Natural Density Distribution of Hermite Normal Forms of Integer Matrices *

Gérard Maze
e-mail: gerard.maze@math.uzh.ch
Mathematics Institute
University of Zürich
Winterthurerstr 190, CH-8057 Zürich, Switzerland

September 27, 2010

Abstract

The Hermite Normal Form (HNF) is a canonical representation of matrices over any principal ideal domain. Over the integers, the distribution of the HNFs of randomly looking matrices is far from uniform. The aim of this article is to present an explicit computation of this distribution together with some applications. More precisely, for integer matrices whose entries are upper bounded in absolute value by a large bound, we compute the asymptotic number of such matrices whose HNF has a prescribed diagonal structure. We apply these results to the analysis of some procedures and algorithms whose dynamics depend on the HNF of randomly looking integer matrices.

Key Words: Natural density, Hermite normal form, integer lattices.
Subject Classification: 15A21, 05A16, 52C07, 15B36.

1 Introduction

Given a principal ideal domain R, the notion of Hermite Normal Form (HNF) of a $n \times m$ matrix with entries in R is well defined. When $R = \mathbb{Z}$, which will be the case in this article, a matrix in HNF can be defined as follows, see e.g. [5, 12]:

Definition 1 (Hermite Normal Form (HNF)) *A $n \times m$ matrix H with integer entries is in Hermite normal form if H is upper diagonal with the following properties:*

1. *The first r rows of H are the non-zero rows of H,*
2. *for each row i, if h_{ij_i} is its first non-zero entry, then $h_{ij_i} > 0$ and $j_1 < j_2 < ... < j_r$,*
3. *for each $1 \leq k < i \leq r$, the entries h_{kj_i} of the j_i^{th} column of H satisfy $0 \leq h_{kj_i} < h_{ij_i}$.*

The positive integers h_{ij_i} are called the pivot of the matrix in HNF.

The main result about HNF, discovered by Charles Hermite, is that for all $n \times m$ integer matrix A, there exists a (possibly non-unique) unimodular $n \times n$ matrix U and a unique $n \times m$ integer matrix H in HNF such that $A = UH$. The left equivalence between A and H means that there is a sequence of elementary row operations that will produce H when applied to A. Note that the definition of the HNF can slightly change in the literature (e.g. lower triangular vs. upper triangular, column operations vs. row operations). Since the matrix H is uniquely defined, we can write without ambiguity $H = \mathrm{HNF}(A)$. Typically, the shape of a matrix in HNF will be the following:

*Preprint submitted for publication

$$\begin{array}{ccccccccc}
* & * & * & * & * & * & * & * & * \\
0 & * & * & * & * & * & * & * & * \\
0 & 0 & 0 & * & * & * & * & * & * \\
0 & 0 & 0 & 0 & * & * & * & * & * \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & * & * \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & * \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0
\end{array}$$

The above example is given with $(n, m, r) = (8, 9, 6)$ and the sequence j_i of column positions of the pivots is $1, 2, 4, 5, 8, 9$. As a matter of fact, only a very small proportion of HNFs of integer matrices has this type of shape. Anyone who had to compute the HNF of arbitrary integer matrices more than once was forced to observe that they do not appear "randomly", that is, the elements h_{ij_i} do not seem to follow an equiprobable law of distribution. For instance, the case $j_i = i$ and $r = \min(n, m)$ appears predominantly, and a strongly recurring structure is that all the pivots h_{ij_i} with $i < r$ are small and increasing with i (typically less than 10, even for matrices with very large entries) and the last pivot h_{ij_r} is large (of the order of $\det A$ when $n = m$). This particular point is intrinsically interesting, but was also used in several occasions (see e.g. [1, 15, 18]) in order to heuristically understand or analyze the behavior of an algorithm.

Our focus in this paper is set on the "probability" that the HNF of a random $n \times m$ integer matrix has a given diagonal. We aim to obtain an explanation of the strong biases mentioned above. For instance, Proposition 6 below gives the frequency of appearance of a given non-zero diagonal in the HNF of a randomly looking matrix. Proposition 6 also shows that the density of HNFs with the above shape is in fact 0. Corollary 7 shows that given strictly positive integers $d_1, d_2, ..., d_{n-1}$, the "probability" that a $n \times n$ integer matrix A has a HNF of the form

$$\mathrm{HNF}\,(A) = \begin{bmatrix}
d_1 & * & * & * & * \\
0 & d_2 & * & * & * \\
0 & 0 & \ddots & * & * \\
0 & 0 & 0 & d_{n-1} & * \\
0 & 0 & 0 & 0 & d
\end{bmatrix}$$

where $d = \frac{\det(A)}{\prod_{i=1..n-1} d_i}$ is given by

$$\left(\zeta(n) \cdot \zeta(n-1) ... \cdot \zeta(2) \cdot d_1^n \cdot d_2^{n-1} \cdot \ldots \cdot d_{n-1}^2\right)^{-1},$$

where ζ is the usual zeta function. Of course, the notion of "probability" and "density" used here have to be made precise. The appropriate concept is the notion of natural density, see e.g. [22]. Different definition of densities appear naturally in analytic number theory with the study of prime numbers and expected values of arithmetic functions, see e.g. [22] and [9] for several examples. As for the *natural density*, the explicit multidimensional aspects of the question appear in e.g. [10, 13, 14] and more implicitly in e.g. [1, 3, 4, 8, 11]. On the more specialized study of density of canonical form of matrices, let us mention the work of Evans [6] where the density of Smith normal form over the ring of integers of a local field is studied. The subject treated in the present article does not seem to have been the object of a publication in the past.

