, I_2$, and thus that $r - r' \in I_1 \cap I_2$.

But this means that $r - r' + I_1 \cap I_2$ is just $0 + I_1 \cap I_2$, using the same argument as before (noting that $r - r' \in I_1 \cap I_2$ and $0 \in I_1 \cap I_2$). Therefore, $r + I_1 \cap I_2 = r' + I_1 \cap I_2$, and thus ϕ is injective.

4. To show this is surjective, suppose that we are given an element in $(R/I_1, R/I_2)$, $(r_1 + I_1, r_2 + I_2)$. We wish to show there is some element $r \in R$ so that $r + I_1 = r_1 + I_1$ and $r + I_2 = r_2 + I_2$; then $r + I_1 \cap I_2$ will be sent by ϕ to this element.

In other words, we wish to construct an element $r \in R$ so that $r - r_1 \in I_1, r - r_2 \in I_2$, as by the arguments in the previous parts this is enough to show that $r+I_1 = r_1+I_1$ and $r+I_2 = r_2+I_2$. To do this, by assumption we have that $1 \in I_1 + I_2$, meaning that there exist elements $i_1 \in I_1, i_2 \in I_2$ such that $1 = i_1+i_2$. From here, let $r = r_2i_1+r_1i_2$. Notice that $r-r_1 = r_1(1-i_2)+r_2i_1 = r_1i_1+r_2i_1 \in I_1$ and $r - r_2 = r_1i_2 + r_2(1 - i_1) = r_1i_2 + r_2i_2 \in I_2$. Therefore, our element $r \in R$ exists, which as we've argued previously is enough to show that ϕ is surjective.

Problem 1.3.4. Using the previous problem, derive the Chinese Remainder Theorem for integers. Namely, show that, given relatively prime integers m, n, show that given residues $r_1 \pmod{m}$ and $r_2 \pmod{n}$, there exists a unique residue $r \pmod{mn}$ so $r \equiv r_1 \pmod{m}$ and $r \equiv r_2 \pmod{n}$.

Solution. First, we show that if m, n are relatively prime, then $m\mathbb{Z} + n\mathbb{Z} = \mathbb{Z}$. To show this, recall that $\mathbb{Z}$ is a PID. However, if the ideal $m\mathbb{Z} + n\mathbb{Z}$ equals the ideal $\langle d \rangle$, then we need d to divide m, n. But then $d = 1$, and so therefore $m\mathbb{Z} + n\mathbb{Z} = \mathbb{Z}$.

Furthermore, note that $m\mathbb{Z} \cap n\mathbb{Z} = mn\mathbb{Z}$. Indeed, if d lies in both $m\mathbb{Z}$ and $n\mathbb{Z}$, then d is divisible by both m, n. But as m, n are relatively prime, it follows that d divides mn, or that $d \in mn\mathbb{Z}$. Therefore, $m\mathbb{Z} \cap n\mathbb{Z} \subset mn\mathbb{Z}$, and as the other inclusion is not hard to see, we find that these two ideals are actually equal.

Now, by the previous problem, we know that $\mathbb{Z}/mn\mathbb{Z}$, is isomorphic to $\mathbb{Z}/m\mathbb{Z} \times \mathbb{Z}/n\mathbb{Z}$, using the map that sends the remainder $r \pmod{mn}$ to the pair $(r \pmod{m}, r \pmod{n})$. This map being injective and surjective means that each pair of residues, one $\pmod{m}$ and one $\pmod{n}$, has exactly one residue modulo mn that is equivalent to the first modulo m and the second modulo n, which is what we wanted to show.

2 Vector Spaces

2.1 Definitions

Problem 2.1.1. Prove the following spaces are vector spaces.

1. The set of polynomials with complex coefficients (with the standard addition and multiplication operations), over the field $\mathbb{C}$.

2. $\mathbb{R}$, (with standard addition and multiplication operations), over the field $\mathbb{Q}$.

Solution. **1.** Since we have previously shown that $\mathbb{C}[x]$ is a ring, we have that addition and multiplication are associative, commutative, and distributive if we multiply and add two polynomials. In particular, these will also hold if we restrict our multiplication to only allowing multiplication of elements in $\mathbb{C}$, namely the constant polynomials. This gives us conditions 1, 2, 3, 7. Similarly, we have conditions 4 and 5 from the fact this is a ring.

Finally, notice that $1 \cdot p(x) = p(x)$ for any polynomial p, since the multiplicative identity in $\mathbb{C}[x]$ is 1, which also lies in $\mathbb{C}$.

2. Again, like in the first part, we are given that $\mathbb{R}$ is a field that contains $\mathbb{Q}$. Therefore, the fact $\mathbb{R}$ is closed under addition and multiplication, and that these operations are associative, commutative, and distributive means that they continue to have these properties if we restrict multiplication so the first element we multiply is a rational. This also gives us conditions 4 and 5, since $\mathbb{R}$ is a field, and ergo a ring. Just like in the previous part, we also have that $1 \in \mathbb{Q} \subset \mathbb{R}$ is the multiplicative identity in $\mathbb{R}$ too, so condition 6 holds. Hence, $\mathbb{R}$ is a vector space over the field $\mathbb{Q}$.

Problem 2.1.2. Determine all possible fields k such that $\mathbb{Z}$ can be made into a vector space over k, using the standard addition operations. In particular, you'll need to consider all possible scalar multiplication operations.

Solution. Suppose that k is a field such that $\mathbb{Z}$ is a vector space over k with the typical addition operation. We will notate the elements of k with the subscript k. First, notice that if $1_k + 1_k = 0_k$, then by the distributivity law we have that $0_k \cdot 1 = 1_k \cdot 1 + 1_k \cdot 1 = 2$. But $0_k \cdot 1 = (0_k \cdot 1) + (0_k \cdot 1)$ implies that $0_k \cdot 1 = 0$, contradiction.

Otherwise, if $1_k + 1_k$ is not 0_k, then by the definition of a field it has a multiplicative inverse, say r_k. But then observe that $r_k \cdot (1_k + 1_k) = 1_k$ implies that $r_k + r_k = 1_k$, by the distributivity of the field operation. Finally, notice that $r_k \cdot 1 + r_k \cdot 1 = (r_k + r_k) \cdot 1 = 1_k \cdot 1 = 1$. In other words, we have an integer that, added with itself, we get 1. But no such integer exists.

Therefore, there is no field where $\mathbb{Z}$ can be made into a vector space over k using the normal addition operation in $\mathbb{Z}$.

2.2 Coordinates and Bases

Problem 2.2.1. Find two distinct bases (the plural of basis) for the vector space of polynomials with real coefficients of degree at most 3, and prove they are bases.

Solution. There are many choices of bases that one could use for this vector space. For this solution, we will show that $\{1, x, x^2, x^3\}$ and $\{1, 1+x, 1+x+x^2, 1+x+x^2+x^3\}$ are both bases.

We first show that the first set is linearly independent. To do this, suppose that a_0, a_1, a_2, a_3 are real, such that $a_0 + a_1x + a_2x^2 + a_3x^3 = 0$. Then, we need a_0, a_1, a_2, a_3 to all be zero. Furthermore, for spanning, any polynomial of degree at most 3 can be written as $a_3x^3 + a_2x^2 + a_1x + a_0$, which is definitely a linear combination of $1, x, x^2, x^3$.

To show the second set is a basis, we first show that this is linearly independent. Suppose there are real coefficients a_0, a_1, a_2, a_3 such that $a_0 + a_1(1+x) + a_2(1+x+x^2) + a_3(1+x+x^2+x^3) = 0$. Then, we have $(a_0 + a_1 + a_2 + a_3) + (a_1 + a_2 + a_3)x + (a_2 + a_3)x^2 + a_3x^3 = 0$. Then, the coefficient of x^3 is zero, so $a_3 = 0$. Meanwhile, the x^2 coefficient requires $a_2 + a_3 = 0$, so $a_2 = 0$. Repeating this for the x coefficient yields $a_1 = 0$, and finally for the constant term we need $a_0 = 0$. This shows that $1, 1+x, 1+x+x^2, 1+x+x^2+x^3$ are linearly independent.

To show they span, suppose we have a polynomial $a_3x^3 + a_2x^2 + a_1x + a_0$. Notice that this equals $a_3(x^3 + x^2 + x + 1) + (a_2 - a_3)(x^2 + x + 1) + (a_1 - a_2)(x + 1) + (a_0 - a_1)$, meaning that each polynomial with real coefficients is a linear combination of the polynomials in the second set. It thus follows that $1, 1+x, 1+x+x^2, 1+x+x^2+x^3$ forms a basis of the vector space of polynomials of degree at most 3, which is what we wanted to show.

Problem 2.2.2. Suppose that S is a spanning set, and v is a vector that doesn't lie in S.

1. Show that $S \cup \{v\}$ is linearly dependent.

2. Suppose furthermore that v is nonzero. Then, show there exists a vector $w \in S$ such that $(S - \{w\}) \cup \{v\}$ is a spanning set.

Solution. **1.** Suppose that S is a spanning set. This means that there exist vectors $v_1, v_2, \ldots, v_n$ and $a_i \in k$ such that $a_1v_1 + \cdots + a_nv_n = v$, or that $a_1v_1 + a_2v_2 + \cdots + a_nv_n - v = 0$. But $-1 \neq 0$ in the field k, meaning that $S \cup \{v\}$ is linearly dependent, which is what we wanted to show.

2. Suppose that S is a spanning set of a vector space V. Then, v can be written as a linear combination of vectors in S, given by $\sum_{i=1}^{n} a_i v_i$, where the v_i lie in S and the a_i lie in the field k. Since v is nonzero, one of the a_i is nonzero, say a_1. We claim that $S' = (S - \{v_1\}) \cup \{v\}$ is a spanning set of the vector space.

Indeed, let $w \in V$. Since S is a spanning set, we can write w as the sum $\sum_{i=1}^{m} b_i w_i$, where the w_i are elements of S. If none of the w_i are v_1, then this is also a linear combination of elements in S'. Otherwise, by swapping labelling we can assume $w_1 = v_1$ and the other vectors are distinct. Then, notice that

$$w = \sum_{i=1}^{m} b_i w_i = b_1 w_1 + \sum_{i=2}^{m} b_i w_i = b_1(\frac{1}{a_1} v - \sum_{i=2}^{n} \frac{a_i}{a_1} v_i) + \sum_{i=2}^{m} b_i w_i.$$

In other words, w is a linear combination of elements in S'. As this holds for all $w \in V$, it follows that S' is a spanning set for V.

Problem 2.2.3. Show that if $L \not\subseteq S$, we can replace a vector in S with one in L so that S remains a spanning set, and $S \cap L$ increases in size by one.

Solution. Given S and L, we know that there exists some vector v that lies in L that doesn't lie in S. Notice that $v \neq 0$, as otherwise L is not a linearly independent set. Then, by the previous problem, there exists a vector $w \in S$ such that $(S - \{w\}) \cup \{v\}$ is also a spanning set.

Furthermore, by the construction in the previous problem, we can pick a w such that $w \notin L$. Otherwise, per the construction in the previous problem, v is a linear combination of vectors in $L \cap S$, contradicting linear independence. Therefore, writing v as a linear combination of vectors in S, there is some vector in this linear combination that lies in S but not L. We can, from the previous problem, remove this vector from S and replace it with v to get another spanning set.

In other words, we can replace w with v to keep S a spanning set. By definition, notice that $S \cap L$ has gained an element, since we removed w from S (which does not change $S \cap L$) and added v to S (which adds an element to $S \cap L$). This is what we wanted to show.

Problem 2.2.4. Using the above procedure, show that L must be finite, and that L must have at most as many elements as S. Conclude that the size of every linearly independent set is at most the size of every spanning set.

Solution. Suppose that we are given a spanning set S which is finite and L which is linearly independent. If L is not contained in S, by the previous problem, we may repeatedly replace elements in S with elements in L such that $|L \cap S|$ strictly increases. It follows that after at most $|L|$ steps of repeating this process, we eventually have a spanning set that contains L. Furthermore, by construction, our spanning set is the same size as S, since we remove an element from S and add a different element at each step.

Therefore, it follows that $|S| \geq |L|$, meaning that L in particular is also a finite subset with at most as many elements of S.

Finally, noting that we can apply this procedure to any linearly independent set and any spanning set. Therefore, this means that for any spanning set and any linearly independent set, the spanning set has at least as many elements as the linearly independent set.

Problem 2.2.5. Prove the following.

1. Any spanning set with finitely many elements can be reduced to a basis. That is, we may remove elements from our spanning set such that the resulting set is a basis.

2. Any linearly independent set can be extended to a basis. That is, we may add elements to our linearly independent set so that the resulting set is a basis.

Solution. **1.** Let $S = \{v_1, v_2, \dots, v_n\}$ be a spanning set with finitely many elements. If S is not linearly independent, then we claim that we can remove a vector from S while still keeping the set a spanning set. Indeed, suppose that there exist $a_i \in k$ such that $\sum\limits_{i=1}^{n} a_i v_i = 0$, where not all of the a_i are zero. Without loss of generality, assume that $a_1 \neq 0$. Then, we have that $v_1 = -\sum\limits_{i=2}^{n} \frac{a_i}{a_1} v_i$.

We show that $\{v_2, v_3, \dots, v_n\}$ is still a spanning set. Indeed, given any vector $v \in V$, our vector space, we may write it as a linear combination $\sum\limits_{i=1}^{n} b_i v_i$. But then notice that this equals $\sum\limits_{i=2}^{n} (\frac{-a_i b_1}{a_1} + b_i) v_i$, from our formula for v_1 above. This proves that $\{v_2, v_3, \dots, v_n\}$ is a spanning set of V.

We can repeat this procedure until eventually we end up with a linearly independent set (as the empty set is by convention a linearly independent set, this procedure must eventually terminate with a linearly independent and spanning set). This set will then be our basis, which is a subset of our spanning set. This proves the first part of the problem.

2. Let L be a linearly independent set. From the previous problem it must be a finite set, say $\{v_1, v_2, \dots, v_k\}$. If L is a spanning set, we are done. Otherwise, there exists a vector $v \in V$ that cannot be written as a linear combination of vectors in L. Consider the set $L \cup \{v\}$. Notice that this set is also linearly independent (otherwise, a linear combination of elements in $L \cup \{v\}$ that equals zero where not all coefficients are zero, say $av + \sum\limits_{i=1}^{k} a_i v_k = 0$, means either $a \neq 0$, so $v = \sum\limits_{i=1}^{k} \frac{-a_i}{a} v_k$, or $a = 0$, and this linear combination tells us that the vectors in L are not linearly independent).

