2023-04-18

Bézout Domains and Elementary Divisor Domains: Are They the Same?

Michael D. Walton
Brigham Young University

Bézout Domains and Elementary Divisor Domains: Are They the Same?

Michael D. Walton

A thesis submitted to the faculty of
Brigham Young University
in partial fulfillment of the requirements for the degree of

Master of Science

Pace Nielsen, Chair
Stephen Humphries
Darrin Doud

Department of Mathematics

Brigham Young University

ABSTRACT

Bézout Domains and Elementary Divisor Domains: Are They the Same?

Michael D. Walton
Department of Mathematics, BYU
Master of Science

This thesis examines the connections between Bézout domains and elementary divisor domains. I establish what both of these domains are, and I provide some clarifying examples of each. I state and prove some key results that have been established already in the literature. I describe a process by which I tried to show a distinction between Bézout domains and elementary divisor domains, and then provide an explicit example which shows that this process as formulated would not lead to an example of a Bézout domain which is not an elementary divisor domain. Throughout the thesis, I also state open questions that could lead to future research in this area.

Keywords: Bézout domain, elementary divisor domain, Smith normal form

Acknowledgements

I would like to thank my advisor, Dr. Pace Nielsen, for his support and assistance. Thank you very much for all the time you were willing to put in to help me research this topic, as well as your help polishing this thesis. I would also like to thank my other committee members, Dr. Stephen Humphries and Dr. Darrin Doud, for their time reviewing my thesis.

I would also like to thank my fellow graduate students Alex Lee and Dahlia Maxwell. Thank you Alex for being my sounding board for the many tricky parts involving entire functions. The problems were not easy or always the most clear, but thanks for listening and providing ideas. Thank you Dahlia for being a really good friend and for helping motivate me to get this done in a timely manner.

I would like to thank my family members, especially my parents Brian and Heather for having my back and for being excited for me. Thanks for literally everything! And thank you to my fiancée Allison Dahl for cheering me on and for being there for me. It has meant so much to me, so thank you so much!

Contents

Chapter 1. Introductory Materials

1.1 Basic Definitions and Results

As we all know, the integers satisfy $ab = 0$ exactly when $a = 0$ or $b = 0$. This idea, that the only way to multiply numbers together and get zero is when at least one of the numbers is zero, is captured by the idea of integral domains.

Definition 1.1.1. A *domain*, or *integral domain*, is a commutative ring with identity $1 \neq 0$ with no nonzero zero-divisors.

We also recall the definition of an ideal in a ring. In this paper, all rings will be assumed to be commutative.

Definition 1.1.2. Let R be a ring, I be a subring of R, and let $r \in R$. Then I is an *ideal* of R if for all $i \in I$, both $ri, ir \in I$.

Definition 1.1.3. An ideal I of a ring is *principal* if it is generated, as an ideal, by a single element.

While the set of all multiples of an element in a ring will always be an ideal of the ring, not all ideals in a generic ring or domain are necessarily principal.

Example 1.1.4. The domain $\mathbb{Z}[x]$ has the following set as an ideal:

$$\{p(x) \in \mathbb{Z}[x] : \text{ the constant term of } p(x) \text{ is even}\}$$

This is an ideal since for any polynomial $q(x) \in \mathbb{Z}[x]$, the polynomial $p(x)q(x)$ has an even constant coefficient. However, this ideal is not principal. Instead, this set is the set of all polynomials that can be written as a sum of multiples of the polynomial x and the polynomial 2. We will denote this as $\langle x, 2 \rangle$ to signify that the ideal is generated by the elements x and 2.

One question that arises from this example is: When are ideals of a ring principal?

As we saw from Example 1.1.4, not every ideal in a generic ring or even domain is principal. The domains in which every ideal is principal have a special name.

Definition 1.1.5. A *principal ideal domain*, or a *PID* for short, is an integral domain in which every ideal is principal. More generally, a ring in which every ideal is principal is a *principal ideal ring*, or a PIR for short.

There are two relatively intuitive ways to weaken the condition that every ideal in a ring is principal, and yet still maintain many useful properties. One option is to require that every ideal be finitely generated, and the other is to require that every finitely generated ideal be principal.

Definition 1.1.6. A commutative ring R with identity is *Noetherian* if every ideal of R is finitely generated.

Definition 1.1.7. A ring is called *Bézout* if every ideal generated by two elements is principal. We refer to a domain which has this property as a *Bézout domain*.

Noetherian rings have been studied extensively. This paper will focus on the less-studied Bézout domains. Note that our definition of a Bézout ring does not *a priori* match our earlier comment that we should require every finitely generated ideal to be principal. However, the following theorem clarifies that the two descriptions are actually equivalent.

Theorem 1.1.8. *In a ring R, every finitely generated ideal is principal if and only if every ideal generated by two elements is principal.*

Proof. The forward direction of this proof is trivial as every ideal generated by two elements is finitely generated. We now approach the reverse direction.

Assume that every ideal generated by two elements is principal, i.e., for any $a_1, a_2 \in R$, there exists some $a \in R$ such that $\langle a_1, a_2 \rangle = \langle a \rangle$. Assume by way of induction that any ideal generated by n elements is principal. Let $I = \langle a_1, a_2, \ldots, a_n, b \rangle$ be an ideal generated by $n+1$ elements. Since $J = \langle a_1, \ldots, a_n \rangle$ is principal, it is equal to $\langle a \rangle$ for some $a \in R$.

Therefore, $I = \langle a, b \rangle = \langle c \rangle$ for some $c \in R$, and is thus principal. Therefore, any finitely generated ideal is principal. □

We can then quickly prove the following corollary.

Corollary 1.1.9. *A ring is a principal ideal ring if and only if it is a Noetherian Bézout ring.*

Proof. Let R be a principal ideal ring. Then every ideal is principal, and thus every ideal is finitely generated. Furthermore, every ideal generated by two elements is principal, so R is a Noetherian Bézout ring.

Conversely, let R be a Noetherian Bézout ring. Since R is Noetherian, every ideal is finitely generated. Since R is Bézout, Theorem 1.1.8 shows that this implies that every ideal is principal. Therefore, R is a principal ideal ring. □

In a Bézout ring, since any ideal generated by two elements $\langle a, b \rangle$ is principal, there is some element d such that $\langle a, b \rangle = \langle d \rangle$. If we were to do this in the domain of integers, we call this element the greatest common factor, or GCD, of a and b. We now extend this definition to any Bézout ring.

Definition 1.1.10. In a ring R, a *GCD* of two elements a and b is any element d such that $\langle a, b \rangle = \langle d \rangle$.

Proposition 1.1.11. *The GCD of two elements in a domain is unique up to a unit, when it exists.*

Proof. Let R be a domain such that for $a, b \in R$, there are two GCDs, namely $d, e \in R$. Then $\langle a, b \rangle = \langle d \rangle = \langle e \rangle$, which implies $d \in \langle e \rangle$ and $e \in \langle d \rangle$. Thus there are elements $u, v \in R$ such that $d = ue$ and $e = vd$. Therefore, $d = ue = uvd$.

If $d = 0$, then $e = 0$ and we are done. If $d \neq 0$, then since we are in a domain, we may cancel the d and set $uv = 1$. Thus u is a unit, and e is a unit multiple of d. Therefore, the GCD of two elements in a domain is unique up to a unit. □

There is another, more general, definition of GCD which says that d is a GCD of a and b if there is a unique minimal principal ideal $\langle d\rangle$ containing $\langle a,b\rangle$. However, in this paper, all GCDs will be assumed to be a GCD in the sense of Definition 1.1.10.

When we later construct examples of Bézout domains which display certain useful properties, we will require some simple results pertaining to integral domains. The following results can easily be found in most texts which introduce the concept of integral domains. As they are fundamental results, they are stated without proof.

Proposition 1.1.12 ([2, Proposition 7.4])**.** *If D is an integral domain and a is an indeterminate, then $D[a]$ is also an integral domain.*

Using induction on the number of indeterminates quickly leads to the following corollary.

Corollary 1.1.13. *If D is an integral domain and $\{a_1, a_2, a_3, \ldots, a_n\}$ $(n \in \mathbb{N})$ are all algebraically independent indeterminates, then $D[a_1, a_2, a_3, \ldots, a_n]$ is an integral domain.*

We will also need the following lemma:

Lemma 1.1.14 ([2, Theorem 7.15])**.** *If D is a domain, then* $\mathrm{Frac}(D) = \{ab^{-1} : a, b \in R, b \neq 0\}$ *is a field.*

1.2 Specific Types of Bézout Domains

Now that we have established what a Bézout domain is, a natural question is whether or not there is a Bézout domain which is not a PID? Since all PIDs are also Bézout domains, many easy-to-construct examples of Bézout domains are also PIDs. One example of a Bézout domain that is not a PID is the ring of all entire functions on the complex plane. To see this, we will first need the following two theorems. As both are fairly fundamental results in complex analysis, I will state them without proof.

Theorem 1.2.1. *(The Weierstrass Factorization Theorem) Let $E \subseteq \mathbb{C}$ be a set with no limit points. For each $\alpha \in E$, associate a positive integer $m(\alpha)$ to α. Then there exists an entire* *function f all of whose zeros are in E, and such that f has a zero of order $m(\alpha)$ at each $\alpha \in E$.*

An improvement of this theorem can be found in Walter Rudin's textbook "Real and Complex Analysis". (I have changed the wording slightly for this second theorem from how it appears in [12] to make it more clear.)

Theorem 1.2.2 ([12, Theorem 15.13])**.** *Let $E \subseteq \mathbb{C}$ be a set with no limit points. For each $\alpha \in E$, associate a nonnegative integer $m(\alpha)$ and complex numbers $\beta_{n,\alpha}$ for each $0 \leq n \leq m(\alpha)$. Then there is an entire function f such that*

$$f^{(n)}(\alpha) = \beta_{n,\alpha}$$

for each $\alpha \in E$ and $0 \leq n \leq m(\alpha)$.

Theorem 1.2.3. *The set of all entire functions is a Bézout domain.*

Proof. Let f, g be entire functions. Clearly, $f + g$ and fg are also entire functions, and if $fg = 0$, then either $f = 0$ or $g = 0$. Thus the set of all entire functions is a domain. We now show that this domain is Bézout.

Let f, g be nonzero entire functions. Let Z_f and Z_g denote the zero sets of f, g respectively. For each $\alpha \in Z_f \cup Z_g$, let $m_f(\alpha)$ be the degree of the zero of f at α and $m_g(\alpha)$ be the degree of the zero of g at α. Let h be any entire function which vanishes at each point $\beta \in Z_f \cap Z_g$ with multiplicity exactly $\min\{m_f(\beta), m_g(\beta)\}$; such a function is guaranteed by Theorem 1.2.1. The function h is unique up to an entire function with no zeros, namely a unit. Clearly this h is a common divisor of both f and g. Furthermore, if h' is also a common divisor of both f and g, then the zero set of h' must be contained in both Z_f and Z_g. Thus, the degree of the zero at each $\alpha \in Z_{h'}$ must be less than or equal to $\min\{m_f(\beta), m_g(\beta)\}$, so h' must divide h. Therefore, h will be the greatest common divisor of f, g as long as we show that h is a linear combination of f and g.

We may reduce to the case where f and g are entire functions such that they do not share a zero, after dividing by h is necessary. Thus it will suffice to prove that the ideal generated by f and g is the ring of all entire functions if $Z_f \cap Z_g = \emptyset$. We want to find entire functions r and s such that $rf + sg = 1$. Equivalently, we need to choose s such that $r := (1-sg)/f$ is an entire function. Thus, we need $m_{1-sg}(\alpha) \geq m_f(\alpha)$. Since g is never zero where f is zero, we can use Theorem 1.2.2 to find an entire function s such that $s(\alpha) = 1/g(\alpha)$ for every $\alpha \in Z_f$ as well as having derivatives at α be chosen in exactly the way so that $(1-sg)/f$ has a removable singularity at every $\alpha \in Z_f$. Thus once these singularities are filled, $(1-sg)/f$ is an entire function and we are done. □

In the previous theorem, where we showed that the set of entire functions is a Bézout domain, we used factorization properties of the entire functions. Similar factorization properties have appeared in other types of Bézout domains, so one natural question seems to be to determine what, if any, factorization properties are inherent in a generic Bézout domain.

One well-known type of domain with some factorization properties, which is connected to Bézout domains, is that of unique factorization domains (or UFDs).

Definition 1.2.4. An integral domain D is a *unique factorization domain*, or a *UFD* when

(i) every nonzero, nonunit element $d \in D$ can be written as a finite product of irreducibles $p_i \in D$ (i.e., $d = p_1 p_2 \cdots p_n$), and

(ii) the factorization of d is unique up to units and order; namely, if $d = q_1 q_2 \cdots q_m$ is another factorization, then $n = m$ and (after reindexing) $p_i = q_i u_i$ with each $u_i \in D$ a unit.

Clearly not all Bézout domains are UFDs since the ring of entire functions has examples of functions which cannot be written as a finite product of irreducibles (such as the infinite product form of $\sin(\pi z)$ referenced in the proof of Theorem 4.3.2). Furthermore, not all UFDs are Bézout domains. To see this, note that $\mathbb{Z}[x]$ is a unique factorization domain, but the ideal $\langle 2, x \rangle$ is not principal. So while there are connections between UFDs and Bézout domains, these connections will not be expounded upon much in this paper.