The article is structured as follows. We address the question of a suitable definition of natural density in $\mathbb{Z}^k$ in Section 2 below. In Section 3 we present some results linking unimodular matrices and natural density of vectors. The main results of the article are stated and proved in Section 4 and in Section 5 we present some applications.

We will used the following notations. The set of primes in $\mathbb{N}^* = \mathbb{N} \setminus \{0\}$ is $\mathbb{P}$, Landau's notations $f(x) = o(x)$ and $g(x) = O(x)$ mean that $\lim_{x\to\infty} f(x)/x = 0$ and $\lim_{x\to\infty} g(x)/x = c$. The Riemann zeta function ζ is $\zeta(s) = \sum_{n\geq 1} n^{-s} = \prod_{p\in\mathbb{P}}(1 - p^{-s})^{-1}$. The cardinality of a set S is $|S|$. We will also use the expression "randomly looking (integer) vector" in an informal way, meaning that the

entries of the vector have been chosen uniformly at random in a large interval $[-B, B[$. The symbol $*$ represents an integer whose value is unimportant depending on the context.

2 Natural density in $\mathbb{Z}^k$

In order to make the intuitive notion of probability in $\mathbb{Z}^k$ precise we first remark that the uniform distribution over $\mathbb{Z}^k$ or over $\mathbb{N}^k$, even when $k = 1$, has little meaning. For this reason researchers often use the concept of *natural density* when stating probability results in $\mathbb{N}$. In the following we briefly explain this concept. Let $S \subset \mathbb{N}$ be a set. Define the upper (respectively lower) natural density as

$$\overline{\mathbb{D}}(S) = \limsup_{B\to\infty} \frac{|S \cap [0, B[\,|}{B} \quad , \quad \underline{\mathbb{D}}(S) = \liminf_{B\to\infty} \frac{|S \cap [0, B[\,|}{B}.$$

When both limits are equal one defines the natural density of the set S as

$$\mathbb{D}(S) := \overline{\mathbb{D}}(S) = \underline{\mathbb{D}}(S).$$

The notion of natural density allows to tackle questions related to the frequency of realization of events concerning randomly looking integers, i.e., for uniformly chosen integers in $[0, B[$, when B goes to infinity. A famous example in $\mathbb{N}$ is that the natural density of square free integers is $6/\pi^2 = \zeta(2)^{-1}$, see e.g. [9]. The extension of the above definition in higher dimension is sometimes implicit in the literature. For example the natural density of n coprime integers, equal to $\zeta(n)^{-1}$, has been studied by several authors, starting with Cesàro in 1884 [4] (1881 for the case $n = 2$ [2, 3]), Lehmer in 1900 [11] and Nymann [17]. For the historical fatherhood of the result see [14]. This natural density means that there are $\zeta(n)^{-1} \cdot B^n + o(B^n)$ n-vector in $[0, B[^n$ whose entries are coprime. An explicit definition of a higher dimensional notion of natural density has been developed in e.g. [10, 13, 14]. In these articles, the notion of natural density of a set S in $\mathbb{Z}^k$ is defined as a "centered symmetric cube" version of the unidimensional definition, i.e., as the limit, when it exists, $D(S) = \lim_{B\to\infty} \frac{|S\cap[-B,B[^k|}{(2B)^k}$. We will however need a stronger definition. In order to see why, let us consider a set S in $\mathbb{N}$ with density δ. Since $|S \cap [0, l[| = \delta \cdot l + o(l)$, any interval $[l, l + B[$ with $l = o(B)$ contains $\delta \cdot B + o(B)$ elements of S. Being able to estimate the local density in non centered cubes does not seem to be always possible in dimension $k > 1$ with the above definition of density. In order to achieve this, we require in the definition that the cubes can lie anywhere in $\mathbb{Z}^k$. In the sequel, we call a cube any set of the form $\prod_{i=1}^k [z_i - B, z_i + B[^k$ for some $z \in \mathbb{Z}^k$ and $B > 0$.

Definition 2 (Natural density in $\mathbb{Z}^k$) Let S be a subset of $\mathbb{Z}^k$. If for all $z \in \mathbb{Z}^k$, the following limit exists

$$\mathbb{D}(S) = \lim_{B\to\infty} \frac{|S \cap \prod_{i=1}^k [z_i - B, z_i + B[\,|}{(2B)^k}$$

and is independent of z, then it is called the natural density of S.

Let us notice that it would have been even possible to extend the definition of natural density from $\mathbb{Z}$ to $\mathbb{Z}^k$ by using k-rectangles instead of cubes (i.e. different B_i for each dimension). However both definition are equivalent since rectangles can be decomposed into smaller cubes. We will not use this property in the sequel. Another direction of generalization is the spherical model. This setting considers centered n-balls instead of n-cubes. Due to the symmetry of the balls around the origin, it is a natural choice in the study of different asymptotic results concerning lattices, integer matrices, and varieties in general see e.g. [16, 19]. This model suffers however from the same problem as noted before and from the fact that the entries of the different objects of study are not independent anymore, i.e., the "random looking aspect" is somehow lost.

