Chinese Remainder Theorem for Cyclotomic Polynomials in $\mathbf{Z}[X]$

Kamalakshya Mahatab*

The Institute of Mathematical Sciences, Chennai

Kannappan Sampath*

Statistics and Mathematics Unit, Indian Statistical Institute, Bangalore-560059.

Abstract

By the Chinese remainder theorem, the canonical map

$$\Psi_n : R[X]/(X^n - 1) \to \oplus_{d|n} R[X]/\Phi_d(X)$$

is an isomorphism when R is a field whose characteristic does not divide n and Φ_d is the dth cyclotomic polynomial. When R is the ring $\mathbf{Z}$ of rational integers, this map is injective but not surjective. In this paper, we give an explicit formula for the elementary divisors of the cokernel of Ψ_n (when $R = \mathbf{Z}$) using the prime factorisation of n. We also give a pictorial algorithm using Young tableaux that takes $O(n^{3+\epsilon})$ bit operations for any $\epsilon > 0$ to determine a basis of Smith vectors (see Definition 3.1) for Ψ_n. In general when R is a unique factorisation domain, we prove that the determinant of the matrix of $\Psi : R[X]/(\prod_j f_j) \to \bigoplus_j R[X]/(f_j)$ written with respect to the standard basis is $\prod_{1 \leqslant i < j \leqslant n} \mathcal{R}(f_j, f_i)$, where f_i's are monic polynomials and $\mathcal{R}(f_j, f_i)$ is the resultant of f_j and f_i.

1. Introduction

Motivation

Let $m_1, \ldots, m_r$ be pairwise coprime elements in a principal ideal domain (PID) R, that is, for $i \neq j$, if $a \mid m_i$ and $a \mid m_j$, then, a is a unit in R. The Chinese remainder theorem (CRT) states that, for $a_1, \ldots, a_r \in R$, the system

*Corresponding Author

Email addresses: kamalakshya@imsc.res.in (Kamalakshya Mahatab), knsam.name@gmail.com (Kannappan Sampath)

(in X)

$$\begin{aligned} X &\equiv a_1 \bmod m_1 \\ X &\equiv a_2 \bmod m_2 \\ &\vdots \\ X &\equiv a_r \bmod m_r \end{aligned} \tag{1.1}$$

has a solution and any two solutions are congruent modulo $\prod_i m_i$. In terms of ideals, the natural map from the ring $R/(\prod_i m_i)$ to the ring $\prod_i R/(m_i)$ is an isomorphism. The surjectivity of the natural map encapsulates the fact that the system (1.1) has a solution and the injectivity encapsulates the fact that any two solutions are congruent modulo $\prod_i m_i$.

However, such a theorem does not hold true over rings which are not PID's. For example, consider the system (in $h(X)$ over $\mathbf{Z}[X]$):

$$\begin{aligned} h(X) &\equiv 1 \bmod (X-1) \\ h(X) &\equiv 0 \bmod (X+1). \end{aligned}$$

This system does not have a solution over $\mathbf{Z}[X]$: to wit, if $f_1(X), f_2(X) \in \mathbf{Z}[X]$ are such that

$$\begin{aligned} h(X) &= f_1(X)(X-1)+1 \\ h(X) &= f_2(X)(X+1), \end{aligned}$$

then, we are led to the absurdity $2f_2(1) = 1$. This phenemenon serves as a motivation for the questions we study in this article.

Setup

Let R be an integral domain which is not a field, so that $R[X]$ is not a PID. Suppose that f is a monic polynomial and

$$f = \prod_{i=1}^{n} f_i$$

where $\{f_i\}_{i=1}^n$ are pairwise coprime polynomials in $R[X]$. Consider the natural map:

$$\begin{aligned} \Psi_f : R[X]/(f) &\to \bigoplus_i R[X]/(f_i) \\ h(X) \bmod f &\mapsto \bigoplus_i h(X) \bmod f_i. \end{aligned}$$

The map becomes injective if R is replaced by its field of fractions; therefore, Ψ_f is injective. However, as we have already remarked in general, Ψ_f is not

surjective. As a measure of the failure of surjectivity, we would like to determine the cokernel $G(f)$ of the map Ψ_f:

$$0 \longrightarrow R[X]/(f) \xrightarrow{\Psi_f} \bigoplus_i R[X]/(f_i) \xrightarrow{\overline{\Psi}_f} G(f) \longrightarrow 0.$$

We would also like to understand when a given element $\alpha \in \bigoplus_i R[X]/(f_i)$ lies in the image of Ψ_f. To the best of our knowledge, it seems to us that problems of this nature have not been explicitly studied elsewhere in the literature.

We shall solve the above problems when $R = \mathbf{Z}$ and $f(X) = X^n - 1$ with its factorisation $\prod_{d|n} \Phi_d(X)$ into cyclotomic polynomials.

Results

Let us consider the map Ψ_n defined by:

$$\begin{aligned} \Psi_n : \mathbf{Z}[X]/\langle X^n - 1\rangle &\to \bigoplus_{d|n} \mathbf{Z}[X]/\langle \Phi_d(X)\rangle \\ f(X) \bmod (X^n - 1) &\mapsto \bigoplus_{d|n} f(X) \bmod \Phi_d(X). \end{aligned}$$

The associated exact sequence is:

$$0 \longrightarrow \mathbf{Z}[X]/(X^n - 1) \xrightarrow{\Psi_n} \bigoplus_{d|n} \mathbf{Z}[X]/(\Phi_d(X)) \xrightarrow{\overline{\Psi}_n} G(n) \longrightarrow 0.$$

The domain and codomain of Ψ_n are free $\mathbf{Z}$-modules of the same rank and therefore, the cokernel $G(n)$ is a finite abelian group. We endow $\mathbf{Z}[X]/(X^n - 1)$ with the basis $(1, \overline{X}, \ldots, \overline{X}^{n-1})$ and $\mathbf{Z}[X]/(\Phi_d(X))$ with the basis $(1, \overline{X}, \ldots, \overline{X}^{\phi(d)-1})$. Denote the matrix of Ψ_n with respect to this basis by A_n. For example, we have

$$A_2 = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}, \quad A_3 = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & -1 \\ 0 & 1 & -1 \end{pmatrix}, \quad A_4 = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & -1 \end{pmatrix}.$$

The structure of the abelian group $G(n)$ is completely determined by the elementary divisors of A_n (see for instance, [10, Theorem 7.7]). For example, the elementary divisors of A_6 are $\{1, 1, 1, 2, 6, 6\}$ and the group $G(6)$ is isomorphic to $\mathbf{Z}/2\mathbf{Z} \oplus \mathbf{Z}/6\mathbf{Z} \oplus \mathbf{Z}/6\mathbf{Z}$. We first reduce the problem of determining the elementary divisors of A_n to that of A_{p^e} for a prime p (Theorem 2.4). For a prime p, the matrix A_{p^e} has the following structure (Lemma 2.5):

$$A_{p^e} = \begin{pmatrix} A_{p^{e-1}} & A_{p^{e-1}} & \ldots & A_{p^{e-1}} & A_{p^{e-1}} \\ I_{p^{e-1}} & 0 & \ldots & 0 & -I_{p^{e-1}} \\ 0 & I_{p^{e-1}} & \ldots & 0 & -I_{p^{e-1}} \\ \vdots & & \ddots & & \vdots \\ 0 & 0 & \ldots & I_{p^{e-1}} & -I_{p^{e-1}} \end{pmatrix} \tag{1.2}$$

with $A_{p^0} = A_1 = (1)$; we exploit this recursive structure in determining the elementary divisors of the matrix A_{p^e}.

From this approach, we deduce that (Theorem 2.18), if $(e_1, \ldots, e_n)$ is the tuple of elementary divisors of A_n with $e_i \mid e_{i+1}$, then the tuple $Q_n = \left(e_1, \frac{e_2}{e_1}, \ldots, \frac{e_n}{e_{n-1}}\right)$ of quotients is a rearrangement of the tuple

$$(\underbrace{p_1, \ldots, p_1}_{\alpha_1 \text{ times}}, \underbrace{p_2, \ldots, p_2}_{\alpha_2 \text{ times}}, \ldots, \underbrace{p_r, \ldots, p_r}_{\alpha_r \text{times}}, \underbrace{1, \ldots, 1}_{n - \sum_i \alpha_i \text{ times}})$$

where $n = p_1^{\alpha_1} \ldots p_r^{\alpha_r}$ and p_i's are distinct primes. Moreover, the sth p_i appears at the index $n - \frac{n}{p_i^s} + 1$ in Q_n. Since $|G(n)| = |\det(A_n)|$, we have that (Corollary 2.9):

$$|G(n)| = \prod_{i=1}^{r} p_i^{\frac{n(1-p_i^{-\alpha_i})}{(p_i-1)}} = \prod_{k=1}^{n} \gcd(k, n).$$

In Appendix A, we prove:

$$\det(A_n) = \prod_{\substack{d_1, d_2 \mid n \\ 1 \leqslant d_1 < d_2 \leqslant n}} \mathcal{R}(\Phi_{d_2}, \Phi_{d_1}) = (-1)^{n-1} \prod_{i=1}^{r} p_i^{\frac{n(1-p_i^{-\alpha_i})}{(p_i-1)}}, \tag{1.3}$$

where $\mathcal{R}(g_1, g_2)$ is the resultant of the polynomials g_1 and g_2. More generally, if f is a monic polynomial over a unique factorisation domain and if $f = \prod_{k=1}^{n} f_k$ is a factorisation of f into pairwise relatively prime polynomials, then (Theorem A.3),

$$\det(\Psi_f) = \prod_{1 \leqslant i < j \leqslant n} \mathcal{R}(f_j, f_i). \tag{1.4}$$

We notice that the group algebra $\mathbf{Z}[G]$ over $\mathbf{Z}$ of a group G isomorphic to the cyclic group $\mathbf{Z}/n\mathbf{Z}$ is $\mathbf{Z}[X]/\langle X^n - 1 \rangle$. From this perspective, the absolute value of the determinant of Ψ_n is the index of the group algebra $\mathbf{Z}[G]$ in $\bigoplus_{j=0}^{n} \mathbf{Z}[X]/\Phi_{p^j}(X)$. Raymond Ayoub and Christine Ayoub determine this index [2, Theorem 7(C)]. They also determine a basis for this quotient $\mathbf{Z}$-module, which is then used to determine a basis of Smith vectors (Definition 3.1) for the group algebra $\mathbf{Z}[G]$, in the case $n = p^e$ for a prime p.

In this paper, we carry out the program of determining a basis of Smith vectors for a general n (Section 4) by a pictorial algorithm involving Young diagrams. A basis of Smith vectors for a general n can be realised as the columns of the matrix U_n^{-1} for some $U_n \in \mathrm{GL}_n(\mathbf{Z})$ for which there exists a $V_n \in \mathrm{GL}_n(\mathbf{Z})$ such that $U_n A_n V_n$ is the Smith normal form of A_n (Lemma 3.2). The best known algorithm [16, Proposition 7.20] for computing the Smith normal form of A_n and the unimodular transformations takes $O(n^{2+\theta+\epsilon})$ bit operations for any $\epsilon > 0$ where $O(n^\theta)$ is the bit complexity in multiplying two $n \times n$ matrices over a ring R. In [18], it is proven that $2 \leqslant \theta \leqslant 2.373$. Our algorithm determines a basis of Smith vectors for a general n in $O(n^{3+\epsilon})$ bit operations for any $\epsilon > 0$ without actually computing these transformation matrices (Theorem 4.10). The output of the algorithm requires $O(n^{3+\epsilon})$ bits space (Lemma 4.9).

Framework

In Section 2, we compute the elementary divisors of A_n. In Section 3, we prove some basic facts required in the algorithm for determining a basis of Smith vectors for n which is followed by a presentation of the algorithm in Section 4. In the appendix that follows, we compute the determinant of A_n in a way that generalises to any factorisation of a monic polynomial over a unique factorisation domain.

2. Smith Normal Form of A_n

To establish a relationship between the Smith normal form of A_{mn} and those of A_m and A_n for relatively prime positive integers m and n, we begin with the following observation:

LEMMA 2.1. *Given relatively prime positive integers n and m, the ring homomorphism $P_{m,n} : \mathbf{Z}[X]/(X^m - 1) \otimes \mathbf{Z}[Y]/(Y^n - 1) \to \mathbf{Z}[t]/(t^{mn} - 1)$ defined by:*

$$P_{m,n}(\overline{X} \otimes 1) = \overline{t}^n \text{ and } P_{m,n}(1 \otimes \overline{Y}) = \overline{t}^m$$

is an isomorphism. Furthermore, with respect to the standard basis, the matrix of $P_{m,n}$ as a $\mathbf{Z}$-module homomorphism is a permutation matrix.