Thus, given any linearly independent set which is not spanning, we may add a vector to it such that the set remains linearly independent. We may keep repeating this until our set is spanning. This procedure terminates since any linearly independent set has at most as many vectors as a spanning set (and we know one spanning set exists which has finitely many vectors). Therefore, we can extend a linearly independent set into a basis, which is what we wanted to show.

Problem 2.2.6. Show that any two bases of our finite dimensional vector space have the same size. This size is known as the **dimension** of the vector space, denoted as $\dim V$.

Solution. By problem 23, every linearly independent set is at most the size of any spanning set. Therefore, in particular, since we know that V has a finite spanning set, every basis (which is a linearly independent set) also is finite length. Furthermore, from the previous problem, by starting with a finite spanning set and removing vectors, we know that every vector space has a basis.

Furthermore, if our bases are B_1 and B_2, then noting that B_1 is linearly independent and B_2 is spanning yields that $|B_1| \leq |B_2|$. But similarly, B_1 is spanning and B_2 is linearly independent, so $|B_1| \geq |B_2|$. Combining these equalities yields that $|B_1| = |B_2|$, which is the dimension of V, as desired.

Problem 2.2.7. Show that if W is a subspace of V, then the dimension of W is at most that of V.

Solution. If W is a subspace of V, we know that a basis of W exists. This set is linearly independent. Therefore, the number of elements in this basis is at most the dimension of V, which is the size of any basis of V (which, in particular, is a spanning set). Therefore, $\dim V \geq \dim W$, which is what we wanted to show.

2.3 Linear Transforms

Problem 2.3.1. Suppose that $k = \mathbb{Q}$, and V, W are vector spaces over $\mathbb{Q}$. Show that if $T : V \to W$ satisfies $T(v_1 + v_2) = T(v_1) + T(v_2)$ for all $v_1, v_2 \in V$, then T is actually linear.

Solution. Suppose that $f(v_1 + v_2) = f(v_1) + f(v_2)$ for all v_1, v_2 in a vector space V over $\mathbb{Q}$. We will show that $f(qv) = qf(v)$ for all vectors $v \in V$.

We can first argue by induction that $f(mv) = mf(v)$ for all $m \in \mathbb{N}$. Indeed, the base case $m = 1$ is given by definition, and given that this holds for $m \leq k$, then $f((k+1)v) = f(v) + f(kv) = f(v) + kf(v) = (k+1)f(v)$. Notice that this can be extended to $m \in \mathbb{Z}$, by observing that $f(mv) + f((-m)v) = f(mv + (-m)v) = f((m + -m)v) = f(0) = 0$, since $f(0) + f(0) = f(0)$, or $f(0) = 0$.

Finally, given a rational number $q = m/n$, notice that $nf(qv) = f(mv) = mf(v)$, or that $f(qv) = m/nf(v) = qf(v)$ for all $v \in V$. Therefore, it follows that $f(qv) = qf(v)$ for all $q \in \mathbb{Q}$ and $v \in V$, or that f is linear, which is what we wanted to show.

Problem 2.3.2. Show that $\ker T$ is a subspace of V.

Solution. In order to show this is a subspace of V, we need to verify that this space contains 0_V and is closed under addition and scalar multiplication. Notice that $T(0) + T(0) = T(0+0) = T(0)$, or that $T(0) = 0$, meaning that $0_V \in \ker T$. Furthermore, given $v_1, v_2 \in V$, we have that $T(v_1 + v_2) = T(v_1) + T(v_2) = 0 + 0 = 0$, so hence $\ker T$ is closed under addition. Finally, given $v \in V$ and $r \in k$, we have that $T(rv) = rT(v) = r \cdot 0_W = 0_W$. But this means that $\ker T$ is closed under scalar multiplication. In other words, we have that $\ker T$ is a subspace of V, which is what we wanted to show.

Problem 2.3.3. Suppose that V and W are finite dimensional vector spaces with the same dimension d. Prove that V, W are **isomorphic**; that is, there exists an isomorphism between them.

Solution. Suppose that V, W both are vector spaces of dimension d. Then, it follows that there exist bases $v_1, v_2, \ldots, v_d$ of V and $w_1, w_2, \ldots, w_d$ of W. Let $T : V \to W$ be the map given by $T(\sum\limits_{i=1}^{n} a_i v_i) = \sum\limits_{i=1}^{n} a_i w_i$.

First, notice that this map is well-defined. Indeed, since $v_1, v_2, \ldots, v_d$ is spanning, we have that this formula defines the map for all $v \in V$. Furthermore, if $\sum\limits_{i=1}^{n} a_i v_i = \sum\limits_{i=1}^{n} a'_i v_i$, where the $a_i, a'_i \in k$, then $\sum\limits_{i=1}^{n} (a_i - a'_i)v_i = 0$. But since the v_i are linearly independent, it follows that $a_i = a'_i$ for each i. But then this map is well-defined, since each v has a unique representation as a linear combination of the vectors v_i.

We verify first that this is a linear transformation. Indeed, given vectors $v = \sum\limits_{i=1}^{n} a_i v_i$ and

$v' = \sum\limits_{i=1}^{n} a_i' v_i$, we notice that

$$T(v+v') = T(\sum_{i=1}^{n}(a_i + a_i')v_i) = \sum_{i=1}^{n}(a_i + a_i')w_i = \sum_{i=1}^{n} a_i w_i + \sum_{i=1}^{n} a_i' w_i = T(v) + T(v').$$

Similarly, we find that given this vector v, and $s \in k$, we have $T(sv) = T(\sum\limits_{i=1}^{n} s a_i v_i) = \sum\limits_{i=1}^{n} s a_i w_i = sT(v)$. Therefore, T is linear.

Furthermore, we can define the linear map $S : W \to V$ by $S(\sum\limits_{i=1}^{n} s_i w_i) = \sum\limits_{i=1}^{n} s_i v_i$, which by the same logic as above is well-defined and a linear map. Finally, notice that for all vectors $v \in V$, if $v = \sum\limits_{i=1}^{n} a_i v_i$, then $S(T(v)) = S(\sum\limits_{i=1}^{n} a_i w_i) = \sum\limits_{i=1}^{n} a_i v_i$, and similarly given $w \in W$, if $w = \sum\limits_{i=1}^{n} b_i w_i$, then $T(S(w)) = T(\sum\limits_{i=1}^{n} b_i v_i) = \sum\limits_{i=1}^{n} b_i w_i$. Therefore, S, T are inverses, so V, W are isomorphic, as desired (noting that maps with inverses are bijective).

Problem 2.3.4. Prove that an infinite dimensional vector space cannot be isomorphic to a finite dimensional vector space.

Solution. Suppose for the sake of contradiction that V, W are isomorphic vector spaces, with V infinite-dimensional and W finite-dimensional, and $T : W \to V$ is our isomorphism. Then, as W is finite-dimensional, it has a finite spanning set $\{w_1, w_2, \ldots, w_n\}$. Now, since T is an isomorphism, it is surjective, meaning that every $v \in V$ can be written as $v = T(w)$ for some $w \in W$. But from our spanning set, $w = \sum\limits_{i=1}^{n} a_i w_i$ for some $a_i \in k$, meaning that, by linearity, $v = T(w) = \sum\limits_{i=1}^{n} a_i T(w_i)$. In particular, this means that every element of V is a linear combination of vectors in the set $T(w_1), T(w_2), \ldots, T(w_n)$, meaning that V has a finite spanning set. But this contradicts the fact that V is infinite dimensional.

Therefore, there cannot be an isomorphism between a finite dimensional vector space and an infinite dimensional vector space, as desired.

Problem 2.3.5. To prove the theorem, prove the following:

1. Show that this basis of $\ker T$ can be extended to a basis of V.
2. Suppose that this extension adds vectors $w_{n+1}, w_{n+2}, \ldots, w_m$. Show that $T(w_{n+1}), T(w_{n+2}), \ldots, T(w_m)$ form a basis for imT, and from here prove the theorem.

Solution. **1.** For the first part, we observe that a basis of $\ker T$ is linearly independent in V. Therefore, by Problem 2.2.5, we can extend this linearly independent set to a basis of V. Say this gives us the vectors $w_1, w_2, \ldots, w_n, w_{n+1}, \ldots, w_m$.

2. We now verify that $T(w_{n+1}), \ldots, T(w_m)$ forms a basis for imT. First, we check that this is spanning. To see this, suppose that $w \in imT$, so then $w = T(v)$ for some $v \in V$. Since $w_1, w_2, \ldots, w_m$ is a basis of V, we can write $v = \sum\limits_{i=1}^{m} a_i w_i$ for some $a_i \in k$. Therefore, $w = T(v) = \sum\limits_{i=1}^{m} a_i T(w_i)$. Since $w_1, w_2, \ldots, w_n \in \ker T$, this equals $\sum\limits_{i=n+1}^{m} a_i T(w_i)$. This shows that $T(w_i)$, for $i = n+1, n+2, \ldots, m$, span imT.

To show that these are linearly independent, suppose that $\sum_{i=n+1}^{m} a_i T(w_i)$, we have that $T(\sum_{i=n+1}^{m} a_i w_i) = 0$. Therefore, $\sum_{i=n+1}^{m} a_i w_i \in \ker T$, so hence $\sum_{i=n+1}^{m} a_i w_i = \sum_{i=1}^{n} (-a_i) w_i$ for some $a_1, a_2, \dots, a_n \in k$, since $w_1, w_2, \dots, w_n$ is a basis for $\ker T$. But this means that $\sum_{i=1}^{m} a_i w_i = 0$, or that all of the a_i are zero. In particular, this means that $T(w_{n+1}), T(w_{n+2}), \dots, T(w_m)$ is linearly independent, and therefore they form a basis for imT.

In particular, $\dim imT = m - n = \dim V - \dim \ker T$, which is enough to prove the theorem.

Problem 2.3.6. Show that two finite dimensional vector spaces are isomorphic if and only if they have the same dimension.

Solution. We have already shown that two finite dimensional vector spaces that are the same dimension are isomorphic to each other. Now, suppose that V, W are isomorphic vector spaces; consider linear transformations $T : V \to W$ that is an isomorphism.

Then, using the above result, we note that $\dim \ker T = 0$, since $\ker T$ only consists of the 0 vector by T being injective (the empty set being a basis for $\ker T$). Furthermore, T is surjective, meaning that $imT = W$. But then $\dim V = \dim \ker T + \dim imT = 0 + \dim W = \dim W$, meaning that V, W are the same dimension, which is what we wanted to show.

2.4 Matrices and Row Reduction

Problem 2.4.1. Show that this above map is well-defined and is an isomorphism between V and k^n.

Solution. We first check that this map is well-defined. However, this follows from the fact that $w_1, w_2, \dots, w_n$ is a basis for V, meaning that the coefficients $a_1, a_2, \dots, a_n$ are uniquely defined by v (as the difference between two such representations of v is a linear combination of the w_i that yields 0, which by linear independence must be zero).

To show that this linear, suppose this map is T. We observe that, given vectors $v, v' \in V$, if $v = \sum_{i=1}^{n} a_i w_i$ and $v' = \sum_{i=1}^{n} a'_i w_i$, then our map sends $v + v'$ to the column vector $\begin{pmatrix} a_1 + a'_1 \\ a_2 + a'_2 \\ a_3 + a'_3 \\ \vdots \\ a_n + a'_n \end{pmatrix} = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \\ \vdots \\ a_n \end{pmatrix} + \begin{pmatrix} a'_1 \\ a'_2 \\ a'_3 \\ \vdots \\ a'_n \end{pmatrix}$, which is equal to $T(v) + T(v')$. In addition, given $a \in k$, we have $T(av) = \begin{pmatrix} aa_1 \\ aa_2 \\ aa_3 \\ \vdots \\ aa_n \end{pmatrix} = a \begin{pmatrix} a_1 \\ a_2 \\ a_3 \\ \vdots \\ a_n \end{pmatrix} = aT(v)$. Hence, this map is linear.

Injectivity is clear: $\ker T$ consists of the vectors that send $\sum\limits_{i=1}^{n} a_i v_i$ to $\begin{pmatrix} a_1 \\ a_2 \\ a_3 \\ \vdots \\ a_n \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$, or that $v = 0 + 0 + \cdots + 0$. Surjectivity is also clear, since $\begin{pmatrix} a_1 \\ a_2 \\ a_3 \\ \vdots \\ a_n \end{pmatrix} = T(\sum\limits_{i=1}^{n} a_i w_i)$, and thus we have shown that T is an isomorphism, which is what we wanted.

Problem 2.4.2. Show that if we multiply the matrix of T with the coordinate representation of $v \in V$, we get the coordinate representation of $T(v)$. In this sense, our notion of matrix multiplication is consistent with the way our linear transformation acts on vectors.

Solution. Suppose that we are given a vector $v = \sum\limits_{i=1}^{n} a_i v_i$. Then, $T(v) = \sum\limits_{i=1}^{n} a_i T(v_i) = \sum\limits_{i=1}^{n} \sum\limits_{j=1}^{m} a_i a_{j,i} w_j$, which has coordinate representation $\begin{pmatrix} \sum\limits_{i=1}^{n} a_i a_{1,i} \\ \sum\limits_{i=1}^{n} a_i a_{2,i} \\ \sum\limits_{i=1}^{n} a_i a_{3,i} \\ \vdots \\ \sum\limits_{i=1}^{n} a_i a_{m,i} \end{pmatrix}$. Meanwhile, if we consider the matrix multiplication of the coordinates of v, we have

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ a_3 \\ \vdots \\ a_n \end{pmatrix} = \begin{pmatrix} \sum\limits_{i=1}^{n} a_i a_{1,i} \\ \sum\limits_{i=1}^{n} a_i a_{2,i} \\ \sum\limits_{i=1}^{n} a_i a_{3,i} \\ \vdots \\ \sum\limits_{i=1}^{n} a_i a_{m,i} \end{pmatrix},$$

which is the same as the coordinate representation of $T(v)$.

Problem 2.4.3. Show that there exists bases for V, W such that the only nonzero entries of T are along the diagonal; that is, only the (i,i)th entries are nonzero for $i = 1, 2, \ldots, r$ for some nonnegative integer r. What is the value of r?