In order to prove our main results, we will need the existence and the value of the natural density of tuples of integers whose greatest common divisor is a given positive integer d. This is treated in Lemma 3 below. As mentioned in the introduction, in the weaker form of density definition given above, this problem has been studied several authors, see e.g. [4, 11].

Lemma 3 *The set* $\{(x_1, \ldots, x_k) \in \mathbb{Z}^k : \gcd(x_i) = d\}$ *has a density equal to* $(\zeta(k) \cdot d^k)^{-1}$.

Proof: Let $x \in \mathbb{Z}^k$ and $S = \{(x_1, \ldots, x_k) \in \mathbb{Z}^k : \gcd(x_i) = d\}$. Then $x \in S$ if and only if $x_i/d \in \mathbb{Z}$ and $\gcd(x_i/d) = 1$. Let $z'_i = z_i/d$, $B' = B/d$ and $S' = \{(x_1, \ldots, x_k) \in \mathbb{Z}^k : \gcd(x_i) = 1\}$. The first equality of the following equations is straightforward.

$$\left| S \cap \prod_{i=1}^{k} [z_i - B, z_i + B[\right| = \left| S' \cap \prod_{i=1}^{k} [z'_i - B', z'_i + B'[\right| = \left| S' \cap \prod_{i=1}^{k} [-B', B'[\right| + o(B^k). \qquad (1)$$

In order to prove that the second equality of Equation (1) is valid, consider an element $(x_1, \ldots, x_k)$ in the set of the left hand side of the equality. Let us fix all the components but the i^{th} one, and consider $t = \gcd_{j \neq i}(x_j)$. The integers $x_1, \ldots, x_k$ are coprime if and only if x_i has no common factor with t. So if P is the set of prime divisors of t, both the interval $[z'_i - B', z'_i + B'[$ and $[-B', B'[$ contains $2B' \prod_{p \in P} \left(1 - \frac{1}{p}\right) + o(2B')$ integer coprime to the fixed x_j. This shows that the error resulting in setting $z_i = 0$ in the mid term of Equality (1) can be adjusted by $o(B') = o(B)$. Taking into account the effect of all dimensions together leads to the correction term $o(B^k)$. Now, as mentioned before, see e.g. [11, 14], we have

$$\left| S' \cap \prod_{i=1}^{k} [-B', B'[\right| = \zeta(k)^{-1}(2B')^k + o((B')^k)$$

which leads to

$$\mathbb{D}(S) = \lim_{B \to \infty} \frac{|S \cap \prod_{i=1}^{k} [z_i - B, z_i + B[|}{(2B)^k} = \lim_{B \to \infty} \frac{\zeta(k)^{-1}(2B')^k + o((B')^k)}{(2B)^k} = (\zeta(k) \cdot d^k)^{-1}$$

□

3 Generalities on unimodular matrices

For a matrix $A \in \mathbb{Z}^{n \times m}$ with $n < m$, the following three properties are equivalent, see e.g. [12]:

1. A can be completed into a $m \times m$ invertible matrix over $\mathbb{Z}$,
2. there exists a $m \times n$ integer matrix B such that $AB = Id_n$,
3. the $n \times n$ minors of A are coprime.

In the square case, i.e. when $A \in \mathbb{Z}^{n \times n}$, the third condition has no real meaning and the first one means that A is invertible. In the sequel we will adopt the usual convention and call a $n \times m$ matrix A over $\mathbb{Z}$ *unimodular* if $n \leq m$ and if it fulfills one of the first two above conditions. Unimodular matrices play a special role with respect to sets with densities as shown in the next lemma:

Lemma 4 *Let* $S \subset \mathbb{Z}^m$ *be a set with density* $\delta > 0$ *and let* V *be a* $m \times m$ *unimodular matrix. Then* $V(S) = \{Vx : x \in S\}$ *has a density equal to* δ.

Proof: Given a cube $\sigma = \prod_{i=1}^{k} [z_i - B, z_i + B[$ in $\mathbb{Z}^k$, let us count the number of points of the set $V(S)$ that lie inside σ. Since V is a bijection, this number is exactly the number of elements of S inside $V^{-1}(\sigma)$. The map V^{-1} is linear, and thus $V^{-1}(\sigma)$ is a k dimensional parallelepiped whose boundary $\partial V^{-1}(\sigma)$ is a union of parallelepipeds of dimension $k-1$. Let us cover $V^{-1}(\sigma)$ with a disjoint union of N cubes of side length B_0 with $B_0 = o(B)$, where B_0 is an unbounded function of B, e.g. $B_0 = \ln(B)$. Since V is unimodular, the volume of $V^{-1}(\sigma)$ is $(2B)^k$, and taking into account the border effect (see Figure 1), we have