Proof. We note that $t^{ni+mj} \equiv t^\alpha \bmod (t^{mn} - 1)$ if and only if $ni + mj \equiv \alpha \bmod mn$. Now, by the Chinese Remainder Theorem for $\mathbf{Z}$, the set

$$\{ni + mj : 0 \leqslant i \leqslant m - 1, 0 \leqslant j \leqslant n - 1\}$$

consists all the residues mod mn, exactly once. Thus, $P_{m,n}$ is a bijection between the standard bases. This proves the lemma. □

LEMMA 2.2. *Suppose that m and n are relatively prime positive integers. Then, the map $T_{m,n} : \mathbf{Z}[X]/\Phi_m(X) \otimes \mathbf{Z}[Y]/\Phi_n(Y) \to \mathbf{Z}[t]/\Phi_{mn}(t)$ defined by*

$$\overline{X}^i \otimes \overline{Y}^j \mapsto \overline{t}^{ni+mj}$$

and extending $\mathbf{Z}$-linearly is a ring isomorphism.

Proof. Consider the maps:

$$\phi : \mathbf{Z}[X]/(\Phi_m(X)) \to \mathbf{Z}[t]/(\Phi_{mn}(t)) \quad \text{and} \quad \psi : \mathbf{Z}[Y]/(\Phi_n(Y)) \to \mathbf{Z}[t]/(\Phi_{mn}(t))$$
$$\overline{X} \mapsto \overline{t}^n \qquad\qquad \overline{Y} \mapsto \overline{t}^m$$

Now, $T_{m,n}$ is composition of the canonical map $\phi \otimes \psi$ with the identification map $\overline{f} \otimes \overline{g} \mapsto \overline{fg} : \mathbf{Z}[t]/(\Phi_{mn}(t)) \otimes \mathbf{Z}[t]/(\Phi_{mn}(t)) \to \mathbf{Z}[t]/(\Phi_{mn}(t))$. Thus, $T_{m,n}$ is a ring homomorphism.

To prove surjectivity, we show that $\overline{t} \in \mathbf{Z}[t]/\Phi_{mn}(t)$. Indeed, since $\overline{t}$ is invertible in $\mathbf{Z}[t]/\Phi_{mn}(t)$ and that $\gcd(m, n) = 1$, there are integers $i, j \in \mathbf{Z}$ such that $t^{ni+mj} \equiv t \bmod \Phi_{mn}(t)$.

We claim this map is also injective: letting K be the kernel of the map $T_{m,n}$, the exact sequence:

$$0 \longrightarrow K \longrightarrow \frac{\mathbf{Z}[X]}{\Phi_m(X)} \otimes \frac{\mathbf{Z}[Y]}{\Phi_n(Y)} \xrightarrow{T_{m,n}} \frac{\mathbf{Z}[t]}{\Phi_{mn}(t)} \longrightarrow 0$$

splits since $\mathbf{Z}[t]/\Phi_{mn}(t)$ is a free $\mathbf{Z}$-module showing:

$$\frac{\mathbf{Z}[X]}{\Phi_m(X)} \otimes \frac{\mathbf{Z}[Y]}{\Phi_n(Y)} \simeq K \oplus \frac{\mathbf{Z}[t]}{\Phi_{mn}(t)}. \tag{2.1}$$

Being a submodule of a free module over the PID $\mathbf{Z}$, K is a free $\mathbf{Z}$-module. A comparison of the rank tells us that K is of rank 0. Thus, $K = \{0\}$, equivalently, $T_{m,n}$ is injective. □

REMARK 2.3. Along the lines of the proof of Lemma 2.2, it may be shown that for relatively prime positive integers m and n, the $\mathbf{Z}$-linear extension of the map

$$\overline{X}^i \otimes \overline{Y}^j \mapsto \overline{t}^{mj+ni} : \frac{\mathbf{Z}[X]}{\Phi_m(X)} \otimes \frac{\mathbf{Z}[Y]}{Y^n - 1} \to \frac{\mathbf{Z}[t]}{\Phi_m(t^n)}, \quad \begin{smallmatrix} 0 \leqslant i \leqslant \phi(m)-1 \\ 0 \leqslant j \leqslant n-1 \end{smallmatrix} \tag{2.2}$$

is a ring isomorphism.

2.1. Smith Equivalence of $A_m \otimes A_n$ and A_{mn}

For a matrix A over the integers, let $S(A)$ denote the Smith normal form of A in which all the elementary divisors are non-negative[1]. We now state and prove one of the main results of this section:

THEOREM 2.4. $S(A_m \otimes A_n) = S(A_{mn})$.

Proof. Consider the following diagram:

$$\begin{array}{ccc} \mathbf{Z}[X]/(X^m-1) \otimes \mathbf{Z}[Y]/(Y^n-1) & \xrightarrow{\Psi_m \otimes \Psi_n} & \bigoplus_{\substack{d_1|m \\ d_2|n}} \mathbf{Z}[X]/\Phi_{d_1}(X) \otimes \mathbf{Z}[Y]/\Phi_{d_2}(Y) \\ \Big\downarrow{\scriptstyle P_{m,n}} & & \Big\downarrow{\scriptstyle T(m,n)} \\ \mathbf{Z}[t]/(t^{mn}-1) & \xrightarrow{\Psi_{mn}} & \bigoplus_{d|mn} \mathbf{Z}[t]/(\Phi_d(t)) \end{array}$$

The map $\Psi_m \otimes \Psi_n$ is the canonical map, defined by:

$$(\Psi_m \otimes \Psi_n)(X^i \otimes Y^j) = \Psi_m(X^i) \otimes \Psi_n(Y^j)$$

and extended $\mathbf{Z}$-linearly. We shall prove that there is an isomorphism $T(m,n)$ that renders the diagram commutative.

[1]For later purposes, we note that $\det(S(A)) = |\det(A)|$ is non-negative.

Indeed, define $T(m,n)$ by

$$T(m,n) = \bigoplus_{\substack{d_1|m \\ d_2|n}} T_{d_1,d_2}$$

Clearly, $T(m,n)$ is an isomorphism. As is seen by the following computation, $T(m,n)$ also renders the above diagram commutative:

$$\begin{aligned}
(\Psi_{mn} \circ P_{m,n})(\overline{X}^i \otimes \overline{Y}^j) &= \Psi_{mn}(\bar{t}^{ni+mj}) \\
&= \oplus_{d|mn} t^{ni+mj} \bmod \Phi_d(t) \\
(T(m,n) \circ (\Psi_m \otimes \Psi_n))(\overline{X}^i \otimes \overline{Y}^j) &= T(m,n)(\Psi_m(\overline{X}^i) \otimes \Psi_n(\overline{Y}^j)) \\
&= T(m,n)(\oplus_{\substack{d_1|m \\ d_2|n}} X^{ni} \bmod \Phi_{d_1}(X) \otimes Y^{mj} \bmod \Phi_{d_2}(Y)) \\
&= \oplus_{d|mn} t^{ni+mj} \bmod \Phi_d(t).
\end{aligned}$$

This completes the proof. □

Motivated by Theorem 2.4, we present our strategy to determine $S(A_n)$: first determine $S(A_{p^\alpha})$ for $p^\alpha \parallel n$; since Kronecker product of diagonal matrices is a diagonal matrix, describe the smith form of a diagonal matrix; finally, use this to determine the elementary divisors of A_n.

2.2. Smith Normal Form of A_{p^e} for a prime p

Let p be a prime. We begin by noting that we have an explicit formula for $\Phi_{p^e}(X)$:

$$\Phi_{p^e}(X) = \sum_{i=0}^{p-1} X^{ip^{e-1}}. \tag{2.3}$$

Using this information, the following lemma determines A_{p^e} recursively:

LEMMA 2.5. *A_{p^e} is a block matrix given by* (1.2).

Proof. Let $A_{p^e} = (B_{ij})_{1 \leqslant i,j \leqslant p}$ where B_{ij} are matrices of size $p^{e-1} \times p^{e-1}$. Since $X^{(k-1)p^{e-1}+i} \equiv X^i \bmod \Phi_{p^j}(X)$ when $0 \leqslant j \leqslant e-1$, $0 \leqslant i \leqslant p^{e-1}$ and $1 \leqslant k \leqslant p$, it follows that $B_{1k} = A_{p^{e-1}}$. Also, X^i is itself the remainder on division by $\Phi_{p^e}(X)$ when $0 \leqslant i \leqslant \phi(p^e) - 1 = p^e - p^{e-1} - 1$. This shows that $(B_{ij})_{\substack{2 \leqslant i \leqslant p \\ 1 \leqslant j \leqslant p-1}}$ is an identity matrix. Finally, using

$$X^{p^{e-1}(p-1)+i} \equiv -X^{p^{e-1}(p-2)+i} - X^{p^{e-1}(p-3)+i} - \cdots - X^{p^{e-1}+i} - X^i \bmod \Phi_{p^e}(X),$$

we see that $B_{ip} = -I$ for $2 \leqslant i \leqslant p$. This completes the proof. □

THEOREM 2.6. *For $e > 0$, the distinct elementary divisors of A_{p^e} are $\{p^i : 0 \leqslant i \leqslant e\}$. The multiplicity of p^i is $\phi(p^{e-i})$.*

For $n \times n$ matrices L and M, let us write $L \sim M$ to mean that L and M are Smith equivalent: that is, $L \sim M$ if and only if there are matrices $P, Q \in \mathrm{GL}_n(\mathbf{Z})$ such that $M = PLQ$.

Proof. We prove this by induction on e.

The case $e = 1$. A_p is a $p \times p$ matrix of the form:

$$A_p = \begin{pmatrix} 1 & 1 & \dots & 1 & 1 \\ 1 & 0 & \dots & 0 & -1 \\ 0 & 1 & \dots & 0 & -1 \\ \vdots & & \ddots & & \vdots \\ 0 & 0 & \dots & 1 & -1 \end{pmatrix}$$

Adding all the columns to the rightmost column, we get the matrix:

$$A_p \sim B_p := \begin{pmatrix} 1 & 1 & \dots & 1 & p \\ 1 & 0 & \dots & 0 & 0 \\ 0 & 1 & \dots & 0 & 0 \\ \vdots & & \ddots & & \vdots \\ 0 & 0 & \dots & 1 & 0 \end{pmatrix} \tag{2.4}$$

Therefore, we have, $|\det(A_p)| = p$, from which the theorem follows.

The Induction Step. Consider the matrix (1.2). Proceeding analogous to the $e = 1$ case, we note that A_{p^e} is Smith equivalent to the matrix

$$A_{p^e} \sim B_{p^e} := \begin{pmatrix} A_{p^{e-1}} & A_{p^{e-1}} & \dots & A_{p^{e-1}} & pA_{p^{e-1}} \\ I_{p^{e-1}} & 0 & \dots & 0 & 0 \\ 0 & I_{p^{e-1}} & \dots & 0 & 0 \\ \vdots & & \ddots & & \vdots \\ 0 & 0 & \dots & I_{p^{e-1}} & 0 \end{pmatrix} \tag{2.5}$$

Now, performing row operations, we may obtain the following matrix, Smith equivalent to A_{p^e}:

$$\begin{pmatrix} 0 & 0 & \dots & 0 & pA_{p^{e-1}} \\ I_{p^{e-1}} & 0 & \dots & 0 & 0 \\ 0 & I_{p^{e-1}} & \dots & 0 & 0 \\ \vdots & & \ddots & & \vdots \\ 0 & 0 & \dots & I_{p^{e-1}} & 0 \end{pmatrix}$$

We now interchange rows to obtain the following form:

$$\begin{pmatrix} I_{p^{e-1}} & 0 & \dots & 0 & 0 \\ 0 & I_{p^{e-1}} & \dots & 0 & 0 \\ \vdots & & \ddots & & \vdots \\ 0 & 0 & \dots & I_{p^{e-1}} & 0 \\ 0 & 0 & \dots & 0 & pA_{p^{e-1}} \end{pmatrix} \sim \begin{pmatrix} I_{p^{e-1}} & 0 & \dots & 0 & 0 \\ 0 & I_{p^{e-1}} & \dots & 0 & 0 \\ \vdots & & \ddots & & \vdots \\ 0 & 0 & \dots & I_{p^{e-1}} & 0 \\ 0 & 0 & \dots & 0 & pS(A_{p^{e-1}}) \end{pmatrix}$$

Since this matrix is in Smith normal form, it must be the Smith normal form of the matrix A_{p^e}. Now, we verify the assertions of the theorem: indeed, the elementary divisors of A_{p^e} are $\{p^i : 0 \leqslant i \leqslant e\}$; the multiplicity of p^{i+1} is $\phi(p^{e-1-i})$ for $0 \leqslant i \leqslant e-1$ (from the induction hypothesis) and 1 appears $p^{e-1}(p-1) = \phi(p^e)$ times. This completes the proof. □

REMARK 2.7. We may actually calculate the determinant of A_{p^e} for $e > 0$ from the proof of Theorem 2.6. Let $I(p,k)$ be the column block matrix

$$I(p,k) = \begin{pmatrix} I_{p^{k-1}} \\ \vdots \\ I_{p^{k-1}} \end{pmatrix}$$

of $p-1$ blocks. Consider the matrix T_{p^e} defined as follows:

$$T_{p^e} = \begin{pmatrix} I_{p^e - p^{e-1}} & I(p,e) \\ 0 & I_{p^{e-1}} \end{pmatrix}. \tag{2.6}$$

Then, it is an easy computation to see that $A_{p^e}T_{p^e} = B_{p^e}$ (see (2.5)). Since $\det(T_{p^e}) = 1$ for all p and e, we see that

$$\det(A_{p^e}) = \det(B_{p^e}) = (-1)^{\phi(p^e)}p^{p^{e-1}}\det(A_{p^{e-1}}). \tag{2.7}$$

This gives us a recursive formula for the determinant of A_{p^e} (indeed, we know $\det(A_1) = \det((1)) = 1$). It now follows that

$$\det(A_{p^e}) = (-1)^{p^e-1}p^{\sum_{k=0}^{e-1}p^k}. \tag{2.8}$$

Thus, for $e > 0$, we have that $\det(A_{p^e})$ is positive for all odd primes p and negative for $p = 2$.