Another type of domain that has some useful factorization properties is what is called an *elementary divisor domain.*

Definition 1.2.5. A commutative ring R is called an *elementary divisor ring* if for every matrix $A \in M_{m,n}(R)$, then there exist invertible matrices $P \in \mathrm{GL}_m(R)$ and $Q \in \mathrm{GL}_n(R)$ such that $PAQ = D$ with $D = (d_{i,j})$ diagonal (i.e., $d_{i,j} = 0$ whenever $i \neq j$) and each $d_{i,i} | d_{i+1,i+1}$. (In this case, we say that A and D are *equivalent* and that A *admits diagonal reduction.*) The matrix D is called a *Smith normal form*, or *SNF*, of A. A domain with this property is called an *elementary divisor domain.*

Theorem 1.2.6. *Every elementary divisor domain is a Bézout domain.*

Proof. Suppose that R is an elementary divisor domain, and let $a, b \in R$. Let A be the 2×2 diagonal matrix with a and b on the diagonal. Then there are invertible matrices $P, Q \in \mathrm{GL}_2(R)$ such that PAQ is in Smith normal form. Writing these multiplications out explicitly, we see that

$$PAQ = \begin{pmatrix} d & 0 \\ 0 & e \end{pmatrix} = \begin{pmatrix} x_1 & y_1 \\ z_1 & w_1 \end{pmatrix} \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix} \begin{pmatrix} x_2 & y_2 \\ z_2 & w_2 \end{pmatrix} = \begin{pmatrix} x_1ax_2 + y_1bz_2 & x_1ay_2 + y_1bw_2 \\ z_1ax_2 + w_1bz_2 & z_1ay_2 + w_1bw_2 \end{pmatrix}.$$

Therefore, d is a linear combination of a and b, so $d \in \langle a, b \rangle$.

Since $d|e$, we can write $e = df$. Then $A = P^{-1}(PAQ)Q^{-1}$ can be written

$$A = \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix} = \begin{pmatrix} x_3 & y_3 \\ z_3 & w_3 \end{pmatrix} \begin{pmatrix} d & 0 \\ 0 & df \end{pmatrix} \begin{pmatrix} x_4 & y_4 \\ z_4 & w_4 \end{pmatrix} = \begin{pmatrix} x_3dx_4 + y_3dfz_4 & x_3dy_4 + y_3dfw_4 \\ z_3dx_4 + w_3dfz_4 & z_3dy_4 + w_3dfw_4 \end{pmatrix}.$$

This means that both a and b are multiples of d, so $a, b \in \langle d \rangle$. Therefore, $\langle a, b \rangle = \langle d \rangle$ and R is a Bézout domain. □

There is a fairly straightforward algorithm for calculating the Smith normal form of a matrix A when the ring is a PID. I will illustrate the algorithm on a 2×2 matrix, and the algorithm can be easily generalized to matrices of arbitrary size.

Algorithm 1.2.7. Let R be a PID, and let $A \in M_2(R)$ be an arbitrary 2×2 matrix over R that is not yet in Smith normal form. The matrix D is found recursively as follows. The following algorithm for computing the Smith normal form is a standard method, and will be referred to later in this paper as the standard SNF algorithm. (This algorithm was described by Smith in his paper introducing the SNF of a matrix in 1861. See [13].)

Step 1: Make the top-left corner of the matrix nonzero.

If the top-left corner a is nonzero, we can move on to Step 2. However, if the top-left corner is zero, then there are three subcases that must be dealt with. For each of these cases, the matrix multiplications to move on to Step 2 are listed.

Step 1a: The bottom-left corner is nonzero.

If $a = 0$ but $c \neq 0$, then swap the rows using the following multiplication:

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & b \\ c & d \end{pmatrix} = \begin{pmatrix} c & d \\ 0 & b \end{pmatrix}.$$

Step 1b: The first column is all zero, but the top-right corner is nonzero.

If $a = c = 0$ but $b \neq 0$, we can swap the columns using the following multiplication:

$$\begin{pmatrix} 0 & b \\ 0 & d \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} b & 0 \\ d & 0 \end{pmatrix}.$$

Step 1c: The only nonzero entry in A is the bottom-right corner.

If $a = b = c = 0$ but $d \neq 0$, we swap both the rows and the columns using the following multiplication:

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 0 & d \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} d & 0 \\ 0 & 0 \end{pmatrix}.$$

Note that if all entries were zero, then we were in Smith normal form already, so these three substeps cover all the cases.

Step 2: Make the top-right corner zero.

If the top-right corner b is zero, we can move on to Step 3. If not, we again have two subcases.

Step 2a: $a|b$**.**

If $a|b$, then $b = ak$ so we will subtract k multiples of column 1 from column 2. Now perform following multiplication:

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} 1 & -k \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} a & 0 \\ c & d - ck \end{pmatrix}.$$

Step 2b: $a \nmid b$**.**

If $a \nmid b$, then since all PID's are Bézout, there is some greatest common divisor e of a and b such that $a = a'e, b = b'e$, and there are $x, y \in R$ such that $a'x + b'y = 1$. This is done using the following multiplication:

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x & -b' \\ y & a' \end{pmatrix} = \begin{pmatrix} e & 0 \\ cx + dy & da' - cb' \end{pmatrix}.$$

(Note that the matrix we multiplied A by has determinant 1, so it is invertible.)

At this point, the first row has a pivot in the top-left corner and the rest of the row is zero.

Step 3: Make the bottom-left corner zero.

This step is accomplished through the transposition of the matrix multiplications performed in Step 2. In other words, we have the two following cases:

Step 3a: $a|c$**.**

If $a|c$, then $c = ak$ for some k and we perform the following multiplication:

$$\begin{pmatrix} 1 & 0 \\ -k & 1 \end{pmatrix} \begin{pmatrix} a & 0 \\ c & d \end{pmatrix} = \begin{pmatrix} a & 0 \\ 0 & d \end{pmatrix}.$$

At this point, we can jump to Step 5 in the algorithm as we have turned the matrix into a diagonal matrix.

Step 3b: $a \nmid c$**.**

If $a \nmid c$, then set $e = \text{GCD}(a, c)$. Write $a = a'e$ and $c = c'e$ for some $a', c' \in R$, and then fix some $x, y \in R$ such that $a'x + c'y = 1$. We then perform the following multiplication:

$$\begin{pmatrix} x & y \\ -c' & a' \end{pmatrix} \begin{pmatrix} a & 0 \\ c & d \end{pmatrix} = \begin{pmatrix} e & dy \\ 0 & a'd \end{pmatrix}.$$

Note that this process may turn the entries that were previously zero into nonzero entries.

Step 4: Make both the top-right and bottom-left corners zero simultaneously.

Repeat Steps 2 and 3 until both the top-right and bottom-left corners are both zero. But how do we know that this process will terminate?

After each iteration of Steps 2b and 3b, the top-left corner is always a proper divisor of the previous value in that position. Since we are in a PID, which is Noetherian, the chain of ideals generated by these divisors must stabilize. (In Theorem 4.3.2, we will look at an example of a matrix in a domain which is an elementary divisor domain but not a PID. Step 4 will never terminate in that example.)

Step 5: Make the matrix be in Smith normal form.

At this point, we have a diagonal matrix, but we still need the top-left entry to divide the bottom-right entry. Later in this paper, we prove in Lemma 1.3.3 that any 2×2 diagonal matrix in a Bézout domain can be put into Smith normal form. This lemma also provides the specific matrices which turn the matrix into Smith normal form.

Remark 1.2.8. Note that each of these steps has very specific matrices which allow us to move to the next step. All of these except Step 4 are guaranteed to work in a generic Bézout domain, so the problem of showing that all Bézout domains are elementary divisor domains can actually be reduced to showing that any triangular 2×2 matrix can reduce down to a diagonal matrix. We will later examine what can occur in this algorithm when our domain is an elementary divisor domain rather than a PID.

It has been proven that there are examples of Bézout rings that are not elementary divisor rings (see [3]). The examples there are the only examples of Bézout rings that are not elementary divisor rings that I could find in the literature. Those Bézout rings are constructed as the rings of continuous functions from specific types of topological spaces to the real numbers. If the extra condition that this ring be an integral domain is also imposed, then the resulting ring is just the ring of constant functions, which is isomorphic to $\mathbb{R}$ and is thus a field. So, without some major reconstruction and using a different class of functions from topological spaces to some other space, the techniques do not seem to yield an easy avenue to constructing a Bézout *domain* that is not an elementary divisor domain. This will be expounded more in Section 1.3.

1.3 More Results on Domains

It has been proven by Kaplansky [7] that a Bézout domain is an elementary divisor domain if and only if it satisfies:

For all $a, b, c \in R$ with $\langle a, b, c\rangle = R$, there exist $p, q \in R$ such that $\langle pa, pb + qc\rangle = R$.

In other words, for all $a, b, c \in R$ with $1 = ax + by + cz$ for some $x, y, z \in R$, there exists $p, q \in R$ with $\langle pa, pb+qc\rangle = R$. This condition will be referred to as the *Kaplansky condition* when referenced later.

Theorem 1.3.1. *Let D be a domain. Then $D[a, b, c, x, y, z : ax+by+cz = 1] = D[a, b, c, x, y, z]/\langle ax+by + cz = 1\rangle$ is also an integral domain.*

Proof. Define

$$R := D[a, b, c, x, y, z : ax + by + cz = 1] \cong D[a, b, c, x, y, z]/\langle ax + by + cz - 1\rangle.$$

Fix $S = D[a, b, c, x, y]$. Note that by Corollary 1.1.13, S is a domain. Since S is a domain, we can then set $Q = \mathrm{Frac}(S)$ to be its field of fractions by Lemma 1.1.14.

Define $f : S[z] \to Q$ by taking $f|_S$ to be the identity map of S, and by taking

$$f(z) = c^{-1}(1 - ax - by),$$

and then by extending naturally using the universal property of polynomial rings to make f a ring homomorphism. Clearly the element $w := ax + by + cz - 1$ is in the kernel of this homomorphism. To show that R is a domain, it suffices to show that everything in the kernel is a multiple of w. To see this, if $\ker f = \langle w\rangle$, then by the first isomorphism theorem for rings [2, Theorem 7, Section 7], $R \cong f(D[a, b, c, x, y, z]) \subseteq Q$. Since R is then isomorphic to a subset of a field, there are no nonzero zero-divisors in R.

To see that $\ker f = \langle w\rangle$, let $s_0 + s_1z + s_2z^2 + \cdots + s_nz^n \in S[z] \cap \ker(f)$. After subtracting multiples of w, we may remove any instances of cz. Thus we may assume that $s_1, \ldots, s_n \in D[a, b, x, y]$ and assume by way of contradiction that $s_n \neq 0$. Since $s_0 + s_1z + \cdots + s_nz^n \in \ker(f)$, we have $f(s_0+s_1z+\cdots+s_nz^n) = 0$. But a direct computation of $f(s_0+s_1z+\cdots+s_nz^n)$

yields

$$s_0 + s_1 c^{-1}(1 - ax - by) + \cdots + s_n c^{-n}(1 - ax - by)^n \in D[a, b, x, y][c, c^{-1}]$$

which is not zero by considering the terms of lowest degree in c, a contradiction. Hence, f is a ring homomorphism from $S[z] \to Q$ with kernel $\langle w \rangle$. Thus, $S[z]/\langle w \rangle \cong R$ is a domain. $\square$

A similar condition to Kaplansky's condition is found in the following unresolved question:

Question 1.3.2. Suppose R is a Bézout domain which satisfies the following property: given α, β, γ with $\mathrm{GCD}(\alpha, \beta) = 1$ and $\gamma \neq 0$, then there exists some δ such that $\mathrm{GCD}(\alpha + \delta\beta, \gamma) = 1$. Is R always an elementary divisor domain?

In their 1974 paper [10], Larsen et al. showed that when looking at elementary divisor domains, we do not need to look at matrices of arbitrary size, but we can focus on the 2×2 matrices. For completeness, I will provide the proof below. To do so, we will need the following lemmas.

Larsen et al. stated the first lemma below without proof (c.f. 3.1 in [10]).

Lemma 1.3.3. *If R is a Bézout domain, then for every diagonal 2×2 matrix $A \in M_2(R)$, there exist invertible matrices $P, Q \in \mathrm{GL}_2(R)$ such that $PAQ = D$ with $D = \begin{pmatrix} d_{1,1} & 0 \\ 0 & d_{2,2} \end{pmatrix}$ and $d_{1,1} | d_{2,2}$.*

Proof. Let R be a Bézout domain, and let $A = \mathrm{diag}(a, b) \in M_2(R)$. Since R is a Bézout domain, we know that there are elements $r, s, x, y \in R$ such that $a = dr$, $b = ds$, and $rx + sy = 1$.

Let $P = \begin{pmatrix} x & 1 \\ -sy & r \end{pmatrix}$ and $Q = \begin{pmatrix} 1 & s \\ y & -rx \end{pmatrix}$. Note that

$$PAQ = \begin{pmatrix} x & 1 \\ -sy & r \end{pmatrix} \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix} \begin{pmatrix} 1 & s \\ y & -rx \end{pmatrix} = \begin{pmatrix} d & 0 \\ 0 & d(-rs) \end{pmatrix}.$$

Furthermore, straightforward matrix multiplication shows that $P^{-1} = \begin{pmatrix} r & -1 \\ sy & x \end{pmatrix}$ and that $Q^{-1} = \begin{pmatrix} rx & s \\ y & -1 \end{pmatrix}$, showing that P and Q are invertible, thus concluding the proof. $\square$

Lemma 1.3.4 ([7, Theorem 5.1]). *If all 2×1 and 2×2 matrices over a ring R admit diagonal reduction, then all matrices admit diagonal reduction and R is an elementary divisor ring.*

Proof. Let A be an $m \times n$ matrix. If A admits diagonal reduction and $PAQ = D$, then $Q^T A^T P^T = D^T$ shows that A^T admits diagonal reduction. Thus, it suffices to handle the case when $m \geq n$. We may suppose by induction that diagonal reduction is possible for all smaller m, and for the given m if n is smaller. It is to be observed that m is at least 3. Write A_1 for the first row of A and A_2 for the remaining $m-1$ rows.

