Integral Matrices and Applications to Algebraic Number Theory

John Peca-Medlin
Department of Mathematics
University of California, Irvine
March 23, 2018

1 Introduction
Matrices with integer values are fundamental objects in linear algebra, with applications to number theory and group theory — of note, the entire theory of finite groups is encapsulated by such matrices. I will review some of the history behind the study of integral matrices, focusing on results that tie directly with some topics covered in Math 232B. I will provide an alternative proof of the fact the cyclotomic ring of integers is monogenic. Additionally I will outline a proof for the Latimer-MacDuffee Theorem, with applications to class number calculations. The first five sections follow closely many results from [10] (with some particular deviations and alternative approaches), although many other results and applications have been omitted. I will later move to random integral matrices, and explore theory that has developed with these objects relating to Number Theory. In particular, I will look at universality results relating to the Cohen-Lenstra Heuristics, along with applications to the study the surjectivity of near-square random matrices.

2 Smith Normal Form
Let $M_{n,m}(\mathbb{Z})$ denote the $n \times m$ integral matrices and $M_n(\mathbb{Z})$ the $n \times n$ integral square matrices. Recall every $A \in M_{m,n}(\mathbb{Z})$ has a unique Smith Normal Form (SNF), say $S$, such that $A = PSQ$ for $P, Q$ two unimodular matrices, that is, $P \in M_n(\mathbb{Z})$, $Q \in M_m(\mathbb{Z})$ and $\det(P), \det(Q) \in \{\pm 1\}$, where $S = \text{diag}(s_1, \ldots, s_k, 0, \ldots, 0)$ with $s_i = \frac{d_i(A)}{d_{i-1}(A)}$ for $d_i(A) = \gcd(\text{all } i \times i \text{ minors of } A)$ (where a $i \times i$ minor is the determinant of the corresponding $i \times i$ submatrix), using the convention $d_0(A) = 1$. Note the final result yields $s_1 \div s_2 \div \cdots \div s_n$.

Example 2.1. We can calculate the SNF for $A = \begin{pmatrix} 3 & 1 \\ 0 & 3 \end{pmatrix}$ as follows: we see $s_1 = d_1(A) = \gcd(0, 1, 3) = 1$ and $d_2(A) = \det(A) = 9$ so that $s_2 = \frac{d_2(A)}{d_1(A)} = 9$, and so $A$ has SNF $\begin{pmatrix} 1 \ & 0 \\ 0 \ & 9 \end{pmatrix}$.
Note the SNF can be used to deduce divisibility results on the integers solutions by rewriting the equation \( Ax = b \) as \( SQx = P^{-1}b \).

The SNF of a matrix is intimately related to the invariant factors form in the classification of finitely generated Abelian groups: If \( G \) is Abelian and generated by \( n \) elements, then there exist integers \( s_1 \mid s_2 \mid \cdots \mid s_k \) such that

\[
G \cong \mathbb{Z}/s_1\mathbb{Z} \times \cdots \times \mathbb{Z}/s_k\mathbb{Z} \times \mathbb{Z}^{n-k}.
\]

In fact, noting \( A \in \mathbb{M}_{n,m}(\mathbb{Z}) \) can be viewed as a linear map \( A : \mathbb{Z}^m \to \mathbb{Z}^n \), we have exactly the cokernel of \( A \) satisfies the relation

\[
\text{coker}(A) := \mathbb{Z}^n / \text{Im}(A) \cong \mathbb{Z}/s_1\mathbb{Z} \times \cdots \times \mathbb{Z}/s_k\mathbb{Z} \times \mathbb{Z}^{n-k}
\]

if \( A \) has SNF \( S = \text{diag}(s_1, \ldots, s_k, 0, \ldots, 0) \). In Example 2.1, we see

\[
\mathbb{Z}^2 / \text{Im}(A) \cong \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/6\mathbb{Z} \cong \mathbb{Z}/12\mathbb{Z}.
\]

It follows then since \( \text{Im}(A) = \text{Span}\{ (3, 0), (1, 3) \} \), we see for any \( (x, y) \in \mathbb{Z}^2 \), then \( 9(x, y) \in \text{Im}(A) \), since

\[
9\begin{pmatrix} x \\ y \end{pmatrix} = (3x - y) \begin{pmatrix} 3 \\ 0 \end{pmatrix} + 3y \begin{pmatrix} 1 \\ 3 \end{pmatrix}.
\]

Example 2.2. As another example, using

\[
B = \begin{pmatrix} 2 & 4 & 4 \\ -6 & 6 & 12 \\ 10 & -4 & -16 \end{pmatrix},
\]

one can show \( B \) has the SNF of

\[
\begin{pmatrix} 2 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & 12 \end{pmatrix},
\]

so that

\[
\text{coker}(B) \cong \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/6\mathbb{Z} \times \mathbb{Z}/12\mathbb{Z}.
\]

In fact, the SNF of a matrix can then determine precisely when \( A \) is surjective: \( A \in \mathbb{M}_{n,m}(\mathbb{Z}) \) is surjective if and only if \( \text{coker}(A) \cong 1 \) if and only if \( S = (I \mid 0) \) is the SNF of \( A \), that is, we have \( s_n = 1 \).

Calculations for increasingly larger matrices get computationally burdensome relatively quickly. The \( 2 \times 2 \) case is trivial, but for the \( 3 \times 3 \) case, as in \( B \) above, we can easily compute \( s_1 \), but for \( s_2 \) one would need to first calculate the 9 \( 2 \times 2 \) minors before taking the gcd of these, and then one can easily compute \( s_3 \) as the ratio of \( |\det(B)| \) and \( d_2(B) \). In [7], Hafner and McCurley give an example of a \( 20 \times 20 \) matrix with all of its entries nonnegative integers less than 10 but whose Hermite normal form (which is further explored in [10]) has entries of more than 5000 digits. Building on a prior result from Domich, Kanan and Trotter [3], Hafner and McCurley developed an algorithm that runs in \( O(m^2nM(m \log ||A||)) \) time, where \( M(t) \) is a monotonic upper bound on the number of bit operations required to multiply two \( t \) bit integers: for nonsingular \( A \), their method starts by choosing some \( h \) that is divisible by \( \det(A) \), then finding the SNF of the matrix \( A \pmod{h} \), and then lifting the result to yield the SNF of \( A \). This result has been improved since using probabilistic techniques: for instance, Labahn and Storjohan first compute the principal minors of \( UAV \) for \( U, V \) randomly chosen unimodular matrices to yield the SNF of \( A \) [9].
3 GL(n, Z) and SL(n, Z)

Recall

\[ \text{GL}(n, \mathbb{Z}) = \{ A \in M_n(\mathbb{Z}) : \det(A) = \pm 1 \} \]
\[ \text{SL}(n, \mathbb{Z}) = \{ A \in M_n(\mathbb{Z}) : \det(A) = 1 \} \]

where we note this definition holds since \( A \in \text{GL}(n, \mathbb{Z}) \) if and only if \( \det(A) \in \mathbb{Z}^\times = \{ \pm 1 \} \). (So, earlier, unimodular matrices are merely elements of \( \text{GL}(n, \mathbb{Z}) \).) Let \( E_{ij} \) denote the \( n^2 \) elementary basis elements that generate \( M_n(\mathbb{Z}) \). We can write the three basic elementary matrices in the form

(i) \( E_{ij}(\alpha) = I + \alpha E_{ij} \) for \( i \neq j \), (ii) \( E_k(\alpha) = I + (\alpha - 1)E_{kk} \) for \( \alpha \neq 1 \), and (iii) \( E(k, \ell) \) the transposition matrix corresponding to \((k \ell)\). Recall then \( \text{GL}(n, \mathbb{Z}) \) is generated by the elementary matrices \( E_{ij}(\pm 1), E_k(-1) \), and \( E(k, \ell) \). In particular, \( \text{GL}(n, \mathbb{Z}) \) is finitely generated.