REMARK 2.8. Denote the totality of column (resp. row) operations needed to bring A_{p^e} to its Smith normal form by V_{p^e} (resp. U_{p^e}) so that U_{p^e} and V_{p^e} satisfy the following:

$$U_{p^e}A_{p^e}V_{p^e} = S(A_{p^e}). \tag{2.9}$$

The matrix V_{p^e} can be read off from the proof of the last proposition to be:

$$V_{p^e} = \begin{pmatrix} I_{p^e - p^{e-1}} & I(p,e)V_{p^{e-1}} \\ 0 & V_{p^{e-1}} \end{pmatrix} = \begin{pmatrix} I_{p^{e-1}} & & & & V_{p^{e-1}} \\ & I_{p^{e-1}} & & & V_{p^{e-1}} \\ & & \ddots & & \vdots \\ & & & I_{p^{e-1}} & V_{p^{e-1}} \\ & & & & V_{p^{e-1}} \end{pmatrix}. \tag{2.10}$$

For later purposes, we note that the following equation sets up a recursion for the matrix $W_{p^e} := A_{p^e} V_{p^e}$ with $W_1 = (1)$:

$$W_{p^e} = \begin{pmatrix} A_{p^{e-1}} & \cdots & A_{p^{e-1}} & pW_{p^{e-1}} \\ I_{p^{e-1}} & & & 0 \\ & \ddots & & \vdots \\ & & I_{p^{e-1}} & 0 \end{pmatrix}. \tag{2.11}$$

See Lemma 3.2 for an interpretation of the columns of W_{p^e}.

Let A be an $n \times n$ matrix and B be an $m \times m$ matrix, then, we have:

$$\det(A \otimes B) = (\det(A))^m (\det(B))^n.$$

By Theorem 2.4 and (2.8) we have the following:

COROLLARY 2.9. *If* $n = p_1^{\alpha_1} \dots p_r^{\alpha_r}$, *then*

$$|\det(A_n)| = \prod_{i=1}^{r} p_i^{\frac{n(1-p_i^{-\alpha_i})}{(p_i-1)}} \tag{2.12}$$

REMARK 2.10. One may cast the expression for $|\det(A_n)|$ in many different forms. For example, by comparing the exponent of primes in both sides, one may prove:

$$|\det(A_n)| = \prod_{k=1}^{n} \gcd(k, n). \tag{2.13}$$

In turn, this yields several nice expressions for the determinant:

$$g(n) := \prod_{k=1}^{n} \gcd(k, n) = \prod_{d|n} d^{\phi\left(\frac{n}{d}\right)} = n^n \prod_{d|n} \frac{1}{d^{\phi(d)}}. \tag{2.14}$$

The arithmetic properties of the function g have been studied in [11]. The author begins by observing that for a multiplicative function h, the function

$$g(n; h) := \prod_{k=1}^{n} h(\gcd(k, n))$$

satisfies a curious relationship for relatively prime positive integers m and n:

$$g(mn; h) = g(m; h)^n g(n; h)^m. \tag{2.15}$$

and concludes that $g(n; h)^{1/n}$ is multiplicative. However, it is now clear that underlying this curiosity is the Kronecker product (Theorem 2.4). It is shown that the Dirichlet series

$$\sum_{n=1}^{\infty} \frac{\log(g(n))}{n^s}$$

converges absolutely for $\Re(s) > 2$ and equals $-\frac{\zeta(s-1)\zeta'(s)}{\zeta(s)}$ where $\zeta(s)$ is the Riemann's zeta function. More intricate connections between the function $g(n)$ and the Riemann's zeta function are established (see Corollary 4, loc. cit.). It is also shown that,

$$\max\left(n^{n/v(n)}, n^{\tau(n)/(2n)}\right) \leqslant g(n) \leqslant 27\left(\frac{\log(n)}{\omega(n)}\right)^{n\omega(n)}$$

where $\tau(n)$ is the number of divisors of n, $v(n)$ is the largest prime power divisor of n and $\omega(n)$ is the number of distinct prime factors of n.

Calculating the sign of this determinant turns out to be quite tricky. We will take a different approach (see Appendix A) to calculate the determinant which will also tell us the sign of $\det(A_n)$.

2.3. Smith Normal Form of a Diagonal Matrix

The next order of business is to work out the Smith normal form of a diagonal matrix, D:

$$D = \begin{pmatrix} a_1 & & & \\ & a_2 & & \\ & & \ddots & \\ & & & a_n \end{pmatrix}$$

Notice that we may first permute the rows of D so that the zero rows of the matrix are the last few rows of D. If D° denotes the maximum principal submatrix of D whose rows are all non-zero, the Smith normal form of D is, simply:

$$S(D) = \begin{pmatrix} S(D^\circ) & \\ & 0 \end{pmatrix}$$

Thus, we may assume that $\{a_1, \ldots, a_n\}$ are all non-zero.

ALGORITHM 1. Let $D = \mathrm{diag}(a_1, \ldots, a_n)$ be a diagonal matrix with $a_i \neq 0$ for all i. Let $\mathcal{P}$ be the set of primes that divides at least one of the a_i's. The algorithm proceeds in two steps:

1. Corresponding to a prime $p_j \in \mathcal{P}$, we may associate the partition $\lambda^{(j)}$ obtained by rearranging the sequence of numbers $(\gamma_1, \ldots, \gamma_i, \ldots, \gamma_n)$ in weakly decreasing order, where $p_j^{\gamma_i} \parallel a_i$. Indeed, a partition associated to a prime this way has atmost n non-zero parts.

2. The elementary divisors of the matrix D are now given by the formulae:

$$e_k = \prod_{j=1}^{r} p_j^{\lambda^{(j)}_{n-k+1}}$$

The fact that $\lambda^{(j)}$ is a sequence of weakly decreasing non-negative integers shows that

$$e_1 \mid \cdots \mid e_n.$$

We shall find it convenient to develop a pictorial language for the algorithm. The partitions naturally suggest Young diagrams:

DEFINITION 2.11 (Young Diagram). The Young diagram associated to a partition $\lambda = (\lambda_1, \ldots, \lambda_l)$ is a left-aligned array of boxes with the ith row of the array containing λ_i boxes.

For example, the Young diagram of the partition $\nu = (2, 2, 1)$ is Figure 1. Notice that by definition, the Young diagrams of the partitions ν, that of $(\nu, 0)$,

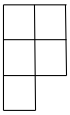

Figure 1: Young Diagram of $(2, 2, 1)$

$(\nu, 0, 0)$ etc. are all the same.

EXAMPLE 2.12. Consider the diagonal matrix

$$D = \begin{pmatrix} 6 & & & \\ & 4 & & \\ & & 7 & \\ & & & 12 \end{pmatrix}$$

The set $\mathcal{P}$ is therefore $\{2, 3, 7\}$. A simple calculation shows that the associated partitions are

$$\begin{aligned} 2 &\leftrightarrow (2, 2, 1, 0) \\ 3 &\leftrightarrow (1, 1, 0, 0) \\ 7 &\leftrightarrow (1, 0, 0, 0) \end{aligned}$$

Therefore, the elementary divisors are:

$$\begin{aligned} e_1 &= 2^0 3^0 7^0 = 1 \\ e_2 &= 2^1 3^0 7^0 = 2 \\ e_3 &= 2^2 3^1 7^0 = 12 \\ e_4 &= 2^2 3^1 7^1 = 84 \end{aligned}$$

It may be helpful to draw the Young diagrams (and this will play a crucial role as we proceed!) on a ruled sheet of paper, see Figure 2.

	2	2		3		7	84
	2	2		3			12
	2						2
							1

Figure 2: The Elementary Divisors of D

THEOREM 2.13. *Algorithm 1 indeed gives the Smith normal form of the diagonal matrix* $D = \mathrm{diag}(a_1, \ldots, a_n)$.

Proof. From [9, Theorem 3.9], we have that a sequence $f_1, \ldots, f_n$ satisfies

$$\prod_{j=1}^{k} f_j = \text{gcd of } k \times k \text{ minors of } D \tag{2.16}$$

for every index $k = 1, \ldots, n$ if and only if $f_1, \ldots, f_n$ are the elementary divisors of D upto units.

Let $e_1, \ldots, e_n$ be the output of the algorithm. We will prove (2.16) with $f_j = e_j$ by backward induction on k. The case when $k = n$ is clear. Note that it suffices to prove that primes and the exponents to which they occur on either sides of (2.16) are equal.

Let $\mathcal{P} = \{p_1, \ldots, p_r\}$ be the set of all primes dividing atleast one of the a_i's. Let us set $a_k = \prod_{j=1}^{r} p_j^{e_{kj}}$.

Since $\gcd(\prod_j p_j^{r_j}, \prod_j p_j^{s_j}) = \prod_j p_j^{\min(r_j, s_j)}$, it suffices to verify the following equality for every j $(1 \leqslant j \leqslant r)$:

$$\begin{aligned} \min\Big\{ \sum_{i \in I} e_{ij} : I \subseteq \{1, \ldots, n\}, |I| = k \Big\} - \lambda^{(j)}_{n-k+1} \\ = \min\Big\{ \sum_{i \in I} e_{ij} : I \subseteq \{1, \ldots, n\}, |I| = k-1 \Big\} \end{aligned}$$

for $1 \leqslant k \leqslant n$ where the notation $\lambda^{(j)}_{n-k+1}$ is as in Algorithm 1. But, this follows since $\min\Big\{ \sum_{i \in I} e_{ij} : I \subseteq \{1, \ldots, n\}, |I| = k \Big\}$ equals the sum of the first k elements when, for a fixed j, the exponents e_{ij}'s are written in ascending order. That is,

$$\min\Big\{ \sum_{i \in I} e_{ij} : I \subseteq \{1, \ldots, n\}, |I| = k \Big\} = \sum_{p=0}^{k-1} \lambda^{(j)}_{n-p}$$

for every $1 \leqslant k \leqslant n$. □

FACT 2.14. Let G be a finite abelian group. The elementary divisors are easily computed from the primary decomposition by step 2 of the algorithm. Conversely, given its elementary divisors, the primary decomposition is the set of

pairs $(p_j, \lambda^{(j)})$ obtained from step 1 of the algorithm, with these elementary divisors as the entries of a diagonal matrix.

Before we can compute the elementary divisors of A_n, we need to compute the partitions and primes in the Kronecker product. To do this, we recall that for the matrix $S(A_{p^\alpha})$, the set $\mathcal{P}$ is singleton $\{p\}$ and the partition associated to p is,

$$p \leftrightarrow (\quad \ldots, \underbrace{\alpha - i, \ldots, \alpha - i}_{\phi(p^i) \text{ times}}, \ldots \quad), \qquad 0 \leqslant i \leqslant \alpha$$

where $\alpha - i$ appears $\phi(p^i)$ times, $0 \leqslant i \leqslant \alpha$ (see Theorem 2.6).