Since A_2 is a smaller dimensional matrix, we can find invertible matrices P_1, Q_1 such that $B = P_1 A_2 Q_1 = \text{diag}(x, \ldots)$ is in Smith normal form. Then also

$$C = \begin{pmatrix} 1 & 0 \\ 0 & P_1 \end{pmatrix} \begin{pmatrix} A_1 \\ A_2 \end{pmatrix} Q_1 = \begin{pmatrix} A_1 Q_1 \\ B \end{pmatrix}.$$

Now write D for the first two rows of C, and E for the remaining rows. Applying induction again we have $F = P_2 D Q_2 = \text{diag}(y, \ldots)$ in Smith normal form, and then

$$H = \begin{pmatrix} P_2 & 0 \\ 0 & I_{m-2} \end{pmatrix} \begin{pmatrix} D \\ E \end{pmatrix} Q_2 = \begin{pmatrix} F \\ EQ_2 \end{pmatrix}.$$

Now y is a divisor of all the elements of F, and since $D = P_2^{-1} F Q_2^{-1}$, then y is also a divisor of all elements of D; in particular y is a divisor of x since x is one of the elements of D. The elements of EQ_2 are linear combinations of those of E, and hence they are divisible by x and thus also by y.

Thus y is a divisor of every element of H. We may now use elementary transformations to eliminate the first column of H and we reach

$$\begin{pmatrix} y & 0 \\ 0 & K \end{pmatrix}$$

where y is still a divisor of every element of K. Applying our inductive hypothesis to K, we complete the reduction. □

Theorem 1.3.5 ([10, Corollary 3.7]). *A ring R is an elementary divisor ring if and only if every 2×2 matrix over R is equivalent to a diagonal matrix.*

Proof. If R is an elementary divisor ring, then the definition immediately implies that 2×2

matrix over R is equivalent to a diagonal matrix.

Now suppose that every 2×2 matrix over R is equivalent to a diagonal matrix. In order to use Lemma 1.3.4, we only need to show that all 2×1 matrices over R admit diagonal reduction. Let $a, b \in R$ be arbitrary. Let M' and M be the matrices

$$M' = \begin{pmatrix} a \\ b \end{pmatrix} \text{ and } M = \begin{pmatrix} a & 0 \\ b & 0 \end{pmatrix}.$$

Since M is equivalent (by hypothesis) to a diagonal matrix, there are $P, Q \in \mathrm{GL}_2(R)$ such that PMQ is in Smith normal form. In other words, there are $p_i, q_i, d, e \in R$ (with $i \in \{1, 2, 3, 4\}$) such that $d|e$ and

$$\begin{pmatrix} p_1 & p_2 \\ p_3 & p_4 \end{pmatrix} \begin{pmatrix} a & 0 \\ b & 0 \end{pmatrix} \begin{pmatrix} q_1 & q_2 \\ q_3 & q_4 \end{pmatrix} = \begin{pmatrix} d & 0 \\ 0 & e \end{pmatrix}.$$

Direct computation of the left-hand side results in

$$\begin{pmatrix} (p_1a + p_2b)q_1 & (p_1a + p_2b)q_2 \\ (p_3a + p_4b)q_1 & (p_3a + p_4b)q_2 \end{pmatrix} = \begin{pmatrix} (p_1a + p_2b)q_1 & 0 \\ 0 & (p_3a + p_4b)q_2 \end{pmatrix}.$$

Since R is a domain and $(p_3a + p_4b)q_1 = 0$, either $p_3a + p_4b = 0$ or $q_1 = 0$. If $q_1 = 0$, then PMQ has a zero in the top-left entry and thus $PMQ = 0$. Therefore, $(p_3a + p_4b)q_2 = 0$ and either $p_3a + p_4b = 0$ or $q_2 = 0$. If $q_2 = 0$, then the matrix Q had a zero row and is not invertible, a contradiction. Therefore, whether or not $q_1 = 0$, we have $p_3a + p_4b = 0$. Thus,

$$\begin{pmatrix} p_1 & p_2 \\ p_3 & p_4 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} \begin{pmatrix} 1 \end{pmatrix} = \begin{pmatrix} p_1a + p_2b \\ 0 \end{pmatrix}$$

and M' is equivalent to a diagonal matrix.

Therefore, by Lemma 1.3.4, R is an elementary divisor ring. □

Therefore, any question of whether a specific Bézout domain is an elementary divisor domain or not centers on the question of whether a 2×2 matrix over that Bézout domain is equivalent to a diagonal matrix.

Remark 1.3.6. The question about whether or not every Bézout *ring* is an elementary divisor *ring* has been answered in the negative. In a 1956 paper by Gillman and Henriksen

[3], they prove that there are Bézout rings which are not elementary divisor rings. The key results are summarized below with notation standardized to the rest of this paper.

Before we are able to see the process by which Gillman and Henriksen found a counterexample, we need the following definitions.

Definition 1.3.7. A ring R is *K-Hermite* if every rectangular matrix over R can be reduced to lower-triangular form. That is, R is K-Hermite if for any rectangular matrix $B \in M_{m,n}(R)$, there exists $Q \in \mathrm{GL}_n(R)$ such that BQ is lower-triangular.

(There have been several slightly different iterations of what a Hermite ring should be. For clarity, I will follow the notation in [9] and call this type of ring a K-Hermite ring after Kaplansky [7] who first dealt with these rings in 1949.)

Definition 1.3.8. If X is a completely regular Hausdorff topological space, let $C(X)$ denote the ring of all continuous functions from X to $\mathbb{R}$. If $C(X)$ is a Bézout ring, we will call X a *Bézout space*. (In [3] it is called an F-space.)

Definition 1.3.9. With X and $C(X)$ as above, if $C(X)$ is K-Hermite, we will call X a *K-Hermite space*. (In [3] it is called a T-space.)

Definition 1.3.10. With X and $C(X)$ as above, if $C(X)$ is an elementary divisor ring, we will call X an *elementary divisor space*. (In [3] it is called an D-space.)

The process by which Gillman and Henriksen were able to find an example of a Bézout ring that is not an elementary divisor ring is to find conditions on the topological space X which make it either a Bézout space, a K-Hermite space, or an elementary divisor space. They then find specific topological spaces X through Stone-Čech compactifications of subsets of $\mathbb{R}^2$ which have specific properties that either satisfy or violate the conditions equivalent to being the different types of spaces. They then construct a Bézout ring which is not a K-Hermite ring [3, Example 3.4] and a ring which is both Bézout and K-Hermite, but not an elementary divisor ring [3, Example 4.11].

It is a difficult problem to modify the arguments used in [3] to find a Bézout domain which is not an elementary divisor domain. For example, imposing the condition that $C(X)$ also be a domain is too restrictive, as the only continuous functions that would then be in $C(X)$ are the constant functions (and are isomorphic to $\mathbb{R}$, a field). To see that, let $f \in C(X)$ be nonconstant. Then there are $x, y \in X$ such that $f(x) = a \neq b = f(y)$ for some $a, b \in \mathbb{R}$. Without loss of generality, assume $a < b$. Then there is some $c \in \mathbb{R}$ such that $a < c < b$ and continuous functions $\phi, \psi : \mathbb{R} \to \mathbb{R}$ such that (1) ϕ restricted to $(-\infty, c]$ is the constant 0 function and is nonzero on (c, ∞), and (2) ψ restricted to $(-\infty, c)$ is nonzero and is the constant 0 function on $[c, \infty)$. Then $\phi \circ f$ and $\psi \circ f$ are nonzero continuous functions from X to $\mathbb{R}$ such that $(\phi \circ f)(\psi \circ f) = 0$.

If $C(X)$ is a Bézout ring, there may be a subset of $C(X)$ that is a Bézout domain. The subset of all constant functions is isomorphic to $\mathbb{R}$, which is a field and thus a Bézout domain. However, trying to find a subset that is a domain but not a field is challenging, since many simple subsets of $C(X)$ are not even rings. The following example demonstrates what can go wrong if we try to pass to simpler subsets of $C(X)$.

Example 1.3.11. If we define $C_0(X)$ to be the subset of all continuous functions with finitely many zeros, then $C_0(X) \cup \{0\}$ is closed under multiplication and has no nonzero zero-divisors. However, it is not a ring since it is not closed under addition. If $X = \mathbb{R}$, then if $f : \mathbb{R} \to \mathbb{R}$ is given by

$$f(x) = \begin{cases} -1 & x \in (-\infty, -1] \\ x & x \in [-1, 1] \\ 1 & x \in [1, \infty) \end{cases}$$

and $g(x) = 1$ for all x, then both $f, g \in C_0(X) \cup \{0\}$. However, $f + g \notin C_0(X) \cup \{0\}$ since $f + g \neq 0$ but there are infinitely many zeros.

These examples of f, g also cause the following two subsets to also not be rings. If $C_1(X)$ is the set of continuous functions with no open sets in their zero sets, then $C_1(X) \cup \{0\}$ is not a ring. Also, if $C_2(X)$ is the set of continuous functions with no limit points in their zero sets, then $C_2(X) \cup \{0\}$ is not a ring.

It is not obvious whether there even should be a Bézout domain that is not an elementary divisor domain. For example, [3, Example 3.4] establishes that not all Bézout rings are K-Hermite rings, however every Bézout domain is K-Hermite, which we prove below. Thus at least one of the results which is true in the case where R is a ring cannot be true in the case where R is a domain.

Lemma 1.3.12. *All Bézout domains are K-Hermite, i.e., if R is a Bézout domain, then every square matrix over R can be reduced to triangular form.*

Proof. Let A be an $n \times n$ matrix whose entries $a_{i,j} \in R$, with R a Bézout domain. We will do induction on the row k and the column l. Let $a = a_{k,k}$ and $b = a_{k,l}$. Since R is a Bézout domain, there are elements $x, y, a', b', d \in R$ such that $a'x + b'y = 1$, $a'd = a$, and $b'd = b$. Let $E = (e_{i,j})$ be the $n \times n$ matrix which is the identity matrix except that we replace the following four entries: $e_{k,k} = x$, $e_{k,l} = -b'$, $e_{l,k} = y$, and $e_{l,l} = a'$. The determinant of this matrix is still 1, so it is still invertible.

Our base case is $k = 1$ and $l = 2$. Here $A = AE$ is a matrix where the new $(1,1)$-entry is d, and the $(1,2)$-entry is 0. Note that for fixed k, as l increases, the entries that were 0 in the k-th row remain 0 as we repeat this multiplication. Then when we increase k by 1 and set $l = k+1$, then multiplication here leaves all the 0s in the rows above as 0. Repeating this process a finite number of times turns all the entries above the diagonal to 0, so R is K-Hermite. □

It turns out that the converse is also true.

Theorem 1.3.13. *A domain R is Bézout if and only if it is K-Hermite.*

Proof. If R is Bézout, then by Lemma 1.3.12, R is also K-Hermite.

Conversely, now assume that R is a K-Hermite domain. Let $a, b \in R$. Since R is K-Hermite, there is an invertible $Q \in \mathrm{GL}_2(R)$ such that

$$\begin{pmatrix} a & b \end{pmatrix} Q = \begin{pmatrix} d & 0 \end{pmatrix}.$$

Therefore, d is an R-linear combination of a and b, and if we multiply both sides on the right by $Q^{-1} \in \mathrm{GL}_2(R)$, then we see that both a and b are multiples of d. Therefore, $\langle a, b \rangle = \langle d \rangle$ and R is Bézout. □

For many more equivalences involving K-Hermite rings and Bézout rings, I recommend going to Section 1.4 in T. Y. Lam's book "Serre's Problem on Projective Modules" (see [9]).

Many of the common examples of Bézout domains are in fact known to also be elementary divisor domains. For example, the ring of all entire functions is a standard example of a Bézout domain that is not a PID, but it is still an elementary divisor domain (see Corollary 2.2.6). The other classic example is the ring of algebraic integers, which is also an elementary divisor domain (see Theorems 2.2.7 and 2.2.8).

Chapter 2. Notable Results from Existing Literature

2.1 Existing Literature on Bézout Domains

We have now established what a Bézout domain is, as well as having seen that different ways of classifying them has led to useful results in the case where we are only looking at Bézout rings. We may wish to find different ways of classifying Bézout domains other than via the definition. One way to characterize a Bézout domain in terms of "simpler" conditions is the following theorem from the literature (from the paper by P. M. Cohn [1]; in this paper, the rings are all implicitly assumed to be domains).

Theorem 2.1.1 ([1, Theorem 2.8]). *A domain R is a Bézout domain if and only if R is integrally closed, every element is primal, and for all finitely generated ideals I, J, K of the ring R if $IJ = IK$ then either $J = K$ or $I = 0$.*

These conditions that were imposed upon the domain to show that it is Bézout have specific names.

Definition 2.1.2. A *Schreier domain* is an integrally closed domain in which every element is primal. An *integrally closed domain* is one where the integral closure of its field of fractions is itself. An element r in a commutative ring R is *primal* if whenever $r|ab$, with $a, b \in R$, then there exist elements $s, t \in R$ such that $r = st$ and $s|a$ and $t|b$.

Definition 2.1.3. A *Prüfer domain* is a domain where for all finitely generated ideals I, J, K of the ring R, if $IJ = IK$ then either $J = K$ or $I = 0$.

Therefore, another way to describe a Bézout domain is it is a domain that is both Schreier and Prüfer.

Remark 2.1.4. There are many equivalent ways to define a Prüfer domain, such as "a Prüfer domain is a domain in which every nonzero finitely generated ideal is invertible". Here, an ideal I of R being invertible means that $I \cdot I^{-1} = R$ where $I^{-1} = \{r \in \text{Frac}(R) : rI \subseteq R\}$. For some equivalent ways of defining a Prüfer domain, see chapter IV of [4]. In fact, Theorem 23.4 in [4] proves that every Prüfer domain is integrally closed.