Note \( \text{SL}(2, \mathbb{Z}) \) is of particular interest (which I’ll return to shortly). From the discussion of the Smith Normal Form, we see \( \text{SL}(2, \mathbb{Z}) \) can explicitly be realized by the sequence of natural maps. Gow and Tamburini in \([6]\) showed also for \( J \) the full Jordan block with characteristic polynomial \((x - 1)^n\), we have:

**Theorem 3.1** (Gow, Tamburini). If \( n \neq 4 \), then \( J, J^T \) generates \( \text{SL}(n, \mathbb{Z}) \). If \( n = 4 \), then the subgroup generated by \( J, J^T \) has index 8 in \( \text{SL}(4, \mathbb{Z}) \).

A lot of attention has been spent on \( \text{SL}(2, \mathbb{Z}) \), and in particular, on the factor group

\[ \text{PSL}(2, \mathbb{Z}) = \frac{\text{SL}(2, \mathbb{Z})}{Z(\text{SL}(2, \mathbb{Z}))}, \]

where we note \( Z(\text{SL}(2, \mathbb{Z})) = \{ \pm I \} \). It can be shown that \( \text{PSL}(2, \mathbb{Z}) \) is isomorphic with the free product \( \mathbb{Z}/2\mathbb{Z} \ast \mathbb{Z}/3\mathbb{Z} \), and hence any group generated by two elements of orders 2 and 3 is then a homomorphic image of \( \text{SL}(2, \mathbb{Z}) \).

**Example 3.1.** Let \( \tau = (1 \ 2)(3 \ 5) \) and \( \sigma = (2 \ 3 \ 5) \), then \( \sigma^\tau = (1 \ 4 \ 2 \ 3 \ 5) \), and so \((\tau, \sigma)\) has order at least \( 2 \cdot 3 \cdot 5 = 30 \), so that \( A_5 = \langle \tau, \sigma \rangle \). It follows \( A_5 \) is a homomorphic image of \( \text{SL}(2, \mathbb{Z}) \), which can explicitly be realized by the sequence of natural maps

\[ \text{SL}(2, \mathbb{Z}) \rightarrow \text{SL}(2, 5) \rightarrow \text{PSL}(2, 5) \]

with \( \text{PSL}(2, 5) \cong A_5 \).

Tamburini, Wilson and Gavioli \([5]\) yield the (remarkable) result: if \( n \geq 28 \) then there is a normal subgroup \( K_n \) of \( \text{SL}(2, \mathbb{Z}) \) such that \( \text{SL}(2, \mathbb{Z})/K_n \cong \text{SL}(n, \mathbb{Z}) \).
4 Latimer-MacDuffee Theorem

For this section, note in addition to following [10], I am also following the particular proof of the Latimer-MacDuffee Theorem outlined in a student project [12] as well as some results mentioned in [11] and [2].

Recall two matrices $A, B \in M_n(\mathbb{Z})$ are integrally similar if there exists some $U \in \text{GL}(n, \mathbb{Z})$ such that $A = U^{-1}BU$. Clearly integrally similar matrices are similar, but the converse does not hold:

Example 4.1. Take

$$A = \begin{pmatrix} 1 & 1 \\ 0 & -1 \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},$$

which are similar (for instance, conjugate $A$ by $\begin{pmatrix} 1 & 1 \\ 0 & -2 \end{pmatrix}$); if $A$ and $B$ were integrally similar, then $A \pmod{2}$ and $B \pmod{2}$ would be similar over $\mathbb{F}_2$, which is not true as these are distinct Jordan forms.

We recall if $K$ is a number field, then $K$ is a finite extension of $\mathbb{Q}$, and we write $\mathcal{O}_K$ for the ring of integers of $K$, that is, the ring of elements that are roots of monic polynomials over $\mathbb{Z}$. We saw then if $[K : \mathbb{Q}] = n$, then $\mathcal{O}_K$ is a free abelian group of rank $n$, and it has a basis $\omega_1, \ldots, \omega_n$, which is also a vector space basis of $K$ over $\mathbb{Q}$. If $\omega'_1, \ldots, \omega'_n$ is another basis of $\mathcal{O}_K$, then we can write

$$\omega'_j = \sum_{i=1}^{n} u_{ij} \omega_j$$

and hence we have $U = (u_{ij}) \in \text{GL}(n, \mathbb{Z})$. Conversely, we can form a different basis of $\mathcal{O}_K$ by using any starting basis and element in $\text{GL}(n, \mathbb{Z})$. It follows then any other nonzero ideal $I$ of $\mathcal{O}_K$ must also contain a basis for $K$ over $\mathbb{Q}$ since $\omega_1, \ldots, \omega_n$ is linearly independent if $\alpha \neq 0$ and is contained in $I$ if $\alpha \in I$. It follows then $\mathcal{O}_K/I$ is finite, and in particular, if $y_1, \ldots, y_n$ is a basis for $I$, then for $y_i = \sum_{j=1}^{n} a_{ij} \omega_j$ and $S = \text{diag}(s_1, \ldots, s_n)$ the SNF of $A = (a_{ij})$ (where we note $S$ needs to be nonsingular since $y_1, \ldots, y_n$ are linearly independent), we have

$$\mathcal{O}_K/I \cong \mathbb{Z}/s_1 \mathbb{Z} \times \cdots \times \mathbb{Z}/s_n \mathbb{Z}.$$

In fact, using these invariants, we can find a new basis $z_1, \ldots, z_n$ for $\mathcal{O}_K$ such that $I$ has basis $s_1 z_1, \ldots, s_n z_n$. Given any finite collection of ideals of $\mathcal{O}_K$, one can choose a basis of $\mathcal{O}_K$ such that we can simultaneously form bases of each ideal in this way.

Also, since $\mathcal{O}_K$ is a Dedekind domain, then every nonzero ideal can be written uniquely as the product of prime ideals, and for $\mathcal{I}(K)$ the group of fractional ideals of $\mathcal{O}_K$ and $\mathcal{P}(K)$ the group of principal fractional ideals, then

$$\text{Cl}(K) = \mathcal{I}(K)/\mathcal{P}(K)$$

is the ideal class group of $K$, and $|\text{Cl}(K)|$ is the class number of $K$. Moreover, $\mathcal{O}_K$ is a PID if and only if $K$ has class number 1.