Now, in the Kronecker Product, $S(A_{p_1^{n_1}}) \otimes S(A_{p_2^{n_2}})$, the set $\mathcal{P}$ of primes is $\{p_1, p_2\}$ and the associated partitions are,

$$p_1 \leftrightarrow (\quad \ldots, \underbrace{n_1 - i, \ldots, n_1 - i}_{\phi(p_1^i)p_2^{n_2} \text{ times}}, \ldots \quad), \qquad 0 \leqslant i \leqslant n_1$$

$$p_2 \leftrightarrow (\quad \ldots, \underbrace{n_2 - j, \ldots, n_2 - j}_{\phi(p_2^j)p_1^{n_1} \text{ times}}, \ldots \quad), \qquad 0 \leqslant j \leqslant n_2$$

The following is easily seen by induction:

THEOREM 2.15. *Suppose that*

$$n = p_1^{\alpha_1} \ldots p_r^{\alpha_r}$$

is the factorisation of a positive integer n. Then, the set $\mathcal{P}$ of primes for the diagonal matrix

$$\bigotimes_j S(A_{p_j^{\alpha_j}})$$

is the set $\{p_1, \ldots, p_r\}$ and the associated partitions are

$$p_j \leftrightarrow (\quad \ldots, \underbrace{\alpha_j - i, \ldots, \alpha_j - i}_{\phi(p_j^i)n/p_j^{\alpha_j} \text{ times}}, \ldots \quad), \qquad 0 \leqslant i \leqslant \alpha_j.$$

for $j = 1, \ldots, r$.

Theorem 2.15 together with Algorithm 1 completely solves the problem of determining the elementary factors of the matrix A_n. We illustrate this in an example:

EXAMPLE 2.16. Let $n = 12 = 2^2 3^1$. Then, the set $\mathcal{P}$ of primes for the diagonal matrix $S(A_4) \otimes S(A_3)$ is $\{2, 3\}$. Associated partitions are:

$$2 \leftrightarrow (2, 2, 2, 1, 1, 1)$$
$$3 \leftrightarrow (1, 1, 1, 1)$$

Now, the formulae for the elementary divisors show that the elementary divisors of A_{18} are $\{1, 1, 1, 1, 1, 1, 2, 2, 6, 12, 12, 12\}$. Pictorially, we have:

	2	2		3	12
	2	2		3	12
	2	2		3	12
	2			3	6
	2				2
	2				2
					1
	⋮				⋮
					1

Figure 3: The Elementary Divisors of A_{12}

2.4. Consequences

To proceed further, we need some more notions related to a Young Diagram. The *height* of a Young diagram Y is the number of rows in Y. A cell in Y is called a *corner* if there is no cell to its right and there is no cell below it.

Suppose that a positive integer n has the factorisation

$$n = p_1^{\alpha_1} \dots p_r^{\alpha_r}$$

with $\alpha_i > 0$ and $p_1 < \cdots < p_r$. Let $E(A_n)$ denote the multiset of elementary divisors of A_n.

PROPOSITION 2.17. *With notations as above, we have:*

1. *The multiplicity of* 1 *in* $E(A_n)$ *is* $n - \frac{n}{p_1}$. *Equivalently, the rank of the cokernel of* Ψ_n *is* $\frac{n}{p_1}$.

2. *The least integer* $d > 1$ *in* $E(A_n)$ *is* p_1 *and its multiplicity in* $E(A_n)$ *is*

$$\begin{cases} \frac{n}{p_1} - \frac{n}{p_2}, & \text{if } p_2 < p_1^2 \\ \frac{n}{p_1} - \frac{n}{p_1^2}, & \text{if } p_2 > p_1^2 \end{cases}$$

3. *The largest integer* $m \in E(A_n)$ *is* n *and its multiplicity in* $E(A_n)$ *is*

$$\frac{n}{\max\{p_i^{\alpha_i} \mid i = 1, \dots, r\}}.$$

In the proof of this proposition, we will make use of the general tableaux diagram found in Figure 4.

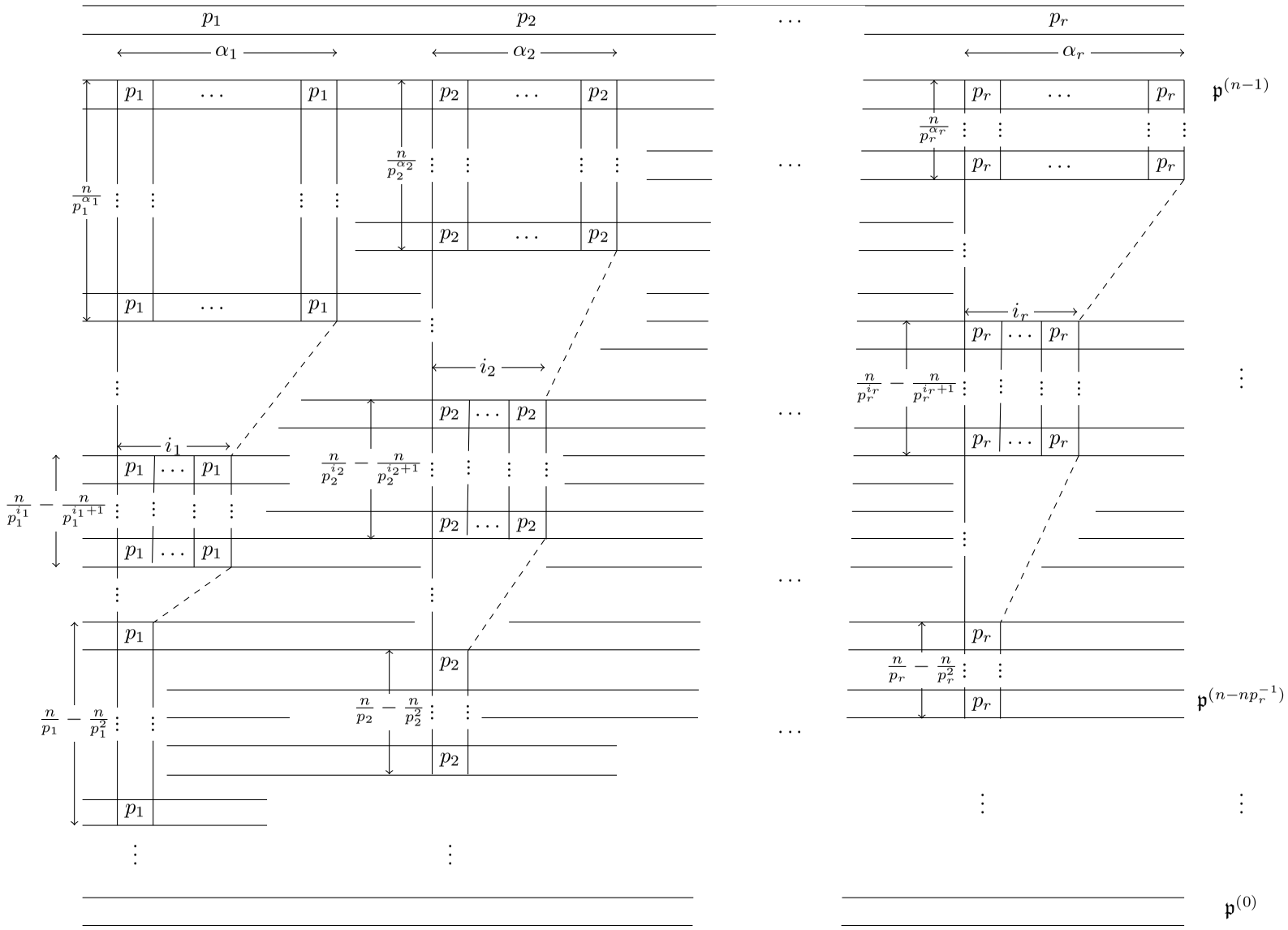

Figure 4: Algorithm seen through Young Diagrams

Proof.

1. The multiplicity of 1 in $E(A_n)$ is

$$\begin{aligned} & n - \max\{\text{height of } p_i \text{ tableau} \mid i = 1, \ldots, r\} \\ &= n - \frac{n}{p_1} \end{aligned}$$

2. Let $d > 1$ be the least elementary divisor of A_n. Then d is the product of p_is in the lowest non-empty row. Since the p_1-tableau is the tallest having one box in the last row, this d must be p_1. The multiplicity of p_1 is

$$\begin{aligned} & \frac{n}{p_1} - \text{Index of the row containing the second last corner} \\ &= \begin{cases} \frac{n}{p_1} - \frac{n}{p_2}, & \text{if } \frac{n}{p_2} > \frac{n}{p_1^2} \\ \frac{n}{p_1} - \frac{n}{p_1^2}, & \text{if } \frac{n}{p_2} < \frac{n}{p_1^2} \end{cases} \end{aligned}$$

   which proves the claim.

3. The largest elementary divisor is the product of the numbers in the first row of Figure 4. This number is clearly n. The multiplicity of n in $E(A_n)$ is the index of the row containing the first corner. We see that this multiplicity is

$$\begin{aligned} & \min\left\{ \frac{n}{p_i^{\alpha_i}} \middle| 1 \leqslant i \leqslant r \right\} \\ &= \frac{n}{\max\{p_i^{\alpha_i} : 1 \leqslant i \leqslant r\}} \end{aligned}$$

   This completes the proof. □

This proposition and its proof suggest that Figure 4, in principle, gives a "formula" for the elementary divisors and their muliplicities, equivalently, the multiset $E(A_n)$.

THEOREM 2.18.

1. *The n-tuple* $\left(e_1, \frac{e_2}{e_1}, \ldots, \frac{e_n}{e_{n-1}}\right)$ *is a permutation of the n-tuple*

$$(\underbrace{p_1, \ldots, p_1}_{\alpha_1 \text{ times}}, \underbrace{p_2, \ldots, p_2}_{\alpha_2 \text{ times}}, \ldots, \underbrace{p_r, \ldots, p_r}_{\alpha_r \text{ times}}, \underbrace{1, \ldots, 1}_{n - \sum_i \alpha_i \text{ times}})$$

2. *The ratio* $\frac{e_j}{e_{j-1}}$ *is* p_i *if and only if* $j = n - \frac{n}{p_i^t} + 1$ *for some* t *satisfying* $1 \leqslant t \leqslant \alpha_i$.

Proof. 1. We need to prove that $\frac{e_i}{e_{i-1}}$ is a prime divisor of n. Note that this claim is equivalent to proving that every row in Figure 4 has atmost one

corner. That is, exactly one of the p_i tableau has a corner. Towards a contradiction, assume that there are two distinct primes p_{i_1} and p_{i_2} whose tableaux for n have a corner each in the same row R:

Thus, there are indices l_1 and l_2 with $0 \leqslant l_1 \leqslant \alpha_{i_1} - 1$ and $0 \leqslant l_2 \leqslant \alpha_{i_2} - 1$ such that

$$\begin{aligned} R &= \sum_{j=0}^{l_1} \frac{n}{p_{i_1}^{\alpha_{i_1}}} \phi(p_{i_1}^j) \\ &= \sum_{j=0}^{l_2} \frac{n}{p_{i_2}^{\alpha_{i_2}}} \phi(p_{i_2}^j) \end{aligned}$$

This equality implies that

$$\frac{n}{p_{i_1}^{\alpha_{i_1} - l_1}} = \frac{n}{p_{i_2}^{\alpha_{i_2} - l_2}}$$

which is a contradiction, since p_{i_1} and p_{i_2} are distinct primes.

2. Notice that index k of the n-tuple contains p_i if the row k contains a corner of the p_i-tableau. These indices are therefore given by

$$\left\{ \frac{n}{p_i^{\alpha_i}} \sum_{j=0}^{r} \phi(p_i^{\alpha_i - j}) + 1 \middle| r = 0, 1, \ldots, \alpha_i - 1 \right\}.$$

Thus, we get, the indices which contain p_i are

$$\left\{ n - \frac{n}{p_i^r} + 1 \middle| r = 1, \ldots, \alpha_i \right\}$$

which completes the proof. □

We leave it to the reader to find another proof of Proposition 2.17 using Theorem 2.18.

COROLLARY 2.19. *If n is square-free, that is, $\alpha_i = 1$ for all $i = 1, \ldots, r$, then*

$$\left(e_1, \frac{e_2}{e_1}, \ldots, \frac{e_n}{e_{n-1}} \right) = (p_r, \ldots, p_1, 1, \ldots, 1)$$

In the following sections, we shall describe a basis for $G(n)$ in terms of the standard basis of $\oplus_{d|n} \mathbf{Z}[x]/\Phi_d(x)$ through a pictorial algorithm: by a basis for $G(n)$ is meant a set of generators

$$\{0 \oplus 0 \oplus \cdots \oplus \underbrace{1}_{i\text{th place}} \oplus 0 \oplus \cdots \oplus 0 : 0 \leqslant i \leqslant n-1\}$$

for the abelian group $\bigoplus \mathbf{Z}/e_n(i)\mathbf{Z}$ where $(e_n(0), \ldots, e_n(n-1))$ is the tuple of elementary divisors for A_n.