A useful result that we will use later, involving Prüfer domains, is the following lemma. (This and other useful results involving Prüfer domains can be found in [6] and [4].)

Lemma 2.1.5. *Let R be a Prüfer domain. Let $a_1, \ldots, a_n \in R$ not be all zero. Let $d = \text{GCD}(a_1, \ldots, a_n)$. Then there are elements $a_{i,j} \in R$ for $i \in \{2, \ldots, n\}$ and $j \in \{1, \ldots, n\}$ such that*

$$\det \begin{pmatrix} a_1 & a_2 & \cdots & a_n \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & & \vdots \\ a_{n,1} & a_{n,2} & \cdots & a_{n,n} \end{pmatrix} = d.$$

Since Theorem 2.1.1 says that a domain R is Bézout exactly when it is both Schreier and Prüfer, then we can show certain domains are not Bézout domains by showing that they are either not Schreier or not Prüfer. For example, we can show that $\mathbb{Z}[\sqrt{5}]$ is not a Bézout domain by showing it is not a Prüfer domain.

Corollary 2.1.6. *The integral domain $\mathbb{Z}[\sqrt{5}]$ is not a Bézout domain. More precisely, it is neither a Prüfer domain nor a Schreier domain.*

Proof. Note first that $\mathbb{Z}[\sqrt{5}] \cong \mathbb{Z}[x]/\langle x^2-5\rangle$. Since x^2-5 is irreducible in $\mathbb{Z}[x]$, then $\langle x^2-5\rangle$ is a prime ideal of $\mathbb{Z}[x]$, so $\mathbb{Z}[x]/\langle x^2-5\rangle$ is an integral domain. Furthermore, the polynomial $x^2-x-1 \in \mathbb{Z}[\sqrt{5}][x]$ does not have roots in $\mathbb{Z}[\sqrt{5}]$ (the roots are $(1\pm\sqrt{5})/2$, which are not in $\mathbb{Z}[\sqrt{5}]$). Thus, $\mathbb{Z}[\sqrt{5}]$ is not integrally closed, so by Theorem 2.1.1, $\mathbb{Z}[\sqrt{5}]$ cannot be a Bézout domain. Furthermore, since it is not integrally closed, it cannot be a Prüfer domain nor a Schreier domain. □

2.2 Existing Literature on Elementary Divisor Domains

Just as the properties of Prüfer domains and Schreier domains can help us determine if a domain is Bézout, there is a closely related type of domain that connects to elementary divisor domains. These domains are called *adequate domains*. The majority of the information on adequate domains comes from a paper by Olaf Helmer [6], where all rings in the paper are implicitly assumed to be domains.

Definition 2.2.1. Let R be an integral domain, with $a, b \in R$ and $a \neq 0$. If there is a factor a_1 of a such that $a = a_1a_2$ and $\mathrm{GCD}(a_1, b) = 1$, and for any nonunit factor a_3 of a_2 it holds that $\mathrm{GCD}(a_3, b) \neq 1$, then we will call the element a_1 a *relatively prime part of a with respect to b*, which we will denote by $RP(a, b)$. (Olaf Helmer proved this element is unique up to a unit [6, Lemma 5].)

A domain is an *adequate domain* if it is a Prüfer domain and for all a and b in the domain with $a \neq 0$, then $RP(a, b)$ exists. Implicitly, this means that any two elements have a GCD.

This definition will be used in the following theorem.

Theorem 2.2.2. *All adequate domains are elementary divisor domains.*

Proof. We know by Theorem 1.3.5 that the ring R is an elementary divisor ring if and only if every 2×2 matrix over R is equivalent to a diagonal matrix. Suppose now that R is an adequate domain. We will show that every 2×2 matrix over R is equivalent to a diagonal matrix.

Let $A \in M_2(R)$ with $A = (a_{ij})$. Let the rank of A be r. If $r = 0$, then A is the zero matrix and is already in Smith normal form.

If $r = 1$, then the two rows of A are linearly dependent in R. If we let

$$A = \begin{pmatrix} \alpha_1 & \beta_1 \\ \alpha_2 & \beta_2 \end{pmatrix},$$

then there exist relatively prime $r, s \in R$ such that $r\alpha_1 + s\alpha_2 = 0$ and $r\beta_1 + s\beta_2 = 0$. Since $\mathrm{GCD}(r, s) = 1$, there exist $t, u \in R$ such that $rt + su = 1$. Then

$$\begin{pmatrix} u & -t \\ r & s \end{pmatrix} \begin{pmatrix} \alpha_1 & \beta_1 \\ \alpha_2 & \beta_2 \end{pmatrix} = \begin{pmatrix} u\alpha_1 - t\alpha_2 & u\beta_1 - t\beta_2 & \\ r\alpha_1 + s\alpha_2 & \alpha_2 & \beta_2 \end{pmatrix} = \begin{pmatrix} u\alpha_1 - t\alpha_2 & u\beta_1 - t\beta_2 \\ 0 & 0 \end{pmatrix} = A'.$$

Let $a := u\alpha_1 - t\alpha_2$ and $b := u\beta_1 - t\beta_2$.

Since R is adequate, there is some $d = RP(a, b)$ such that $a = da_2$, $\mathrm{GCD}(d, b) = 1$, and $\mathrm{GCD}(a_3, b) \neq 1$ for all nonunits $a_3 | a_2$. If a_2 is not a unit, then $\mathrm{GCD}(a_2, b) = e \neq 1$, so $a_2 = ea_3$ and $b = eb_3$ and $\mathrm{GCD}(a_3, b_3) = 1$. Since $\mathrm{GCD}(d, b) = 1$ we have $\mathrm{GCD}(d, b_3) = 1$, so $\mathrm{GCD}(da_3, b_3) = 1$. Thus, there are $x, y \in R$ such that $1 = da_3x + b_3y$. Since the matrix

$$V_1 = \begin{pmatrix} x & -b_3 \\ y & da_3 \end{pmatrix}$$

has determinant $1 = da_3x + b_3y$, it is invertible. Therefore,

$$\begin{pmatrix} a & b \\ 0 & 0 \end{pmatrix} \begin{pmatrix} x & -b_3 \\ y & da_3 \end{pmatrix} = \begin{pmatrix} da_3e & b_3e \\ 0 & 0 \end{pmatrix} \begin{pmatrix} x & -b_3 \\ y & da_3 \end{pmatrix} = \begin{pmatrix} da_3ex + b_3ey & 0 \\ 0 & 0 \end{pmatrix}$$

and $IA'V_1$ is in Smith normal form.

If $r = 2$, then we need the following modification of [6, Theorem 1]. I will state the modification, but not prove it, as it is immediate from Helmer's work simplified to the 2×2 case.

Lemma 2.2.3 ([6, Theorem 1])**.** *Let R be an adequate domain, and let*

$$M = \begin{pmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{pmatrix} \in M_2(R).$$

If $\mathrm{GCD}(a_{1,1}, a_{1,2}, a_{2,1}, a_{2,2}) = 1$, *there exists* $t \in R$ *such that* $\mathrm{GCD}(a_{1,1}t + a_{2,1}, a_{1,2}t + a_{2,2}) = 1$.

We now continue the proof of Theorem 2.2.2. For the rank 2 matrix $A \in M_2(R)$, let $d = \text{GCD}(a_{i,j})$. Then we can write

$$A = dB = d\begin{pmatrix} b_{1,1} & b_{1,2} \\ b_{2,1} & b_{2,2} \end{pmatrix}$$

with $\text{GCD}(b_{i,j}) = 1$. Therefore, there is some $t \in R$ such that $\text{GCD}(b_{1,1}t+b_{2,1}, b_{1,2}t+b_{2,2}) = 1$. Hence, there are $x, y \in R$ such that $(b_{1,1}t + b_{2,1})x + (b_{1,2}t + b_{2,2})y = 1$.

Let $P, Q \in M_2(R)$ be the matrices

$$P = \begin{pmatrix} t & 1 \\ (xb_{1,1} + yb_{1,2})t - 1 & (xb_{1,1} + yb_{1,2}) \end{pmatrix}$$

and

$$Q = \begin{pmatrix} x & -(b_{1,2}t + b_{2,2}) \\ y & b_{1,1}t + b_{2,1} \end{pmatrix}.$$

Note that both have determinant 1, and thus are invertible. Then, since $PAQ = dPBQ$, we can compute the product to be

$$\begin{aligned}
&d\begin{pmatrix} t & 1 \\ (xb_{1,1} + yb_{1,2})t - 1 & (xb_{1,1} + yb_{1,2}) \end{pmatrix}\begin{pmatrix} b_{1,1} & b_{1,2} \\ b_{2,1} & b_{2,2} \end{pmatrix}\begin{pmatrix} x & -(b_{1,2}t + b_{2,2}) \\ y & b_{1,1}t + b_{2,1} \end{pmatrix} \\
=&d\begin{pmatrix} tb_{1,1} + b_{2,1} & tb_{1,2} + b_{2,2} \\ (xb_{1,1} + yb_{1,2})(tb_{1,1} + b_{2,1}) - b_{1,1} & (xb_{1,1} + yb_{1,2})(tb_{1,2} + b_{2,2}) - b_{1,2} \end{pmatrix}\begin{pmatrix} x & -(b_{1,2}t + b_{2,2}) \\ y & b_{1,1}t + b_{2,1} \end{pmatrix} \\
&=d\begin{pmatrix} (tb_{1,1} + b_{2,1})x + (tb_{1,2} + b_{2,2})y & 0 \\ (xb_{1,1} + yb_{1,2})((b_{1,1}t + b_{2,1})x + (b_{1,2}t + b_{2,2})y) - xb_{1,1} - yb_{1,2} & b_{1,1}b_{2,2} - b_{1,2}b_{2,1} \end{pmatrix} \\
&=d\begin{pmatrix} (tb_{1,1} + b_{2,1})x + (tb_{1,2} + b_{2,2})y & 0 \\ (xb_{1,1} + yb_{1,2}) - xb_{1,1} - yb_{1,2} & b_{1,1}b_{2,2} - b_{1,2}b_{2,1} \end{pmatrix} \\
&=d\begin{pmatrix} 1 & 0 \\ 0 & b_{1,1}b_{2,2} - b_{1,2}b_{2,1} \end{pmatrix}.
\end{aligned}$$

Note that since this last matrix is clearly in Smith normal form, we are done in the case where the rank of A is 2. Since these are all of the cases for 2×2 matrices, all adequate domains are elementary divisor domains. □

Question 2.2.4. While we know that all adequate domains are elementary divisor domains, the reverse question remains. Are all elementary divisor domains adequate domains? And if not, what is an example of an elementary divisor domain which is not an adequate domain?

Remark 2.2.5. Note that the three types of domains we have listed listed (Schreier, Prüfer, and adequate) lead to one possible avenue of future research. In trying to determine if there is a Bézout domain that is not an elementary divisor domain, one could look at the properties of Schreier domains which are not adequate domains. If the properties of a Schreier domain along with those of a Prüfer domain somehow also force the existence of $RP(a, b)$, then all Bézout domains would be elementary divisor domains.

A corollary of Theorem 2.2.2 is the following:

Corollary 2.2.6. *The ring of entire functions is an elementary divisor domain.*

Proof. By Theorem 2.2.2, in order to show that the ring of entire functions is an elementary divisor domain, we need only to show that it is an adequate domain.

Let f and g be entire functions with $f \neq 0$. Let Z_f be the zero multiset of f. Let Z'_f and Z'_g be the zero sets of f and g respectively, not counting multiplicities. Let Z_k be the multiset of all elements in $Z'_f \cap Z'_g$, but with the multiplicities that appear in Z_f. Let f_2 be the entire function (up to a unit) guaranteed by the Weierstrass factorization theorem with zero multiset Z_k, and let f_1 be the entire function such that $f = f_1 f_2$. By construction, f_1 has no nonunit factors in common with g, so $(f_1, g) = 1$. Furthermore, any nonunit factor of f_2 shares a zero with g, so f_1 is a relatively prime part of f with respect to g. Therefore, the ring of entire functions is an adequate domain. □

See [14] for an early proof of this fact that does not use this approach, or see [5] for a more comprehensive treatment. The ring of entire functions is an important example of a Bézout domain that is not a PID, and any proof of the statement "all Bézout domains are elementary divisor domains" would need to apply to the ring of entire functions. On the other hand, any counterexample to the statement would need to be sufficiently different than this ring, and in particular be a nonadequate ring.

Along with the ring of entire functions, the other classic example of a Bézout domain that is not a PID is the ring of (all) algebraic integers. One very concise proof of this from the literature can be found in Irving Kaplansky's textbook "Commutative Rings" [8]. We following his proof with only minor modifications.

Theorem 2.2.7 ([8, Theorem 102])**.** *The ring of all algebraic integers is a Bézout domain.*

Proof. The algebraic integers are a subring of the complex numbers, and thus are a domain.

Let L be the algebraic closure of $\mathbb{Q}$, and let A be the integral closure of $\mathbb{Z}$ in L. Let $a, b \in A$ and $I = \langle a, b \rangle$ be an ideal in A. Then a and b generate a finite-dimensional extension L_0 of $\mathbb{Q}$. Let A_0 be the ring of integers in L_0 and $I_0 = \langle a, b \rangle$ in A_0. Since the class group of the ring of integers in an algebraic number field is finite, A_0 has a torsion class group (see [8]). Thus, some power of I_0 is principal; say $I_0^k = dA_0$ for $d \in A_0$. Let $c \in L$ be a k-th root of d. In the ring A_1 of integers in $L_0(c)$ we have $(I_0A_1)^k = \langle c^k \rangle$. Since A_1 is a Dedekind domain, any nonzero ideal is uniquely a product of prime ideals [8, Theorem 97]. Thus, $I_0A_1 = \langle c \rangle$. Hence the ideal I in A is also principal, and the ring of all algebraic integers is Bézout. □

So, is the ring of algebraic integers an example of a Bézout domain which is not an elementary divisor domain? The answer is no, the ring of algebraic integers is an elementary divisor domain. The proof of this fact is rather lengthy, see [11]. The proof given there is five pages long and uses results involving modules.