By the primitive element theorem, we can write $K = \mathbb{Q}(\theta)$ for some $\theta \in \mathcal{O}_K$ and $\mathbb{Z}[\theta]$ is an order of $\mathcal{O}_K$, that is, a subring of $\mathcal{O}_K$ of finite index. Similarly, we can define the ideal class groups and class numbers for orders of $\mathcal{O}_K$. Let $f(x) \in \mathbb{Z}[x]$ be the minimal polynomial for $\theta$, so that $\deg f = n$ and $f$ is monic. For $\omega_1, \ldots, \omega_n$ a basis of $\mathcal{O}_K$, let $B = (b_{ij}) \in M_n(\mathbb{Z})$ denote multiplication on this basis by $\theta$, that is, $\theta \omega_i = \sum_{j=1}^{n} b_{ij} \omega_j$. Then $B \omega = \theta \omega$ for $\omega = (\omega_1, \ldots, \omega_n)^T$ so that $\theta$ is an eigenvalue of $B$ and hence $B$ has characteristic polynomial $f$ since then $\theta$ is a root of the characteristic polynomial of $B$, which is of degree $n = \deg f$ and is divisible by $f(x)$. 
This hints at a nice relationship between the ideal class group and the equivalence classes of integrally similar matrices:

**Theorem 4.1** (Latimer-MacDuffee). Let \( f(x) \in \mathbb{Z}[x] \) be an irreducible monic polynomial of degree \( n \). Then there is a one-to-one correspondence between the integral similarity classes of matrices \( A \in M_n(\mathbb{Z}) \) with characteristic polynomial \( f(x) \) and ideal classes in the order \( \mathbb{Z}[\theta] \) of \( \mathbb{Q}(\theta) \) for \( \theta \in \mathbb{C} \) a root of \( f \).

Some immediate consequences based on our knowledge of class numbers follow:

**Corollary 4.1.1.** There are only finitely many integral similarity classes of matrices that have irreducible characteristic polynomial \( f \).

**Corollary 4.1.2.** If \( K = \mathbb{Q}(\theta) \) has class number 1 for \( \theta \in \mathcal{O}_K \) with minimal polynomial \( f \), then any integral matrix with characteristic polynomial \( f \) is integrally similar with the companion matrix of \( f \).

First, before proving this theorem, recall the result:

**Lemma 4.2.** Let \( A, B \in M_n(K) \) have irreducible characteristic polynomials \( \sigma_A, \sigma_B \in K[x] \). Then \( A \) and \( B \) are similar if and only if \( \sigma_A = \sigma_B \).

**Proof.** Clearly similar matrices have the same characteristic polynomials (even if they are reducible). Conversely, suppose \( \sigma_A = \sigma_B = f \). Let \( x \in K^n \) be a nonzero vector, then we have \( A = \{x, Ax, A^2 x, \ldots, A^{n-1} x\} \) is linearly independent: \( K^n \) has no non-trivial \( A \)-stable subspaces (since otherwise the characteristic polynomial of the restriction of \( A \) to this subspace must divide \( f \)) and if \( A^k x \in W = \text{Span}\{A^j x : j = 0, \ldots, k-1\} \), then \( W \) is \( A \)-stable. Similarly \( B = \{x, Bx, \ldots, B^{n-1} x\} \) is linearly independent, and so \( A \) and \( B \) are bases of \( K^n \), and multiplication by \( A \) and \( B \) on each respective basis yields the companion matrix of \( f \). It follows then for \( U \) the change of basis matrix sending \( A \) to \( B \), we have \( B = U^{-1} A U \). \( \square \)

Note the irreducible condition cannot be dropped by considering different sizes of Jordan blocks with diagonals of 1.

We see that this theorem does not carry over to \( M_n(\mathbb{Z}) \) instead of over \( M_n(\mathbb{Q}) \):

**Example 4.2.** Consider

\[
A = \begin{pmatrix} 0 & -13 \\ 1 & 0 \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} -1 & -7 \\ 2 & 1 \end{pmatrix}.
\]

We have \( \sigma_A(x) = \sigma_B(x) = x^2 + 13 \), which is irreducible over \( \mathbb{Q} \) (by noting, say, it’s Eisenstein), and so these are similar over \( \mathbb{Q} \) by the prior lemma. But we see that these are not integrally similar: let \( C = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \), then we have four linear equations in \( a, b, c, d \) by considering \( CA = BC \), which has solutions of the form \( a = -\frac{1}{2}(c - d) \) and \( b = -\frac{1}{2}(13c + d) \) (with \( c, d \) free) and hence

\[
\det(C) = ad - bc = \frac{1}{2} (13c^2 + d^2) = \pm 1,
\]

which has no rational solutions.

Let \( K = \mathbb{Q}(\sqrt{-13}) \) for \( \theta = \sqrt{-13} \), with then \( \mathcal{O}_K = \mathbb{Z}[\theta] \) since \(-13 \not\equiv 1 \pmod{4} \). By the Latimer-MacDuffee theorem and Example 4.2, we now know \( K \) has class number at least 2 without doing any calculations. In particular, \( \mathcal{O}_K \) is not a PID.
In fact, we actually see $A$ and $B$ are the representatives of the only two integral similarity classes in $M_2(\mathbb{Z})$ of matrices with characteristic polynomial $x^2 + 13$ since $K$ has class number exactly 2. This class number calculation is a standard result from 232A: We have Minkowski constant $M_K \approx 9.2$, and so we need only consider $f(x)$ reduced modulo 2, 3, 5, 7. Since $f(x) = (x + 1)^2$ over $\mathbb{F}_2$ and $f(x) = (x - 1)(x + 1)$ over $\mathbb{F}_7$, while $f(x)$ is irreducible over $\mathbb{F}_3$ and $\mathbb{F}_5$, then by Kummer’s theorem we have $\text{Cl}(K) = \langle [p_2], [p_7] \rangle$ for $p_2 = (2, \theta + 1)$, $p_7 = (7, \theta - 1)$; and since $p_2 p_7 = (14, 2(\theta - 1), 7(\theta + 1), -14) = (\theta - 1)$, then $\text{Cl}(K) = \langle [p_2] \rangle$. We further see $p_2$ is not principal since $a^2 + 13b^2 = \pm 2$ has no integer solutions, and $p_2^2 = (2)$ yields $\text{Cl}(K) \cong \mathbb{Z}/2\mathbb{Z}$.

Last, note the following result:

**Lemma 4.3.** Suppose $f(x) = x^n + c_{n-1}x^{n-1} + \cdots + c_0 \in K[x]$ is a separable irreducible polynomial and $\theta$ a root of $f$. If $A \in M_n(K)$ has characteristic polynomial $f$, then any eigenvector for $\theta$ has linearly independent components (in particular, they are nonzero).

*Proof.* Since $f$ has distinct roots, $A$ is diagonalizable and $L^n$ is the direct sum of the one dimensional eigenspaces of $A$ for $L$ the splitting field of $f$.