$$N = \frac{(2B)^k + O(B_0 B^{k-1})}{(2B_0)^k}.$$

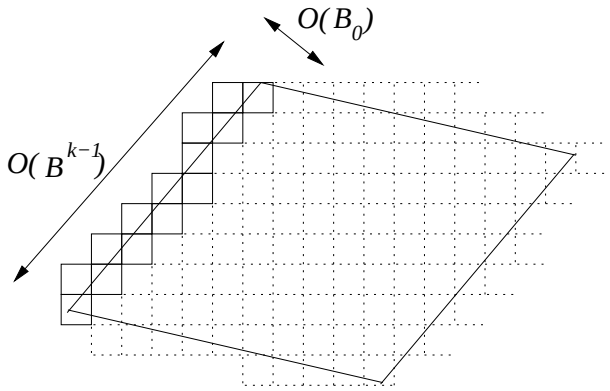

Figure 1: The border effect on $V^{-1}(\sigma)$

Each of the cubes contains $\delta(2B_0)^k + o(B_0^k)$ points of S. Therefore,

$$\begin{aligned} |S \cap V^{-1}(\sigma)| &= N \cdot (\delta(2B_0)^k + o(B_0^k)) \\ &= \delta(2B)^k + O(B_0 B^{k-1}) + \big((2B)^k + O(B_0 B^{k-1})\big) \cdot \frac{o(B_0^k)}{(2B_0)^k}. \end{aligned}$$

Finally, using the conditions on B_0, we see that

$$\lim_{B\to\infty} \frac{|V(S) \cap \sigma|}{(2B)^k} = \lim_{B\to\infty} \frac{|S \cap V^{-1}(\sigma)|}{(2B)^k} = \delta$$

which finishes the proof of the lemma. □

The previous lemma can be used to prove that unimodular matrices keep the density of vectors with entries of given greatest common divisor invariant. More precisely, we have the following proposition.

Proposition 5 *Let $U \in \mathbb{Z}^{n\times m}$, $n \leq m$ be a unimodular matrix and $d \in \mathbb{N}^*$. Then*

$$\mathbb{D}(x \in \mathbb{Z}^m \,:\, \gcd((Ux)_i \,:\, i = 1, \ldots, n)) = d) = (\zeta(n) \cdot d^n)^{-1}.$$

Proof: Consider $S = \{(y_1, \ldots, y_m) \in \mathbb{Z}^m \,:\, \gcd(y_i \,:\, i = 1, \ldots n) = d\}$. Because of Lemma 3 above, the set S has a density equal to $(\zeta(n) \cdot d^n)^{-1}$. Since U is unimodular, there exists a square $m \times m$ unimodular matrix W whose first n rows equal the n rows of U. The result follows by applying Lemma 4 with $V = W^{-1}$ to S since $x \in V(S)$ if and only if $Wx = y \in S$. □

4 Distribution of Hermite normal forms

We start this section by noticing that the pivots h_{ij_i} of the HNF of a matrix A are determined by the greatest common divisor of the $i \times i$ minors of the matrix that consists in the j_l^{th} columns of A, for $l = 1, \ldots, i$. This is true because these gcd's are left invariant when A is multiplied on the left by any unimodular matrix, and because the gcd of the $i \times i$ minors of the matrix that consists in the j_l^{th} columns of the HNF of A, for $l = 1, \ldots, i$, is precisely equal to $\prod_{l=1,\ldots,i} h_{ij_i}$ (all the other determinants are zero due to the shape of the HNF of A). Note that when $j_i = i$ the above minors are simply the $i \times i$ minors of the first i columns of A.

This property can be used as a basis of a basic algorithm to compute the HNF of A. We start by computing the greatest common divisor h_1 of the entries of the first non-zero column of A. Using

the extended Euclidean algorithm we can express h_1 as a linear combination of the entries of the column. In a matrix form, this means that there exists a sequence of row operations, i.e., there exists a unimodular matrix U_1, such that the first non-zero column of $U_1 A$ is $[h_1, 0, \ldots, 0]^t$. This process can be repeated recursively as follows. There exists a unimodular matrix U_k such that the first k columns of $U_k A$ form a matrix in HNF, as follows:

$$U_k A = \begin{bmatrix} * & \cdots & * & * & * & * & * & * \\ 0 & \ddots & * & * & * & * & * & * \\ & 0 & * & * & * & * & * & * \\ & & 0 & x_1 & * & * & * & * \\ & & \vdots & \vdots & & & & \vdots \\ & & 0 & x_s & * & * & * & * \end{bmatrix}. \tag{2}$$

Let $[*, \ldots, *, x_1, \ldots, x_s]^t$ be the $(k+1)$-th column of $U_k A$. If $x_i = 0$, $i = 1, \ldots, s$, then this column is disregarded. The next column for which one of the x_i is non-zero is selected. Using the previous remark, the next pivot h_{ij_i} is given by $h_{ij_i} = \gcd(x_i)$, and using elementary row operations, there exists a unimodular matrix U_{k+1} such that the corresponding column of $U_{k+1} A$ is $[*, \ldots, *, h_{ij_i}, 0, \ldots, 0]^t$. Appropriate elementary row operations can modify U_{k+1} and force the $*$ elements of the column to satisfy the conditions of the HNF, i.e., to belongs to $[0, h_{ij_i}[$. At the end of the process, the resulting matrix is clearly in HNF and must therefore be HNF (A). This algorithmic approach will be useful in the proof of Proposition 6 below. The key point is that we can can construct the HNF of A column after column, from left to right, via a sequence of left multiplications by unimodular matrices.

In order to simplify the statement of our results, let us use the following notation. For any $n \times m$ matrix A, the diagonal diag (A) of A is the list of elements $(a_{ii})_{i=1,\ldots,\min(n,m)}$. For given n and m, if $d_1, d_2, ..., d_k$ are integer, we write

$$\Delta_{d_1,d_2,\ldots,d_k} = \left\{ A \in \mathbb{Z}^{n \times m} \, : \, \text{diag}\left(\text{HNF}\left(A\right)\right) = (d_1, ..., d_k, *, ..., *) \right\}$$

whenever $k \leq \min(n, m)$.