Solution. Consider the subspace $\ker T \subset V$; this vector space has a basis $v_1, v_2, v_3, \ldots, v_k$. Extend this to a basis of V by $v_1, v_2, \ldots, v_k, v_{k+1}, v_{k+2}, \ldots, v_n$. Relabel this basis, furthermore, such that

$v_{n-k+1}, v_{n-k+2}, \ldots, v_n$ are the basis vectors of $\ker T$. Now, consider the vectors $w_1 = T(v_1), w_2 = T(v_2), \ldots, w_{n-k} = T(v_{n-k})$. By Problem 31, these vectors form a basis for imT, and therefore are linearly independent in W. Hence, we may extend this to a basis $w_1, w_2, \ldots, w_m$.

Consider the matrix for T with respect to these bases. Then, observe that $T(v_i) = w_i$ for $1 \leq i \leq n-k$, and $T(v_i) = 0$ otherwise, by construction. Therefore, the matrix of T looks like

$$\begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 0 \end{pmatrix},$$

which is what we wanted to show.

We observe that r, the value above, is equal to $n-k = \dim V - \dim \ker T = \dim imT$, which is the rank of T.

Problem 2.4.4. Show that if we apply one of our row reduction operations to a matrix for T, we get another matrix for T, using a different basis for W (but the same basis for V). How do you relate the old basis to the new basis?

Solution. Suppose that we are given a matrix for T, with entries $a_{i,j}$, with respect to the bases $v_1, v_2, \ldots, v_n$ for V and $w_1, w_2, \ldots, w_m$ for W. By definition, for each i, we have $T(v_i) = \sum_{j=1}^{m} a_{ji} w_j$. We consider each of the choices of row operation.

1. Consider scaling up the rth row of the matrix by $c \neq 0$. Then, notice that if the entries of this new matrix are $a'_{i,j}$, notice that

$$T(v_i) = \sum_{j=1}^{m} a_{ji} w_j = \sum_{j=1}^{r-1} a'_{ji} w_j + a'_{ri} \frac{w_r}{c} + \sum_{j=r+1}^{m} a'_{ji} w_j.$$

   Notice that this holds for each i, so it suffices to check that $w_1, w_2, \ldots, w_{r-1}, w_r/c, w_{r+1}, \ldots, w_m$ is a basis for W. This follows as this is still a spanning set (we can write any vector as a linear combination of $w_1, w_2, \ldots, w_m$, which gives us a linear combination of our new basis, scaling up the coefficient of the rth vector by $c \neq 0$), and is still linearly independent (if a linear combination of these vectors is 0, then this is a linear combination of our original basis, with the coefficient of the rth vector scaled down by $1/c$; these must be zero, so the original coefficients were zero too).

2. Consider swapping the rth row with the sth row, $r \neq s$. Then, if our new matrix has entries $a'_{i,j}$, then $a'_{i,j} = a_{i,j}$ if $i \neq r, s$, $a'_{r,j} = a_{s,j}$, and $a'_{s,j} = a_{r,j}$. We can easily check that

$$T(v_i) = \sum_{j=1}^{m} a_{ji} w_j = \sum_{\substack{1 \leq j \leq m \\ j \neq r,s}} a'_{ji} w_j + a'_{ri} w_s + a'_{si} w_r.$$

   This corresponds, then, to the basis with w_r, w_s swapped positions (which does not affect spanning or linear independence).

3. Consider adding c times the rth row to the sth row, where $r \neq s$. Then, if our new matrix has entries $a'_{i,j}$, then $a'_{i,j} = a_{i,j}$ if $i \neq s$ and $a'_{s,j} = a_{s,j} + ca_{r,j}$. We again see that

$$T(v_i) = \sum_{j=1}^{m} a_{ji} w_j = \sum_{\substack{1 \leq j \leq m \\ j \neq s}} a'_{ji} w_j + (a'_{si} - ca'_{ri}) w_s = \sum_{\substack{1 \leq j \leq m \\ j \neq r}} a'_{ji} w_j + a'_{ri}(w_r - cw_s).$$

This corresponds to the basis $w_1, w_2, \ldots, w_r - cw_s, w_{r+1}, \ldots, w_m$.

We need to check that this is a basis. For spanning, given a vector $w \in W$, we may write it as $\sum_{i=1}^{m} a_i w_i$, which is also equal to $(a_s + ca_r)w_s + a_r(w_r - cw_s) + \sum_{\substack{1 \leq i \leq m \\ i \neq r,s}} a_i w_i$. Hence, every vector in W is a linear combination of vectors in our new set, so our set is actually spanning. For linear independence:

$$a_r(w_r - cw_s) + \sum_{\substack{1 \leq i \leq m \\ i \neq r}} a_i w_i = (a_s - ca_r)w_s + \sum_{\substack{1 \leq i \leq m \\ i \neq s}} a_i w_i = 0$$

implies that $a_i = 0$ for $i \neq s$, and $a_s - ca_r = 0$. But $a_r = 0$ implies that $a_s = 0$, which means our vectors are linearly independent. Therefore, our set is actually a basis.

Therefore, each row reduction step turns the matrix of T into a different matrix of T with respect to a different basis for W, which is what we wanted to show.

We can relate the old bases to the new bases in each type of operation through what we computed above.

1. For the first, if we scale the rth row by c, this corresponds to scaling the rth basis vector by $1/c$.
2. For the second, swapping rows r, s corresponds to swapping the rth and sth basis vectors.
3. For the third, adding c times row r to row s corresponds to adding $-c$ of the sth basis vector to the rth basis vector.

Problem 2.4.5. Reduce the following matrices to reduced row echelon form.

1. $\begin{pmatrix} 1 & 2 & 3 \\ 6 & 5 & 4 \end{pmatrix}$
2. $\begin{pmatrix} 4 & 2 & -1 & -3 \\ 1 & 0 & -5 & 2 \\ 0 & 1 & 0 & 2 \end{pmatrix}$

Solution. **1.** We start with the matrix $\begin{pmatrix} 1 & 2 & 3 \\ 6 & 5 & 4 \end{pmatrix}$. Adding -6 times the first row to the second row yields the matrix $\begin{pmatrix} 1 & 2 & 3 \\ 0 & -7 & -14 \end{pmatrix}$. Dividing the second row by -7 then yields the matrix $\begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \end{pmatrix}$. Finally, subtracting two times the second row from the first row yields $\begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 2 \end{pmatrix}$, which is our desired reduced row echelon form.

2. We start with the matrix $\begin{pmatrix}4 & 2 & -1 & -3\\ 1 & 0 & -5 & 2\\ 0 & 1 & 0 & 2\end{pmatrix}$. Swap the first and second rows to get the matrix $\begin{pmatrix}1 & 0 & -5 & 2\\ 4 & 2 & -1 & -3\\ 0 & 1 & 0 & 2\end{pmatrix}$. Subtract four times the first row from the second to get $\begin{pmatrix}1 & 0 & -5 & 2\\ 0 & 2 & 19 & -11\\ 0 & 1 & 0 & 2\end{pmatrix}$. Swap the second and third row to get $\begin{pmatrix}1 & 0 & -5 & 2\\ 0 & 1 & 0 & 2\\ 0 & 2 & 19 & -11\end{pmatrix}$. Subtract two times the second row from the third row to get $\begin{pmatrix}1 & 0 & -5 & 2\\ 0 & 1 & 0 & 2\\ 0 & 0 & 19 & -15\end{pmatrix}$. Divide the third row by 19 to get $\begin{pmatrix}1 & 0 & -5 & 2\\ 0 & 1 & 0 & 2\\ 0 & 0 & 1 & -\frac{15}{19}\end{pmatrix}$. Finally, add 5 times the third row to the first row to get $\begin{pmatrix}1 & 0 & 0 & \frac{-37}{19}\\ 0 & 1 & 0 & 2\\ 0 & 0 & 1 & -\frac{15}{19}\end{pmatrix}$. This is our desired reduced row echelon form.

Problem 2.4.6. Show that for any i, using the row operations on the matrix for T, if column i had at least one nonzero entry initially, then there is a sequence of row operations such that the resulting matrix only has one nonzero entry in column i, and it is a 1.

Solution. Suppose that the matrix for T has a nonzero entry in column i at row j. Say that the entries of T are a_{kl} in row k and column l, and suppose T has m columns and n rows. Then, for each $k \neq j$, if we add $\frac{-a_{ki}}{a_{ji}}$ times the jth row to the kth row, notice that the entry in row k, column i is $a_{ki} + \frac{-a_{ki}}{a_{ji}}a_{ji} = 0$. Finally, if we scale the jth row by $\frac{1}{a_{ji}}$, the jth row has a 1 in column i.

In particular, using these row operations, we can change the matrix of T such that the ith column only has one nonzero entry, namely 1 (in the jth row), which is what we wanted.

Problem 2.4.7. Using the above subprocedure, show that any matrix can be reduced to reduced row echelon form.

Solution. We inductively show that, for a matrix T with m rows and n columns, for a given positive integer i, by performing row operations, we can reduce our matrix to a form such that

1. Among the first i rows, for each row with a nonzero entry, their leftmost nonzero entry is a 1, which are pivots.
2. These i pivots are the only nonzero entry in their column.
3. The pivot of the kth row is left of the pivot of the jth row if $k < j$.
4. If the pivot of the last nonzero row among the first i rows is in column k, then the leftmost nonzero entry of any other row is right of column k.
5. Any rows with all zeroes at at the bottom of the matrix.

Our base case is $i = 1$. If there is no row with nonzero entries, then the matrix is the zero matrix, and there is nothing we need to do. Otherwise, we first swap rows such that all of the rows that are entirely zero are at the bottom; in particular, by assumption, the first row is nonzero. Consider the first column k with a nonzero entry (k being the smallest index); suppose that row l has a nonzero entry in column k. We can then swap this with row 1, so row 1 has a nonzero entry in column k. Using the subprocedure from the previous problem, we can make it such that row 1 has a 1 in column k, and all the other rows have a zero in column k. Furthermore, in row 1, there are only zeros to the left of this 1, and by construction there are no nonzero entries to the left of column k. Swapping the rows around to get all of the zero rows at the bottom gives us a matrix satisfying the base case.

Suppose we have shown this for $i-1$, and $i \leq m$. If all of the rows outside of these rows are zero, then this clearly satisfies the conditions for i. Otherwise, we know that row i is not all zeros by construction, and furthermore rows $1, 2, \ldots, i-1$ all have pivots (and are nonzero). Among the nonzero rows that are below the $(i-1)$st row, consider the leftmost nonzero entries of each row. Let row l be the row with the leftmost such entry, and say it is in column k. Swap row l with row i. Next, perform the subprocedure such that row i has a 1 in column k, and and no other row has a nonzero entry in column k. Finally, again swap rows such that all rows with all zeroes are at the bottom of the matrix.

By assumption of which case we are in, rows 1 to $i-1$ had pivots, and these pivots are not changed by the subprocedure, since in the columns of those pivots row i has a zero, by our inductive hypothesis. Row i, then, also has a pivot by our procedure, and furthermore we have made it such that the pivot is the only nonzero entry in the column. For the other $i-1$ pivots, we have not changed this, since in the columns of those pivots, row i had a zero, and so the column does not change.

For the third condition, suppose that the pivot of row $i-1$ is in column k'. Then, by the inductive hypothesis fourth condition, $k > k'$, meaning that the pivot of row i is right of every pivot in rows $1, 2, \ldots, i-1$. For the fourth condition, by assumption, in all the other rows they only have zeros in columns left of column k, and from our subprocedure they have only zeros in column k. Finally, the fifth condition holds by our swapping at the end (which does not affect rows 1 to i, and therefore still preserves the first four conditions).

This finishes the inductive step, and thus we have shown that the condition above holds for each $i \leq m$. In particular, if $i = m$, we obtain the reduced row echelon form.

Problem 2.4.8. Show that the number of pivots of T is equal to the rank of T, and the number of columns without pivots is equal to the nullity of T, without using the rank-nullity theorem. (One can prove the rank-nullity theorem by analyzing the reduced row echelon form of a matrix).

Solution. First, notice that the rank of $T : V \to W$ is the dimension of the image, and furthermore the rank does not change under row operations, since these row operations simply correspond to changing bases (and not the underlying vectors underneath). So it suffices to compute the rank and nullity corresponding to the reduced row echelon form for the matrix.

Suppose that T has k pivots, and these are in columns $i_1, i_2, \ldots, i_k$. Suppose that this matrix corresponds to bases $\{v_1, v_2, \ldots, v_n\}$ for V and $\{w_1, w_2, \ldots, w_n\}$ for W. Then, notice that the image contains w_j for $j = 1, 2, \ldots, k$, since by definition v_{i_j} gets sent to w_j. Furthermore, notice that since T has k pivots, and each nonzero row has a pivot, rows $k+1, k+2, \ldots, m$ are all zeros, meaning that the image only contains vectors that are linear combinations of $w_1, w_2, \ldots, w_k$. In particular, the image has dimension k ($w_1, w_2, \ldots, w_k$ all lie in the image, and they linearly independent and span the image).

We now consider the kernel. Suppose that $v = \sum\limits_{i=1}^{n} x_i v_i$ gets sent to zero. Then, by considering the matrix multiplication, we note that if the matrix for T has entries a_{ij}, then the jth coordinate of Tv is $\sum\limits_{i=1}^{n} a_{ji}x_i$. Notice for $j > k$ that this is automatically zero. Meanwhile, for $j \leq k$, this equation is of the form $x_{i_j} + \sum\limits_{i=i_j+1}^{n} a_{ji}x_i = 0$, or that $x_{i_j} = -\sum\limits_{i=i_j+1}^{n} a_{ji}x_i$. Let $S = \{i_1, i_2, \ldots, i_k\}$.

In particular, this means that $v = \sum\limits_{i=1}^{n} x_i v_i$ is therefore equal to

$$v = -\sum_{j=1}^{k} \sum_{i=i_j+1}^{n} a_{ji}x_i v_{i_j} + \sum_{\substack{1 \leq i \leq n \\ i \notin S}} x_i v_i.$$

Notice that $a_{ji_l} = 0$ for $l \neq j$, meaning that we can rewrite this as $\sum\limits_{\substack{1 \leq i \leq n \\ i \notin S}} \left(x_i v_i - \sum\limits_{j=1}^{k} x_i a_{ji} v_{i_j}) \right)$. In other words, every element in the kernel has to be a linear combination of the $n-k$ vectors $v_i' = v_i - \sum\limits_{j=1}^{k} a_{ji}v_{i_j}$, where $i \notin S$. Furthermore, we can manually check here that these vectors all lie in the kernel: the jth coordinate of $T(v_i - \sum\limits_{j=1}^{k} a_{ji}v_{i_j})$ is equal to $-a_{ji} + a_{ji} = 0$ if $j < k$, and 0 otherwise.