First consider $A = C$, the companion matrix

$$
\begin{pmatrix}
0 & -c_0 \\
1 & -c_1 \\
\vdots & \vdots \\
1 & -c_{n-1}
\end{pmatrix}
$$

Solving for the eigenvector $v$ that satisfies the equation $Cv = \theta v$ yields the system of linear equations:

$$
-c_0 v_n = \theta v_1 \\
v_1 - c_1 v_n = \theta v_2 \\
\vdots \\
v_{n-2} - c_{n-2} v_n = \theta v_{n-1} \\
v_{n-1} - c_{n-1} v_n = \theta v_n
$$

Backward substitution yields $v_{n-1} = (c_{n-1} + \theta) v_{n-1} = c_{n-2} v_n + \theta v_{n-1} = (c_{n-2} + c_{n-1} \theta + \theta^2) v_n$, ..., $v_i = (c_i + c_{i+1} \theta + \cdots + c_{n-1} \theta^{n-i-1} + \theta^{n-1}) v_n$. Now write

$$
g_i(x) = c_i + c_{i+1} x + \cdots + c_{n-1} x^{n-i-1} + x^{n-i}
$$

for $1 \leq i \leq n - 1$, and $g_n(x) = 1$, so that $v_i = g_i(\theta) v_n$. Hence, any eigenvector of $C$ for $\theta$ is a scalar multiple of

$$
\begin{pmatrix}
g_1(\theta) \\
g_2(\theta) \\
\vdots \\
g_n(\theta)
\end{pmatrix}
$$

(since the associated eigenspace has dimension 1). Note we see $g_i(\theta) \neq 0$ since deg $g_i < \deg f$, and, moreover, we see $g_1(\theta), \ldots, g_n(\theta)$ are linearly independent in $K(\theta)$ over $K$ since any linear combination is of degree at most $n - 1$ in $\theta$. It follows then $v_i = 0$ for some $i$ if and only if $v_n = 0$ if and only if $v = 0$. Hence, any eigenvector of $C$ has nonzero components.

Now suppose $A$ has characteristic polynomial $f$. By Lemma 4.2, there exists $U \in \text{GL}(n, K)$ such that $A = UCU^{-1}$, and hence $Uv$ is an eigenvector of $A$ for $\theta$ for some associated eigenvector $v$ of
Since the components of $v$ are linearly independent over $K$ (and nonzero), it follows then the components of $Uv$, which are then non-trivial linear combinations of the components of $v$, must also be nonzero and linearly independent. 

Now we can prove the Latimer-MacDuffee Theorem. Note Latimer and MacDuffee first introduced this result in 1933, before Olga Taussky Todd gave a refinement of their argument over a decade later (so this is sometimes referred to as the Latimer-MacDuffee-Taussky Theorem).

Proof of Latimer-MacDuffee. Suppose $A \in M_n(\mathbb{Z})$ has characteristic polynomial $f$. Let $v$ be an eigenvector for $A$ with respect to $\theta$, and we can further assume $v = (v_1, \ldots, v_n)^T \in \mathbb{Z}[\theta]^n$. (By scaling, we can assume $v_j = 1$ for some $j$, and $Av = \theta v$ is a system of linear equations in the remaining $v_i$ with coefficients in $\mathbb{Z}[\theta]$, and hence $v_i \in \mathbb{Q}(\theta)$ for all $i$; so again scaling, we can assume now $v_i \in \mathbb{Z}[\theta]$.) Let $I$ be the additive subgroup (i.e., $\mathbb{Z}$-module) generated by $v_1, \ldots, v_n$, where we note this has rank $n$ by Lemma 4.3. Since $Av = \theta v$ and $A$ has integer entries, then $\theta v_i \in I$ for all $i$ and so $I$ is a fractional ideal of $\mathbb{Z}[\theta]$. Rescaling $v$ gives a fractional ideal in the same ideal class as $I$. If $U \in \text{GL}(n, \mathbb{Z})$, then
\[(UAU^{-1})(Uv) = \theta(Uv),\]
and the entries of $Uv$ form another basis for $I$. This yields a well-defined map from the matrix conjugacy classes to ideal classes.

Conversely, suppose $I$ is a fractional ideal of $\mathbb{Z}[\theta]$. Since $\mathbb{Z}[\theta]$ has rank $n$, then $I$ also has rank $n$ as a $\mathbb{Z}$-module. Let $\omega_1, \ldots, \omega_n$ be a basis for $I$ and let $A$ represent multiplication on this basis by $\theta$. It follows $A \in M_n(\mathbb{Z})$ and $A\omega = \theta \omega$ for $\omega = (\omega_1, \ldots, \omega_n)^T$ so $\sigma_A(\theta) = 0$ and hence $f | \sigma_A$ and $\deg f = \deg \sigma_A$, so that $f = \sigma_A$ since both are monic. To see we have a well-defined inverse correspondence, suppose $\alpha I = \beta I'$, with $A, B$ the corresponding matrices that represent multiplication by $\theta$ of the respective bases $\omega_1, \ldots, \omega_n$ and $\omega'_1, \ldots, \omega'_n$ for $I, I'$, and let $\omega = (\omega_1, \ldots, \omega_n)^T$, $\omega' = (\omega'_1, \ldots, \omega'_n)^T$ so that $A\omega = \theta \omega$ and $B\omega' = \theta \omega'$. Let $U \in \text{GL}(n, \mathbb{Z})$ be the change of basis matrix that sends $\beta \omega'_1, \ldots, \beta \omega'_n$ to $\alpha \omega_1, \ldots, \alpha \omega_n$, so that $\alpha \omega = U(\beta \omega')$. It follows
\[\beta(AU)\omega' = A(U\beta \omega') = A(\alpha \omega) = \theta(\alpha \omega) = \theta(U\beta \omega') = \beta(UB)\omega'\]
so that
\[(AU - UB)\omega' = 0.\]
Since the entries of $\omega'$ are linearly independent over $\mathbb{Z}$ by Lemma 4.3, it follows $AU - UB = 0_{n \times n}$, so that $B = U^{-1}AU$. 

Note this theorem also yields a method of computing corresponding similarity classes and ideal classes: the algebraic components of an eigenvector for a given matrix generate the corresponding ideal class, and similarly starting with a basis of an ideal class, we can construct a corresponding matrix that represents multiplication on this basis by $\theta$. In particular, we see the companion matrix generates the identity in the ideal class group.

Example 4.3. Since $\mathbb{Z}[i]$ is a PID, then $\mathbb{Z}[i]$ has class number 1, and so by the Latimer-MacDuffee theorem, any two integral matrices with characteristic polynomial $x^2 + 1$ are integrally similar. Similarly, any two integral matrices with characteristic polynomial $\Phi_3(x) = x^2 + x + 1$ are integrally similar.

Example 4.4. Returning to Example 4.2 looking at $K = \mathbb{Q}(\theta)$ for $\theta = \sqrt{-13}$, where I showed $\text{Cl}(K) = \langle p_2 \rangle$ for $p_2 = (2, \theta + 1)$, then we have $\mathcal{O}_K = \mathbb{Z}[\theta]$ and $p_2$ are both ideal class representatives in $\text{Cl}(K)$. We can find respective representatives of integrally similar matrices in $M_2(\mathbb{Z})$ with minimal polynomial $x^2 + 13$ by taking bases for each representative ideal and multiplying by $\theta$: Since $\{1, \theta\}$ is a basis for $\mathcal{O}_K$, multiplying this by $\theta$ yields $\{\theta, -13\}$, and so we have a representative matrix for
multiplication of this basis by \( \theta \) of \( A = \begin{pmatrix} 0 & -13 \\ 1 & 0 \end{pmatrix} \). Similarly using \( \{2, 1 + \theta\} \) as the basis of \( p_2 \), multiplication by \( \theta \) yields \( \{2\theta, \theta - 13\} = \{-1(2) + 2(1 + \theta), -7(2) + 1(1 + \theta)\} \), and hence we have the matrix \( B = \begin{pmatrix} -1 & -7 \\ 2 & 1 \end{pmatrix} \).