Proposition 6 *Let n, m, k be positive integers and let $d_1, d_2, ..., d_k \in \mathbb{N}$.*

1. *Suppose n, m, k satisfy $k \leq m$ if $m < n$ and $k < n$ otherwise. If $d_k = 0$,*

$$\mathbb{D}\left(\Delta_{d_1,d_2,\ldots,d_k}\right) = 0.$$

If $d_i \neq 0, \forall i = 1, ..., k$, then

$$\mathbb{D}\left(\Delta_{d_1,d_2,\ldots,d_k}\right) = \left(\zeta(n) \cdot \zeta(n-1) ... \cdot \zeta(n-k+1) \cdot d_1^n \cdot d_2^{n-1} \cdot \ldots \cdot d_k^{n-k+1}\right)^{-1}.$$

2. *Suppose $n \leq m$ and let $0 \leq r < d \in \mathbb{N}$. Then*

$$\mathbb{D}\left(A \in \Delta_{d_1,d_2,\ldots,d_{n-1},a} \, : \, a \equiv r \mod d\right) = \frac{1}{d} \cdot \mathbb{D}\left(\Delta_{d_1,d_2,\ldots,d_{n-1}}\right).$$

The first point of the previous proposition clearly shows that the example of HNF given in the introduction (with a 0 in the diagonal) will only rarely appear. The powers d_i^{n-i} appearing in the expression of the density explain the decreasing expectation to see a randomly looking $n \times m$ matrix having elements on the top of the diagonal of its HNF larger than 1. Let us now prove the proposition.

Proof: First, if $d_k = 0$ for some k, then the k^{th} column vector v of $U_{k-1} A$, see Equation (2) above, is $v = [*, \ldots, *, 0, \ldots, 0]^t$. This means that there are at most $O(B^{k-1})$ choices for v in any cube of volume $(2B)^n$. Since $n > k$, this implies that $\mathbb{D}\left(\Delta_{d_1,d_2,\ldots,0}\right) = 0$. Let us now focus on the case where

$d_i \neq 0, \forall i = 1, ..., k$. We prove by induction on k that the expression of the density is valid. For $k = 1$, the claim is true as a consequence that h_{11} is the greatest common divisor of the entries of the first column vector, and this case is Proposition 5 above, when U is the identity. The induction step is as follows. For a bound B, the number of $n \times (k-1)$ matrices A' with entries in a cube of side length $2B$ such that $\operatorname{diag}(\operatorname{HNF}(A')) = (d_1, ..., d_{k-1})$ is

$$(2B)^{n(k-1)} \cdot \left(\zeta(n) \cdot \zeta(n-1) ... \cdot \zeta(n-k+2) \cdot d_1^n \cdot d_2^{n-1} \cdot .... \cdot d_{k-1}^{n-k+2}\right)^{-1} + o(B^{n(k-1)}).$$

For each of these matrices, there exists an $n \times n$ unimodular matrix U such that UA' is upper diagonal and $\operatorname{diag}(UA') = (d_1, ..., d_{k-1})$. For any cube σ of side length $2B$ and dimension n, using Proposition 5, we see that for each A', there are $(2B)^n \cdot (\zeta(n-k+1) \cdot d^{n-k+1})^{-1} + o(B^n)$ vectors v in σ such that

$$U[A'|v] = \begin{bmatrix} d_1 & \cdots & * & (Uv)_1 \\ 0 & \ddots & * & \vdots \\ & 0 & d_{k-1} & (Uv)_{k-1} \\ & & 0 & (Uv)_k \\ & & \vdots & \vdots \\ & & 0 & (Uv)_n \end{bmatrix}$$

with $\gcd((Uv)_i : i = k, \ldots, n)) = d_k$. Based on the algorithmic description of the HNF given earlier, the diagonal of the HNF of $[A'|v]$ is $(d_1, ..., d_{k-1}, d_k)$. The number of such matrices is then, up to an error of order $o(B^{n(k-1)+n})$,

$$(2B)^{n(k-1)+n} \cdot \left(\zeta(n) \cdot ... \cdot \zeta(n-k+2) \cdot d_1^n \cdot .... \cdot d_{k-1}^{n-k+2}\right)^{-1} \cdot \left(\zeta(n-k+1) \cdot d^{n-k+1}\right)^{-1}.$$

Since $n(k-1)+n = kn$, the claim is correct. This argument can be continued as long as Proposition 5 can be applied, i.e. until $k \leq m$ if $m < n$ or $k < n$ otherwise. This finishes the proof of the first statement of the proposition.