Theorem 2.2.8 ([11, Theorem 5])**.** *The ring of all algebraic integers is an elementary divisor domain.*

Chapter 3. Creating Bézout Domains

A first guess for a method to create a Bézout domain that is not an elementary divisor domain might be to look at a subring of a domain that has the desired property. However, when working with Bézout domains, it is important to remember that not all subrings of Bézout domains are Bézout. For example, Corollary 2.1.6 tells us that $\mathbb{Z}[\sqrt{5}]$ is not Bézout, even though it is a subring of the Bézout domain $\mathbb{R}$. Thus, a subring of a Bézout domain is not guaranteed to be a Bézout domain. In fact, even the intersection of two Bézout domains is not always a Bézout domain.

Theorem 3.0.1. *The intersection of Bézout domains need not be Bézout.*

Proof. Recall from Theorem 2.2.7 that the ring of algebraic integers $\mathbb{A}$ is a Bézout domain. Further, recall that every field is a PID and thus Bézout. In particular, the field $\mathbb{Q}(\sqrt{-5})$ is a Bézout domain. Since $-5 \equiv 3 \pmod 4$, we have $\mathbb{A} \cap \mathbb{Q}(\sqrt{-5}) = \mathbb{Z}[\sqrt{-5}]$. We now show that $\mathbb{Z}[\sqrt{-5}]$ is not a Bézout domain.

Suppose by way of contradiction that $\mathbb{Z}[\sqrt{-5}]$ is a Bézout domain. Then there must be some Bézout GCD of the elements 3 and $1+\sqrt{-5}$. Under the norm $N(a+b\sqrt{-5}) = a^2+5b^2$, then the norm of the GCD must divide the norms of 3 and $1+\sqrt{-5}$, which are 9 and 6 respectively. Then the GCD must have norm 1 or 3. Since $a^2+5b^2 \neq 3$ for any $a, b \in \mathbb{Z}$, the GCD must have norm 1, hence it is a unit. Therefore, up to a unit, the GCD is 1.

Thus, there are elements $a+b\sqrt{-5}$ and $c+d\sqrt{-5}$ such that

$$3(a+b\sqrt{-5}) + (1+\sqrt{-5})(c+d\sqrt{-5}) = 1.$$

Therefore, by multiplying the terms out, we get that $3a+c-5d = 1$ and $(3b+c+d)\sqrt{-5} = 0$. Solving for c we get $c = -3b-d$, and then substituting that in for the c in the first equality, we get $3a-6d-3b = 1$. But $a, b, c, d \in \mathbb{Z}$, so 1 is an integer multiple of 3, which is a contradiction. Therefore, there is no Bézout GCD in $\mathbb{Z}[\sqrt{-5}]$, so it cannot be a Bézout domain. Therefore, the intersection of Bézout domains need not also be a Bézout domain. □

Note that $\mathbb{Z}[\sqrt{-5}]$ is a GCD-domain in the more general notion of a GCD mentioned after Definition 1.1.10, but not when we are using Bézout-type GCDs.

Therefore, rather than beginning with a Bézout domain and then looking at subrings that have the desired property, we might make some progress if we begin with a ring or domain with the desired properties and then try and add Bézout GCDs to extend the ring to a Bézout domain.

3.1 Constructing a Bézout GCD

Now that we have seen a few examples of Bézout domains, as well as domains which display some desired traits yet are not Bézout domains, we will now discuss ways to create a Bézout domain from a domain that is not yet Bézout.

Let R be an integral domain. Let $a, b \in R$ be two elements that do not have a GCD in R. Note that this implies that both a and b are nonzero and nonunits.

We want to know if there is some minimal/free construction of a ring which does these three things:

(1) it contains R as a unital subring,

(2) it is an integral domain, and

(3) it contains a GCD for a and b.

One might consider starting with the ring

$$S_1 = R[g, c, d, r, s : a = gc, b = gd, ra + sb = g].$$

The relations force g to divide both a and b, and to be a linear combination of them, so it meets criteria (3) of being a GCD. However, this ring fails condition (2) as there exist nonzero zero-divisors in S_1. For example, since $a = gc$, $b = gd$, and $ra + sb = g$, we get

$$g = ra + sb = rgc + sgd \implies g(1) = g(rc + sd) \implies g(1 - rc - sd) = 0.$$

Since $a, b \in R$ are nonzero, g cannot be zero. Since we do not necessarily have $rc + sd = 1$, there are zero-divisors in this ring.

However, the computation leading to a zero-divisor does show the following. Suppose that a ring T satisfies conditions (1), (2), and (3). Furthermore, suppose that we have $\mathrm{GCD}_T(a, b) = g \in T$. If we take $c \in T$ to be the unique element satisfying $a = gc$ and similarly take $d \in T$ to be the unique element satisfying $b = gd$, then $\mathrm{GCD}_T(c, d) = 1$, so we can write $rc + sd = 1$ for some (generally nonunique) elements $r, s \in T$. Note that we also have the equality $bc = gdc = gcd = ad$. Moreover, we can solve for the variable g as $g = ra + sb$.

So perhaps a better place to start is the ring

$$S_2 = R[c, d, r, s : bc = ad, rc + sd = 1].$$

Notice that in this ring we have

$$(ra + sb)c = rac + s(bc) = rac + sad = a(rc + sd) = a.$$

Similarly, $(ra + sb)d = b$. So $ra + sb$ is a common divisor of both a and b in S_2. Moreover, $ra + sb$ is clearly a linear combination of a and b. Thus the ring S_2 satisfies condition (3). The ring S_2 is also much easier to describe than the ring S_1.

However, even though S_2 avoids the zero-divisors we found in S_1, it is still possible that S_2 has nontrivial zero-divisors. To give a simple example, suppose a and b have a common divisor $e \in R$ which is not a GCD. The equality $bc = ad$ can then be rewritten as

$$e\left(\frac{b}{e}c - \frac{a}{e}d\right) = 0.$$

Thus to guarantee condition (2) holds, we would need to add the relation $(b/e)c = (a/e)d$ as a new defining condition to S_2, for each common divisor $e \in R$ of a and b. In fact, we can describe other possible nontrivial zero-divisors as follows.

Let $x \in S_2$, and write x as a sum of monomials in the variables c, d, r, s with coefficients from R. Let m be a nonnegative integer. If m is large enough, we can write $d^m x$ as a sum of monomials in just the variables c, d, r (with no instances of s) by repeatedly using the equality $sd = 1 - rc$. Similarly, let n be a nonnegative integer. If n is large enough, we can write $b^n d^m x$ as a sum of monomials in just the variables d, r (with no instances of s or c) by repeatedly using the equality $bc = ad$.

When written this way, if $b^n d^m x$ is zero, then we are forced to add the new relation $x = 0$, since both b and d are nonzero in any ring satisfying condition (1).

Remark 3.1.1. This result leads to some questions. First, does this process capture every new relation we need? Second, after we add all these new relations, are any additional relations forced? Third, writing x as we did above is not necessarily unique. Does that process of rewriting a multiple of x by powers of b's and d's in terms of just the variables d and r sometimes give us zero, and other times result in a nonzero expression?

We can answer these questions by noting that any commutative domain embeds in its field of fractions. Fortunately, this process does capture all of the needed relations, it doesn't force any new relations, and if the rewriting process doesn't end with zero, it will never yield zero under any rewriting. To see this, we consider the ring $R[r,d]$ which is a domain (see Corollary 1.1.13). Any sum of monomials in the variables in just the variables r and d in this ring is nonzero, except the zero expression. Now the ring we want is

$$S_3 = \big(R[r,d]\big)[adb^{-1}, (1 - radb^{-1})d^{-1}] \subseteq \mathrm{Frac}(R[r,d]).$$

Setting $c := adb^{-1}$ and $s := (1 - rc)d^{-1}$, we have the two relations $bc = ad$ and $rc + sd = 1$. Moreover, S is a ring satisfying all three of the desired conditions. Finally, as we multiply any $x \in S_3$ by powers of d's and b's to clear denominators, the only time x is zero is if and only if x becomes a zero expression in $R[r,d]$. Thus S_3 is the ring we want.

Remark 3.1.2. We could alternatively have described S_3 as being (isomorphic to) a subring of $\mathrm{Frac}(R[s,c])$ by symmetry considerations. We can also treat S_3 as a $\mathbb{Z}$-graded ring by giving c, d grades 1, giving r, s grades -1, and giving elements of R grade 0.

3.2 Freeness

Note that the way that we constructed S_3 will not work if a and b are arbitrary, since it requires both a and b to be nonzero. However, even the assumption that a and b do not have a GCD in R is not a sufficient assumption to prevent a unit multiple of $\mathrm{GCD}_{S_3}(a,b)$ from being in R. To see this, we look at the following example.

Example 3.2.1. Let $R = \mathbb{Z}[a, b : a^2 + ab + b^2 = 0]$. We will show that

(i) R is a domain,

(ii) a and b do not have a GCD in R,

(iii) in any ring T which satisfies conditions (1), (2), and (3) from Section 3.1, then $\text{GCD}_T(a,b) = a$ (up to a unit), and

(iv) T always includes the cube root of unity ab^{-1}.

In order to prove the statements claimed in this example, we will need the following lemma:

Lemma 3.2.2. *Let D be a UFD, and let $g \in D$ be an irreducible element of D. Then $D/\langle g\rangle$ is an integral domain.*

Proof. For notational convenience, let $I = \langle g\rangle$. Since I is an ideal of D, we know D/I is a ring. Since g is irreducible, it is not a unit, so we have $1 + I$ as the multiplicative identity in the ring D/I. Assume that $a + I, b + I \in D/I$ satisfy $(a + I)(b + I) = 0 + I$. Then $ab + I = 0 + I$ so $ab \in I$ as an element of D. Therefore, $ab = gk$ for some $k \in D$. Since g is irreducible, g is in the irreducible factorization of ab (up to a unit). Therefore, without loss of generality, $a = gl$ and so $a + I = 0 + I$. Therefore, there are no nonzero zero-divisors in D/I, so D/I is an integral domain. □

We now prove the statements claimed in Example 3.2.1.

Lemma 3.2.3. *The ring $R = \mathbb{Z}[a, b : a^2 + ab + b^2 = 0]$ is a domain.*

Proof. We note that $\mathbb{Z}[a, b]$ is a UFD (see Sections 8.3 and 9.3 of [2] for more information on UFD's), as well as the fact that $a^2 + ab + b^2$ is irreducible in $\mathbb{Z}[a, b]$. By Lemma 3.2.2, $\mathbb{Z}[a, b]/\langle a^2 + ab + b^2\rangle \cong \mathbb{Z}[a, b : a^2 + ab + b^2 = 0] = R$ is a domain. □

Lemma 3.2.4. *The elements $a, b \in R$ do not have a Bézout GCD in R.*

Proof. Suppose by way of contradiction that there is a Bézout GCD of a and b in R. Since a and b are both of degree 1, the GCD is either 1 or a. If $\text{GCD}(a, b) = a$, then for some $b' \in R$, we have $0 = a^2 + ab + b^2 = a^2(1 + b' + (b')^2)$. Since both a and b are degree 1, b' is degree 0 and thus constant. Since there is no constant b' in $\mathbb{Z}$ such that $(1 + b' + (b')^2) = 0$, then $a^2 = 0$ and $a = 0$, a contradiction.

If $\mathrm{GCD}(a, b) = 1$, then there are $x, y \in R$ such that $ax+by = 1$. Then since $a^2+ab+b^2 = 0$, we have

$$1 = (ax+by)^2 = a^2x^2+2abxy+b^2y^2 = (-ab-b^2)x^2+2abxy+b^2y^2 = b(-ax^2-bx^2+2axy+by^2).$$

Therefore, b is a unit. Since $a^2 + ab + b^2 = 0$ is a homogeneous relation, the ring R is graded, with a and b having degree 1. Since b is degree 1, its inverse must be degree -1, which does not exist, which is a contradiction.

Therefore, the elements $a, b \in R$ do not have a Bézout GCD in R. □

Lemma 3.2.5. *In any ring T which satisfies conditions (1), (2), and (3) from Section 3.1, then* $\mathrm{GCD}_T(a, b) = a$.

Proof. Suppose that T is a domain with R as a unital subring and that there is a $d \in T$ such that $d = \mathrm{GCD}_T(a, b)$. Then there are $a', b', x, y \in T$ such that $a = a'd$, $b = b'd$, and $a'x + b'y = 1$. Since $a^2 + ab + b^2 = 0$, then $d^2((a')^2 + a'b' + (b')^2) = 0$. Since T is a domain, either $d = 0$ or $(a')^2 + a'b' + (b')^2 = 0$. Since $a, b \neq 0$, then $d \neq 0$, which then implies $(a')^2 + a'b' + (b')^2 = 0$.