Moreover, taking \( C = \begin{pmatrix} 3 & 11 \\ -2 & -3 \end{pmatrix} \), this has characteristic polynomial \( f \) with eigenvector \( (-3-\theta) \) for \( \theta \), and so \( C \) is integrally similar to \( B \) since \( (2, -3 - \theta) = (2, 1 + \theta) \). One can easily construct this similarity by conjugating \( C \) by \( \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix} \) \( \in \text{GL}(2, \mathbb{Z}) \) to yield \( B \).

**Example 4.5.** Now take \( K = \mathbb{Q}(\sqrt{-23}) \) with \( \mathcal{O}_K = \mathbb{Z}[\omega] \) for \( \omega = \frac{1+\sqrt{-23}}{2} \), which has minimal polynomial \( f(x) = x^2 - x + 6 \); note then \( \omega^2 = \omega - 6 \). From the final homework in 23A2, we saw \( \text{Cl}(K) = \langle \langle p_2 \rangle \rangle \cong \mathbb{Z}/3\mathbb{Z} \) for \( p_2 = (2, \omega) \), with \( p_2^2 = (4, \omega + 2) \), so that \( \mathcal{O}_K, p_2, p_2^2 \) are representatives of the three distinct ideal classes in \( \text{Cl}(K) \). Now taking the basis \( \{1, \omega\} \) of \( \mathcal{O}_K \), we see multiplication by \( \omega \) yields \( \{\omega, \omega^2\} = \{\omega, \omega - 6\} \) and so we have one integral similarity representative \( A = \begin{pmatrix} 0 & -6 \\ 1 & 1 \end{pmatrix} \) which we can easily verify has characteristic polynomial \( f \). Next, taking the basis \( \{2, \omega\} \) of \( p_2 \), we see multiplication by \( \omega \) yields \( \{2\omega, \omega^2\} = \{2\omega, -3(2) + \omega\} \), and so we have corresponding integral similarity representative \( B = \begin{pmatrix} 0 & -3 \\ 2 & 1 \end{pmatrix} \). And last, using the basis \( \{4, \omega + 2\} \) of \( p_2^2 \), we see multiplication by \( \omega \) yields \( \{4\omega, \omega^2 + 2\omega\} = \{4\omega, 3\omega - 6\} = \{-2(4) + 4(\omega + 2), -3(4) + 3(\omega + 2)\} \), with corresponding integral similarity representative \( C = \begin{pmatrix} -2 & -3 \\ 4 & 3 \end{pmatrix} \). A straightforward check shows no two of \( A, B, C \) are integrally similar.

**Example 4.6.** Continuing on, look at \( K = \mathbb{Q}(\sqrt{-47}) \), where \( \mathcal{O}_K = \mathbb{Z}[\omega] \) for \( \omega = \frac{1+\sqrt{-47}}{2} \), with minimal polynomial \( f(x) = x^2 - x + 12 \), so that \( \omega^2 = \omega - 12 \). Again, in the final homework of 23A2, we saw \( \text{Cl}(K) = \langle \langle p_2 \rangle \rangle \cong \mathbb{Z}/5\mathbb{Z} \) for \( p_2 = (2, \omega) \), with \( \mathcal{O}_K, p_2, p_2^2 = (4, \omega), p_3^2 = (8, \omega + 4) \), and \( p_5^2 = (16, \omega + 4) \) representatives of the distinct ideal classes of \( \mathcal{O}_K \). One can then similarly construct distinct integral similarity class representatives

\[
\begin{pmatrix} 0 & -12 \\ 1 & 1 \end{pmatrix}, \begin{pmatrix} 0 & -6 \\ 2 & 1 \end{pmatrix}, \begin{pmatrix} 0 & -3 \\ 4 & 1 \end{pmatrix}, \begin{pmatrix} -4 & -4 \\ 8 & 5 \end{pmatrix}, \begin{pmatrix} -4 & -2 \\ 16 & 5 \end{pmatrix}
\]

that each have characteristic polynomial \( f \).

**Example 4.7.** Note for \( \alpha = 3 + \sqrt{10} \) we have \( K = \mathbb{Q}(\alpha) = \mathbb{Q}(\sqrt{10}) \) and \( \mathcal{O}_K = \mathbb{Z}[\sqrt{10}] = \mathbb{Z}[\alpha] \). Also, \( f(x) = (x - 3)^2 - 10 = x^2 - 6x - 1 \) is irreducible with root \( \alpha \). One can again easily show \( K \) has class number 2 (we have Minkowski constant \( M_K < 9 \), and since \( x^2 - 10 \) factors as \( x^2 \) over \( \mathbb{F}_2 \) and \( \mathbb{F}_5 \), and for \( p_2 = (2, \sqrt{10}) \) and \( p_5 = (5, \sqrt{10}) \), we see \( p_2p_5 = (\sqrt{10}) \), and \( p_2^2 = (2) \), then \( \text{Cl}(K) = \langle \langle p_2 \rangle \rangle \cong \mathbb{Z}/2\mathbb{Z} \)). Now we can find corresponding representatives of integral similarity classes with characteristic polynomials \( f(x) \) by consider the representatives \( \mathcal{O}_K, p_2 \): \( \mathcal{O}_K \) has basis \( \{1, \sqrt{10}\} \), and when multiplied by \( \alpha \), we get \( \{3 + \sqrt{10}, 10 + 3\sqrt{10}\} \) and hence we have the matrix

\[
\begin{pmatrix} 3 & 10 \\ 1 & 3 \end{pmatrix}
\]

that \( p_2 \) has basis \( \{2, \sqrt{10}\} \), and when multiplied by \( \alpha \), we get \( \{3 \cdot 2 + 2\sqrt{10}, 5 \cdot 2 + 3\sqrt{10}\} \), and hence corresponding matrix

\[
\begin{pmatrix} 3 & 5 \\ 2 & 3 \end{pmatrix}
\]

**Example 4.8.** Now consider \( f(x) = x^2 - m^2d \) for \( d \) squarefree and \( m \geq 2 \), so that \( f \) is irreducible with root \( \alpha = m\sqrt{d} \). Consider now \( \mathbb{Z}[\alpha] \) and \( \mathbb{Z}[\sqrt{d}] \), which do not lie in the same \( \mathbb{Z}[\alpha] \)-ideal
classes $(\mathbb{Z}[\alpha])$ is principal but $\mathbb{Z}[\sqrt{d}]$ is not since if $\mathbb{Z}[\sqrt{d}] = (a + b\alpha)$ for some $a, b \in \mathbb{Z}$ then $\sqrt{d} = (s + t\sqrt{d})(a + b\sqrt{d}) = (sa + btm^2d) + (sb + at)m\sqrt{d}$ for some $s, t \in \mathbb{Z}$ and so $(sb + at)m = 1$, but $m \geq 2$ yields $|sb + at| \leq \frac{1}{2}$ and $sb + at \in \mathbb{Z}$ and so $sb + at = 0$, a contradiction. Taking the basis $\{1, \sqrt{d}\}$ of $\mathbb{Z}[\sqrt{d}]$, we see under multiplication by $\alpha$ we have $\{0 \cdot 1 + m \cdot \sqrt{d}, md + 0\sqrt{d}\}$ and so corresponding matrix $\begin{pmatrix} 0 & md \\ m & 0 \end{pmatrix}$, and now considering the basis $\{1, \alpha\}$ of $\mathbb{Z}[\alpha]$, we get the companion matrix $\begin{pmatrix} 0 & m^2d \\ 1 & 0 \end{pmatrix}$ under multiplication by $\alpha$. It follows these are not integrally similar. (I am not saying anything about all ideal classes of $\mathbb{Z}[\alpha]$, just these two.)