Finally, these vectors are linearly independent, since if $\sum\limits_{i \notin S} a_i v_i' = 0$, then notice that, expanding this in terms of the v_j, the only v_i' with a nonzero v_j term, for $j \notin S$, is v_j', and the coefficient is 1, meaning that we need $a_j = 0$ for each $j \notin S$. Thus, these vectors are linearly independent, and therefore they form a basis for the kernel. In particular, the nullity of T is $n - k$, which is also the number of columns without a pivot.

3 Modules

Problem 3.0.1. Show that for any ring R and ideal I of R, I is an $R-$module under the addition and multiplication operations of the ring R.

Solution. We check the properties of a module. We note for an ideal I that property 1 follows from the definition of an ideal, and properties 2, 3, 6, 7 follow from the properties for a ring R. For property 4, we notice that in a ring R, $0 \cdot r = 0$, so in particular $0 \in I$ (by multiplying an element of I by 0). Property 5 follows from the fact that, if (-1) is the additive inverse of 1, then $(-1) \cdot r$ is the additive inverse of r, since $r + (-1) \cdot r = 1 \cdot r + (-1) \cdot r = (1 + (-1)) \cdot r = 0 \cdot r = 0$. As all of the properties have been checked, it follows that I is a module over R.

Problem 3.0.2. Show that any finitely generated free module is isomorphic to R^n for some $n \in \mathbb{N}$.

Solution. Suppose that M is a finitely generated free module over a ring R. We know that, as it is free, it has a free basis, and furthermore that this free basis is finite from finite generation (pick a finite set of generators, and write them as linear combinations of the free basis: only finitely many elements in the basis are used, so every element can be written as a linear combination of elements among these finitely many. In particular, there can be no other basis elements by linear independence).

Now, suppose this free basis is $m_1, m_2, \ldots, m_n$. Consider the map $\phi : M \to R^n$ defined by sending $\sum\limits_{i=1}^{n} r_i m_i$ to $(r_1, r_2, \ldots, r_n) \in R^n$. First, this is well-defined, since from the definition of free basis, every element has such a form, and furthermore this form is unique by linear independence (the difference of two possible forms is a linear combination of zero, so all of the coefficients have to be zero by linear independence). We can easily verify that this map satisfies the properties for a module homomorphism, since

$$\phi(\sum_{i=1}^{n} r_i m_i + \sum_{i=1}^{n} r_i' m_i) = \phi(\sum_{i=1}^{n} (r_i + r_i') m_i) = (r_1 + r_1', r_2 + r_2', \ldots, r_n + r_n') = \phi(\sum_{i=1}^{n} r_i m_i) + \phi(\sum_{i=1}^{n} r_i' m_i)$$

and for $r \in R$, we have

$$\phi(r \cdot \sum_{i=1}^{n} r_i m_i) = \phi(\sum_{i=1}^{n} r r_i m_i) = (rr_1, rr_2, \ldots, rr_n) = r\phi(\sum_{i=1}^{n} r_i m_i).$$

We can check that this is surjective, since any element $(r_1, r_2, \ldots, r_n) \in R^n$ arises from the element $\sum\limits_{i=1}^{n} r_i m_i$. For injectivity, if an element maps to $(0, 0, \ldots, 0) \in R^n$, then notice that it maps from $\sum\limits_{i=1}^{n} 0 m_i = 0$. Therefore, this map ϕ is a homomorphism which is injective and surjective, and therefore an isomorphism.

Problem 3.0.3. Given a submodule N of an $R-$module M, consider the map $\kappa_{M,N} : M \to M/N$ that sends m to $m + N$. Show that this map is a surjective homomorphism. What is the kernel of $\kappa_{M,N}$?

Solution. We first check that this is a homomorphism of modules. Suppose that $m, m' \in M$. Then, $\phi(m + m') = (m + m') + N = \{m + m' + n | n \in N\}$. But by definition, $(m + m') + N = (m + N) + (m' + N) = \phi(m) + \phi(m')$. Furthermore, if $r \in R$, then $\phi(rm) = rm + N = r(m + N) = r\phi(m)$, so this is a homomorphism.

To check that it is surjective, we know that every element in M/N is of the form $m + N$ for some $m \in M$; but then this means that $\phi(m) = m + N$, so this is surjective. For the kernel, we claim that the kernel is N. Indeed, if $m + N = N$, then in particular we require $m + 0 \in N$. But then $m \in N$. This in turn implies that for all $n \in N$, we have $m + n \in N$, so $m + N \subset N$. Furthermore, for all $n \in N$, $n - m \in N$, so $m + (n - m) \in m + N$, and thus $N \subset m + N$, or that $N = m + N$, as desired.

Problem 3.0.4. Given a homomorphism between $R-$modules M, N :

1. Show that there exists a homomorphism $\bar{\phi} : M/\ker\phi \to N$ such that

$$\bar{\phi}(\kappa_{M,\ker\phi}(m)) = \phi(m)$$

   for all $m \in M$.

2. Show that $M/\ker\phi$ is isomorphic to imϕ.

Solution. **1.** Consider the map $\bar{\phi}$ defined by sending $m + \ker\phi$ to $\phi(m)$. This satisfies the desired property that

$$\bar{\phi}(\kappa_{M,\ker\phi}(m)) = \phi(m),$$

since $\bar{\phi}(\kappa_{M,\ker\phi}(m)) = \bar{\phi}(m + \ker\phi) = \phi(m)$. We need to check that this is a well-defined homomorphism.

Suppose that $m + \ker\phi = m' + \ker\phi$. Then, in particular, as $0 \in \ker\phi$, we have that $m \in m' + \ker\phi$, so there exists some $k \in \ker\phi$ so that $m = m' + k$. But then $\phi(m) = \phi(m' + k) = \phi(m') + \phi(k) = \phi(m')$, and so $\bar{\phi}(m + \ker\phi) = \bar{\phi}m' + \ker\phi)$, so this is well-defined.

To check that this is a homomorphism, we verify that

$$\bar{\phi}((m + \ker\phi) + (m' + \ker\phi)) = \bar{\phi}((m + m') + \ker\phi) = \phi(m + m')$$

$$= \phi(m) + \phi(m') = \bar{\phi}(m + \ker\phi) + \bar{\phi}(m' + \ker\phi),$$

and

$$\bar{\phi}(r(m + \ker\phi)) = \bar{\phi}(rm + \ker\phi) = \phi(rm) = r\phi(m) = r\bar{\phi}(m + \ker\phi).$$

This shows that we have the desired homomorphism.

2. To show that the modules are isomorphic, we claim that $\bar{\phi}$ is our desired isomorphism. We note that our above homomorphism is a well-defined homomorphism from $M/\ker\phi$ to $\mathrm{im}\phi$. We need to check that it is injective and surjective.

For surjectivity, by definition, for all $n \in \mathrm{im}\phi$, $n = \phi(m)$ for some $m \in M$. But then we have that $\bar{\phi}(m + \ker\phi) = \phi(n)$. We just need to show that this map is injective.

Suppose that $m + \ker\phi$ gets sent by $\bar{\phi}$ to zero. Then, by definition, $\phi(m) = 0$, meaning that $m \in \ker\phi$. But then $m + \ker\phi = \ker\phi$, which is enough to prove injectivity.

4 The PID Structure Theorem

4.1 Noetherian Rings and Modules

Problem 4.1.1. Prove the following.

1. Show that every ideal can be generated by a finite set of elements in a Noetherian ring.
2. For any chain $I_1 \subset I_2 \subset I_3 \subset \dots$ of ideals, show that $\bigcup_{i=1}^{\infty} I_i$ is an ideal.
3. Suppose that ring R is such that every ideal can be generated by a finite set of elements. Prove that R is Noetherian. As a hint, consider the previous part, and consider a finite set that generates the union of ideals in the chain. Where do each of the elements in this finite set live?
4. Conclude that every PID is Noetherian.

Solution. **1.** Suppose for the sake of contradiction that some ideal I cannot be generated by a finite set of elements. We show that R is not Noetherian.

To do this, we define the ideals I_j inductively, as follows. First, let i_0 be an element in I, and let $I_0 = \langle i_0 \rangle$. Now, given the elements $i_0, i_1, \dots, i_n$, and ideals $I_j = \langle i_0, i_1, \dots, i_j \rangle$ for $j = 0, 1, \dots, n$, we know that $I_n \neq I$ since I cannot be generated by a finite set of elements. Thus, there exists an element $i_{n+1} \in I - I_n$. From here, let $I_{n+1} = \langle i_0, i_1, \ldots, i_{n+1} \rangle$.

Notice that this inductive construction yields a chain of ideals $I_0 \subset I_1 \subset I_2 \subset \ldots$. However, notice that by construction, $I_j = \langle i_0, i_1, \ldots, i_j \rangle$ is not equal to I_{j+1}, since I_{j+1} contains i_{j+1}, but I_j does not (by definition). This is a contradiction of the fact that R is Noetherian.

Therefore, it follows that every ideal I can be generated by a finite set of elements, which is what we wanted to show.

2. Given a chain of ideals $I_1 \subset I_2 \subset \ldots$, we show that $\bigcup_{i=1}^{\infty} I_i$ is an ideal. First, suppose that r_1, r_2 lie in this set. It follows that there exist $i_1, i_2 \in \{1, 2, \ldots\}$ such that $r_j \in I_{i_j}$ for $j = 1, 2$. But then if i is the maximum of i_1, i_2, from our chain we have that $r_1, r_2 \in I_i$, so $r_1 + r_2 \in I_i \subset \bigcup_{i=1}^{\infty} I_i$.

Furthermore, for all $r \in \bigcup_{i=1}^{\infty} I_i$ and $s \in R$, there exists an $i \in \{1, 2, \ldots\}$ such that $r \in I_i$. Then, $sr \in I_i \subset \bigcup_{i=1}^{\infty} I_i$, and we are done.

3. Suppose that every ideal can be generated by a finite set of elements. Then, consider any ascending chain of ideals in R, say $I_1 \subset I_2 \subset I_3 \subset \ldots$. Let I be their union; from the previous part, we know this is an ideal, so it can be generated by a finite set of elements $\{r_1, r_2, \ldots, r_k\}$. By definition, each r_j lies in a union of the I_i, so there exists some i_j for each j so that $r_j \in I_{i_j}$. Let i be the maximum of the i_j values, so $\{r_1, r_2, \ldots, r_k\} \subset I_i$. Then, notice that for all $i' \geq i$, we have that $I_i \subset I_{i'} \subset I$. But $I = \langle r_1, r_2, \ldots, r_k \rangle \subset I_i$, since for all $s_1, s_2, \ldots, r_k \in R$, by I_i being an ideal, $\sum_{j=1}^{k} s_j r_j \in I_i$. Therefore, $I_i \subset I_{i'} \subset I \subset I_i$, meaning that all of these subsets are equalities.

In particular, $I_{i'} = I_i$ for all $i' \geq i$. As this holds for all ascending chains of ideals, it follows that R is a Noetherian ring.

4. By definition, every ideal in a PID can be generated by a single element, and so thus it can be generated by a finite number of elements. But by the previous part, it must be Noetherian, which is what we wanted to show.

Problem 4.1.2. Prove the following statements.

1. Show that if M is a Noetherian module, then every submodule of M is finitely generated. (Hint: can you think of what a submodule generated by elements would be?)
2. Suppose that every submodule of M is finitely generated. Prove that M is Noetherian.

Solution. **1.** Suppose that N is a submodule of M, and suppose for the sake of contradiction that N was not finitely generated. Then, we can construct a sequence of elements $n_1, n_2, \ldots$ that lie in N such that for each i, the module generated by $n_1, n_2, \ldots, n_k$ does not contain n_{k+1}. Let N_k be the submodule of M consisting of all linear combination of elements $\{n_1, n_2, \ldots, n_k\}$. Then, by construction, $N_k \subset N_{k+1}$ for each $k \in \mathbb{N}$, but $n_{k+1} \notin N_k$, meaning that $N_k \neq N_{k+1}$.

We therefore have an increasing chain of submodules in M, given by $N_1 \subset N_2 \subset \ldots$, which does not stabilize by construction. This is a contradiction. Hence, N must be finitely generated, which is what we wanted to show.

2. Suppose that M is a module where every submodule of M is Noetherian, and suppose $N_1 \subset N_2 \subset \ldots$ is an increasing chain of submodules of M. Let N be the union of these modules. Notice that it is also a submodule of M; indeed, if $n_1, n_2 \in N$, there is some i_1, i_2 so $n_j \in N_{i_j}$

for $j = 1, 2$. But then letting i be the maximum of the i_j, we have that n_1, n_2 lie in N_i, and so $n_1 + n_2, rn_1 \in N_i \subset N$ for all $r \in R$.

Since this is a submodule of M, it is finitely generated, say by $n_1, n_2, \ldots, n_k$. Furthermore, each of these generators has to lie in some module in the increasing chain. Therefore, there is some i such that $n_1, n_2, \ldots, n_k \in N_i$. However, this means that every linear combination of the n_j lie in N_i too. But these linear combinations are precisely N, by definition, meaning that $N \subset N_i$. In particular, for $l \geq i$, we have $N_i \subset N_l \subset N \subset N_i$, or that $N_l = N_i$ for $l \geq i$, which is what we wanted to show.

Problem 4.1.3. Given an $R-$module M and a submodule N of M, show that if M is Noetherian, then $N, M/N$ are both Noetherian. (Hint: consider the map $\kappa_{M,N}$, and use the previous problem).

Solution. Suppose that M is Noetherian. By the previous problem, this means that every submodule of M is finitely generated. In particular, this means that every submodule of N is also finitely generated, meaning that N is Noetherian. Furthermore, suppose that L' is a submodule of M/N. Consider the set $\kappa_{M,N}^{-1}(L')$, the set of elements which get sent to L' under the $\kappa_{M,N}$ map. Then, observe that, since $\kappa_{M,N}$ is a homomorphism, and furthermore L' is a submodule of M/N, that this is also a submodule: for instance, $l_1, l_2 \in \kappa_{M,N}^{-1}(L')$ means that $\kappa_{M,N}(l_1), \kappa_{M,N}(l_2) \in L'$, or that $\kappa_{M,N}(l_1) + \kappa_{M,N}(l_2) = \kappa_{M,N}(l_1 + l_2) \in L'$, or that $l_1 + l_2 \in \kappa_{M,N}^{-1}(L')$, and similarly with scalar multiplication.