Let us concentrate now on the second statement. There exists a $n \times n$ unimodular matrix U such that the first $(n-1)$ column of UA are in HNF. Clearly, if $a = (UA)_{n,n}$ then $a = \operatorname{HNF}(A)_{n,n}$, i.e., $a = u \cdot \alpha$, where u is the last row of U and α is the last column of A. For any cube σ of side length $2B$ and dimension n, we want to find the number of n-vectors α in σ such that $u \cdot \alpha \equiv a \mod d$. Since the entries of u are coprime, at least one is coprime to d, say u_i. For each $(2B)^{n-1}$ choices of α_j in σ, $j \neq i$, there are $\frac{2B}{d} + o(B)$ $\alpha_i \in \sigma$ such that $u_i \alpha_i \equiv a - \sum_{j \neq i} u_j \alpha_j \mod d$. In other words, the density of the α's is $\frac{1}{d}$. The result follows by applying the same counting argument as before and by using the previous expression of the density of $\Delta_{d_1, d_2, ..., d_{n-1}}$. This finishes the proof of the proposition. □

Corollary 7 *Let $d_1, d_2, ..., d_{n-1} \in \mathbb{N}^*$. The natural density of $n \times n$ integer matrices whose HNF has diagonal $\left(d_1, d_2, \ldots, d_{n-1}, \frac{\det A}{\prod_{i=1..n-1} d_i}\right)$ is*

$$\left(\zeta(n) \cdot \zeta(n-1) ... \cdot \zeta(2) \cdot d_1^n \cdot d_2^{n-1} \cdot .... \cdot d_{n-1}^2\right)^{-1}.$$

The natural density of unimodular rectangular $n \times m$ integer matrices, with $n > m$, has been computed in [14], with the weak definition of natural density presented in Section 2. Proposition 6 allows to extend the result to the stronger natural density defined in this article. With the material in hand, the proof is straightforward.

Corollary 8 *The set of $n \times m$ unimodular integer matrices, with $n > m$, has a natural density equal to $(\zeta(n) \cdot \zeta(n-1) ... \cdot \zeta(n-m+1))^{-1}$.*

5 Applications

5.1 Selection of Random Lattices in Cryptology

In the following, we discuss the consequences of Proposition 6 above to the various shapes of lattice bases that arise in lattice based cryptology.

An integer lattice $\mathcal{L}$ is a discrete $\mathbb{Z}$-module of dimension n in $\mathbb{R}^m$ with $\mathcal{L} = \mathbb{Z}b_1 + \ldots + \mathbb{Z}b_n$, where $b_i \in \mathbb{Z}^m$ and $\mathrm{Vol}(\mathcal{L}) = \det([b_i \cdot b_j]_{i,j})^{1/2} \neq 0$. A matrix B whose row vectors b_i are independent and generate $\mathcal{L}$ is called a basis of the lattice. Any matrix $B' = UB$ with U unimodular is a basis of $\mathcal{L}$. We refer the reader to, e.g., [16, Chapter 3] and [20] for the use of lattices in cryptology. Several types of lattice bases naturally appear in lattice based cryptology. Among them, we find the knapsack $n \times (n+1)$ bases (a), the NTRU $2n \times 2n$ bases (b) and the so-called random $n \times n$ lattice basis (c).

$$\underset{(a)}{\begin{bmatrix} I_n & x \end{bmatrix}} \quad \underset{(b)}{\begin{bmatrix} I_n & H_n \\ 0_n & qI_n \end{bmatrix}} \quad \underset{(c)}{\begin{bmatrix} I_{n-1} & x \\ 0 & q \end{bmatrix}}$$

A direct consequence of the previous proposition is that the density of integer matrices A with HNF of the form (a) is 0. The density of integer matrices A with HNF of the form (c) is given by $(\zeta(n) \cdot \ldots \cdot \zeta(2))^{-1}$. Since $\zeta(n)$ converges rapidly towards 1, the above density converges rather fast to the limit d with $\mathrm{d} = \left(\prod_{j=2}^{\infty} \zeta(j)\right)^{-1} = 0.43575707677....$ This translates into the facts that the random lattice bases of type (c) have a positive density in the set of lattices with corresponding dimension. The strict positivity of this density has been know since the work of Goldstein and Mayer [7] (see also [1] for an elementary proof). This density being equal to d, this shows that the process of selecting random lattice by selecting random row matrices of type (c) and large determinant q covers almost 44% of all possible cases of randomly looking matrices.

In the case of NTRU bases (b), $q = 2^s$, where s is a small integer and Proposition 6 suggests that the density of such lattice bases is roughly equal to $\mathrm{d} \cdot 2^{-N}$, with $N = \frac{n^2}{2}s$. Here again, the density is strictly positive, but much smaller than in the random case (c).

5.2 Distribution of $\gcd(\det([A|x], \det[A|y]))$

Using the weak notion of density presented in Section 2, Hafner, Sarnack and McCurley have computed the probability that two randomly looking $n \times n$ matrices are coprime [8]. The situation where the randomly looking matrices differs in one column only turns out to be interesting as well.

Let A be a randomly looking $n \times (n-1)$ integer matrix and x, y be two randomly looking n-vectors. The distribution of the greatest common divisor $g = \gcd(\det([A|x], \det[A|y]))$ has been used in order to predict the behavior of fast algorithms that compute the HNF of an integer matrix, see [15, 18]. Miccaccio and Warinschi [15] notice that g is "typically very small for randomly chosen matrices", and Pernet and Stein [18], based on numerical simulation, provide an histogram of the distribution of the g's. We propose here to exactly compute this distribution based on the natural density distribution of Proposition 6. Suppose $\mathrm{diag}(\mathrm{HNF}(A)) = (d_1, \ldots, d_{n-1})$, with $UA = \mathrm{HNF}(A)$, U unimodular. Then