Since $a'x + b'y = 1$, then (similarly to part of the proof of the last lemma)

$$1 = (a'x + b'y)^2 = b'(-a'x^2 - b'x^2 + 2a'xy + b'y^2)$$

and b' is a unit. Similarly, a' is a unit. Thus, $\mathrm{GCD}_T(a, b) = d$ where d is a unit multiple of a. Thus, $\mathrm{GCD}_T(a, b) = a$. □

Lemma 3.2.6. *The ring T always includes the cube root of unity ab^{-1}.*

Proof. Since we have $\mathrm{GCD}_T(a, b) = a \neq 0$, then $b = ab'$ and thus

$$0 = a^2 + ab + b^2 = a^2(1 + b' + (b')^2)$$

and $1 + b' + (b')^2 = 0$. Thus, b' is a unit with inverse $(b')^{-1} = -1 - b'$. Furthermore, since $1 + b' + (b')^2 = 0$, then b' must be a 3rd root of unity. Note that $ab^{-1} = aa^{-1}(b')^{-1} = (b')^{-1}$

which is also a cube root of unity. Therefore, T always contains the cube root of unity ab^{-1}. □

Thus, starting with the ring R in Example 3.2.1, then the ring S_3 is not a minimal extension of R where a and b have a GCD. We instead want the proper subring of S_3 generated by R and ab^{-1}. (By Lemma 3.2.6, the ring S_3 contains ab^{-1}. Starting with a different domain R with $a, b \in R$, we do not expect that $ab^{-1} \in S_3$.) To see that the subring is proper, we note that the free variable $r \notin \langle R, ab^{-1}\rangle \subseteq S_3$.

Question 3.2.7. Starting with any domain R, is there always a a unique subring of S_3 that is minimal with respect to satisfying conditions (1), (2), and (3) from Section 3.1?

Question 3.2.8. Can we find an example of a domain R which contains elements a, b, c, x, y, z which satisfies the following three properties?

(i) There is the relation $ax + by + cz = 1$.

(ii) The elements a, b, and c are not "elementary divisor elements" meaning that the matrix
$$A = \begin{pmatrix} a & b \\ 0 & c \end{pmatrix} \in M_2(R)$$
is such that there are no invertible matrices P and Q such that PAQ is in Smith normal form (see Definition 1.2.5).

(iii) After adding a GCD of a and b in the almost free way described in Section 3.1 to get
$$S_3 = \big(R[r, d]\big)[adb^{-1}, (1 - rc)d^{-1}],$$
then a, b, and c are then elementary divisor elements.

Finding such a ring is more challenging than it might appear at first glance. Part (ii) in particular is especially challenging as showing that there are not invertible matrices P and Q such that PAQ is in Smith normal form is more involved than simply showing that the SNF algorithm does not terminate, as we will see in Section 4.3.

3.3 Iteration

In the process outlined in Section 3.1, we gave a method to construct a GCD for two elements that did not have a GCD already. However, this process introduces new elements which may or may not have a GCD between them. In order to eventually end up with a Bézout domain where all elements have a GCD, we will need to iterate this process.

Question 3.3.1. Begin with $R = \mathbb{Z}[a, b, c, x, y][(1 - ax - by)c^{-1}]$. Choose two elements of R which do not have a Bézout GCD, and then adjoin a GCD in the almost free way described in Section 3.1. Repeat this process (perhaps transfinitely) for every pair of elements with no Bézout GCD until every pair of elements have a Bézout GCD. We know that the final ring (call it T) will be a Bézout domain, but what else can be shown? Is T a unique factorization domain? A Noetherian ring? A PID? An elementary divisor domain?

If we establish that T is Noetherian, then T is also a PID. Since all PIDs are UFDs and elementary divisor domains, then this would show that any Bézout domain constructed this way is also an elementary divisor domain. While we do not know if T is Noetherian, the following lemma can be used to show that the ring R is Noetherian. Since this is a well-known result for Noetherian rings, I state it without proof, although a straightforward proof can be found in Section 15 of [2]. (Note that there are similar results for Bézout and elementary divisor rings, which I will expound on in the next section in Theorems 3.4.1 and 3.4.3.)

Lemma 3.3.2. *If S is Noetherian and I is an ideal of S, then S/I is a Noetherian ring. Furthermore, the homomorphic image of a Noetherian ring is Noetherian.*

There is a natural surjective homomorphism $\varphi : \mathbb{Z}[a, b, c, x, y, z]/\langle 1 - ax - by - cz \rangle \to R$ where $z \mapsto (1 - ax - by)c^{-1}$. Thus R is the homomorphic image of a Noetherian ring, and thus is Noetherian.

Therefore, in determining whether T is Noetherian, we need to determine if adding GCD's preserves Noetherianity. Recall from Section 3.1 that if we want to introduce a GCD of two elements $a, b \in R$ such that $ra + sb = \mathrm{GCD}(a, b)$, then we extend to the ring

$$S_3 := \big(R[r, d]\big)[adb^{-1}, (1 - radb^{-1})d^{-1}].$$

If $g := \mathrm{GCD}(a, b)$ in S_3, then $a = gc$, $b = gd$, and $rc + sd = 1$. Since there is the natural surjective homomorphism $\psi : \big(R[r, d, c, s]/\langle ad-bc\rangle\big)/\langle 1-rc-sd\rangle \to S_3$ where $s \mapsto (1-rc)d^{-1}$ and $c \mapsto adb^{-1}$, and we know that $R[r, d, c, s]/\langle ad - bc\rangle$ is Noetherian, then Lemma 3.3.2 implies that S_3 is Noetherian.

While this does imply that each domain after adjoining finitely many GCDs is Noetherian, it does not imply that T is Noetherian, so our question from before remains unanswered.

Perhaps one avenue to show that Bézout domains and elementary divisor domains are different would be to begin with a domain in which there is a matrix

$$A = \begin{pmatrix} a & b \\ 0 & c \end{pmatrix} \in M_2(R)$$

such that there are no invertible matrices P and Q such that PAQ is in Smith normal form, and then go through the process of freely adding Bézout GCDs and iterating until we have a Bézout domain. This seems like it could be a promising avenue to show that Bézout domains and elementary divisor domains are different (assuming that is the case) without having to find a topological argument as was required for showing that Bézout rings were not elementary divisor rings (see [3]). However, it is not yet clear how it would be shown that the resulting domain would not be an elementary divisor domain, if that even were the case.

3.4 Creating Bézout Domains from other Bézout Domains

One way to construct a Bézout domain is to start with a Bézout ring and then mod out by a prime ideal.

Theorem 3.4.1. *If R is a Bézout ring, and $I \subseteq R$ is a prime ideal, then R/I is a Bézout domain.*

Proof. Since I is a prime ideal, we know that R/I is an integral domain. Let $\overline{a}, \overline{b} \in R/I$.

Suppose that $\overline{a}$ lifts to some $a \in R$ and $\overline{b}$ lifts to some $b \in R$. Since R is Bézout, then there is some $d = \mathrm{GCD}(a, b)$ such that $a = da_0$, $b = db_0$, and $d = ax + by$ for some $a_0, b_0, x, y \in R$. Then

$$\overline{d} = \overline{ax + by} = \overline{ax} + \overline{by} = \overline{a}\overline{x} + \overline{b}\overline{y},$$

$\overline{a} = \overline{d}\overline{a_0}$, and $\overline{b} = \overline{db_0}$. Therefore, $\overline{d}$ is a linear combination of $\overline{a}$ and $\overline{b}$ as well as a divisor of both. Therefore, $\overline{a}$ and $\overline{b}$ have a Bézout GCD in R/I. Since $\overline{a}, \overline{b}$ were arbitrary, then R/I is a Bézout domain. □

This theorem suggests that there might be a Bézout ring that is not an elementary divisor ring (perhaps constructed similarly to those constructed in [3]) where we could mod out by a prime ideal and be left with a domain which is not an elementary divisor domain. This begs the following question.

Question 3.4.2. Is there a Bézout ring R and a prime ideal I such that R/I is not an elementary divisor ring?

It turns out that any such Bézout ring would already need to fail to be an elementary divisor ring due to the following theorem.

Theorem 3.4.3. *If R is an elementary divisor ring, and $I \subseteq R$ is a prime ideal, then R/I is an elementary divisor domain.*

Proof. Since I is a prime ideal, we know that R/I is an integral domain. Let $\overline{A} \in M_2(R/I)$ be an arbitrary 2×2 matrix. We can lift $\overline{A}$ to some matrix $A \in M_2(R)$. Since R is an elementary divisor ring, there are invertible matrices $P, Q \in M_2(R)$ such that PAQ is in Smith normal form. Since P and Q are invertible, so are $\overline{P}$ and $\overline{Q}$. Thus, $\overline{PAQ}$ is also in Smith normal form. Since $\overline{A}$ was arbitrary, then R/I is an elementary divisor domain. □

Note that these two theorems together imply that finding a Bézout domain which is not an elementary divisor domain using this method would require the Bézout ring R to already fail to be an elementary divisor ring. We know that this is possible (see [3]), but the rings they produce are complicated and do not have obvious prime ideals that might lead to such a result.

Question 3.4.4. Is there a prime ideal I in one of the Bézout rings R that is not an elementary divisor ring, as found in [3], such that R/I is not an elementary divisor domain?

Another way to construct Bézout domains from other Bézout rings is to build up as in the following theorem.

Theorem 3.4.5. *Let I be a totally ordered index set. For each $i \in I$, let R_i be a Bézout domain. Assume that R_i is a subring of R_j for every $j > i$. Set $R := \cup_{i \in I} R_i$. Then R is a Bézout domain, and furthermore, if every R_i is an elementary divisor domain, then so is R.*

Proof. Since every R_i is a domain, each R_i contains an identity, and since the $\{R_i\}_{i \in I}$ are totally ordered, this identity is the same in all R_i, so R is a unital ring.

Suppose that $x, y \in R$ with $xy = 0$. Since $x, y \in R$, we have $x \in R_i$ and $y \in R_j$ for some i, j. Without loss of generality, let $i \leq j$. Then $x, y \in R_j$ and $xy = 0$ in R_j. Thus either x or y was 0 and R is a domain.

Lastly, let $a, b \in R$. Then there exist $i, j \in I$ such that $a \in R_i$ and $b \in R_j$. Without loss of generality, let $i \leq j$. Then $a, b \in R_j$. Since R_j is Bézout, the ideal $\langle a, b \rangle$ is generated by some $d \in R_j$. Therefore, $\langle a, b \rangle = \langle d \rangle$ in R, so R is a Bézout domain.

Let $M = (m_{i,j})$ be any finite matrix over R. Then there is some R_k such that all $m_{i,j} \in R_k$. Since R_k is an elementary divisor domain, there are matrices P, Q over R_k such that $PMQ = D$ where D is a diagonal matrix with the elementary divisor condition from Definition 1.2.5. Since P, Q are also valid matrices over R, then R must be an elementary divisor domain since M is arbitrary. □

Since there are not yet any known Bézout domains which are not elementary divisor domains, this method does not appear to directly lead to a clear way to construct a Bézout domain that is not an elementary divisor domain.

Since any matrix in an elementary divisor domain has a Smith normal form, one method to show a domain is an elementary divisor domain is to show that the SNF algorithm always terminates. However, as we shall see in the next chapter, showing that the SNF algorithm does not terminate is not a proof that the domain is not an elementary divisor domain.

Chapter 4. A Domain in which the Smith Normal Form Algorithm Never Terminates

4.1 Setting up the Process

In Algorithm 1.2.7, we described the standard SNF algorithm. Note that after one iteration of the algorithm, there is always a zero either in the top-right or bottom-left corner. Therefore, we will assume that there is already a zero in the bottom-left corner. If we introduce a transposition after making the top-right corner a zero, we can define the following map that removes the need for Step 3 (by essentially transforming it back to Step 2). Let R be a Bézout domain. Let $e_0, y_0, a_0 \in R$. If $y_0 = e_0 k$ for some $k \in R$, then

$$\varphi \begin{pmatrix} e_0 & y_0 \\ 0 & a_0 \end{pmatrix} = \begin{pmatrix} e_0 & 0 \\ 0 & a_0 \end{pmatrix}.$$

If $e_0 \nmid y_0$, then there are elements e_1, x_1, y_1, a_1, b_1 such that $e_0 = a_1 e_1$, $y_0 = b_1 e_1$, and $a_1 x_1 + b_1 y_1 = 1$. Then take

$$\varphi \begin{pmatrix} e_0 & y_0 \\ 0 & a_0 \end{pmatrix} = \begin{pmatrix} e_1 & a_0 y_1 \\ 0 & a_0 a_1 \end{pmatrix}.$$

(We could fix, once and for all, such elements for each triple (e_0, y_0, a_0) so that φ is unique.)

In this chapter, we focus on the following two interesting questions:

Question 4.1.1. Is is possible to construct a Bézout domain R such that for some matrix $A = \begin{pmatrix} e_0 & y_0 \\ 0 & a_0 \end{pmatrix}$ with $e_0, y_0, a_0 \in R$, there is no $n \in \mathbb{N}$ such that $\varphi^n(A)$ is diagonal? Furthermore, is constructing such a Bézout domain equivalent to showing that said Bézout domain is not an elementary divisor domain?

In the following sections of this chapter, we will show that there is a Bézout domain where $\varphi^n(A)$ is never diagonal for some matrix A, but that this construction is not equivalent to showing the Bézout domain is not an elementary divisor domain.

4.2 Computations of the SNF Algorithm in the Ring of Entire Functions

We have shown in Corollary 2.2.6 that the ring of entire functions is an elementary divisor domain. I will now show the first few iterations of the SNF algorithm on the matrix

$$A_0 = \begin{pmatrix} \frac{\sin(\pi z)}{\pi z} & \sqrt{\pi}(z-i)\frac{\sin(\pi z)}{\pi z(1-z^2)} \\ 0 & \pi z \end{pmatrix}.$$

Note that $e_0 := \sin(\pi z)/\pi z$ has roots at $z = \pm 1, \pm 2, \ldots$ and that

$$y_0 := \sqrt{\pi}(z-i)\frac{\sin(\pi z)}{\pi z(1-z^2)}$$

has roots at $z = i, \pm 2, \pm 3, \ldots$. Therefore, the top-left corner does not divide the top-right corner. Therefore, by Step 2b of Algorithm 1.2.7, there are elements e_1, a_1, b_1, x_1, y_1 such that $e_0 = a_1 e_1$, $y_0 = b_1 e_1$, and $a_1 x_1 + b_1 y_1 = 1$.