Now, notice that $\kappa_{M,N}^{-1}(L')$ is a submodule of M, so is finitely generated by $m_1, m_2, \ldots, m_k$. Consider the elements $\kappa_{M,N}(m_1), \kappa_{M,N}(m_2), \ldots, \kappa_{M,N}(m_k)$; we claim that these are generators of L'. For each element $l' \in L'$, since $\kappa_{M,N}$ is surjective, there is some element $l \in M$ such that $\kappa_{M,N}(l) = l'$, and furthermore l lies in $\kappa_{M,N}^{-1}(L')$. Therefore, it is a linear combination $\sum\limits_{i=1}^{k} r_i m_i$ for $i = 1, 2, \ldots, k$.

But then we have that $l' = \sum\limits_{i=1}^{k} r_i \kappa_{M,N}(m_i)$. This proves that L' is finitely generated. As this holds for all submodules, we have that M/N is Noetherian.

Problem 4.1.4. Prove the following.

1. Given a module M and a submodule N of M suppose that $M_1 \subset M_2$ are submodules such that $M_1 \cap N = M_2 \cap N$ and $\kappa_{M,N}(M_1) = \kappa_{M,N}(M_2)$. Show that $M_1 = M_2$.

2. Using the above result, show that if N and M/N are Noetherian, then M is also Noetherian.

Solution. **1.** Suppose we are given two submodules of M, say M_1, M_2, such that $M_1 \subset M_2$. If $M_1 \cap N = M_2 \cap N$ and $\kappa_{M,N}(M_1) = \kappa_{M,N}(M_2)$, we claim that $M_1 = M_2$.

To prove this, suppose for the sake of contradiction that this was not the case, so there is some $m \in M_2$ so $m \notin M_1$. Now, $\kappa_{M,N}(m) \in \kappa_{M,N}(M_1)$, so there is some $m' \in M_1$ such that $\kappa_{M,N}(m) = \kappa_{M,N}(m')$. But this means that $\kappa_{M,N}(m - m') = 0$, or that $m - m' \in N$, since N is the kernel of $\kappa_{M,N}$.

However, $m' \in M_2, m \in M_2$ means that $m - m' \in M_2 \cap N = M_1 \cap N$, and therefore $m = m - m' + m' \in M_1$, contradiction. Therefore, we have that $M_1 = M_2$.

2. Suppose that $N, M/N$ are Noetherian modules. Consider any chain $M_1 \subset M_2 \subset \ldots$ of submodules of M. Consider the chains $M_1 \cap N \subset M_2 \cap N \subset \ldots$ and $\kappa_{M,N}(M_1) \subset \kappa_{M,N}(M_2) \subset \ldots$. Note that for a submodule M' of M that $M' \cap N$ and $\kappa_{M,N}(M')$ are submodules. For instance,

$m_1, m_2 \in M' \cap N$ means that $m_1, m_2 \in M'$ and N, and so therefore $m_1 + m_2 \in M'$ and N, so $m_1 + m_2 \in M' \cap N$. A similar argument holds for the other three properties we would need to check to get the two submodules.

By $N, M/N$ Noetherian, it follows that there is an i_1 such that $j \geq i_1$ implies that $M_j \cap N = M_{i_1} \cap N$, and an i_2 so $j \geq i_2$ implies that $\kappa_{M,N}(M_j) = \kappa_{M,N}(M_{i_2})$. Let i be the maximum of i_1, i_2, and $j \geq i$. Then, notice that $M_i \subset M_j$, and $M_i \cap N = M_{i_1} \cap N = M_j \cap N$, and $\kappa_{M,N}(M_j) = \kappa_{M,N}(M_{i_2}) = \kappa_{M,N}(M_i)$, so hence by our above claim we have that $M_j = M_i$. This holds for all chains of submodules of M, meaning that M is Noetherian, which is what we wanted to show.

Problem 4.1.5. Let R be a Noetherian ring, and let M be a module over R.

1. Show that if M is Noetherian, then M is finitely generated.

2. Show that R is an Noetherian $R-$module (hint: what are submodules of R?)

3. Show that R^n/R^m is isomorphic to R^{n-m}, for positive integers $n \geq m$, where we embed R^m into R^n as the elements where the last $m - n$ coordinates are zero.

4. Using the previous part, show that R^n is a Noetherian $R-$module for every positive integer n.

5. Show therefore that if M is finitely generated, then M is a Noetherian $R-$module.

Solution. **1.** We know that M is Noetherian if and only if all of the submodules are finitely generated. But then M in particular is also finitely generated.

2. It is enough to show that all submodules M of R are finitely generated over R. However, notice that M is an ideal, since M is closed under addition and multiplication by elements of R. Therefore, since R is Noetherian, every ideal is finitely generated, so in particular M is finitely generated as an ideal. Say its generators are $m_1, m_2, \ldots, m_k$. Notice, however, that every element $m \in M$ can be written as a linear combination of the elements $m_1, m_2, \ldots, m_k$, which in particular means that these k elements generate M as a module, so M is finitely generated as an $R-$module.

Thus, we have that R is a Noetherian $R-$module, which is what we wanted to show.

3. Let the free basis for R^n be $e_1, e_2, \ldots, e_n$, and R^m have free basis generated by the first m elements. Say that R^{n-m} has a free basis generated by $v_1, v_2, \ldots, v_{n-m}$. Consider the linear map $\phi : R^n \to R^{n-m}$ by $\phi(\sum\limits_{i=1}^{n} a_i e_i) = \sum\limits_{i=1}^{n-m} a_{m+i} v_i$. Notice that this is a well-defined homomorphism, and furthermore the kernel consists precisely of the elements where $a_{m+1}, a_{m+2}, \ldots, a_n$ are zero, which is just R^m. Furthermore, this is surjective, since any element of R^{n-m} is generated by $v_1, v_2, \ldots, v_{n-m}$.

Therefore, R^n/R^m is isomorphic to the image of ϕ, which is R^{n-m}. This is what we wanted to show.

4. We prove that R^n is a Noetherian $R-$module by induction on n. The base case, $n = 1$, is given by induction.

Now, suppose that this has been shown for $n \leq k$. Then, consider R^{k+1}; we know that R^k (the subspace, say, where the last coordinate is zero) is Noetherian by our inductive hypothesis. Meanwhile, by the previous part, we have that $R^{k+1}/R^k \simeq R$ is also Noetherian. Therefore, by the previous problem, it follows that R^{k+1} is Noetherian.

This finishes the inductive step, and therefore proves the claim.

5. Suppose that M is finitely generated, generated by $m_1, m_2, \dots, m_n$. Then, consider the map ϕ from R^n to M by $\sum_{i=1}^{n} a_i e_i$ to $\sum_{i=1}^{n} a_i m_i$. This can be quickly checked, in the same manner as the previous problems, to be a surjective homomorphism.

Therefore, it follows that $R^n/\ker\phi$ is isomorphic to M. However, notice that R^n is a Noetherian module by the previous part, and $\ker\phi$ is a submodule, meaning that $R^n/\ker\phi$ is a Noetherian $R-$module. But then it follows that M is also a Noetherian $R-$module (using this isomorphism, if N is a submodule of M, then $\overline{\phi}^{-1}(N)$ is a submodule of $R^n/\ker\phi$ and is finitely generated, so N is finitely generated too by the images under ϕ of these finite generators). This is what we wanted to show.

4.2 Smith Normal Form

Problem 4.2.1. Suppose that we multiply A on the right by an $m \times m$ invertible matrix S to get the matrix A'. Show that $N(A') = N(A)$. (Hint: show that an element on the left-hand side lies on the right-hand side. Use invertibility).

Solution. Suppose that $A' = AS$, where S is an invertible $m \times m$ matrix. We show that $N(A') \subset N(A)$. This is enough, since applying this argument with the equality $A = A'S^{-1}$ will give us the other inclusion, which will finish the proof of the above statement. Suppose the (i,j)th entry of A' is $a'_{i,j}$, the (i,j)th entry of A is $a_{i,j}$, and the (i,j)th entry is $s_{i,j}$.

Suppose we are given some vector $v \in N(A')$. Then, we know that v can be written as $\sum_{j=1}^{m} b_j \sum_{i=1}^{n} a'_{i,j} e_i$, where the b_j lie in R. However, from the definition of matrix multiplication, we have that $a'_{i,j} = \sum_{k=1}^{m} a_{i,k} s_{k,j}$. Substituting this into the above expression for v yields the following expression:

$$\sum_{j=1}^{m} b_j \sum_{i=1}^{n} \sum_{k=1}^{m} a_{i,k} s_{k,j} e_i.$$

But this in turn is equal to

$$\sum_{k=1}^{m} \sum_{j=1}^{m} b_j s_{k,j} \sum_{i=1}^{n} a_{i,k} e_i,$$

which lies in the submodule generated by $\sum_{i=1}^{n} a_{i,k} e_i$ for $k = 1, 2, \dots, m$. Therefore, $v \in N(A)$, and we are done.

Problem 4.2.2. Suppose that we multiply A on the left by an $n \times n$ invertible matrix S to get the matrix A'. Show that $R^n/N(A)$ and $R^n/N(A')$ are isomorphic. (Hint: what is an isomorphism between the two modules?)

Solution. Suppose the (i,j)th entry of A' is $a'_{i,j}$, the (i,j)th entry of A is $a_{i,j}$, and the (i,j)th entry is $s_{i,j}$. We identify the element $v = \sum_{i=1}^{n} v_i e_i$ with column vector $\begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix}$, which by the same logic as in the case of vector spaces is a module isomorphism.

Consider the module homomorphism from R^n onto $R^n/N(A')$ given by first sending the element $v = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix}$ to Sv, before applying the quotient map $\kappa_{R^n,N(A')}$. Note that this map is surjective. To see this, for each $\overline{w} \in R^n/N(A')$, since $\kappa_{R^n,N(A')}$ is surjective, there is some $w \in R^n$ so $\overline{w} = \kappa_{R^n,N(A')}(w)$. From here, note that S is invertible, and so $S^{-1}w$ is well-defined, and satisfies that $\kappa_{R^n,N(A')}(S(S^{-1}w)) = \kappa_{R^n,N(A')}(w) = \overline{w}$.

From here, we consider the kernel of this map. Note that $\kappa_{R^n,N(A')}(Sv) = 0$ holds if and only if $Sv \in N(A')$, from the computation of the kernel of the κ map in Problem 3.0.3. However, by definition, $N(A')$ consists of the vectors of the form $\sum\limits_{j=1}^{m} b_j \sum\limits_{i=1}^{n} a'_{i,j}e_i$; in other words, the vectors of the form $A'b$ for some vector $b = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}$. However, $Sv = A'b$ for some b holds if and only if, multiplying by the inverse of S, $v = S^{-1}A'b = Ab$ for some vector b, which in turn holds if and only if $v \in N(A)$.

Therefore, the kernel of this combined map is $N(A)$, and as it is surjective, it follows that this map is an isomorphism from $R^n/N(A)$ to $R^n/N(A')$, which is what we wanted to show.

Problem 4.2.3. As a first step, we want to reduce a column down to having a single nonzero entry in it, using row operations.

1. Suppose we have a matrix $\begin{pmatrix} r_1 \\ r_2 \end{pmatrix}$, where $r_1, r_2 \neq 0$. Show there exists a 2×2 invertible matrix S such that $S\begin{pmatrix} r_1 \\ r_2 \end{pmatrix} = \begin{pmatrix} r \\ 0 \end{pmatrix}$ for some element $r \in R$. How is r related to $r_1, r_2, \ldots, r_n$?

2. Suppose now we have a matrix $\begin{pmatrix} r_1 \\ r_2 \\ \vdots \\ r_n \end{pmatrix}$, where not all the r_i are zero. Show that there exists an $n \times n$ invertible matrix S such that $S\begin{pmatrix} r_1 \\ r_2 \\ \vdots \\ r_n \end{pmatrix} = \begin{pmatrix} r \\ 0 \\ \vdots \\ 0 \end{pmatrix}$, for some $r \in R$. How is r related to $r_1, r_2, \ldots, r_n$?

3. Given an $n \times m$ matrix A, for $n, m \geq 0$, show that there exists an $n \times n$ invertible matrix S and an $m \times m$ invertible matrix T so that SAT has no nonzero entries in the first row or first column, except for the $(1,1)$ entry. (Hint: suppose that the $(1,1)$ entry doesn't divide every entry in the first row or first column. Apply part b). Repeat, and utilize the fact that R is a PID, ergo Noetherian.)

Solution. **1.** Since R is a PID, the ideal $\langle r_1, r_2\rangle$ is equal to the ideal $\langle r\rangle$. Then, it follows that there exist elements s_1, s_2 such that $s_1r_1 + s_2r_2 = r$. We claim that the matrix $\begin{pmatrix} s_1 & s_2 \\ r_2/r & -r_1/r \end{pmatrix}$ is an invertible matrix in R. To see this, first note that, since $\langle r_1, r_2\rangle = \langle r\rangle$, it follows that r_1, r_2 are divisible by r, so this matrix has entries in R.

Furthermore, note that this matrix has inverse $\begin{pmatrix} r_1/r & s_2 \\ r_2/r & -s_1 \end{pmatrix}$. Finally, notice that $\begin{pmatrix} s_1 & s_2 \\ r_2/r & -r_1/r \end{pmatrix}\begin{pmatrix} r_1 \\ r_2 \end{pmatrix} = \begin{pmatrix} r \\ 0 \end{pmatrix}$, which is what we wanted to show.