$$\begin{aligned} g &= \gcd(\det([A|x], \det[A|y])) = \gcd(\det([UA|Ux], \det[UA|Uy])) \\ &= \prod_{i=1}^{n-1} d_i \cdot \gcd(u \cdot x, u \cdot y) \end{aligned}$$

where u is the last row of U and $u \cdot x$ (resp. $u \cdot y$) is the scalar product of u and x (resp. y). Note that since U is unimodular, we have $\gcd(u_i) = 1$. The natural distribution of $\gcd(u \cdot x, u \cdot y)$ in such a case can be computed as follows. The reader will readily check that for a given modulus d, the distribution of $(u \cdot x \mod d, u \cdot y \mod d)$ is uniform in $(\mathbb{Z}/d\mathbb{Z})^2$. This means that the proportion of pairs $(u \cdot x, u \cdot y)$ that are divisible by d is d^{-2}, and the proportion of pairs $(u \cdot x, u \cdot y)$ that are

not $(0,0)$ modulo a prime p is $1-p^{-2}$. This suggests that the density of n-vectors x, y such that $d = \gcd(u \cdot x, u \cdot y)$ is given by

$$\frac{1}{d^2} \prod_{p \in \mathbb{P}} \left(1 - p^{-2}\right) = \left(d^2 \zeta(2)\right)^{-1}.$$

This development can be made rigorous by using the methods used in Section 2 and the localization methods presented in [14]. Finally, the natural density $D_n(g)$ of $n \times (n-1)$ integer matrices A and n-vectors x, y such that $g = \gcd(\det([A|x], \det[A|y]))$ is given by

$$D_n(g) = \frac{1}{\zeta(2) \cdot \prod_{k=2}^{n} \zeta(k)} \sum_{d_1 \cdot \ldots \cdot d_n = g} \frac{1}{d_1^n \cdot d_2^{n-1} \cdot \ldots \cdot d_{n-2}^3 \cdot d_{n-1}^2 \cdot d_n^2}. \tag{3}$$

If $\sigma_{-k}(g) = \sum_{d|g} d^{-k}$, then using the Dirichlet's convolution product $*$ of arithmetic functions, we obtain

$$D_n = \frac{1}{\zeta(2) \cdot \prod_{k=2}^{n} \zeta(k)} \cdot \left(\sigma_{-n} * \sigma_{-n+1} * \ldots * \sigma_{-3} * \sigma_{-2} * \sigma_{-2}\right).$$

Since the Dirichlet series associated to σ_{-k} is $\zeta(s+k) \cdot \zeta(s)$ (see e.g. [22]), i.e., $\sum_{g \geq 1} \frac{\sigma_{-k}(g)}{g^s} = \zeta(s) \cdot \zeta(s+k)$, the Dirichlet series associated to D_n is given by

$$\sum_{g \geq 1} \frac{D_n(g)}{g^s} = \left(\zeta(s)\right)^n \cdot \frac{\zeta(s+2)}{\zeta(2)} \cdot \prod_{k=2}^{n} \frac{\zeta(s+k)}{\zeta(k)}\ , \quad \Re(s) > 1.$$

If we write $f_n = \zeta(2) \cdot \prod_{k=2}^{n} \zeta(k) \cdot D_n$, Equation (3) above shows that the arithmetic function f_n is multiplicative, i.e., $f_n(gh) = f_n(g) \cdot f_n(h)$ when $\gcd(g,h) = 1$. It is therefore sufficient to compute $f_n(p^\alpha)$ for $p \in \mathbb{P}$, $\alpha \geq 1$ in order to determine D_n explicitly. Equation (3) with $d_1 = p^i$ gives

$$f_n(p^\alpha) = \sum_{i=0}^{\alpha} \frac{1}{p^{ni}} f_{n-1}(p^{\alpha - i}),$$

together with $f_1(p^\alpha) = \frac{1}{p^{2\alpha}}$. Based on this recurrence relation, we can compute f_n for the first value of n, e.g., $f_2(p^\alpha) = \frac{\alpha+1}{p^{2\alpha}}$ and prove that f_n converges rapidly to a limit function f that satisfies

$$f(1) = 1\ , \quad f(p) = \frac{2p-1}{p^2(p-1)}\ , \quad f(p^2) = \frac{3p^3 - p^2 - 2p + 1}{p^4(p-1)^2(p+1)}$$

and in general for all $\alpha \geq 3$,

$$f(p^\alpha) = \frac{\alpha+1}{p^{2\alpha}} + \frac{\alpha}{p^{2\alpha+1}} + \frac{2\alpha-1}{p^{2\alpha+2}} + \frac{3\alpha-3}{p^{2\alpha+3}} + \frac{5\alpha-7}{p^{2\alpha+4}} + o\left(\frac{1}{p^{2\alpha+5}}\right).$$

The first values of

$$D(g) = \lim_{n \to \infty} D_n(g) = \lim_{n \to \infty} \left(\zeta(2) \cdot \prod_{k=2}^{n} \zeta(k)\right)^{-1} f_n(g) = \frac{\mathrm{d}}{\zeta(2)} \cdot f(g),$$

where d is the constant defined in Section 5.1, are given via

$$f(2) = \frac{3}{4}\,,\ f(3) = \frac{5}{18}\,,\ f(4) = \frac{17}{48}\,,\ f(5) = \frac{9}{100}\,,\ f(6) = \frac{5}{24}\,,\ f(7) = \frac{13}{276}.$$