## 5 Cyclotomic Ring of Integers

Recall a square matrix $A$ is non-derogatory if its characteristic polynomial and minimal polynomial coincide. Equivalently, the Jordan normal form of $A$ is the direct sum of Jordan blocks for distinct eigenvalues, which again is equivalent to each eigenvalue having geometric multiplicity 1. Now we can consider the following result relating to monogenic rings of integers, again following [10]:

**Theorem 5.1.** Let $f(x) \in \mathbb{Z}[x]$ be a monic irreducible polynomial of degree $n$ and let $\theta$ be a root of $f$. Let $K = \mathbb{Q}(\theta)$ and $\mathcal{O}_K$ its ring of integers. Then the following are equivalent:

(i) $\mathcal{O}_K = \mathbb{Z}[\theta]$

(ii) whenever $p$ is a rational prime and $A \in M_n(\mathbb{Z})$ has characteristic polynomial $f$, then $A \pmod{p}$ is non-derogatory.

**Proof.** Suppose $A \in M_n(\mathbb{Z})$ has characteristic polynomial $f$ and $A \pmod{p}$ is derogatory, so that $A \pmod{p}$ has minimal polynomial of degree strictly less than $n$. It follows then there exist $a_i \in \mathbb{Z}$ such that not all $a_i$ are divisible by $p$ so that

$$(a_0I + a_1A + a_2A^2 + \cdots + a_{n-1}A^{n-1}) \pmod{p} = 0_{n \times n}$$

and hence

$$a_0I + a_1A + a_2A^2 + \cdots + a_{n-1}A^{n-1} = pB$$

for some $B \in M_n(\mathbb{Z})$. Now for $\mathbf{v}$ an eigenvector of $A$ corresponding to $\theta$ with values in $\mathbb{Z}[\theta]$ (justified as before), we see then

$$pB\mathbf{v} = (a_0I + a_1A + a_2A^2 + \cdots + a_{n-1}A^{n-1})\mathbf{v} = (a_0 + a_1\theta + a_2\theta^2 + \cdots + a_{n-1}\theta^{n-1})\mathbf{v}$$

and hence

$$\alpha = a_0 + a_1\theta + a_2\theta^2 + \cdots + a_{n-1}\theta^{n-1}$$

is an eigenvalue of $B \in M_n(\mathbb{Z})$ and is hence an algebraic integer as it is a root of $\sigma_B \in \mathbb{Z}[x]$. Moreover, we see $\alpha \notin \mathbb{Z}[\theta]$ since $1, \theta, \theta^2, \ldots, \theta^{n-1}$ is linearly independent and $\frac{\alpha}{p} \notin \mathbb{Z}$ for some $i$ since $p$ did not divide all $a_i$.

Conversely, if $\mathcal{O}_K \neq \mathbb{Z}[\theta]$, then for some $p$ we can find $b_i \in \mathbb{Z}$ such that $\alpha = (b_0 + b_1\theta + \cdots + b_{n-1}\theta^{n-1})/p$ is in $\mathcal{O}_K$ but not in $\mathbb{Z}[\theta]$ (so again, $p$ does not divide all $b_i$). If $\omega_1, \ldots, \omega_n$ is a basis for $\mathcal{O}_K$ and $A$ represents multiplication on this basis by $\theta$, then

$$B = \frac{b_0I + b_1A + b_2A^2 + \cdots + a_{n-1}A^{n-1}}{p}$$


9
represents multiplication on this basis by \( \alpha \), so \( B \in M_n(\mathbb{Z}) \). It follows then
\[
(b_0I + b_1A + b_2A^2 + \cdots + a_{n-1}A^{n-1}) \pmod{p} = 0_{n \times n}
\]
so that \( A \pmod{p} \) is derogatory. \( \square \)

Note next this simple lemma:

**Lemma 5.2.** Suppose \((m, n) = 1 \) and \( \ell = \varphi(mn) = \varphi(m)\varphi(n) \). If \( A \in M_\ell(\mathbb{Z}) \) has characteristic polynomial \( \Phi_{mn} \), then there exist \( B, C \in M_\ell(\mathbb{Z}) \) such that \( A = BC = CB \) and \( B, C \) respectively have minimal polynomials of \( \Phi_m \) and \( \Phi_n \) (and hence characteristic polynomials \( \Phi_m^{\varphi(n)} \) and \( \Phi_n^{\varphi(m)} \)), and \( A, B, C \) have the same eigenvectors.

**Proof.** Suppose \( A \) has characteristic polynomial \( \Phi_{mn} \), then \( B = A^n \) has eigenvalues precisely \( \zeta_{mn}^k = \zeta_m^k \) for \((k, mn) = (k, m) = 1 \) and so \( B \) has minimal polynomial \( \Phi_m \), and \( B \) then commutes with \( A \).

Similarly, setting \( C = A^{1-n} \), clearly then \( A = BC = CB \) and \( C \) has eigenvalues precisely \( \zeta_{mn}^{(1-n)k} = \zeta_m^k \zeta_n^{-k} = \zeta_m^k \) for \((k, mn) = 1 \), so \( C \) has minimal polynomial \( \Phi_n \).

Also, by construction, \( A, B, C \) have the same eigenvectors as these are all eigenvectors for \( A \).
\( \square \)

Now we can provide an alternative proof of the following result:

**Theorem 5.3.** For \( K = \mathbb{Q}(\zeta_n) \) with \( \zeta_n \) a primitive \( n^{\text{th}} \) root of unity, then \( \mathcal{O}_K = \mathbb{Z}[\zeta_n] \).

**Proof.** By the prior theorem, for \( \Phi_n \) the \( n^{\text{th}} \) cyclotomic polynomial and \( m = \varphi(n) \), it suffices to show if \( A \in M_m(\mathbb{Z}) \) has characteristic polynomial \( \Phi_n \), then \( A \pmod{p} \) is non-derogatory for any rational prime \( p \).

Suppose first \( n = p^r \) and \( r \geq 1 \), where we note \( \Phi_n(x) = \Phi_p(x^{p^{r-1}}) \); in particular, \( \Phi_n(1) = \Phi_p(1) = p \). Suppose further \( A \pmod{p} \) is derogatory for some prime \( q \). Since \( x^{p^r} - 1 \) is separable over \( \mathbb{F}_q \) if \( q \neq p \), then we must necessarily have \( q = p \). Since \( \Phi_p(x) = (x - 1)^m \) over \( \mathbb{F}_p \), then the Jordan form of \( A \) over \( \mathbb{F}_p \) consists only of Jordan blocks with eigenvalue 1. Since \( A \pmod{p} \) is derogatory, then \((A - I)^{m-1} \pmod{p} = 0_{m \times m} \) and hence \((A - I)^{m-1} = pB \) for some \( B \in M_m(\mathbb{Z}) \). Since \( \zeta_n \) is a root of \( \Phi_n \), we then have some eigenvector \( v \) of \( A \) corresponding to \( \zeta_n \), which yields \( pBv = (A - I)^{m-1}v = (\zeta_n - 1)^{m-1}v \), and so \( (\zeta_n - 1)^{m-1} \pmod{p} \) is an eigenvalue of \( B \in M_m(\mathbb{Z}) \) and is hence in \( \mathcal{O}_K \) since it is a root of \( \sigma_B \in \mathbb{Z}[x] \). It follows the product of all such numbers is an algebraic integer, but
\[
\prod_{d\mid m} \frac{(\zeta_n^d - 1)^{m-1}}{p^{d-1}} = \frac{\pm \Phi_n(1)^{m-1}}{p^m} = \frac{\pm p^{m-1}}{p^m} = \frac{\pm 1}{p} 
\]
is not an algebraic integer (since \( \mathcal{O}_\mathbb{Q} = \mathbb{Z} \)), a contradiction. It follows then \( A \pmod{q} \) is non-derogatory for all primes \( q \).