2. We proceed by induction on n, showing that r is such that

$$\langle r\rangle = \langle r_1, r_2, \ldots, r_n\rangle.$$

The base case is $n = 2$, where we have shown that this holds by the previous part. Suppose that we have shown this now for $n \leq k - 1$, and consider the vector $\begin{pmatrix} r_1 \\ r_2 \\ \vdots \\ r_n \end{pmatrix}$. If r_n is zero, then the inductive hypothesis lets us conclude that there exists some invertible $n - 1 \times n - 1$ matrix S such that $S\begin{pmatrix} r_1 \\ r_2 \\ \vdots \\ r_{n-1} \end{pmatrix} = \begin{pmatrix} r \\ 0 \\ \vdots \\ 0 \end{pmatrix}$, and

$$r = \langle r_1, r_2, \ldots, r_{n-1}\rangle = \langle r_1, r_2, \ldots, r_{n-1}, r_n\rangle.$$

If the entries of S are s_{ij}, whose inverse T has entries t_{ij} then the matrix

$$S' = \begin{pmatrix} s_{11} & s_{12} & \cdots & s_{1(n-1)} & 0 \\ s_{21} & s_{22} & \cdots & s_{2(n-1)} & 0 \\ \vdots & \vdots & \ddots & \vdots & 0 \\ s_{(n-1)1} & s_{(n-1)2} & \cdots & s_{(n-1)(n-1)} & 0 \\ 0 & 0 & \cdots & 0 & 1 \end{pmatrix},$$

which we can denote as $\begin{pmatrix} S & 0 \\ 0 & 1 \end{pmatrix}$, is also invertible, with inverse

$$\begin{pmatrix} T & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} t_{11} & t_{12} & \cdots & t_{1(n-1)} & 0 \\ t_{21} & t_{22} & \cdots & t_{2(n-1)} & 0 \\ \vdots & \vdots & \ddots & \vdots & 0 \\ t_{(n-1)1} & t_{(n-1)2} & \cdots & t_{(n-1)(n-1)} & 0 \\ 0 & 0 & \cdots & 0 & 1 \end{pmatrix}$$

and $S'\begin{pmatrix} r_1 \\ r_2 \\ \vdots \\ r_{n-1} \\ r_n \end{pmatrix} = \begin{pmatrix} r \\ 0 \\ \vdots \\ 0 \\ 0 \end{pmatrix}$.

Otherwise, note that there exists an invertible 2×2 matrix S, with entries s_{ij}, such that $S\begin{pmatrix} r_{n-1} \\ r_n \end{pmatrix} = \begin{pmatrix} r \\ 0 \end{pmatrix}$, where $\langle r \rangle = \langle r_{n-1}, r_n \rangle$. Furthermore, considering the $n \times n$ matrix

$$\begin{pmatrix} 1 & 0 & \cdots & 0 & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & 0 & \\ 0 & 0 & \cdots & 1 & 0 & 0 \\ 0 & \cdots & 0 & 0 & s_{11} & s_{12} \\ 0 & 0 & \cdots & 0 & s_{21} & s_{22} \end{pmatrix},$$

note this sends $\begin{pmatrix} r_1 \\ r_2 \\ r_3 \\ \vdots \\ r_{n-1} \\ r_n \end{pmatrix}$ to $\begin{pmatrix} r_1 \\ r_2 \\ r_3 \\ \vdots \\ r \\ 0 \end{pmatrix}$, where by the inductive hypothesis and our argument above we can apply another invertible $n \times n$ matrix such that we get the vector $\begin{pmatrix} r' \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$, where $\langle r' \rangle = \langle r_1, r_2, \ldots, r_{n-2}, r \rangle$, and $\langle r \rangle = \langle r_{n-1}, r_n \rangle$. Note that $\langle r_1, r_2, \ldots, r_n \rangle \subset \langle r_1, r_2, \ldots, r_{n-2}, r \rangle$ (since r_{n-1}, r_{n-2} lie in the latter), which in turn lies in $\langle r' \rangle$, and furthermore $\langle r' \rangle = \langle r_1, r_2, \ldots, r_{n-2}, r \rangle \subset \langle r_1, r_2, \ldots, r_{n-2}, r_{n-1}, r_{n-2} \rangle$ (since r lies in the latter), showing that these two ideals are equal. This finishes the inductive step and hence proves the claim.

3. Suppose we are given an $n \times m$ matrix A. Suppose that either the first column or first row are not entirely zero (if they are entirely zero, we are done). Without loss of generality, suppose that the first column is not entirely zero. By part 2, there exists an invertible matrix S such that SA only has one nonzero entry in the first column, and it is in the $(1, 1)$ entry. Say this $(1, 1)$ entry is a_1.

Suppose that the first row has more than one nonzero entry, and that some nonzero entry is not divisible by a_1. Then, by the same argument as in parts 1 and 2, there exists some $m \times m$ matrix T such that $(SA)T$ has no nonzero entry in the first row, other than the $(1, 1)$ entry. Notice that from part 2 that the new $(1, 1)$ entry a_2 divides a_1, since $\langle a_2 \rangle$ contains a_1. Furthermore, since some entry in row 1 is not divisible by a_1, we have that $\langle a_2 \rangle$ is strictly larger than $\langle a_1 \rangle$.

If the first column has a nonzero entry not divisible by a_2, we can repeat, finding an invertible matrix S' such that $S'SAT$ has no nonzero entries in the first column other than the $(1, 1)$ entry, which is a_3, which divides a_2. We then repeat this, alternating between rows and columns.

Suppose this procedure does not terminate. We then get a strictly increasing sequence of ideals, which is a contradiction of the fact that R is a PID, ergo a Noetherian ring. Thus, at some point, we arrive at a point where every entry in the first row and column is divisible by the $(1, 1)$ entry. The invertible row operations of adding some multiple of the first row/column to each other row/column, respectively, then give us our desired form for the matrix.

Problem 4.2.4. Show that, given an $n \times m$ matrix A, there exists an $n \times n$ invertible matrix S

and an $m \times m$ invertible matrix T such that the nonzero entries of SAT lie on the diagonal (so $a_{i,j} \neq 0$ implies that $i = j$).

Solution. We induct on the minimum of m, n. The base case, when $\min\{m, n\} = 1$, is given to us by Problem 53. Suppose we have shown this for all matrices A that are $m \times n$ where $\min\{m, n\} \leq k$. Suppose that $\min\{m, n\} = k + 1$.

Then, by the previous problem, there exist invertible matrices S, T such that SAT has no nonzero entries in the first row or first column, other than the $(1, 1)$ entry. In particular, we know that SAT has the following form:

$$\begin{pmatrix} a'_{1,1} & 0 & 0 & \dots & 0 \\ 0 & a'_{2,2} & a'_{2,3} & \dots & a'_{2,m} \\ 0 & a'_{3,2} & a'_{3,3} & \dots & a'_{3,m} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & a'_{n,2} & a'_{n,3} & \dots & a'_{n,m} \end{pmatrix}.$$

Now, we know that there exist matrices S', T', by the inductive hypothesis, such that if A' is the matrix

$$\begin{pmatrix} a'_{2,2} & a'_{2,3} & \dots & a'_{2,m} \\ a'_{3,2} & a'_{3,3} & \dots & a'_{3,m} \\ \vdots & \vdots & \ddots & \vdots \\ a'_{n,2} & a'_{n,3} & \dots & a'_{n,m} \end{pmatrix}.$$

So then the matrices $S'' = \begin{pmatrix} 1 & 0 \\ 0 & S' \end{pmatrix}$ and $T'' = \begin{pmatrix} 1 & 0 \\ 0 & T' \end{pmatrix}$ are such that $S''SATT''$ has its nonzero entries only along the diagonal, and we know that $S''S$ and TT'' are invertible matrices. This finishes the inductive step and hence proves the claim.

Problem 4.2.5. Show that we can further reduce the matrix so that the nonzero entries are $a_{1,1}, a_{2,2}, \dots, a_{k,k}$ for some positive integer $1 \leq k \leq m$, and so that $a_{i,i}$ divides $a_{i-1,i-1}$ for $i = 2, 3, \dots, k$. This is known as **Smith normal form.**

Solution. By the previous problem, along with swapping rows and columns appropriately, we can reduce our matrix to the form such that the only nonzero entries are on the diagonal, and these are $a_{1,1}, a_{2,2}, \dots, a_{k,k}$.

Given i and j, suppose that $\langle a_{i,i}, a_{j,j} \rangle = \langle d \rangle$, and say s_1, s_2 are such that $s_1 a_{i,i} + s_2 a_{j,j} = d$. Then, consider the following operations.

1. Swap rows i and j.
2. Add s_1 times row j to row i.
3. Add s_2 times column j to column i.
4. Subtract $a_{i,i}/d$ times row i from row j.
5. Subtract $a_{j,j}/d$ times column i from column j.

Notice that this only affects the (i,j)th, (j,i)th, (i,i)th, and (j,j)th entries, since all the other entries in either row i or column j are zero and remain zero at each step of the process (as either the swapping nor the addition operations change this fact, since by assumption the only nonzero entries at the beginning are along the diagonal). We can then just analyze what this does to these four entries, giving us the following table. Now, run the following procedure: apply this above subprocedure first with $i = k$, and $j = k-1, k-2, \ldots, 1$. Notice that after each one of these steps, we make the (k,k)th entry a divisor of the (j,j)th entry and the previous (k,k)th entry. But by induction, we can see that this new entry will be divisible by the (l,l)th entry for $l = j+1, j+2, \ldots k-1$. Hence, after running all of these, we see that the (k,k)th entry divides the (j,j)th entry for $j < k$. Say this entry is $a'_{k,k}$.

	(i,i)th entry	(i,j)th entry	(j,i)th entry	(j,j)th entry
After Step 1	0	$a_{j,j}$	$a_{i,i}$	0
After Step 2	$s_1 a_{i,i}$	$a_{j,j}$	$a_{i,i}$	0
After Step 3	d	$a_{j,j}$	$a_{i,i}$	0
After Step 4	d	$a_{j,j}$	0	$-a_{i,i}/d$
After Step 5	d	0	0	$-a_{i,i}/d$

Run this procedure now with $i = k-1$, and $j = k-2, k-3, \ldots 1$. Notice that at each step that the new $(k-1,k-1)$st entry of the matrix will be divisible by the (k,k)th entry, since from our above procedure this entry lies in the ideal generated by the old diagonal entries (which, in particular, by our above observations, lies in $\langle a'_{k,k} \rangle$), and furthermore by the same argument as the above the $(k-1,k-1)$st entry, after running the procedure on $k-2, k-3, \ldots, j$, divides the new (l,l)th entry for $l = j, j+1, \ldots, k-2$. So after running all of these, if we end up with the entry $a'_{k-1,k-1}$ in the $(k-1,k-1)$ entry, then $a'_{k,k}$ divides $a'_{k-1,k-1}$, and this entry divides all of the other entries.

Repeating this for $i = k-2, k-3, \ldots, 2$ eventually yields our desired form for the matrix.

4.3 Proof of the Theorem

Problem 4.3.1. Show that there exists a surjective map ϕ from a free module F of finite rank to M.

Solution. Since M is finitely generated, there exists a finite set of generators $m_1, m_2, \ldots, m_k$. Then, consider the map $\phi : R^k \to M$ by sending $(r_1, r_2, \ldots, r_k)$ to $\sum\limits_{i=1}^{k} r_i m_i$. Note that this map can be checked to be linear.

Furthermore, since $m_1, m_2, \ldots, m_k$ are a generating set for M, each element of M can be written as $\sum\limits_{i=1}^{k} a_i m_i$ for some $a_i \in R$. But then $(a_1, a_2, \ldots, a_k) \in R^k$, under ϕ, gets sent to $m \in M$.

Problem 4.3.2. Show that $\ker \phi$ is finitely generated.

Solution. Since R is a PID, it is Noetherian by Problem 4.1.1. But then R^k is a finitely generated module over a Noetherian ring, and therefore is a Noetherian module by Problem 4.1.5. Therefore, the submodule $\ker \phi$ is Noetherian by Problem 4.1.2.

Problem 4.3.3. Suppose that the only nonzero entries of A are along the diagonal (that is, $a_{i,j} \neq 0$ implies that $i = j$). Show that M is then isomorphic to a module of the form

$$R^r \oplus \bigoplus_{i=1}^{k} R/\langle d_i \rangle,$$

and describe how you obtain the values d_i and r.

Solution. Suppose that $e_1, e_2, \ldots, e_n$ are the standard basis for R^n; that is e_i is the tuple with the only nonzero entry being a 1 in the ith coordinate. Then, since A only has nonzero entries along the diagonal, it follows that the kernel of ϕ is generated by $a_{1,1}e_1, a_{2,2}e_2, \ldots, a_{m,m}e_m$. Let $d_i = a_{i,i}$.

Consider the map $\psi : R^n \to (\bigoplus_{i=1}^m R/\langle d_i \rangle) \oplus R^{n-m}$ by sending the tuple $(r_1, r_2, \ldots, r_n)$ to $(\overline{r_1}, \overline{r_2}, \ldots, \overline{r_m}, r_{m+1}, r_{m+2}, \ldots, r_n)$, where $\overline{r_i}$ is the image of r_i under the quotient map $R \to R/\langle d_i \rangle$. One can quickly verify that this is a surjective homomorphism: indeed, given $(r_1, r_2, \ldots, r_n)$ and $(r_1', r_2', \ldots, r_n') \in R^n$, we have

$$\psi((r_1, r_2, \ldots, r_n) + (r_1', r_2', \ldots, r_n')) = (\overline{r_1 + r_1'}, \overline{r_2 + r_2'}, \ldots, \overline{r_m + r_m'}, r_{m+1} + r_{m+1}', \ldots, r_n + r_n'),$$

which is equal to

$$(\overline{r_1}, \overline{r_2}, \ldots, \overline{r_m}, r_{m+1}, \ldots, r_n) + (\overline{r_1'}, \overline{r_2'}, \ldots, \overline{r_m'}, r_{m+1}', \ldots, r_n') = \psi((r_1, r_2, \ldots, r_n)) + \psi((r_1', r_2', \ldots, r_n')),$$

and similarly for any $r \in R$, we have

$$\psi(r(r_1, r_2, \ldots, r_n)) = (\overline{rr_1}, \overline{rr_2}, \ldots, \overline{rr_m}, rr_{m+1}, \ldots, rr_n) = r\psi((r_1, r_2, \ldots, r_n)).$$

For surjectivity, given $(\overline{r_1}, \overline{r_2}, \ldots, \overline{r_m}, r_{m+1}, r_{m+2}, \ldots, r_n) \in (\bigoplus_{i=1}^m R/\langle d_i \rangle) \oplus R^{n-m}$, there exist r_i for $i = 1, 2, \ldots, m$ so the image of r_i in $R/\langle d_i \rangle$ is $\overline{r_i}$, so then $(r_1, r_2, \ldots, r_n)$ maps to our above tuple.