3. Setup for the Algorithm

In this section, we will state the definitions and prove some basic lemmas that are instrumental to the algorithm in the next section.

To determine a basis for $G(n)$, it suffices to find a basis $\{\mathfrak{p}^{(j)} : 0 \leqslant j \leqslant n-1\}$ for $\bigoplus_{d|n} \mathbf{Z}[X]/\Phi_d(X)$ so that

$$\{\overline{\Psi}_n(\mathfrak{p}^{(j)}) : e_n(j) > 1\} \tag{3.1}$$

is a set of generators for the abelian group $G(n)$ with respect to which the relations are the simplest possible:

$$e_n(j)\overline{\Psi}_n(\mathfrak{p}^{(j)}) = 0. \tag{3.2}$$

This idea is captured by the following definition:

DEFINITION 3.1. Given a positive integer n, let $(e_n(0), \ldots, e_n(n-1))$ be the tuple of elementary divisors of A_n. We say that $(\mathfrak{p}^{(j)} : 0 \leqslant j \leqslant n-1)$ is a **Smith vector** for n if:

1. $\left\{\mathfrak{p}^{(j)} : 0 \leqslant j \leqslant n-1\right\}$ is a $\mathbf{Z}$-basis of $\bigoplus_{d|n} \mathbf{Z}[X]/\Phi_d(X)$ and
2. $a_j\mathfrak{p}^{(j)} \in \mathrm{Im}(\Psi_n)$ if and only if $e_n(j) \mid a_j$.

The following lemma will tell us how to compute Smith vector for n:

LEMMA 3.2. *If $(\mathfrak{p}^{(j)})$ is a Smith vector for n, then there exists $U_n, V_n \in GL_n(\mathbf{Z})$ such that $S(A_n) = U_nA_nV_n$ and $A_nV_n(\overline{X}^j) = e_n(j)\mathfrak{p}^{(j)}$ where*

$$(e_n(0), \ldots, e_n(n-1))$$

is the tuple of elementary divisors of A_n.

Proof. We introduce a notation for the standard basis of the direct sum $\oplus_{d|n}\mathbf{Z}[X]/\langle\Phi_d(X)\rangle$; for a divisor d of n, and for every i such that $0 \leqslant i \leqslant \phi(d) - 1$, put:

$$g_{i,d}(X) := 0 \oplus \cdots \oplus 0 \oplus X^i \bmod (\Phi_d(X)) \oplus 0 \oplus \cdots \oplus 0.$$

Let U_n be the endomorphism of $\oplus_{d|n}\mathbf{Z}[X]/\langle\Phi_d(X)\rangle$ that exchanges the basis underlying the given Smith vector with the standard basis:

$$\mathfrak{p}^{(j)} \mapsto g_{i,d} \text{ where } j = \sum_{d'<d, d'|n} \phi(d') + i.$$

Thus, U_n is invertible. Since $e_n(j)\mathfrak{p}^{(j)}$ is in the image of A_n, it follows that there are vectors $h^{(j)} \in \mathbf{Z}[X]/\langle X^n - 1\rangle$ such that

$$A_n(h^{(j)}) = e_n(j)\mathfrak{p}^{(j)}.$$

Define the map $V_n : \mathbf{Z}[X]/\langle X^n - 1\rangle \to \mathbf{Z}[X]/\langle X^n - 1\rangle$ as follows:

$$\overline{X}^j \mapsto h^{(j)}, \quad 0 \leqslant j \leqslant n-1$$

Clearly, $U_n A_n V_n = S(A_n)$. It suffices to check that V_n is an isomorphism, that is, $\det(V_n) = \pm 1$:

$$\begin{aligned} U_n A_n V_n = S(A_n) &\Rightarrow \det(U_n)\det(A_n)\det(V_n) = \det(S(A_n)) \\ &\Rightarrow \det(A_n)\det(V_n) = \pm|\det(A_n)| \text{ (since } \det(U_n) = \pm 1) \\ &\Rightarrow \det(V_n) = \pm 1 \text{ (since } \det(A_n) \neq 0). \end{aligned}$$

This completes the proof. □

From Lemma 3.2, we see that $\{e_n(j)\mathfrak{p}^{(j)} : 0 \leqslant j \leqslant n-1\}$ is a basis for the image of Ψ_n and we have the following isomorphism of $\mathbf{Z}$-modules:

$$G(n) \simeq \bigoplus_{e_n(j)>1} \langle \overline{\Psi}_n(\mathfrak{p}^{(j)})\rangle \simeq \bigoplus_{e_n(j)>1} \mathbf{Z}/e_n(j)\mathbf{Z}. \tag{3.3}$$

In Lemma 2.4, for relatively prime positive integers m and n, we have shown that $S(A_{mn}) = S(A_m \otimes A_n)$. It is now natural to ask if Smith vectors for m and n can be coaxed to produce a Smith vector for mn. In the commutative diagram of maps in Figure 5, since both the rows are exact and $P_{m,n}$ and $T(m,n)$ are isomorphisms, a straightforward diagram chasing proves that $f_{m,n}$ is an isomorphism (see also [10, Lemma 7.1]). In the next lemma, we consider

$$\begin{array}{ccccccccc} 0 & \longrightarrow & \frac{\mathbf{Z}[X]}{\langle X^m-1\rangle} \otimes \frac{\mathbf{Z}[Y]}{\langle Y^n-1\rangle} & \xrightarrow{\Psi_m\otimes\Psi_n} & \bigoplus_{\substack{d_1|m\\ d_2|n}} \frac{\mathbf{Z}[X]}{\langle \Phi_{d_1}(X)\rangle} \otimes \frac{\mathbf{Z}[Y]}{\langle \Phi_{d_2}(Y)\rangle} & \xrightarrow{\overline{\Psi_m\otimes\Psi_n}} & G(m,n) & \longrightarrow & 0 \\ & & \downarrow P_{m,n} & & \downarrow T(m,n) & & \downarrow f_{m,n} & & \\ 0 & \longrightarrow & \frac{\mathbf{Z}[t]}{\langle t^{mn}-1\rangle} & \xrightarrow{\Psi_{mn}} & \bigoplus_{d|mn} \frac{\mathbf{Z}[t]}{\langle \Phi_d(t)\rangle} & \xrightarrow{\overline{\Psi_{mn}}} & G(mn) & \longrightarrow & 0 \end{array}$$

Figure 5: Tensor Product of Smith Vectors

the tensor product of Smith vectors in the top row of Figure 5 and study its properties in the bottom row.

LEMMA 3.3. *Let m and n be relatively prime positive integers. Let $\left\{(\mathfrak{p}^{(j)}) : 0 \leqslant j \leqslant m-1\right\}$ and $\left\{(\mathfrak{q}^{(j)}) : 0 \leqslant j \leqslant n-1\right\}$ be Smith vectors for m and n respectively. Then, in the group $G(mn)$,*

1. *the order of the element $\overline{\Psi}_{mn}(\mathfrak{p}^{(i)}(\overline{t}^n)\mathfrak{q}^{(j)}(\overline{t}^m))$ is $e_m(i)e_n(j)$, where $e_k(\iota)$ denotes the ιth elementary divisor of A_k.*

2. *suppose that* $\overline{\Psi}_{mn}(\mathfrak{p}^{(i_1)}(\bar{t}^n)\mathfrak{q}^{(j_1)}(\bar{t}^m))$ *and* $\overline{\Psi}_{mn}(\mathfrak{p}^{(i_2)}(\bar{t}^n)\mathfrak{q}^{(j_2)}(\bar{t}^m))$ *are non-zero. Then, the intersection*

$$\langle\overline{\Psi}_{mn}(\mathfrak{p}^{(i_1)}(\bar{t}^n)\mathfrak{q}^{(j_1)}(\bar{t}^m))\rangle \cap \langle\overline{\Psi}_{mn}(\mathfrak{p}^{(i_2)}(\bar{t}^n)\mathfrak{q}^{(j_2)}(\bar{t}^m))\rangle$$

of subgroups is non-trivial if and only if $i_1 = i_2$ *and* $j_1 = j_2$.

In this lemma, we interpret $\mathfrak{p}^{(j)}(\bar{t}^n)$ as the element $\bigoplus_{d|n}\mathfrak{p}_d^{(j)}(\bar{t}^n)$ which belongs to the direct sum $\bigoplus_{d|n}\mathbf{Z}[t]/\Phi_d(t)$. And, the product $\mathfrak{p}^{(i)}\mathfrak{q}^{(j)}$ is to be interpreted as the element $\bigoplus_{\substack{d_1|m\\d_2|n}}\mathfrak{p}_{d_1}^{(i)}\mathfrak{q}_{d_2}^{(j)}$ which belongs to $\bigoplus_{\substack{d_1|m\\d_2|n}}\mathbf{Z}[t]/\Phi_{d_1d_2}(t)$.

Proof. The set $\left\{\mathfrak{p}^{(i)}\otimes\mathfrak{q}^{(j)} : \substack{0\leqslant i\leqslant m-1\\0\leqslant j\leqslant n-1}\right\}$ is a basis for $\bigoplus_{\substack{d_1|m\\d_2|n}}\frac{\mathbf{Z}[X]}{\Phi_{d_1}(X)}\otimes\frac{\mathbf{Z}[Y]}{\Phi_{d_2}(Y)}$ since $\mathfrak{p}^{(i)}$ and $\mathfrak{q}^{(j)}$ are Smith vectors for m and n respectively. Since $T(m,n)$ is an isomorphism, we get that the set $\left\{\mathfrak{p}^{(i)}(\bar{t}^n)\mathfrak{q}^{(j)}(\bar{t}^m) : \substack{0\leqslant i\leqslant m-1\\0\leqslant j\leqslant n-1}\right\}$ is a $\mathbf{Z}$-basis for the codomain of Ψ_{mn}. From the linear independence of these vectors, (2) follows.

By Lemma 3.2, there are isomorphisms V_m and V_n such that:

$$A_mV_m(\overline{X}^i) = e_m(i)\mathfrak{p}^{(i)} \text{ and } A_nV_n(\overline{Y}^j) = e_n(j)\mathfrak{q}^{(j)} \tag{3.4}$$

for $0\leqslant i\leqslant m-1$ and $0\leqslant j\leqslant n-1$. Since $\overline{X}^i\otimes\overline{Y}^j$ is a basis for $\frac{\mathbf{Z}[X]}{(X^m-1)}\otimes\frac{\mathbf{Z}[Y]}{(Y^n-1)}$, we note that $\mathfrak{p}^{(i)}\otimes\mathfrak{q}^{(j)}$ is a basis for the image of $\Psi_m\otimes\Psi_n$. Now, using the fact that $T(m,n)$ and $P_{m,n}$ are isomorphisms and the commutativity of Figure 5, we have the set $\{e_m(i)e_n(j)\mathfrak{p}^{(i)}(\bar{t}^n)\mathfrak{q}^{(j)}(\bar{t}^m) : \substack{0\leqslant i\leqslant m-1\\0\leqslant j\leqslant n-1}\}$ is a basis for the image of Ψ_{mn} from which (1) follows. □

Remark 3.4. More generally, we note that if k_1 and k_2 are two relatively prime positive integers, then, we may obtain the young diagram for k_1k_2 by repeating k_2 times, the rows of the Young diagram for k_1 and k_1 times, the rows of the Young diagram for k_2. To see this, suppose that

$$k_1 = p_1^{\alpha_1}\dots p_r^{\alpha_r} \text{ and } k_2 = q_1^{\beta_1}\dots q_s^{\beta_s}$$

where $\alpha_i,\beta_j > 0$ and $\{p_i : 1\leqslant i\leqslant r\}$ and $\{q_j : 1\leqslant j\leqslant s\}$ are disjoint. Then, in the p_j tableau for k_1, the row with e boxes (where $0\leqslant e\leqslant\alpha_j$) appears $\phi(p_j^{\alpha_j-e})k_1p_j^{-\alpha_j}$ times, while this row appears $\phi(p_j^{\alpha_j-e})k_1k_2p_j^{-\alpha_j}$ in the p_j tableau for k_1k_2. This proves our claim.

From the lemma, we infer that tensoring Smith vectors does not work, because $\left\{e_m(i)e_n(j) : \substack{0\leqslant i\leqslant m-1\\0\leqslant j\leqslant n-1}\right\}$ is not the set of elementary divisors of A_{mn}. However, we will see how to get over this difficulty in the next section. We conclude this section with the following lemma which we will need in the next section and is interesting in its own right.