Now for general \( n \), suppose the result holds for \( k < n \). Suppose again \( A \in M_m(\mathbb{Z}) \) has characteristic polynomial \( \Phi_n(x) \) and we have some prime \( p \) such that \( A \pmod{p} \) is derogatory. It follows again necessarily \( p \nmid n \) since \( \Phi_n \) is separable over \( \mathbb{F}_q \) if \( q \nmid n \). Write \( n = p^r n' \) such that \( p \nmid n' \). For convenience, write \( s = \varphi(p^r) = p^r - p^{r-1} \) and \( t = \varphi(n') \). By Lemma 5.2, we can write \( A = BC = CB \) where \( \Phi_p \) is the minimal polynomial for \( B \) and \( \Phi_{n'} \) is the minimal polynomial for \( C \), and all have the same eigenvectors. We can further find \( U_j, V_j \in \text{GL}(n, \mathbb{Z}) \) such that \( U_1^{-1}BU_1 \) and \( V_1^{-1}CV_1 \) are in upper triangular block form with each block, say, respectively \( X_i, Y_j \), having characteristic polynomial \( \Phi_{p^i} \) and \( \Phi_{n'} \).

By the inductive hypothesis (and the prior case), \( X_i \pmod{p} \) is non-derogatory and so \( X_i \) is similar over \( \mathbb{F}_p \) to \( J_i \), the full \( s \times s \) Jordan block with eigenvalue 1 since \( \Phi_{p^i}(x) = (x - 1)^s \) over \( \mathbb{F}_p \). Also by the inductive hypothesis, we have \( Y_j \pmod{p} \) is non-derogatory, and so each eigenvalue of
Y_j (mod p) has geometric multiplicity 1 and each Y_j is similar to the t × t companion matrix of Φ_n', say M, by Lemma 4.2. It follows then B is similar over \( \mathbb{F}_p \) to a direct sum of t identical full s × s Jordan blocks \( J_i = J \), that is,

\[
P = \text{diag}(J_1, \ldots, J_t).
\]

Also, C is similar over \( \mathbb{F}_p \) to a diagonal block matrix with s blocks being identically M (the companion matrix for \( \Phi_n' \)). Since \( \Phi_n' \) splits over \( \mathbb{F}_p \), we have M is similar over its splitting field (viz., \( \Phi_n'(\zeta_n') \)) to the diagonal matrix \( \text{diag}(\lambda_1, \ldots, \lambda_t) \) for \( \lambda_i \) a distinct primitive \( n' \)-th root of unity, and hence C is similar to the diagonal block matrix

\[
Q = \text{diag}(\lambda_1 I, \ldots, \lambda_t I),
\]

where I is the s × s identity matrix; in particular, Q is diagonal.

Now since BC = CB, then for U that diagonalizes C as above, we have

\[
A^U = (BC)^U = B^U C^U = B^U Q = QB^U.
\]

Again, now as \( B^U \) and Q commute, then they preserve eigenspaces of one another. In particular, since \( V_{\lambda_i} = \text{Span}\{e_{i(1)} + 1, \ldots, e_{i(s)}\} \) is the corresponding eigenspace for Q, then \( B^U V_{\lambda_i} = V_{\lambda_i} \) implies the first s columns of \( B^U \) must be only nonzero in the top s × s submatrix, and so on, so that \( B^U \) is block diagonal, with \( B^U = \text{diag}(X'_1, \ldots, X'_t) \) and hence \( A^U = B^U Q = \text{diag}(\lambda_1 X'_1, \ldots, \lambda_t X'_t) \). Since B was also similar to \( \text{diag}(J_1, \ldots, J_t) \), then it follows \( X'_j \) is similar to \( J_k \) for each \( i \), in particular, \( (\lambda_i I - \lambda_i X'_1)k = X'_1(I - X'_1)^k \neq 0 \) for \( k < s \), so that each eigenvalue of \( A^U \) and hence of A has geometric multiplicity 1. It follows then A (mod \( \phi \)) is also non-derogatory (and derogatory), a contradiction. Hence, necessarily A (mod \( q \)) is non-derogatory for each prime \( q \), and so \( \mathcal{O}_K = \mathbb{Z}[^{\zeta_n}] \).

6 Random Integral Matrices

Now we can take a break from the deterministic world and explore matrices whose entries are integer valued random variables. Some seemingly straightforward questions one might ask include whether such a matrix is singular. For instance, for A sampled from the \( n \times n \) Bernoulli 0, 1 ensemble with probability \( \frac{1}{2} \), it is easy to calculate \( p_n = P(\text{det}(A) = 0) \) for small n: obviously \( p_1 = \frac{1}{2} \) while

\[
p_2 = 1 - \frac{|\text{GL}(2, 2)|}{2^4} = \frac{5}{8} = 0.625
\]

(since |det(A)| ≤ 1 and so det(A) ≠ 0 iff det(\( A \)) = ±1 iff \( A \in \text{GL}(2, 2) \)). To calculate \( p_3 \), we can check

\[
p_3 = 1 - \frac{|\text{GL}(3, 2)|}{2^9} - \frac{\# \{ A : \text{det}(A) = \pm 2 \}}{2^9} = \frac{169}{256} = 0.6602
\]

using now the fact |det(A)| ≤ 3, and noting also if A has a column with 2 zeros then |det(A)| ≤ 1 (by the \( n = 2 \) case) and similarly if A has a column with three ones. This leads to considering only matrices with only 3 nonzero entries, which can be easily verified to have |det(A)| = 2, and there are 3 choices in the first column for the first 0 and then only 2 choices in the second column for the second 0, which determines exactly where the last 0 goes. It gets more complicated for \( n ≥ 4 \).

Even if we switch approaches (which seems necessary since Hadamard’s bound of a 0, 1 matrix shows

\[
|\text{det}(A)| ≤ \frac{(n + 1)(n + 1)/2}{2^n}
\]

and this bound can be sharp when 4 | n — that is, if there is a Hadamard matrix) by considering whether each successive column is in the span of the prior columns. We could try calculating \( p_n \).
by going through the cases iteratively one column at a time (using inclusion-exclusion along with independence), and it seems one might be en route to give a proper asymptotic of $p_n$. But exactly this has eluded us thus far, and this asymptotic of $p_n$ is still an open problem.