Notice that the kernel of this homomorphism is precisely the set of tuples $(r_1, r_2, \ldots, r_n)$ where $r_i \in \langle d_i \rangle$ for $i = 1, 2, \ldots, m$, and are 0 for the others. However, such tuples are precisely those of the form $\sum\limits_{i=1}^{m} a_i d_i e_i$, which is precisely the kernel of ϕ, as noted above. Therefore, we have the isomorphisms

$$\left(\bigoplus_{i=1}^{m} R/\langle d_i \rangle \right) \oplus R^{n-m} \simeq R^n / \ker \psi = R^n / \ker \phi \simeq M.$$

Finally, noting that $R/\langle 0 \rangle \simeq R$ (since the identity map on R has kernel 0 and is surjective) and that the map permuting the tuples is an isomorphism, it follows that $M \simeq R^r \oplus \bigoplus_{i=1}^k R/\langle d_i' \rangle$, where k is the number of nonzero d_i, the d_i' are precisely the nonzero diagonal entries in A (with some ordering), and $r = n - k$.

Problem 4.3.4. Prove Theorem 4.0.1, and show that a choice of d_i can be made such that the d_i are all powers of prime elements.

Solution. Suppose that M is a finitely generated R-module. We know there exists a module homomorphism $\phi : R^n \to M$ which is surjective. Consider $\ker \phi$, which is finitely generated by some $w_1, w_2, \ldots, w_m$. Using the above construction, we know that $\ker \phi$ is equal to $N(A)$ for some $n \times m$ matrix A, and thus $R/N(A) \simeq M$.

However, by Problem 4.2.5, there exist invertible matrices S, T such that SAT is a matrix whose nonzero entries lie on the diagonal, and by the previous problem we know that $R^n/N(SAT)$ takes the form given in Theorem 4.0.1. However, by Problems 4.2.1 and 4.2.2, we know that $M \simeq R/N(A) \simeq R/N(SA) \simeq R/N(SAT) \simeq M \simeq R^r \oplus \bigoplus_{i=1}^k R/\langle d_i \rangle$. This proves the theorem.

To show that all the d_i can be made into powers of prime elements, recall from Theorem 1.2.7 that PIDs are UFDs, so we have unique factorization of d_i into prime powers, say by $d_i = p_1^{e_1} p_2^{e_2} \dots p_l^{e_l}$, where the p_i are prime elements that are pairwise not multiples of each other by units. Now, we wish to apply Problem 1.3.3. We consider $\langle p_i^{e_i} \rangle + \langle p_j^{e_j} \rangle = \langle p_i^{e_i}, p_j^{e_j} \rangle$. Say this is $\langle d \rangle$, so d divides both $p_i^{e_i}$ and $p_j^{e_j}$. However, by uniqueness of factorization, this requires that d be a unit, since the first divisibility condition requires that the only primes in d's factorization are unit multiples of p_i, and the second requires that the only primes in d's factorization are unit multiples of p_j.

Thus, $\langle p_i^{e_i} \rangle + \langle p_j^{e_j} \rangle = \langle p_i, p_j \rangle = R$. Furthermore, note that $\langle p_i^{e_i} \rangle \cap \langle p_j^{e_j} \rangle = \langle p_i^{e_i} p_j^{e_j} \rangle$. The right-hand side lies in the left-hand side, since any multiple of $p_i^{e_i} p_j^{e_j}$ is a multiple of each of the prime powers. As for the other direction, since we are working in a PID (ergo, a UFD), x being divisible by $p_i^{e_i}, p_j^{e_j}$ means that the unique factorization of x contains $p_i^{e_i}$ and $p_j^{e_j}$ in this factorization. But then x is divisible by $p_i^{e_i} p_j^{e_j}$. Thus, by Problem 1.3.3, we know that $R/\langle d_i \rangle \simeq \bigoplus_{i=1}^l R/\langle p_j^{e_j} \rangle$. Applying this for each d_i (and noting that if d_i is a unit, $R/\langle d_i \rangle$ is equal to $R/\langle 1 \rangle$, with 1 a power of a prime) thus gives us a form where each diagonal entry is a power of a prime.

5 Applications and Asides

5.1 A Counterexample

Problem 5.1.1. Find a $\mathbb{Z}[\sqrt{-5}]-$module that is not isomorphic to a module of the form

$$R^r \oplus \bigoplus_{i=1}^k R/\langle d_i \rangle,$$

where $R = \mathbb{Z}[\sqrt{-5}]$.

Solution. We claim that $I = \langle 2, 1 + \sqrt{-5} \rangle$ is not isomorphic to any module of the above form.

To see this, suppose that there existed such an isomorphism $\phi : R^r \oplus \bigoplus_{i=1}^k R/\langle d_i \rangle \to I$ to some module of the above form. First, suppose that $k \geq 1$, and one of the d_i is not a unit. Let d be one such element. Then, we know that $R/\langle d \rangle$ consists of elements such that, when multiplied by d, they become zero. Say m is one such nonzero element. Then, $dm = 0$, so $\phi(dm) = d\phi(m) = 0$. Thus, $\phi(m)$ is nonzero and vanishes upon multiplication by d. But $\mathbb{Z}[\sqrt{-5}]$ is a domain, so this is impossible.

Hence, the only d_i are units. Suppose that $r \geq 2$. Then, there exist two linearly independent elements s_1, s_2 in $R^r \oplus \bigoplus_{i=1}^k R/\langle d_i \rangle$ (namely, taking the tuple with a 1 only in the first coordinate, and the tuple with only a 1 in the second coordinate). Note, however, that $\phi(s_1), \phi(s_2)$ must be also linearly independent, as $r_1\phi(s_1) + r_2\phi(s_2) = 0$ implies that $\phi(r_1 s_1 + r_2 s_2) = 0$, or that $r_1 s_1 + r_2 s_2 = 0$, or that r_1, r_2 are zero. But $\phi(s_2)\phi(s_1) - \phi(s_1)\phi(s_2) = 0$, and $\phi(s_1), \phi(s_2)$ are nonzero, so this is impossible.

Thus, we would need I to be isomorphic to a module of the form $R \oplus \bigoplus\limits_{i=1}^k R/\langle u_i \rangle$, where the u_i

are units. But note that this is isomorphic to R, since $R/\langle u_i\rangle$ are all isomorphic to 0, and we have the isomorphism from $R \oplus \bigoplus_{i=1}^{k} R/\langle u_i\rangle$ to R sending $(r, 0, 0, \ldots, 0)$ to $r \in R$.

This would thus require that $I \simeq R$. But then I is generated by $\phi(1)$. But this is impossible: $\phi(1)$ would divide both 2 and $1 + \sqrt{-5}$, meaning that if $\phi(1) = a + b\sqrt{-5}$, then there exists $c + d\sqrt{-5} \in \mathbb{Z}[\sqrt{-5}]$ such that $(a + b\sqrt{-5})(c + d\sqrt{-5}) = 2$. But then note that $(a + b\sqrt{-5})(c + d\sqrt{-5})(a - b\sqrt{-5})(c - d\sqrt{-5}) = (a^2+5b^2)(c^2+5d^2) = 4$, as $(a-b\sqrt{-5})(c-d\sqrt{-5}) = ac-5bd-(ad+bc)\sqrt{-5} = 2$. But this is only possible when $a = 1, c = 2$, or $a = 2, c = 1$ (with the rest 0). And as 2 does not divide $1 + \sqrt{-5}$, it follows that $a = 1$, so $\phi(1) = 1$. But $1 \notin I$, as one can quickly check.

Thus, I cannot be isomorphic to any module of the above form, which is what we wanted to show.

5.2 Abelian Groups

Problem 5.2.1. Consider the group $\mathbb{Z}/8\mathbb{Z}$. How many elements are there of each order? What if you replace 8 with any positive integer n?

Solution. For $\mathbb{Z}/8\mathbb{Z}$, we can check each element's order: we have the elements represented by $0, 1, 2, 3, 4, 5, 6, 7$. We note that any odd integer will have order 8 : indeed, if x is odd, then for $nx \equiv 0 \pmod 8$, or 8 to divide nx, we need n to be divisible by 8, and furthermore this is sufficent. Thus, there are 4 elements of order 8. Similarly, if an element x is divisible by 2 but not 4, then $nx \equiv 0 \pmod 8$ implies that $8|nx$, or that $4|n$, so two elements have order 4. Finally, 0 has order 1 and 4 has order 2, so we have 1 element of order 1, 1 of order 2, 2 of order 4, and 4 of order 8.

In general, suppose we have $\mathbb{Z}/n\mathbb{Z}$. Consider the order of an element $x \in \mathbb{Z}/n\mathbb{Z}$. We have that n divides dx if and only if $\frac{n}{\gcd(n,x)}$ divides d. Furthermore, the order is equal to d precisely if $\frac{n}{\gcd(n,x)} = d$. Thus, for any d not dividing n, there are no elements of order d. Meanwhile, for $d|n$, note that we are looking at the set of elements $x \in \{1, 2, \ldots, n\}$ so $\gcd(n, x) = n/d$, or that $\gcd(d, dx/n) = 1$, and x is divisible by n/d. So this set is in bijection with the integers in $\{1, 2, \ldots, d\}$ that are relatively prime to d. But this is just $\varphi(d)$, where φ is the Euler totient function.

Therefore, there are $\varphi(d)$ elements of order d in $\mathbb{Z}/n\mathbb{Z}$.

Problem 5.2.2. Show that $\mathbb{Z}/n\mathbb{Z}$ is cyclic for every positive integer n.

Solution. We show that $\mathbb{Z}/n\mathbb{Z}$ is generated by the element $\overline{1}$. Indeed, notice that any element $\overline{m} \in \mathbb{Z}/n\mathbb{Z}$, by definition, is equal to the residue of m modulo n. We can, by adding more ns to m, assume that $m > 0$. However, notice that m is equal to $1 + 1 + \cdots + 1$, where we have m 1s. Therefore, $\overline{m}$ is equal to the sum of m $\overline{1}$s. But this can be done for each $\overline{m} \in \mathbb{Z}/n\mathbb{Z}$, meaning that $\overline{1}$ generates $\mathbb{Z}/n\mathbb{Z}$, which is what we wanted to show.

Problem 5.2.3. Show that $\mathbb{Z}[x]$, the set of polynomials with coefficients in $\mathbb{Z}$, is not a finitely generated abelian group under addition. (Hint: suppose you had a finite subset of $\mathbb{Z}[x]$. What elements can be written as a sum of these elements?)

Solution. We suppose for the sake of contradiction that $\mathbb{Z}[x]$ was a finitely generated abelian group under addition. Then, we have some set $\{f_1, f_2, \ldots, f_k\}$ of generators for $\mathbb{Z}[x]$ as an abelian group. But note that the sum of any of these polynomials is going to have degree at most the maximum of the degrees of the f_i. Indeed, given any polynomials f, g, if both of their degrees are at most n, then $f + g$ also has degree at most n, and similarly for $f - g$.

Let N be the maximum of the degrees of the f_i, and consider x^{N+1}. For $\mathbb{Z}[x]$ to be generated by $f_1, f_2, \ldots, f_k$, we need for $x^{N+1} = g_1 + g_2 + \ldots + g_m$, where each g_i is either one of the f_i or the additive inverse of an f_i. But the right-hand side is a polynomial of degree at most N, which is a contradiction.

Thus, $\mathbb{Z}[x]$ cannot be a finitely generated abelian group, which is what we wanted to show.

Problem 5.2.4. Show that every abelian group is a $\mathbb{Z}-$module, where one of the operations is $+$. What is the scalar multiplication?

Solution. Let G be an abelian group. We claim that this is a $\mathbb{Z}-$module with scalar multiplication given by $n \cdot g = g + g + \cdots + g$, with n gs, if $n > 0$, and $n \cdot g = -g + (-g) + \cdots + (-g)$, with $-g$ the additive inverse of g and there being $-n$ $(-g)$s. Finally, we let $0 \cdot g = 0$.

We check the properties of being a $\mathbb{Z}-$module. Properties 1, 3, 4, 5 follow from the fact that G is an abelian group, and the associativity of $+$ also follows. For property 6, this follows by definition of our module structure, since $1 \cdot g = g$.

For the associativity of $\cdot$, we first note that $(-n) \cdot g = -1 \cdot (n \cdot g) = n \cdot (-1 \cdot g)$ if $n > 0$. Indeed, note that $(-n) \cdot g$ is equal to $-g + (-g) + \cdots + (-g)$, where there are n $-g$s. Meanwhile, the right-hand side is equal to the additive inverse of $n \cdot g = g + g + \cdots + g$. However, $(n \cdot g) + (-n \cdot g) = g + g + \cdots + g + (-g) + (-g) + \cdots + (-g)$, where there are n gs and n $(-g)$s. Each of these cancel out, and so $(n \cdot g) = (-1) \cdot (n \cdot g)$. Finally, by definition, $(-n) \cdot g$ is equal to $n \cdot (-1 \cdot g)$, which is the sum of n $-1 \cdot g = (-g)$s.

From here, note that for $n = 0$ that both sides are equal to 0, and for $n < 0$, we have that $(-1) \cdot ((-1) \cdot ((-n) \cdot g))$ is equal to $(-n) \cdot g$ (since the two (-1)s corresponding to taking additive inverse and then applying additive inverse again, which cancel each other out). Thus, by the $n > 0$ case, this is also equal to $(-1) \cdot (n \cdot g)$, which is what we wanted. Again, note that $(-n) \cdot g$ is the sum of $-n$ gs, which is also what $n \cdot (-1 \cdot g)$ is, as the additive inverse of the additive inverse of g is g.

From here, given $m, n \in \mathbb{Z}$, let s_m, s_n be the sign of m, n, respectively. Then, note that $m \cdot (n \cdot g) = s_m \cdot (|m| \cdot (s_n \cdot (|n| \cdot g)))$. But by the above, this is equal to $s_m \cdot (s_n \cdot (|m| \cdot (|n| \cdot g)))$. Furthermore, $|m| \cdot (|n| \cdot g)$ is just the sum of $|mn|$ gs (by having $|m|$ sums of $|n|$ gs), so this equals $s_m \cdot (s_n \cdot (|mn| \cdot g)) = s_m s_n |mn| \cdot g = mn \cdot g$. This proves associativity, and hence property 2.