LEMMA 3.5. *Let P be a permutation matrix and m and n be relatively prime positive integers. Then, there are diagonal matrices D_1 and D_2 satisfying:*

$$\begin{aligned} \gcd(\det(D_1), m) &= 1, \\ \gcd(\det(D_2), n) &= 1 \text{ and} \\ \det(nD_1 + mD_2P) &= 1. \end{aligned}$$

Proof. Let P be a $k \times k$ permutation matrix whose associated permutation is π. We shall present an algorithm to find integers $\{a_i\}_{i=1}^k$ and $\{b_j\}_{j=1}^k$ such that

$$D_1 = \operatorname{diag}(a_1, \ldots, a_k) \quad \text{and} \quad D_2 = \operatorname{diag}(b_1, \ldots, b_k)$$

satisfy the requirements of lemma.

Step 1. If π is a k-cycle, then, one may compute $\det(nD_1 + mD_2P)$ by the usual formula:

$$\det(nD_1 + mD_2P) = \sum_{\sigma \in S_k} (-1)^{\operatorname{sgn}(\sigma)} x_{1\sigma(1)} \ldots x_{k\sigma(k)}$$

where x_{ij} are the entries of the matrix $nD_1 + mD_2P$.

For a permutation σ

$$x_{1\sigma(1)} \ldots x_{k\sigma(k)} \neq 0 \Rightarrow \ \forall\, 1 \leqslant i \leqslant k \quad (\sigma(i) = i \text{ or } \sigma(i) = \pi(i)). \tag{3.5}$$

Let σ be not the identity permutation such that $x_{1\sigma(1)} \ldots x_{k\sigma(k)} \neq 0$. By (3.5) there exists j such that $\sigma(j) = \pi(j) \neq j$. Set $S = \{\sigma^\ell(j) : 1 \leqslant \ell \leqslant k\}$. No element of S is fixed by σ, because if $\sigma(\sigma^\ell(j)) = \sigma^\ell(j)$, then $\sigma(j) = j$, a contradiction. Thus, $\sigma|_S = \pi|_S$. But, $\{\pi^\ell(j) : 1 \leqslant \ell \leqslant k\} = \{1, \ldots, k\}$ since π is a k-cycle. Since $\sigma^\ell(j) = \pi^\ell(j)$, the set S is all of $\{1, \ldots, k\}$. Thus, σ must be π. Hence:

$$\det(nD_1 + mD_2P) = n^k \prod_{l=1}^k a_l + (-1)^{k-1} m^k \prod_{l=1}^k b_l.$$

Since $\gcd(m^k, (-1)^{k-1}n^k) = 1$, there exists u and v such that

$$m^k u + (-1)^{k-1} n^k v = 1.$$

Then, it is easy to verify that the following choices

$$\begin{aligned} a_1 &= u \\ a_l &= 1 \text{ for all } l \neq 1 \\ b_1 &= v \\ b_l &= 1 \text{ for all } l \neq 1 \end{aligned}$$

meet the requirements of the lemma.

Step 2. If π is not a cycle, let the cycle decomposition of π be $\pi_1 \ldots \pi_r$. We may now repeat Step 1 on each of the cycles π_i and determine the scalars a_j and b_j for those j not fixed by π_i. □

For an alternative proof of this lemma, see [12].

4. An Algorithm for determining the Smith Vector for n

Given a positive integer n and its prime factorisation

$$n = p_1^{\alpha_1} \dots p_r^{\alpha_r},$$

we derived a formula for the elementary divisors of Ψ_n using r Young tableaux, one for each prime p_i. Now, we will use the same set of tableaux diagrams to give an algorithm to find a Smith vector for n. The algorithm will use the following three modules:

$SV(p^e)$ Construct a Smith vector for p^e, a prime power.

$TSV(k_1, k_2, \mathfrak{p}, \mathfrak{q})$ Given Smith vectors $\mathfrak{p}$ and $\mathfrak{q}$, respectively, for relatively prime positive integers k_1 and k_2, construct a Smith vector for k_1k_2.

$SV(n)$ Construct a Smith vectors for n inductively on prime powers.

We will illustrate these modules with the example $n = 6$. These algorithms are also implemented in SAGE (`http://www.sagemath.org`) and the code is available from:

`https://www.dropbox.com/s/0vtzxc8stb6la56/final_smith_form.sage`.

We will calculate the bit complexity of each module in Section 4.4.

4.1. $SV(p^e)$

Recall from Lemma 3.2 that upto scaling by elementary divisors, the columns of the matrix $W_n = A_nV_n$ give a Smith vector for n. We have established that W_{p^e} has a simple recursive form (see Remark 2.8). Therefore, a Smith vector $\mathrm{SV}(p^e)$ for p^e can be computed from this recursive formula.

REMARK 4.1. Observe that non-zero coefficients in the Smith vector $\mathrm{SV}(p)$ are ± 1. Since the non-zero entries of the matrix A_{p^e} are ± 1, it follows by induction on e that the non-zero coefficients in $\mathrm{SV}(p^e)$ are ± 1. We will use this fact on the calculation of bit complexity of the algorithm.

Examples.

$$\mathrm{SV}(2) = (\overline{1} \oplus 0, \overline{1} \oplus \overline{1}) \tag{4.1}$$

$$\mathrm{SV}(3) = (\overline{1} \oplus 0, \overline{1} \oplus \overline{t}, \overline{1} \oplus \overline{1}) \tag{4.2}$$

4.2. $TSV(k_1, k_2, \mathfrak{p}, \mathfrak{q})$

This module is the most crucial part of our algorithm. Let us first do a pictorial construction and make some observations about it.

Construction. To construct the tableaux diagram for k_1k_2 from the tableaux diagram for k_1 and that of k_2, we need k_2 repetitions of the rows of the tableaux diagram for k_1 and k_1 repetitions of the rows of the tableaux diagram for k_2

(See Remark 3.4). Throughout this subsection, it will be convenient to assume that the rows of the tableaux diagram for n are numbered from bottom to top with indices between 0 and $n-1$. We will also index repetition of a row from the bottom with indices between 0 and $k_i - 1$. So, the components of the Smith vector attached to the rows of the tableaux diagram for k_1 and k_2 must change as follows:

- to the j_1th repetition of i_1th row of tableaux diagram of k_1, we attach the vector $k_2\mathfrak{p}^{(i_1)}(\bar{t}^{k_2})\mathfrak{q}^{(j_1)}(\bar{t}^{k_1})$,
- to the i_2th repetition of j_2th row of the tableaux diagram of k_2, we attach the vector $k_1\mathfrak{p}^{(i_2)}(\bar{t}^{k_2})\mathfrak{q}^{(j_2)}(\bar{t}^{k_1})$.

Here, we have multiplied by k_i so that the image of the corresponding element under $\overline{\Psi}_{mn}$ has order dividing k_{3-i}. Finally, we will juxtapose these tableaux diagrams for k_1k_2 so that the row indices match. This construction is demonstrated in Figure 6. In this tableaux diagram for k_1k_2 the vectors attached to the ℓth row are $k_2\mathfrak{P}_\ell := k_2\mathfrak{p}^{(i_1)}(\bar{t}^{k_2})\mathfrak{q}^{(j_1)}(\bar{t}^{k_1})$ and $k_1\mathfrak{Q}_\ell := k_1\mathfrak{p}^{(i_2)}(\bar{t}^{k_2})\mathfrak{q}^{(j_2)}(\bar{t}^{k_1})$ where (i_1, j_1) and (i_2, j_2) are uniquely determined by:

$$\begin{cases} \ell = k_2i_1 + j_1 = k_1j_2 + i_2 \\ \qquad 0 \leqslant j_s \leqslant k_{3-s} - 1, \quad s \in \{1, 2\}. \end{cases} \tag{4.3}$$

Observations. We now make the following observations about the construction:

1. The diagram in Figure 6 is the tableaux diagram associated to k_1k_2.
2. For $0 \leqslant \ell \leqslant k_1k_2 - 1$, we have $e_{k_1k_2}(\ell) = e_{k_1}(i_1)e_{k_2}(j_2)$ where i_1 and j_2 are determined by (4.3).

   Proof. The ℓth elementary divisor of k_1k_2 is the product of the entries in the ℓth row of Figure 6: clearly, the tableaux diagram for k_1 contributes $e_{k_1}(i_1)$ and that of k_2 contributes $e_{k_2}(j_2)$. □

3. With notations as in (2), if we can find $d_{1,\ell}, d_{2,\ell}$ such that $\gcd(d_{1,\ell}, k_1) = \gcd(d_{2,\ell}, k_2) = 1$ and $\{d_{1,\ell}k_2\mathfrak{P}_\ell + d_{2,\ell}k_1\mathfrak{Q}_\ell : 0 \leqslant \ell \leqslant k_1k_2 - 1\}$ is a basis for the codomain of $\Psi_{k_1k_2}$, then, $(d_{1,\ell}k_2\mathfrak{P}_\ell + d_{2,\ell}k_1\mathfrak{Q}_\ell : 0 \leqslant \ell \leqslant k_1k_2 - 1)$ is a Smith vector for k_1k_2.

   Proof. From Lemma 3.3, it is clear that the order of $\overline{\Psi}_{k_1k_2}(k_1\mathfrak{P}_\ell)$ is $e_{k_1}(i_1)$ and that of $\overline{\Psi}_{k_1k_2}(k_2\mathfrak{Q}_\ell)$ is $e_{k_2}(j_2)$. Since $\gcd(d_{s,\ell}, k_s) = 1$, the order of $\overline{\Psi}_{k_1k_2}(k_s(\cdot)_\ell)$ remains unchanged after multiplication by $d_{s,\ell}$. Therefore, the order of $\overline{\Psi}_{k_1k_2}(d_{1,\ell}k_1\mathfrak{P}_\ell + d_{2,\ell}k_2\mathfrak{Q}_\ell)$ equals $e_{k_1}(i_1)e_{k_2}(j_2)$. This equals $e_{k_1k_2}(\ell)$, by (2) and we are done. □

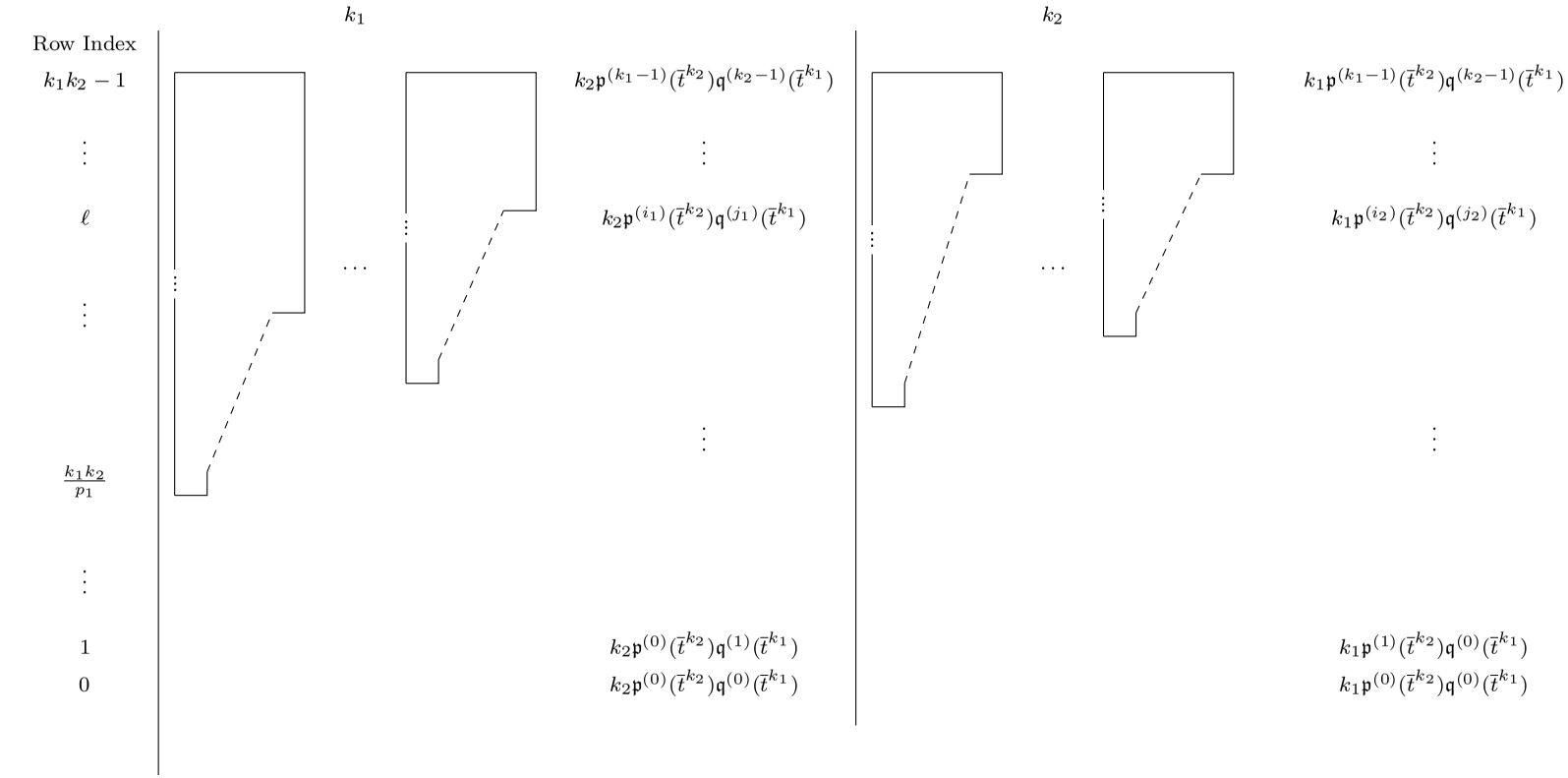

Figure 6: Pictorial Construction for k_1 and k_2

4. Define the following bijections between $[k_1k_2]$ and $[k_1] \times [k_2]$ where $[n] := \{0, 1, \ldots, n-1\}$:

$$\iota(\ell) = (i_1, j_1)$$
$$\iota'(\ell) = (i_2, j_2)$$

where (i_1, j_1) and (i_2, j_2) are determined from ℓ using (4.3). Let L and L' be total orderings on $[k_1] \times [k_2]$ defined by transferring the order from $[k_1k_2]$ via ι and ι' respectively. Then, L is in lexicographic order and L' is in reverse lexicographic order. Let σ_{k_1,k_2} the permutation $\iota'^{-1}\iota$ of the set $[k_1k_2]$.