Numerical simulation showed that already for small dimension n, say $n > 5$, the above values of D give very good approximations of the density D_n. We end up this section by noticing that even though the above remark of Miccancio is true, the expected size of $\gcd(\det([A|x], \det[A|y]))$ is unbounded. Indeed, the real numbers $D(g)$, $g \in \mathbb{N}^*$, define a probability distribution on $\mathbb{N}^*$, i.e., $\sum_{g \in \mathbb{N}^*} D(g) = 1$ and since $D(p) > C/p^2$ for some $C > 0$ and $\sum_{p \in \mathbb{P}} 1/p = \infty$, the expectation of the positive integers under this distribution law is $\sum_{g \in \mathbb{N}^*} gD(g) > \sum_{p \in \mathbb{P}} C/p = \infty$. Notice that $D(1) = \frac{\mathrm{d}}{\zeta(2)} = 0.266014...$ which is not far from 30%, as noted in [15].

6 Conclusion

Numerical experiments indicate that for randomly looking integer matrices, their Hermite normal forms are not uniformly distributed among the upper diagonal matrices. The frequency of apparition of the different diagonals is highly structured. In this paper, we explain this phenomenon, and we exactly compute these frequencies in terms of natural density. On the way, we define a multidimensional extension of the usual natural density over $\mathbb{N}$. We use this analysis in order to shed light on the following two different situations where the expected form of the HNF of randomly looking matrices play a role. First, the densities of three types of lattice bases that naturally appear in lattice based cryptology has been computed. Second, a probability distribution over the positive integer appearing in some HNF algorithms has been explicitly evaluated.

References

[1] Buchmann, J.A., Lindner, R. Density of ideal lattices. *Algorithm and Number Theory*, Dagstuhl Seminar Proceedings 09221, 2009.

[2] Cesàro, E. Question proposée 75. *Mathesis*, 1 (1881), p. 184.

[3] Cesàro, E. Question 75 (Solution). *Mathesis*, 3 (1883), pp. 224–225.

[4] Cesàro, E. Probabilité de certains faits arithméthiques. *Mathesis* 4 (1884), pp. 150–151.

[5] Cohen, H. *A course in computational algebraic number theory*, Springer Graduate Texts In Mathematics, 1995.

[6] Evans, S.N. Elementary divisors and determinant of random matrices over a local field. *Stochastic Processes and its Applications*, Vol. 102 (1) 2002, pp. 89–102.

[7] Goldstein, A. and Mayer, A. On the equidistribution of hecke points *Forum Mathematicum 2003*, 15:2, (2003), pp. 165–189.

[8] Hafner, J. L., Sarnak, P. and McCurley, K. Relatively Prime Values of Polynomials, in Knopp, M. and Seingorn, M., *A Tribute to Emil Grosswald: Number Theory and Related Analysis*, 1993, Providence, RI, Amer. Math. Soc.

[9] Hardy, G. H., Wright, E. M. *An Introduction to the Theory of Numbers*, 5d ed. Oxford University Press, 1979.

[10] Hetzel, A. J., Liew, J. S. and Morrison, K. E. The probability that a matrix of integers is diagonalizable. *American Mathematical Monthly*, Volume 114, Number 6, June-July 2007, pp. 491–499.

[11] Lehmer, D.N. Asymptotic evaluation of certain totient sums. *American Journal of Mathematics*, Vol. 22, No. 4 (Oct., 1900), pp. 293–335.

[12] MacDuffie, C. C. *The Theory of Matrices*, Chelsea Publ. Co., New York, 1946.

[13] Martin, G. and Wong, E. The number of 2×2 integer matrices having a prescribed integer eigenvalue, *Algebra & Number Theory* 2 (2008), no. 8, pp. 979–1000.

[14] Maze, G., Rosenthal, J., and Wagner, U. Natural Density of Rectangular Unimodular Integer Matrices. Submitted, preprint available at http://arxiv.org/abs/1005.3967v1

[15] Micciancio, D. and Warinschi, B. A linear space algorithm for computing the Hermite normal form. *International Symposium on Symbolic and Algebraic Computation - ISSAC 2001.* London, Canada, pp. 231–236 (July 2001)

[16] Nguyen, P. and Vallée, B., editors. *The LLL algorithm : survey and applications,* 496 pp., Information security and cryptography text and monographs, Springer, 2010.

[17] Nymann, J. E. On the probability that k positive integers are relatively prime. *J. Number Th.*, 7 (1975), pp. 406–412.

[18] Pernet, C. and Stein, W. Fast computation of Hermite normal forms of random integer matrices. *J. Number Th.*, 130 (2010), pp. 1675–1683

[19] Sarnak, P., Duke, W. and Rudnick, Z. Density of Integral Points on Affine Homogeneous Varieties, *Duke Math. Jnl.*, 71 (1993), pp. 143–179.

[20] Silverman, J. (editor) *Cryptography and lattices*. Proceedings of the 1st international conference (CaLC 2001) held in Providence, RI, March 29-30, 2001, Lecture Notes in Computer Science 2146, Springer.

[21] Sylvester, J.J. On certain inequalities relating to prime numbers. *Nature*, No. 38, 259-262, 12 July 1888.

[22] Tenenbaum, G. *Introduction to Analytic and Probabilistic Number Theory.* Cambridge studies in advanced mathematics, 46 (1995).