One choice for these elements is

$$\begin{aligned} e_1 :=& \frac{\sin(\pi z)}{\pi z(1-z^2)} \\ a_1 :=& 1 - z^2 \\ b_1 :=& \sqrt{\pi}(z-i) \\ y_1 :=& \frac{3}{4\sqrt{\pi}}(z+i)\frac{\sin(\pi z)}{\pi z(1-z^2)(1-z^2/4)} \\ x_1 :=& \frac{1 - b_1 y_1}{a_1} = \frac{1 - (3/4)(z^2+1)\frac{\sin(\pi z)}{\pi z(1-z^2)(1-z^2/4)}}{1-z^2}. \end{aligned}$$

The first four of these are easy to see as entire functions. However, x_1 might have a simple pole at $z = \pm 1$. Yet we have by l'Hopital's rule

$$\lim_{z \to \pm 1}\left(1 - \frac{3}{4}(z^2+1)\frac{\sin(\pi z)}{\pi z(1-z^2)(1-z^2/4)}\right) = 1 - \frac{3}{2}\lim_{z \to \pm 1}\frac{\pi\cos(\pi z)}{(\pi/4)\big(4 - 15x^2 + 5x^4\big)} = 1 - 1 = 0.$$

Thus at $z = \pm 1$, the function x_1 is analytic.

Therefore, the matrix

$$E_1 := \begin{pmatrix} x_1 & -b_1 \\ y_1 & a_1 \end{pmatrix} = \begin{pmatrix} \frac{1-(3/4)(z^2+1)\frac{\sin(\pi z)}{\pi z(1-z^2)(1-z^2/4)}}{1-z^2} & -\sqrt{\pi}(z-i) \\ \frac{3}{4\sqrt{\pi}}(z+i)\frac{\sin(\pi z)}{\pi z(1-z^2)(1-z^2/4)} & 1-z^2 \end{pmatrix}$$

is invertible. Hence, by Step 2b in Algorithm 1.2.7,

$$A_0E_1 = \begin{pmatrix} \frac{\sin(\pi z)}{\pi z} & \sqrt{\pi}(z-i)\frac{\sin(\pi z)}{\pi z(1-z^2)} \\ 0 & \pi z \end{pmatrix} \begin{pmatrix} \frac{1-(3/4)(z^2+1)\frac{\sin(\pi z)}{\pi z(1-z^2)(1-z^2/4)}}{1-z^2} & -\sqrt{\pi}(z-i) \\ \frac{3}{4\sqrt{\pi}}(z+i)\frac{\sin(\pi z)}{\pi z(1-z^2)(1-z^2/4)} & 1-z^2 \end{pmatrix}$$

$$= \begin{pmatrix} \frac{\sin(\pi z)}{\pi z(1-z^2)} & 0 \\ \pi z\frac{3}{4\sqrt{\pi}}(z+i)\frac{\sin(\pi z)}{\pi z(1-z^2)(1-z^2/4)} & \pi z(1-z^2) \end{pmatrix}.$$

We can then take the transpose of this matrix, and we again are in a situation where we can use Step 2b of Algorithm 1.2.7. We can then go through a similar process to find an invertible matrix E_2 and find the next matrix through the algorithm. If we say that $A_{n-1}E_n = A_n^T$ for each $n \in \mathbb{N}$, then here are choices for the first few E_n:

$$E_1 = \begin{pmatrix} \frac{1-(3/4)(z^2+1)\frac{\sin(\pi z)}{\pi z(1-z^2)(1-z^2/4)}}{1-z^2} & -\sqrt{\pi}(z-i) \\ \frac{3}{4\sqrt{\pi}}(z+i)\frac{\sin(\pi z)}{\pi z(1-z^2)(1-z^2/4)} & 1-z^2 \end{pmatrix}$$

$$E_2 = \begin{pmatrix} \frac{1+\frac{i}{6}(z+i)(z-4i)\frac{\sin(\pi z)}{(1-z^2)(1-z^2/4)(1-z^2/9)}}{1-z^2/4} & -\sqrt{\pi}z\frac{3}{4}(z+i) \\ \frac{2i}{9\sqrt{\pi}}(z-4i)\frac{\sin(\pi z)}{\pi z(1-z^2)(1-z^2/4)(1-z^2/9)} & 1-z^2/4 \end{pmatrix}$$

$$E_3 = \begin{pmatrix} \frac{1+\frac{7i}{1440\pi}(z-4i)(4z-9i)\frac{\sin(\pi z)}{(1-z^2/4)(1-z^2/9)(1-z^2/16)}}{1-z^2/9} & -\sqrt{\pi}z(1-z^2)\frac{-2i}{9}(z-4i) \\ \frac{7}{320\sqrt{\pi}}(4z-9i)\frac{\sin(\pi z)}{\pi z(1-z^2)(1-z^2/4)(1-z^2/9)(1-z^2/16)} & 1-z^2/9 \end{pmatrix}$$

$$E_4 = \begin{pmatrix} \frac{1-\frac{7i}{67400\pi}(4z-9i)(9z-64i)\frac{\sin(\pi z)}{(1-z^2/9)(1-z^2/16)(1-z^2/25)}}{1-z^2/16} & \frac{-7\sqrt{\pi}z(1-z^2)(1-z^2/4)(4z-9i)}{320} \\ \frac{8i}{1685\sqrt{\pi}}(9z-64i)\frac{\sin(\pi z)}{\pi z(1-z^2)(1-z^2/4)(1-z^2/9)(1-z^2/16)(1-z^2/25)} & 1-z^2/16 \end{pmatrix}.$$

Under these choices of E_n, we have the following matrices A_n:

$$A_1 = \begin{pmatrix} \frac{\sin(\pi z)}{\pi z(1-z^2)} & \frac{3}{4\sqrt{\pi}}(z+i)\frac{\sin(\pi z)}{(1-z^2)(1-z^2/4)} \\ 0 & \pi z(1-z^2) \end{pmatrix}$$

$$A_2 = \begin{pmatrix} \frac{\sin(\pi z)}{\pi z(1-z^2)(1-z^2/4)} & \frac{-2i}{9\sqrt{\pi}}(z-4i)\frac{\sin(\pi z)}{(1-z^2/4)(1-z^2/9)} \\ 0 & \pi z(1-z^2/4)(1-z^2/9) \end{pmatrix}$$

$$A_3 = \begin{pmatrix} \frac{\sin(\pi z)}{\pi z(1-z^2)(1-z^2/4)(1-z^2/9)} & \frac{7}{320\sqrt{\pi}}(4z-9i)\frac{\sin(\pi z)}{(1-z^2/9)(1-z^2/16)} \\ 0 & \pi z(1-z^2)(1-z^2/4)(1-z^2/9) \end{pmatrix}$$

$$A_4 = \begin{pmatrix} \frac{\sin(\pi z)}{\pi z(1-z^2)(1-z^2/4)(1-z^2/9)(1-z^2/16)} & \frac{8i}{1685\sqrt{\pi}}(9z-64i)\frac{\sin(\pi z)}{(1-z^2/16)(1-z^2/25)} \\ 0 & \pi z(1-z^2)(1-z^2/4)(1-z^2/9)(1-z^2/16) \end{pmatrix}.$$

These matrices A_n and E_n get rather cumbersome to write, but I claim that there are always choices for the E_n such that this process never terminates. I will prove this in Theorem 4.3.2 in the next section.

4.3 An Elementary Divisor Domain where the SNF Algorithm Never Terminates

There is an elementary divisor domain with a 2×2 matrix that does not reduce to a diagonal matrix under the standard Smith Normal Form algorithm. Namely, the domain is the ring of entire functions, and the matrix is A_0 from the previous section. Before proving this result, we need the following small technical lemma.

Lemma 4.3.1. *Let $f(z)$ be a line with a purely imaginary nonzero root. If we define g to be the unique line passing through the points $\left(n, \frac{d}{f(n)}\right)$ and $\left(-n, \frac{-d}{f(-n)}\right)$ for some $n \in \mathbb{R}-\{0\}$ and $d \in \mathbb{C}-\{0\}$, then $g(z)$ is also a line with a purely imaginary nonzero root.*

Proof. Since $n \neq 0$, the points $\left(n, \frac{d}{f(n)}\right)$ and $\left(-n, \frac{-d}{f(-n)}\right)$ are distinct. As f is a line with a purely imaginary root, we may write f as $f(z) = c(z - ai)$ for some $c \in \mathbb{C}-\{0\}$ and some $a \in \mathbb{R}-\{0\}$. Define $g(z)$ to be the line

$$g(z) = \frac{adi}{cn(n^2+a^2)}\left(z - \frac{n^2}{a}i\right).$$

Clearly, this line has a zero at $n^2 i/a$, which is purely imaginary as a and n are both nonzero and real. Since there is only one line passing through two distinct points, the proof is complete upon showing that g passes through the two specified points, namely that

$$g(n) = \frac{adi}{cn(n^2+a^2)}\Big(n - \frac{n^2}{a}i\Big) = \frac{ad}{cn(n^2+a^2)}\Big(\frac{n^2}{a} + ni\Big) = \frac{d(n+ai)}{c(n^2+a^2)} = \frac{d}{c(n-ai)} = \frac{d}{f(n)}$$

and

$$g(-n) = \frac{adi}{cn(n^2+a^2)}\Big(-n - \frac{n^2}{a}i\Big) = \frac{d(n-ai)}{c(n^2+a^2)} = \frac{-d}{c(-n-ai)} = \frac{-d}{f(-n)}.$$

Therefore, g has the stated properties. □

We now prove the main result.

Theorem 4.3.2. *There is an elementary divisor domain with a 2×2 matrix that does not reduce to a diagonal matrix under the standard Smith normal form algorithm.*

Proof. I claim that in the ring R of entire functions, there is a matrix A_0 that (under certain choices for invertible matrices) never reduces to a diagonal matrix under the standard SNF algorithm. We know by Corollary 2.2.6 that the ring of entire functions is an elementary divisor domain. (Throughout this proof, entire functions will be in terms of the variable z and i is a square root of -1.)

A well-known result from complex analysis is that the function $\sin(\pi z)$ has the infinite product decomposition

$$\sin(\pi z) = \pi z \prod_{n=1}^{\infty}\Big(1 - \frac{z^2}{n^2}\Big).$$

Thus, $\sin(\pi z)$ divided by any of these factors will also be an entire function.

Let $A_0 \in M_2(R)$ be the matrix

$$A_0 = \begin{pmatrix} \frac{\sin(\pi z)}{\pi z} & \sqrt{\pi}(z-i)\frac{\sin(\pi z)}{\pi z(1-z^2)} \\ 0 & \pi z \end{pmatrix}.$$

We will now explicitly find the invertible matrices needed to perform the algorithm.

Let $a_0 := \pi z$ and for each $n \geq 1$, let $a_n := 1 - z^2/n^2$. For each $n \geq 0$, let

$$e_n := \frac{\sin(\pi z)}{\prod_{k=0}^{n} a_k} = \frac{\sin(\pi z)}{\pi z(1-z^2)\cdots(1-z^2/n^2)}.$$

Let $b_1 := \sqrt{\pi}(z-i)$ and $b_2 := \frac{3}{4\sqrt{\pi}}(z+i)$. Then recursively, we define b_{n+1} for each $n \geq 2$ to be the unique line which at the two z values of $\pm n$ satisfies

$$b_{n+1}(n) = \lim_{z\to n} \frac{\left(1-\frac{z^2}{(n-1)^2}\right)\left(1-\frac{z^2}{(n+1)^2}\right)}{b_n(z)} \frac{\left(1-\frac{z^2}{n^2}\right)}{\sin(\pi z)} = \frac{4n^2-1}{(n^2-1)^2}\frac{1}{b_n(n)}\frac{2(-1)^n}{\pi n}, \qquad (4.3.3)$$

$$b_{n+1}(-n) = \lim_{z\to -n} \frac{\left(1-\frac{z^2}{(n-1)^2}\right)\left(1-\frac{z^2}{(n+1)^2}\right)}{b_n(z)} \frac{\left(1-\frac{z^2}{n^2}\right)}{\sin(\pi z)} = \frac{4n^2-1}{(n^2-1)^2}\frac{1}{b_n(-n)}\frac{2(-1)^{n+1}}{\pi n}. \quad (4.3.4)$$

In order for the recursive definition of b_{n+1} to make sense, we must guarantee that $b_n(\pm n) \neq 0$ for each n. We prove this inductively as follows.

Assume that b_n is a line with a purely imaginary nonzero root. (This is clearly true for $n = 1$ and $n = 2$.) If we let

$$d = \frac{4n^2-1}{(n^2-1)^2}\frac{2(-1)^n}{\pi n},$$

then we can write $b_{n+1}(n) = d/b_n(n)$ and $b_{n+1}(-n) = -d/b_n(-n)$. By Lemma 4.3.1, b_{n+1} is a line with a purely imaginary nonzero root. Therefore, $b_n(\pm n) \neq 0$ as desired, and the induction is done.

From how these b_n are defined, we have the equality $1 - a_0a_1\cdots a_{n-2}b_nb_{n+1}e_{n+1} = 0$ when z is evaluated at n and at $-n$. Since a_n has simple zeros at $\pm n$, the function

$$x_n := \frac{1 - a_0a_1\cdots a_{n-2}b_nb_{n+1}e_{n+1}}{a_n}$$

has a simple hole at the points $\pm n$. Therefore, filling in these holes makes x_n an entire function. Furthermore, $(a_n)(x_n) + (a_0a_1\cdots a_{n-2}b_n)(b_{n+1}e_{n+1}) = 1$ for each n.