EXAMPLE 4.2. Take $k_1 = 2, k_2 = 3$. Then, L and L' are increasing along rows in the table below. In cycle notation, $\sigma_{2,3} \equiv (1\ 2\ 4\ 3)$.

ℓ	0	1	2	3	4	5
L	(0, 0)	(0, 1)	(0, 2)	(1, 0)	(1, 1)	(1, 2)
L'	(0, 0)	(1, 0)	(0, 1)	(1, 1)	(0, 2)	(1, 2)

5. Let $\mathfrak{P} = (\mathfrak{P}_\ell : 0 \leqslant \ell \leqslant k_1k_2 - 1)$ and P be the permutation matrix $\sigma_{k_1,k_2}I$. By Lemma 3.5, there are diagonal matrices D_1 and D_2 such that $\det(k_2D_1 + k_1D_2P) = 1$. Since $\{\mathfrak{P}_\ell : 0 \leqslant \ell \leqslant k_1k_2 - 1\}$ is a basis for the codomain of $\Psi_{k_1k_2}$, we have that the components of $(k_2D_1 + k_1D_2P)\mathfrak{P}^t$ forms a basis for the codomain of $\Psi_{k_1k_2}$. Now, setting $D_1 = \mathrm{diag}(d_{1,1}, \ldots, d_{1,k_1k_2-1})$ and $D_2 = \mathrm{diag}(d_{2,1}, \ldots, d_{2,k_1k_2-1})$, by (3), we have that

$$\mathrm{TSV}(k_1, k_2, \mathfrak{p}, \mathfrak{q}) := (d_{1,\ell}k_2\mathfrak{P}_\ell + d_{2,\ell}k_1\mathfrak{Q}_\ell : 0 \leqslant \ell \leqslant k_1k_2 - 1)$$

is a Smith vector for k_1k_2.

ALGORITHM 2. Given the above observations, we are now ready to present the algorithm for this module.

Step 1 Construct the permutation matrix $P = \sigma_{k_1,k_2}I$.

Step 2 Construct diagonal matrices D_1 and D_2 as in the proof of Lemma 3.5.

Step 3 Construct the vector $\mathfrak{P} = (\mathfrak{P}_\ell : 0 \leqslant \ell \leqslant k_1k_2 - 1)$ by the formula:

$$\mathfrak{P}_\ell = \mathfrak{p}^{(i_1)}(\bar{t}^{k_2})\mathfrak{q}^{(j_1)}(\bar{t}^{k_1})$$

where i, j and ℓ are related by (4.3).

Step 4 From the entries of the vector $k_2D_1\mathfrak{P}^t + k_1D_2P\mathfrak{P}^t$, contruct the Smith vector $\mathrm{TSV}(k_1, k_2, \mathfrak{p}, \mathfrak{q})$ as in observation (5).

Example. Let us consider the example $k_1 = 2, k_2 = 3, \mathfrak{p} = \mathrm{SV}(2)$ and $\mathfrak{q} = \mathrm{SV}(3)$. Figure 7 illustrates the construction we carried out in this subsection.

$k_1 = 2$ $k_2 = 3$

Row Index		k_1			k_2		TSV($k_1,k_2,\mathfrak{p},\mathfrak{q}$)
5		2	$3(\bar{1}\oplus\bar{1}\oplus\bar{1}\oplus\bar{1})$		3	$2(\bar{1}\oplus\bar{1}\oplus\bar{1}\oplus\bar{1})$	$\bar{1}\oplus\bar{1}\oplus\bar{1}\oplus\bar{1}$
4		2	$3(\bar{1}\oplus\bar{1}\oplus\bar{t}^2\oplus\bar{t}^2)$		3	$2(\bar{1}\oplus\bar{1}\oplus\bar{0}\oplus\bar{0})$	$\bar{5}\oplus\bar{3}\oplus-3\bar{t}-\bar{1}\oplus3\bar{t}-\bar{3}$
3		2	$3(\bar{1}\oplus\bar{1}\oplus\bar{0}\oplus\bar{0})$			$2(\bar{1}\oplus\bar{0}\oplus\bar{t}^2\oplus\bar{0})$	$\bar{5}\oplus\bar{5}\oplus-2\bar{t}-\bar{2}\oplus2\bar{t}-\bar{2}$
2			$3(\bar{1}\oplus\bar{0}\oplus\bar{1}\oplus\bar{0})$			$2(\bar{1}\oplus\bar{1}\oplus\bar{t}^2\oplus\bar{t}^2)$	$\bar{5}\oplus\bar{0}\oplus-2\bar{t}+\bar{1}\oplus\bar{0}$
1			$3(\bar{1}\oplus\bar{0}\oplus\bar{t}^2\oplus\bar{0})$			$2(\bar{1}\oplus\bar{0}\oplus\bar{1}\oplus\bar{0})$	$\overline{13}\oplus\overline{10}\oplus-3\bar{t}-\bar{3}\oplus\bar{0}$
0			$3(\bar{1}\oplus\bar{0}\oplus\bar{0}\oplus\bar{0})$			$2(\bar{1}\oplus\bar{0}\oplus\bar{0}\oplus\bar{0})$	$\bar{1}\oplus\bar{0}\oplus\bar{0}\oplus\bar{0}$

Figure 7: Pictorial Construction for $k_1 = 2$ and $k_2 = 3$

4.3. $\boldsymbol{SV(n)}$

This algorithm is a recursion using the modules $\boldsymbol{SV(p^e)}$ and $\boldsymbol{TSV(k_1, k_2, \mathfrak{p}, \mathfrak{q})}$.

ALGORITHM 3.

Step 1 Factorise n as $p_1^{e_1} \dots p_r^{e_r}$. For each $p_i^{e_i}$, calculate $\mathrm{SV}(p_i^{e_i})$.

Step 2 Having calculated $\mathrm{SV}(p_1^{e_1} \dots p_i^{e_i})$, we calculate $\mathrm{SV}(p_1^{e_1} \dots p_{i+1}^{e_{i+1}})$ by the formula

$$\mathrm{SV}(p_1^{e_1} \dots p_{i+1}^{e_{i+1}}) = \mathrm{TSV}(p_1^{e_1} \dots p_i^{e_i}, p_{i+1}^{e_{i+1}}, \mathrm{SV}(p_1^{e_1} \dots p_i^{e_i}), SV(p_{i+1}^{e_{i+1}})).$$

4.4. Analysis of Algorithms

To calculate the number of bit operations needed to output $\mathrm{SV}(n)$, we compute the bit complexities of the modules

$$\boldsymbol{TSV(p_1^{e_1} \dots p_\ell^{e_\ell}, p_{\ell+1}^{e_{\ell+1}}\, \mathbf{SV}(p_1^{e_1} \dots p_\ell^{e_{\ell+1}}), \mathbf{SV}(p_{\ell+1}^{e_{\ell+1}}))}$$

for $l = 0, \dots, r-1$, where $n = p_1^{e_1} \dots p_r^{e_r}$ is the prime factorisation of n. We observe that the bit complexity in the case $l = r-1$ dominates and so the total bit complexity is atmost r times this complexity.

Firstly, we need the following definitions and some lemmas.

Recall that for a polynomial $a(X) = \sum_{i=0}^{n} a_i X^i$ with integer coefficients, its height $\mathrm{ht}(a)$ is defined to be $\max\{|a_i| : 0 \leqslant i \leqslant n\}$.

DEFINITION 4.3. Given a vector $(\mathfrak{p}^{(j)} : 0 \leqslant j \leqslant n-1)$ having entries in $\bigoplus_{d|n} \mathbf{Z}[X]/\Phi_d(X)$ such that $\mathfrak{p}^{(j)} = \bigoplus_{d|n} p_d^{(j)} \bmod \Phi_d(X)$ where $p_d^{(j)}$ is the unique representative of degree atmost $\phi(d) - 1$, we define its height to be:

$$\mathrm{ht}((\mathfrak{p}^{(j)} : 0 \leqslant j \leqslant n-1)) = \max\{\mathrm{ht}(p_d^{(j)}) : d \mid n, 0 \leqslant j \leqslant n-1\}.$$

NOTE 4.4. Given a positive integer n, denote the bit length of n by $\mathcal{B}(n)$. Let $\mu(n_1, n_2)$ denote the number of bit operations required to multiply a n_1-bit number with an n_2-bit number. We will also set $\mu(n) := \mu(n, n)$.

(i) We have $\mu(n_1, n_2) \leq \mu(n)$ where $n = \max(n_1, n_2)$. The standard multiplication would suggest $\mu(n) = O(n^2)$ while FFT-methods in Schönage-Strassen algorithm [15] indicate that $\mu(n) = O(n \log n \log\log n)$. This bound was later improved by Fürer [7]. In contrast, addition of a n_1-bit number to a n_2-bit number takes $O(n_1 + n_2)$ bit operations.

(ii) Recall that in Step 1 of the proof of Lemma 3.5, we calculate integers u and v such that $um^k + v(-1)^{k-1}n^k = 1$. By extended Euclidean algorithm, this computation takes $O(k^2 \log(n) \log(m))$ bit operations. Also, $\mathcal{B}(u)$ and $\mathcal{B}(v)$ are $O(k \max(\log(n), \log(m)))$ (see [19, Section 2.2]).

(iii) Let $a(X), b(X) \in \mathbf{Z}[X]$ of degree m and n with bit length of the heights equal to τ_1 and τ_2 respectively. The standard polynomial multiplication algorithm to multiply $a(X)$ and $b(X)$ takes $O(mn\mu(\tau_1, \tau_2))$ bit operations. However, FFT based algorithms improves this to $O(d \log(d)\mu(\tau_1 + \tau_2 + \log d))$ where $d = \max(m, n)$. For more details, see [17, Corollary 8.27] and [13, Lemma 17].

We begin with some lemmas:

LEMMA 4.5. *Let $a(X), b(X) \in \mathbf{Z}[X]$ be such that $b(X)$ is monic and $\deg(a) \geqslant \deg(b)$. Suppose that $\mathrm{ht}(a) = M_1$ and $\mathrm{ht}(b) = M_2$. Let $a(X) = b(X)q(X)+r(X)$ where $q(X)$ and $r(X)$ are quotient and remainder respectively. Then,*

$$\mathrm{ht}(r(X)) \leqslant M_1(1 + M_2)^{\deg(a)-\deg(b)+1}.$$

Therefore, we have that $\mathcal{B}(\mathrm{ht}(r(X)) = O(\log(M_1) + (\deg(a)) \log(M_2))$.