For property 7, we note that $(-1) \cdot (g_1 + g_2)$ is the additive inverse of $g_1 + g_2$. But this is then equal to $(-g_1) + (-g_2)$, since if the additive inverse is e, then $e = e + g_1 + g_2 + (-g_1) + (-g_2) = -g_1 + -g_2$. From here, we note that if s_n is the sign of n, then $n \cdot (g_1 + g_2) = s_n \cdot (|n| \cdot (g_1 + g_2)) = s_n \cdot (|n| \cdot g_1 + |n| \cdot g_2) = n \cdot g_1 + n \cdot g_2$.

Furthermore, given m, n, we note that if both m, n are positive, then $(m+n) \cdot g = m \cdot g + n \cdot g$, and the above argument also covers the case if m, n are both negative (by factoring out the -1 first). If $m + n > 0$ but one of them is nonpositive, suppose without loss of generality that $m > 0, n \leq 0$. Then, $(m+n) \cdot g$ is equal to the sum of $m+n$ gs. But this is equal to m gs plus $|n|$ $(-g)$s, or $m - |n|$ gs. The above observation, giving the distributivity of (-1), gives us the case for when $m + n < 0$ as well, which proves property 7 and thus shows that G can be given a $\mathbb{Z}-$module structure with this notion of scalar multiplication.

Problem 5.2.5. Suppose that G is a finitely generated abelian group. Show that G is isomorphic (as groups) as

$$\mathbb{Z}^r \oplus \bigoplus_{i=1}^{n} \mathbb{Z}/d_i\mathbb{Z},$$

for some positive integers $d_1, d_2, \ldots, d_r$ and nonnegative integers r, n.

Solution. Given such a finitely generated abelian group G generated by $\{g_1, g_2, \ldots, g_k\}$, recall from the previous problem that G can then be given the structure of a $\mathbb{Z}$-module. Since $\mathbb{Z}$ is a PID, and G is a finitely generated $\mathbb{Z}$-module as well (noting that the generators $g_1, g_2, \ldots, g_k$ also generate G as a $\mathbb{Z}$-module).

Therefore, by the PID structure theorem, it follows that G is isomorphic, as $\mathbb{Z}$-modules, to

$$\mathbb{Z}^r \oplus \bigoplus_{i=1}^{n} \mathbb{Z}/d_i\mathbb{Z},$$

for some positive integers $d_1, d_2, \ldots, d_r$ and nonnegative integers r, n. Let ϕ be this isomorphism. But this isomorphism, in particular, satisfies $\phi(g_1 + g_2) = \phi(g_1) + \phi(g_2)$, since ϕ is a module homomorphism. Therefore, ϕ is also a group homomorphism, and as it is bijective, it follows that ϕ is an isomorphism of abelian groups. This proves the theorem.

5.3 Jordan Canonical Form

Problem 5.3.1. Show that V is a $\mathbb{C}[x]-$module, where for a polynomial $p(x) = \sum\limits_{i=0}^{n} a_i x^i$ we define $p(x) \cdot v = \sum\limits_{i=0}^{n} a_i T^i v$ for each $v \in V$.

Solution. We check through the properties of a module. Again, we know that, since V is a vector space, properties 3, 4, 5 hold. Property 1 follows since T, by assumption, sends V to V, and so $p(x) \cdot v$, for any $p \in \mathbb{C}[x]$, is equal to $\sum\limits_{i=0}^{n} a_i T^i v \in V$. For property 2, suppose that $p(x) = \sum\limits_{i=0}^{n} a_i x^i$ and $q(x) = \sum\limits_{i=0}^{m} b_i x^i$. Then,

$$q(x) \cdot (p(x) \cdot v) = q(x) \cdot (\sum_{i=0}^{n} a_i T^i v) = \sum_{j=0}^{m} b_j T^j (\sum_{i=0}^{n} a_i T^i v).$$

However, by linearity, this is equal to $\sum\limits_{j=0}^{m} \sum\limits_{i=0}^{n} a_i b_j T^{i+j} v$. However, note that this is equal to $q(x)p(x) \cdot v$, meaning that property 2 follows.

For property 6, note that by definition that $1 \cdot v = T^0 v = v$, since T^0 is the identity. We finally check property 7: one direction of distributivity follows from the fact that T is linear.

For the other direction, suppose we are given polynomials $p(x) = \sum\limits_{i=0}^{n} a_i x^i$ and $q(x) = \sum\limits_{i=0}^{m} b_i x^i$. We can suppose that $m = n$. Then, note that

$$p(x) \cdot v + q(x) \cdot v = \sum_{i=0}^{n} a_i T^i v + \sum_{i=0}^{n} b_i T^i v = \sum_{i=0}^{n} (a_i + b_i) T^i v.$$

However, note that this equals $(p(x) + q(x)) \cdot v$, giving us property 7. Therefore, V forms a $\mathbb{C}[x]$-module with the given scalar multiplication, as desired.

Problem 5.3.2. Show that, as $\mathbb{C}[x]-$modules and as vector spaces, V is isomorphic to

$$\bigoplus_{i=1}^{n} \mathbb{C}[x]/(x-\lambda_i)^{r_i},$$

for some complex numbers $\lambda_1, \lambda_2, \ldots, \lambda_n$ and positive integers $r_1, r_2, \ldots, r_n$. Again, you will need the Fundamental Theorem of Algebra (see Problem 1.1.5).

Solution. First, by the PID Structure theorem, it follows that as $\mathbb{C}[x]$ modules that V is isomorphic to

$$\mathbb{C}[x]^r \oplus \bigoplus_{i=1}^{n} \mathbb{C}[x]/(p(x))^{r_i},$$

where $p(x)$ are polynomials in $\mathbb{C}[x]$. Furthermore, we know that by Problem 4.3.4 that such a choice can be made where the $p(x)$ are prime elements. But the only such prime elements in $\mathbb{C}[x]$, by Problem 1.1.5, are of the form $x - \lambda$. So we thus have that V is isomorphic to

$$\mathbb{C}[x]^r \oplus \bigoplus_{i=1}^{n} \mathbb{C}[x]/(x-\lambda_i)^{r_i}$$

as $\mathbb{C}[x]$-modules via the homomorphism ϕ, and in particular as vector spaces (since to be a vector space homomorphism, one must only check that $\phi(ax + by) = a\phi(x) + b\phi(y)$ for $x, y \in V$ and $a, b \in \mathbb{C}$, rather than $\mathbb{C}[x]$).

Now, suppose that $r \geq 1$. Then, we know that V is isomorphic to $\mathbb{C}[x]^r \oplus \bigoplus_{i=1}^{n} \mathbb{C}[x]/(x-\lambda_i)^{r_i}$ as vector spaces, and by assumption V is finite dimensional, say dimension k. Then, the right-hand side must be as well. But if $r \geq 1$, then in particular we have that $1, x, x^2, \ldots, x^{k+1}$ (or rather, the tuples with these in the first coordinate and 0 in all others) are linearly independent over $\mathbb{C}$, so the dimension of the right-hand side is more than k, contradiction. Hence, $r = 0$, and so we require that $V \simeq \bigoplus_{i=1}^{n} \mathbb{C}[x]/(x-\lambda_i)^{r_i}$, as desired.

Problem 5.3.3. Prove the following.

1. Show that for each λ_i there exists a set of vectors $v_1, v_2, \ldots, v_{r_i} \in V$ such that $Tv_1 = \lambda_i v_1$, and for $j = 2, 3, \ldots, r_i$, we have that $Tv_j = \lambda_i v_j + v_{j-1}$, and that $v_1, v_2, \ldots, v_{r_i}$ are linearly independent.

2. Show furthermore that these sets can be chosen such that, if we combine all of these sets together, the resulting set of vectors is also linearly independent.

3. Show that this set of vectors must therefore be a basis for V.

Solution. **1.** Consider the element $1 \in \mathbb{C}[x]/(x-\lambda_i)^{r_i}$; for the sake of readability for this part we drop the bars (and understand all the polynomials as residues). We furthermore treat each of these elements as the tuples in $\bigoplus_{i=1}^{n} \mathbb{C}[x]/(x-\lambda_i)^{r_i}$ where all other coordinates are zero. Notice that $1, (x-\lambda_i), \ldots, (x-\lambda_i^{r_i-1}$ are linearly independent, since $a_0 + a_1(x-\lambda_i) + \ldots + a_{r_i-1}(x-\lambda_i)^{r_i-1} = 0$ implies, by translating the polynomialm that $a_0 + a_1 x + \ldots + a_{r_i-1}x^{r_i-1} = 0$, or that all of the a_i are zero. Consider now the vector $\phi^{-1}(1), \phi^{-1}(x-\lambda_i), \ldots, \phi^{-1}(x-\lambda_i^{r_i-1})$. Let $v_j = \phi^{-1}(x-\lambda_i^{r_i-j})$, for $j = 1, 2, \ldots, r_i$. Then, note that $\phi((T-\lambda_i)v_1) = \phi((x-\lambda_i)\cdot v_1) = (x-\lambda_i)(x-\lambda_i)^{r_i-1} = 0$ (working in $\mathbb{C}[x]/(x-\lambda_i)^{r_i}$). In other words, by ϕ being an isomorphism, we have that $(T-\lambda_i)v_1 = 0$, or $Tv_1 = \lambda_i v_1$. Furthermore, we have that

$$\phi((T-\lambda_i)v_j) = \phi((x-\lambda_i)\cdot v_j) = (x-\lambda_i)(x-\lambda_j)^{r_i-j} = \phi(v_{j-1}),$$

or that $Tv_j = \lambda_i v_j + v_{j-1}$. Finally, notice that the v_j are linearly independent, since the $(x-\lambda_i)^j$ are too, proving this first part.

2. Let X_i be the collection of vectors that we chose corresponding to coordinate i in the direct sum $\bigoplus_{i=1}^n \mathbb{C}[x]/(x-\lambda_i)^{r_i}$. Then, notice that for any linear combination of vectors in $\bigcup\limits_{i=1}^n X_i$ to be zero, the linear combination must be zero in each coordinate. But in coordinate i, the only vectors with nonzero coordinate are those in X_i, and these are presumed to be linearly independent. So the coefficients on the vectors in X_i are zero, and this holds for each i. Thus, our choices can be made such that the set of vectors is linearly independent.

3. To show that this is a basis for V, we argue that the corresponding elements in $\bigoplus_{i=1}^n \mathbb{C}[x]/(x-\lambda_i)^{r_i}$ are a basis (which is enough, since ϕ is a vector space isomorphism and thus preserves linear combinations: so every vector $v \in V$ being a linear combination of vectors v_i holds if and only each vector in $\phi(v)$ is a linear combination of vectors $\phi(v_i)$).

Suppose we have a vector $(p_1, p_2, \ldots, p_n) \in \bigoplus_{i=1}^n \mathbb{C}[x]/(x-\lambda_i)^{r_i}$. Then, note that each p_i is equivalent modulo $(x-\lambda_i)^{r_i}$ to some vector that is of degree at most $r_i - 1$, by the normal polynomial division algorithm. So we may suppose that each p_i is degree at most $r_i - 1$. Then, by repeatedly dividing quotients by $x - \lambda_i$, we find that we may write

$$p_i(x) = q_1(x)(x-\lambda_i)+a_0 = q_2(x)(x-\lambda_i)^2+a_1(x-\lambda_i)+a_0 = \ldots = a_{r_i-1}(x-\lambda_i)^{r_i-1}+a_{r_i-2}(x-\lambda_i)^{r_i-2}+\cdots+a_0.$$

This gives us that $(p_1, p_2, \ldots, p_n)$ can be written as the linear combination

$$\begin{gathered} a_{10}(1,0,0,\ldots,0) + a_{11}(x-\lambda_1, 0, 0, \ldots, 0) + \cdots + a_{20}(0,1,0\ldots,0) \\ + \cdots + a_{n0}(0,1,0,\ldots,0) + \cdots + a_{n(r_n-1)}(0,0,\ldots,(x-\lambda_n)^{r_n-1}). \end{gathered}$$

But these are just the vectors in $\bigcup\limits_{i=1}^n X_i$. Hence, it follows that this set is linearly independent and spanning, and therefore is a basis. Thus, it follows that the corresponding vectors in V (obtained by mapping each vector in $\bigcup\limits_{i=1}^n X_i$ to V via ϕ^{-1}) form a basis for V, which is what we wanted to show.

Problem 5.3.4. Show that for any linear transformation T on a $\mathbb{C}$−vector space V, there exists a basis $v_1, v_2, \ldots, v_n$ such that, with respect to this basis, T has block matrix form

$$\begin{pmatrix} J_1 & 0 & \cdots & 0 \\ 0 & J_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & J_k \end{pmatrix},$$

where each of the J_i has the following form for some $\lambda_i \in \mathbb{C}$:

$$\begin{pmatrix} \lambda_i & 1 & 0 & \cdots & 0 \\ 0 & \lambda_i & 1 & \cdots & 0 \\ 0 & 0 & \lambda_i & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \lambda_i \end{pmatrix},$$

with λ_i along the main diagonal and 1s along the diagonal immediately above it. This is known as the **Jordan canonical form** for a linear transformation/matrix.

Solution. Given a linear transformation T on a $\mathbb{C}$-vector space, we know that V, as a $\mathbb{C}[x]$-module, is isomorphic to $\bigoplus_{i=1}^{n} \mathbb{C}[x]/(x - \lambda_i)^{r_i}$, for some λ_i, and from the previous problems we can find a basis $v_{1,1}, v_{1,2}, \ldots, v_{1,r_1}, v_{2,1}, \ldots, v_{2,r_2}, \ldots, v_{n,1}, \ldots, v_{n,r_n}$ satisfying the properties above. Recall that, by construction, for each i, we have that $Tv_{i,1} = \lambda_i v_{i,1}$ and for $j = 2, 3, \ldots, r_i$ that $Tv_{i,j} = \lambda_i v_{i,j} + v_{i,j-1}$. In other words, the matrix of T with respect to this basis has the above block form, where the block corresponding to $v_{i,1}, v_{i,2}, \ldots, v_{i,r_i}$ is given by

$$\begin{pmatrix} \lambda_i & 1 & 0 & \cdots & 0 \\ 0 & \lambda_i & 1 & \cdots & 0 \\ 0 & 0 & \lambda_i & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda_i \end{pmatrix}.$$

This is what we wanted to show.

6 Acknowledgements + Resources

Thank you to Aleksa Milojevic, Nancy Xu, and Oliver Thakar for their feedback on drafts of the power round.

For more information about the material in the power round, see Sheldon Axler's Linear Algebra Done Right (Chapters 1-3) and Michael Artin's Algebra (Chapters 11, 14).