We now verify that the matrix

$$A_0 = \begin{pmatrix} e_0 & b_1e_1 \\ 0 & a_0 \end{pmatrix} = \begin{pmatrix} \frac{\sin(\pi z)}{\pi z} & \sqrt{\pi}(z-i)\frac{\sin(\pi z)}{\pi z(1-z^2)} \\ 0 & \pi z \end{pmatrix}$$

does not reduce to a diagonal matrix under the Smith normal form algorithm. Recursively, given

$$A_{n-1} = \begin{pmatrix} e_{n-1} & a_0a_1\cdots a_{n-2}b_ne_n \\ 0 & a_0a_1\cdots a_{n-2}a_{n-1} \end{pmatrix}$$

for some $n \geq 1$, then define E_n to be the invertible matrix

$$E_n = \begin{pmatrix} x_n & -a_0 a_1 \cdots a_{n-2} b_n \\ b_{n+1} e_{n+1} & a_n \end{pmatrix}.$$

Then

$$\begin{aligned}
A_{n-1} E_n &= \begin{pmatrix} e_{n-1} & a_0 a_1 \cdots a_{n-2} b_n e_n \\ 0 & a_0 a_1 \cdots a_{n-2} a_{n-1} \end{pmatrix} \begin{pmatrix} x_n & -a_0 a_1 \cdots a_{n-2} b_n \\ b_{n+1} e_{n+1} & a_n \end{pmatrix} \\
&= \begin{pmatrix} e_{n-1} x_n + a_0 a_1 \cdots a_{n-2} b_n e_n b_{n+1} e_{n+1} & -a_0 a_1 \cdots a_{n-2} b_n e_{n-1} + a_0 a_1 \cdots a_{n-2} b_n e_n a_n \\ 0 + a_0 a_1 \cdots a_{n-2} a_{n-1} b_{n+1} e_{n+1} & a_0 a_1 \cdots a_{n-2} a_{n-1} a_n \end{pmatrix} \\
&= \begin{pmatrix} e_{n-1} x_n + a_0 a_1 \cdots a_{n-2} b_n e_n b_{n+1} e_{n+1} & -a_0 a_1 \cdots a_{n-2} b_n e_{n-1} + a_0 a_1 \cdots a_{n-2} b_n e_{n-1} \\ a_0 a_1 \cdots a_{n-1} b_{n+1} e_{n+1} & a_0 a_1 \cdots a_{n-1} a_n \end{pmatrix} \\
&= \begin{pmatrix} a_n e_n \left(\frac{1 - a_0 a_1 \cdots a_{n-2} b_n b_{n+1} e_{n+1}}{a_n} \right) + a_0 a_1 \cdots a_{n-2} b_n e_n b_{n+1} e_{n+1} & 0 \\ a_0 a_1 \cdots a_{n-1} b_{n+1} e_{n+1} & a_0 a_1 \cdots a_{n-1} a_n \end{pmatrix} \\
&= \begin{pmatrix} e_n & 0 \\ a_0 a_1 \cdots a_{n-1} b_{n+1} e_{n+1} & a_0 a_1 \cdots a_{n-1} a_n \end{pmatrix} = A_n^T.
\end{aligned}$$

Since the top-right corner of A_n is a product of nonzero entire functions, it is never 0. Thus, none of the A_n are diagonal.

Further, since the top-left corner of A_n is (the analytic continuation of)

$$e_n = \frac{\sin(\pi z)}{\pi z \prod_{k=1}^{n} (1 - z^2/k^2)},$$

it has zeros at all points in the set $\{\pm(n+1), \pm(n+2), \ldots\}$. Then the top-right corner of A_n is

$$a_0 \cdots a_{n-1} b_{n+1} e_{n+1} = \frac{\pi z \cdots (1 - z^2/(n-1)^2) b_{n+1} \sin(\pi z)}{\pi z \prod_{k=1}^{n+1} (1 - z^2/k^2)} = \frac{b_{n+1} \sin(\pi z)}{(1 - z^2/n^2)(1 - z^2/(n+1)^2)},$$

so it has zeros where b_{n+1} has a zero (which is some nonzero point on the imaginary axis), as well as at $\{0, \pm 1, \ldots, \pm(n-1), \pm(n+2), \pm(n+3), \ldots\}$. Note that for one entire function to divide another, the former's zero multiset must be a subset of the latter's zero multiset. Since that is not the case here, the top-left corner does not divide the top-right corner. Hence Step 2a does not apply and the matrix does not reduce to a diagonal matrix by fiat.

(As an aside, note that as we construct each A_n matrix, the bottom-right entry never divides the top-right entry. While not required for the SNF algorithm, it prevents a simple modification of the algorithm from reducing A_n to a diagonal matrix. Such a modification might be to check if the bottom-right corner divides the top-right corner, and if so, add the appropriate multiples of the bottom row to the top row.)

Therefore, the matrix

$$\begin{pmatrix} \frac{\sin(\pi z)}{\pi z} & \sqrt{\pi}(z-i)\frac{\sin(\pi z)}{\pi z(1-z^2)} \\ 0 & \pi z \end{pmatrix}$$

is a matrix in the elementary divisor domain of all entire functions which does not reduce to a diagonal matrix under the standard Smith Normal Form algorithm. □

Since the domain of all entire functions is an elementary divisor domain, the existence of a matrix which does not reduce to a diagonal matrix under the SNF algorithm is not enough to conclude that a Bézout domain is not an elementary divisor domain. Therefore, this technique of constructing a domain which is not an elementary divisor domain and then imposing relations upon it so that it becomes a Bézout domain will require a more careful construction than simply constructing a domain where the SNF algorithm fails to yield the desired matrices if it is to construct the desired counterexample.

4.4 Questions that Result from this Example

There are still a few questions that arise after this proof and specific example of a matrix which does not reduce under the SNF algorithm. In constructing this particular example, we chose the b_n and x_n in a very particular way. Now just as in the integers, where there are multiple linear combinations of 6 and 15 which result in their GCD of 3 (such as $3 = 3(6) + (-1)(15) = -2(6) + (1)(15)$), there are infinitely many choices of b_n and x_n which could be chosen to get the desired relations and lead to an invertible matrix.

Question 4.4.1. Are there different choices of invertible matrices which would convert the matrix

$$\begin{pmatrix} \frac{\sin(\pi z)}{\pi z} & \sqrt{\pi}(z-i)\frac{\sin(\pi z)}{\pi z(1-z^2)} \\ 0 & \pi z \end{pmatrix}$$

to Smith normal form under the standard SNF algorithm? And then if so, is there a different 2×2 matrix which does not terminate no matter which choices are made in following the

SNF algorithm?

The answer to the first question is yes, but I do not have an example of a 2×2 matrix that never terminates under any choice in the SNF algorithm. In the following example, I find invertible matrices using the SNF algorithm which convert the above matrix into Smith normal form.

Example 4.4.2. Just as in the previous section, let A_0 be the matrix

$$\begin{pmatrix} \frac{\sin(\pi z)}{\pi z} & \sqrt{\pi}(z-i)\frac{\sin(\pi z)}{\pi z(1-z^2)} \\ 0 & \pi z \end{pmatrix}.$$

Let E be the matrix

$$\begin{pmatrix} \frac{1-(1+z^2)\frac{\sin(\pi z)}{\pi z(1-z^2)}}{1-z^2} & -\sqrt{\pi}(z-i) \\ \frac{1}{\sqrt{\pi}}(z+i)\frac{\sin(\pi z)}{\pi z(1-z^2)} & 1-z^2 \end{pmatrix}.$$

We only need to check that the top-left entry is entire, and then E is an invertible matrix over the ring of entire functions.

We know that the function $1-(1+z^2)\frac{\sin(\pi z)}{\pi z(1-z^2)}$ is entire, and we also know that $1-z^2$ has simple roots at $z = \pm 1$. Thus to show that the top-left entry of E is entire, we need only show that $1-(1+z^2)\frac{\sin(\pi z)}{\pi z(1-z^2)}$ is zero at $z = \pm 1$. This can be calculated through a simple limit, and we see that

$$\lim_{z\to 1} 1-(1+z^2)\frac{\sin(\pi z)}{\pi z(1-z^2)} = 1 - \frac{2}{(1\pi)(1+1)}\lim_{z\to 1}\frac{\sin(\pi z)}{1-z} = 1 - \frac{1}{\pi}\left(\frac{-\pi}{-1}\right) = 0$$

and

$$\lim_{z\to -1} 1-(1+z^2)\frac{\sin(\pi z)}{\pi z(1-z^2)} = 1 - \frac{2}{(-1\pi)(1-(-1))}\lim_{z\to -1}\frac{\sin(\pi z)}{1+z} = 1 + \frac{1}{\pi}\left(\frac{-\pi}{1}\right) = 0.$$

Therefore, E is an invertible matrix.

A direct computation then finds that

$$AE = \begin{pmatrix} \frac{\sin(\pi z)}{\pi z(1-z^2)} & 0 \\ \frac{\pi z}{\sqrt{\pi}}(z+i)\frac{\sin(\pi z)}{\pi z(1-z^2)} & \pi z(1-z^2) \end{pmatrix}.$$

At this point, the top-left corner of AE divides the bottom-left corner of AE, so Step 3a of the SNF algorithm applies. Using the specified matrix from that step then gives us

$$\begin{pmatrix} 1 & 0 \\ -\frac{\pi z}{\sqrt{\pi}}(z+i) & 1 \end{pmatrix}\begin{pmatrix} \frac{\sin(\pi z)}{\pi z(1-z^2)} & 0 \\ \frac{\pi z}{\sqrt{\pi}}(z+i)\frac{\sin(\pi z)}{\pi z(1-z^2)} & \pi z(1-z^2) \end{pmatrix} = \begin{pmatrix} \frac{\sin(\pi z)}{\pi z(1-z^2)} & 0 \\ 0 & \pi z(1-z^2) \end{pmatrix}.$$

This is a diagonal matrix, and then Step 5 provides the matrices to convert this into Smith normal form, which for this matrix is

$$\begin{pmatrix} 1 & 0 \\ 0 & \sin(\pi z) \end{pmatrix}.$$

Thus, it is possible for some choices of invertible matrices to lead the SNF algorithm to never terminate, while other choices lead to the algorithm terminating.

Another question that arises is the fact that we know that the domain of entire functions is an elementary divisor domain, which implies that there are invertible matrices P and Q such that

$$P \begin{pmatrix} \frac{\sin(\pi z)}{\pi z} & \sqrt{\pi}(z-i)\frac{\sin(\pi z)}{\sqrt{\pi} z(1-z^2)} \\ 0 & \pi z \end{pmatrix} Q = \begin{pmatrix} 1 & 0 \\ 0 & \sin(\pi z) \end{pmatrix}.$$

In this particular case, we know what the Smith normal form of the matrix should be, yet we were unable to arrive at it using the standard SNF algorithm with our original choices of invertible matrices.

Question 4.4.3. Given an elementary divisor domain R and a matrix $A \in M_2(R)$, how do we find the Smith normal form S of A? Furthermore, how do we find the matrices P and Q such that $PAQ = S$? Is there a different algorithm that would always work in an arbitrary elementary divisor domain rather than just PIDs?

Chapter 5. Conclusion

While we still do not yet know if every Bézout domain is an elementary divisor domain, we do now know that there are several ways to construct a Bézout domain. We also know that showing a domain to not be an elementary divisor domain requires a different method than showing that the SNF algorithm does not terminate. There are also several questions noted in this paper that could lead to new results.

Bibliography

[1] P. M. Cohn. Bézout rings and their subrings. *Mathematical Proceedings of the Cambridge Philosophical Society*, 64:251–264, 1968.

[2] David S. Dummit and Richard M. Foote. *Abstract algebra.* John Wiley & Sons, Inc., Hoboken, NJ, third edition, 2004.

[3] Leonard Gillman and Melvin Henriksen. Rings of continuous functions in which every finitely generated ideal is principal. *Transactions of the American Mathematical Society*, 82:366–391, 1956.

[4] Robert Gilmer. *Multiplicative Ideal Theory.* Pure and Applied Mathematics, No. 12. Marcel Dekker, Inc., New York, 1972.

[5] Marek Golasiński and Melvin Henriksen. Residue class rings of real-analytic and entire functions. *Colloquium Mathematicum*, 106:85–97, 2006.

[6] Olaf Helmer. The elementary divisor theorem for certain rings without chain condition. *Bulletin of the American Mathematical Society*, 49:225–236, 1943.

[7] Irving Kaplansky. Elementary divisors and modules. *Transactions of the American Mathematical Society*, 66:464–491, 1949.

[8] Irving Kaplansky. *Commutative Rings.* The University of Chicago Press, Chicago and London, revised edition, 1974.

[9] T. Y. Lam. *Serre's Problem on Projective Modules.* Springer Monographs in Mathematics, Berlin, Heidelberg: Springer-Verlag, 2006.

[10] Max D. Larsen, William J. Lewis, and Thomas S. Shores. Elementary divisor rings and finitely presented modules. *Transactions of the American Mathematical Society*, 187:231–248, 1974.

[11] Morris Newman and Robert C. Thompson. Matrices over rings of algebraic integers. *Linear Algebra and its Applications*, 145:1–20, 1991.

[12] Walter Rudin. *Real and Complex Analysis.* McGraw-Hill Book Company, New York, third edition, 1987.

[13] Henry J. Stephen Smith. On systems of linear indeterminate equations and congruences. *Philosophical Transactions of the Royal Society of London*, 151:293–326, 1861.

[14] J. H. M. Wedderburn. On matrices whose coefficients are functions of a single variable. *Transactions of the American Mathematical Society*, 16:328–332, 1915.

Index