With the top few calculations, one may naively conclude that $p_n \rightarrow 0$ as $n \rightarrow \infty$. However, this is false. One way (in line with the above approach, noting $A \in \text{GL}(n, 2)$ if and only if $\det(A)$ is odd) is to note

$$p_n \leq 1 - \frac{|\text{GL}(n, 2)|}{2^n} := 1 - K_n$$

(where now $K_n$ denotes the probability $\det(A)$ is odd). Since

$$K_n = 2^{-n} \prod_{j=0}^{n-1} (2^n - 2^j) = \prod_{j=1}^{n} (1 - 2^{-j}) = 1 \frac{3}{2} \frac{7}{4} \cdot \cdots \cdot (1 - 2^{-n})$$

and we see

$$\sum_{j \geq 1} \log(1 - 2^{-j}) = - \sum_{j \geq 1} \sum_{k \geq 1} 2^{-jk} - \frac{1}{k} \sum_{k \geq 1} 2^{-jk} = - \sum_{k \geq 1} \frac{1}{k} \cdot \frac{1}{2^k - 1}$$

> $\frac{-2}{2} \sum_{k \geq 1} \frac{2^{-k}}{k} = \log(2^{2})$, 

then we have $K_n > \frac{1}{4}$, so that $p_n < \frac{3}{4}$. (In fact, it can be shown $K_n = (1/2; 1/2)_\infty \approx 0.288788095$, using the $q$-Pochhammer symbol, as noted on oeis.org/A048651.)

What’s more, we actually have $p_n \rightarrow 0$: Komlós [8] in the 1960s showed $p_n = O(n^{-1/2})$, which has been improved in the decades since to exponential decay, such as by Tao and Vu [14] who showed $p_n \leq (3/4 + o(1))^n$ using tools from additive combinatorics and recently by Bourgain, Vu and Philip Matchett Wood [1], who showed $p_n \leq (1/\sqrt{2} + o(1))^n$. It has been conjectured that $p_n = (1/2 + o(1))^n$, noting $p_n \geq 1 + o(1)n^221^{-n} = (1/2 + o(1))^n$ by considering the lower bound of the probability that two rows or columns are the same [14].

Now suppose $m = n + u$ for $u \geq 0$. If a $n \times m$ matrix $A$ over the integers has full rank, then one may next ask whether such a matrix is surjective as a $\mathbb{Z}$-module homomorphism $A : \mathbb{Z}^m \rightarrow \mathbb{Z}^n$. This is equivalent to asking whether the cokernel $\text{coker}(A) = \mathbb{Z}^n/\text{Im}(A)$ is trivial, and hence also similar to asking when $s_n = 1$ in the random SNF of the random matrix $A$. In the case $m = n$, we have $A$ is not surjective with high probability: In fact, as found in [15] and referenced in [13], Melanie Matchett Wood gives the result:

**Theorem 6.1.** Let $u \geq 0$ be fixed and $A_n \in M_{n,n+u}(\xi)$ be a random matrix with entries being iid copies of an $\varepsilon$-balanced random variable $\xi$ of fixed $\varepsilon > 0$. Let $P$ be a finite set of primes, then

$$\lim_{n \rightarrow \infty} \mathbb{P}(\text{coker}(A_n) \not\equiv 1) = \prod_{p \in P} \prod_{k=1}^{\infty} (1 - p^{-k-u}),$$

where $G_p = \prod_{p \in P} G_p$ for $G_p$ the $p$-Sylow subgroup of abelian $G$ and $\xi$ is $\varepsilon$-balanced if

$$\max_{x \in \mathbb{F}_p} \mathbb{P}(\xi = x) \leq 1 - \varepsilon$$

for any prime $p$.

Now note as $P$ increases, the right hand side probability in the prior theorem gets arbitrarily small, which yields:

$$\lim_{n \rightarrow \infty} \sup_{\text{coker}(A_n) \not\equiv 1} \leq \inf_{P} \lim_{n \rightarrow \infty} \mathbb{P}(\text{coker}(A_n) \not\equiv 1) = 0,$$
and so square random matrices are not surjective with high probability for large \( n \), as claimed.

Rectangular matrices are a different story, and Nguyen and Paquette give the very recent result that \( A_n \in M_{n,n+u}(\xi) \) with iid \( \varepsilon \)-balanced entries is surjective on the integral lattice with high probability if \( u = O(\log^{1+o(1)}(n)) \) [13].

The prior result that Wood gave is related to the the Cohen-Lenstra Heuristics, which is a conjecture relating the distribution of the \( p \)-Sylow subgroups of class groups of quadratic number fields with negative discriminants:

**Conjecture 1** (Cohen and Lenstra). Let \( p \) be an odd prime. Let \( S_X^{-1} \) be the set of negative fundamental discriminants \( d \geq -X \). Let \( B \) be a finite abelian \( p \)-group. Then

\[
\lim_{X \to \infty} \frac{\# \{ d \in S_X^{-1} : \text{Cl}(\mathbb{Q}(\sqrt{d}))_p \cong B \}}{|S_X^{-1}|} = \prod_{k=1}^{\infty} \frac{1 - p^{-k}}{|\text{Aut}(B)|}.
\]

(Cohen and Lenstra also conjecture similar results for class groups with positive discriminants and cokernels of rectangular entries, as mentioned earlier.) Results have been shown in support of this conjecture using random matrices \( X_n \) drawn with respect to Haar measure on \( M_n(F_p) \) [4], which gives:

\[
\lim_{n \to \infty} \mathbb{P}(\text{coker}(X_n) \cong B) = \prod_{k=1}^{\infty} \frac{1 - p^{-k}}{|\text{Aut}(B)|}.
\]

In fact, any class group can be realized then as the cokernel of a random matrix: since we have for the map \( \alpha \mapsto (\alpha) \) then

\[
\text{Cl}(K) = \text{coker}(\mathcal{O}_K^S \times \rightarrow I_K^S)
\]

for \( S \) a generating set of ideals in \( \text{Cl}(K) \), \( \mathcal{O}_K^S \) the \( S \)-units in \( \mathcal{O}_K \) and \( I_K^S \) the abelian group of fractional ideals generated by elements of \( S \), then we also have

\[
\text{Cl}(K)_p = \text{coker}(\mathcal{O}_K^S \times \otimes \mathbb{Z}/m\mathbb{Z} \rightarrow I_K^S \otimes \mathbb{Z}/m\mathbb{Z}),
\]

(where we recall \( \mathbb{Z}/m\mathbb{Z} \otimes \mathbb{Z}/n\mathbb{Z} \cong 0 \) if \( (m, n) = 1 \) and hence \( \text{Cl}(K)_p \) is the cokernel of a random \( p \)-adic square matrix (depending on \( d \)). In particular, the prior conjecture then gives a statement about the distribution of cokernels of certain random matrices.

In [15], Wood further generates Cohen-Lenstra results by weakening some conditions on independence of matrix entries, and using different starting distributions. This leads to a universality result for the Cohen-Lenstra Heuristics, akin to results in Random Matrix Theory. Universality relates to how (local or global) asymptotic statistics of certain ensembles are independent of the starting distributions of their entries. For example, other widely known universality results include the Central Limit Theorem from Probability Theory (showing scaled and centered sums of iid random variables converge weakly, almost surely, to the Gaussian distribution) and the Semicircle Law from Random Matrix Theory (showing the empirical spectral distribution of a Wigner random matrix, that is, square \( M \) such that \( M = M^* \), with \( M_{ij} \) iid for \( i > j \), and \( M_{ii} \) iid, independent of the off diagonals, with bounded variance, converges to the semicircle distribution \( \sigma_{sc}(dx) = \frac{1}{2\pi} (4 - x^2)^{1/2} dx \) weakly, almost surely.

Wood and many of her collaborators (including at least one local luminary) continue to work toward positively proving this conjecture.

**References**