Proof. In each step of the Standard division algorithm, we subtract a multiple of $b(X)$ from $a(X)$ so as to reduce the degree of $a(X)$. Suppose that at the ith stage, we are left with a polynomial $a_i(X)$ of height h_i where $0 \leqslant i \leqslant \deg(a) - \deg(b) + 1$. Then, $h_i \leqslant h_{i-1} + M_2 h_{i-1}$ with $h_0 = M_1$. So, the height of $r(X)$ is atmost $M_1(1 + M_2)^{\deg(a)-\deg(b)+1}$. □

LEMMA 4.6 ([13, Theorem 21]). *Let $a(X), b(X) \in \mathbf{Z}[X]$ be such that $b(X)$ is monic of degree n and $\deg(a) \leqslant 2\deg(b)$. Set $\mathcal{B}(\mathrm{ht}(a)) = \tau_1$ and $\mathcal{B}(\mathrm{ht}(b)) = \tau_2$. Then, the number of bit operations required to compute the quotient and remainder on dividing $a(X)$ by $b(X)$ is $O(n \log^2(n)\mu(n\tau_2 + \tau_1))$.*

LEMMA 4.7. *Let $a(X), b(X) \in \mathbf{Z}[X]$ be of degree m and n respectively with $m \geqslant n$. Suppose also that b is monic. Set $\mathcal{B}(\mathrm{ht}(a)) = \tau_1$ and $\mathcal{B}(\mathrm{ht}(b)) = \tau_2$. Then, the number of bit operations required to compute the quotient and remainder on dividing $a(X)$ by $b(X)$ is $O(m \log^2(m)\mu(m\tau_2 + \tau_1))$.*

Proof. If $m \leqslant 2n$, then the assertion follows from the Lemma 4.6.

So, let $m > 2n$. Choose the least k_1 such that the polynomials $a(X)$ and $X^{k_1}b(X)$ satisfy the hypotheses of the Lemma 4.6, that is,

$$2(k_1 + n) \geqslant m \geqslant 2(k_1 - 1 + n).$$

Now, proceeding as in Lemma 4.6, we may obtain polynomials $q_1(X)$ and $r_1(X)$ with $\deg(r_1) < k_1 + n$ such that $a(X) = X^{k_1}b(X)q_1(X) + r_1(X)$. Now, if $\deg(r_1) < n$, then, r_1 is indeed the remainder on dividing $a(X)$ by $b(X)$. If not, we divide $r_1(X)$ by $X^{k_2}b(X)$ for k_2 chosen so that

$$2(k_2 + n) \geqslant \deg(r_1) \geqslant 2(k_2 + n - 1).$$

Continuing this way, we obtain the remainder $r(X)$ on dividing $a(X)$ by $b(X)$.

First note that $\deg(r_i) \leqslant k_i + n - 1$. Since $k_1 + n - 1 \leqslant \frac{m}{2}$, and $k_i + n - 1 \leqslant \frac{k_{i-1}+n-1}{2}$ for all $i \geqslant 2$, we get

$$\deg(r_i) \leqslant k_i + n - 1 \leqslant \frac{m}{2^i}.$$

Secondly, using Lemma 4.5, we have:

$$\begin{aligned}
\mathcal{B}(\mathrm{ht}(r_i)) &= O\bigg(\mathcal{B}(\mathrm{ht}(r_{i-1})) + (\deg(r_{i-1}) - (k_i + n) + 1)\tau_2\bigg) \\
&= O\bigg(\mathcal{B}(\mathrm{ht}(r_{i-1})) + \frac{m}{2^{i-1}}\tau_2\bigg) \\
&= O(\tau_2 m + \tau_1)
\end{aligned}$$

Using Lemma 4.6, the bit complexity in obtaining q_i and r_i is

$$O((n + k_i)\log^2(n + k_i)\mu((n + k_i)\tau_2 + m\tau_2 + \tau_1)).$$

Thus, the total bit complexity is:

$$\begin{aligned}
&O\left(\sum_{i\geqslant 1}(n + k_i)\log^2(n + k_i)\mu((n + k_i)\tau_2 + m\tau_2 + \tau_1)\right) \\
&= O\left(\sum_{i\geqslant 1}\frac{m}{2^i}\log^2(m)\mu(2m\tau_2 + \tau_1)\right) \\
&= O(m\log^2(m)\mu(m\tau_2 + \tau_1)).
\end{aligned}$$

This completes the proof. □

We need the following estimate for coefficients of the cyclotomic polynomials due to Bateman [3].

LEMMA 4.8 (Bateman). *The height of the cyclotomic polynomial Φ_n is $O(\exp(n^{C/\log\log n}))$ for some absolute constant C.*

Thus, $\mathcal{B}(\mathrm{ht}(\Phi_n))$ is $O(n^\epsilon)$ for any $\epsilon > 0$.

LEMMA 4.9. *The bit length of the height of* $\mathrm{SV}(n)$ *is* $O(n^{1+\epsilon})$ *for any* $\epsilon > 0$.

Proof. Suppose that the prime factorisation of n is $p_1^{e_1} \dots p_r^{e_r}$. We prove this result by induction on r. By Remark 4.1, we have that

$$\mathrm{ht}(\mathrm{SV}(p_1^{e_1})) = O(1),$$

and therefore, the bit length $\mathcal{B}(\mathrm{ht}(\mathrm{SV}(p_1^{e_1}))$ of the height is $O(1)$.

Let $\mathfrak{p} := \mathrm{SV}\left(\frac{n}{p_r^{e_r}}\right)$ and $\mathfrak{q} := \mathrm{SV}(p_r^{e_r})$. The bit length of the height of the polynomial $\mathfrak{p}_{d_1}^{(i)}(t^{p_r^{e_r}})\mathfrak{q}_{d_2}^{(j)}(t^{np_r^{-e_r}})$ (before reducing modulo $\Phi_{d_1 d_2}$) is

$$O\left(\log(\max_{j,d_2}(\deg(q_{d_2}^{(j)}))) + \mathcal{B}(\mathrm{ht}(\mathfrak{p}))\right)$$

since the coefficients in the product is given by convolution. By induction hypothesis for the height of $\mathfrak{p}$, this equals $O(\log p_r^{e_r} + (np_r^{-e_r})^{1+\epsilon})$ for any $\epsilon > 0$.

Using Lemmas 4.5 and 4.8, after reduction $\mathrm{mod}\Phi_{d_1 d_2}$, we get:

$$\begin{aligned}\mathcal{B}(\mathrm{ht}(\mathfrak{p}_{d_1}^{(i)}(\bar{t}^{p_r^{e_r}})\mathfrak{q}_{d_2}^{(j)}(\bar{t}^{np_r^{-e_r}}))) &= O(\log p_r^{e_r} + (np_r^{-e_r})^{1+\epsilon} + n \cdot n^\epsilon) \\ &= O(n^{1+\epsilon})\end{aligned}$$

Now, scaling the coefficients of $\mathfrak{p}$ and $\mathfrak{q}$ by $p_r^{e_r} D_1$ and $np_r^{-e_r} D_2 P$ as in Observation (5) of Section 4.2 contributes to an addition of atmost $O(n \log n)$ bits to the total height (see Note 4.4(ii)). This finishes the proof. □

Now, we are ready to calculate the bit complexity of each of these modules. We will use the soft-Oh notation $O^\sim(\cdot)$ which drops out polylogarithmic factors. For functions $f, g : \mathbf{R}^s \to \mathbf{R}$, we say that $f = O^\sim(g)$ if there is a constant $c > 0$ such that $f = O(g \log^c(g))$. We will let ϵ be an arbitrary positive real number.

***SV*(p^e).** Since W_{p^e} is given by a recursive formula, calculation of Smith vector for p^e takes $O(p^{2e})$ steps.

***TSV*$(k_1, k_2, \mathfrak{p}, \mathfrak{q})$.** We will calculate the bit complexity of this algorithm by going over each step of the algorithm:

- In Step 1, the algorithm needs $O((k_1 k_2)^2)$ steps.
- For Step 2, in determining the diagonal matrices D_1 and D_2 as in Lemma 3.5, we need the cycles of the permutation σ_{k_1,k_2}. The construction of $\sigma_{k_1 k_2}$ and its cycle decomposition takes $O((k_1 k_2)^2)$ bit operations. Suppose that the cycle lengths of σ are $c_1, \dots, c_s$ so that $\sum_{i=1}^s c_i = k_1 k_2$. By Note 4.4 (ii), the bit complexity of this step is

$$O\left((k_1 k_2)^2 + \sum_{i=1}^{s} c_i^2 \log(k_1) \log(k_2)\right) = O^\sim\left((k_1 k_2)^2\right).$$

- In Step 3, to construct $\mathfrak{P}_\ell$, we need to find $\mathfrak{p}_{d_1}^{(i_1)}(t^{k_2})\mathfrak{q}_{d_2}^{(j_1)}(t^{k_1}) \bmod \Phi_{d_1d_2}(t)$ for every $d_1 \mid k_1$ and $d_2 \mid k_2$. Since $\mathfrak{p}_{d_1}^{(i_1)}$ and $\mathfrak{q}_{d_2}^{(j_1)}$ has degree atmost $\phi(d_1)$ and $\phi(d_2)$ respectively, the standard polynomial multiplication (Note 4.4 (iii)) costs $O(\phi(d_1d_2)(\mu_h(\mathfrak{p},\mathfrak{q}) + \log(\phi(d_1d_2))))$ bit operations where

$$\mu_h(\mathfrak{p},\mathfrak{q}) := \mu(\mathcal{B}(\mathrm{ht}(\mathfrak{p})), \mathcal{B}(\mathrm{ht}(\mathfrak{q}))).$$

To reduce modulo $\Phi_{d_1d_2}(t)$, using Lemma 4.7 and Lemma 4.8 we need $O^\sim\big(\beta(d_1,d_2)^2(d_1d_2)^\epsilon + \beta(d_1,d_2)\mu_h(\mathfrak{p},\mathfrak{q})\big)$ bit operations where $\beta(d_1,d_2) = k_1\phi(d_2) + k_2\phi(d_1)$. Since $\phi(d_1d_2) \leqslant \beta(d_1,d_2)$, the total bit complexity involved in finding $\mathfrak{p}_{d_1}^{(i_1)}(t^{k_2})\mathfrak{q}_{d_2}^{(j_1)}(t^{k_1}) \bmod \Phi_{d_1d_2}(t)$ is

$$O^\sim\big(\beta(d_1,d_2)^2(d_1d_2)^\epsilon + \beta(d_1,d_2)\mu_h(\mathfrak{p},\mathfrak{q})\big).$$

Now, we have:

$$\begin{aligned}\sum_{\substack{d_1\mid k_1\\ d_2\mid k_2}} \beta(d_1,d_2) &= k_1k_2(\tau(k_1)+\tau(k_2))\\ &= O((k_1k_2)^{1+\epsilon})\end{aligned}$$

since $\tau(n)$ (the number of divisors of n) is $O(n^{1+\epsilon})$ ([8, Theorem 315]). Also

$$\sum_{\substack{d_1\mid k_1\\ d_2\mid k_2}} \beta(d_1,d_2)^2 \leqslant \Big(\sum_{\substack{d_1\mid k_1\\ d_2\mid k_2}} \beta(d_1,d_2)\Big)^2.$$

Thus, the bit complexity to compute $\mathfrak{P}_\ell$ is:

$$\begin{aligned}&O^\sim\Big(\sum_{\substack{d_1\mid k_1\\ d_2\mid k_2}} \beta(d_1,d_2)^2(d_1d_2)^\epsilon + \beta(d_1,d_2)\mu_h(\mathfrak{p},\mathfrak{q})\Big)\\ &= O^\sim\Big((k_1k_2)^\epsilon \sum_{\substack{d_1\mid k_1\\ d_2\mid k_2}} \beta(d_1,d_2)^2 + \mu_h(\mathfrak{p},\mathfrak{q}) \sum_{\substack{d_1\mid k_1\\ d_2\mid k_2}} \beta(d_1,d_2)\Big)\\ &= O^\sim\Big((k_1k_2)^{2+\epsilon} + \mu_h(\mathfrak{p},\mathfrak{q})(k_1k_2)^{1+\epsilon}\Big)\end{aligned}$$

So the total bit complexity in Step 3 is

$$O^\sim((k_1k_2)^{3+\epsilon} + \mu_h(\mathfrak{p},\mathfrak{q})(k_1k_2)^{2+\epsilon}).$$

- Using the facts in Note 4.4 (ii), one may prove that the bit lengths of the entries of the matrices k_2D_1 and k_1D_2 are $O^\sim(k_1k_2)$. Since there are $(k_1k_2)^2$ terms in $\vec{\mathfrak{P}}$, calculating $(k_2D_1 + k_1D_2P)\mathfrak{P}^t$, which is the Step 4, needs $O^\sim((k_1k_2)^{2+\epsilon}\mu_h(\mathfrak{p},\mathfrak{q}) + (k_1k_2)^{3+\epsilon})$ bit operations.

Summing the bit complexities of each step, we conclude that the bit complexity of this module is

$$O^{\sim}((k_1k_2)^{2+\epsilon}\mu_h(\mathfrak{p},\mathfrak{q}) + (k_1k_2)^{3+\epsilon}) \text{ for any } \epsilon > 0. \tag{4.4}$$

SV(n). The bit complexity of Step 1 is $O(n^2)$. For Step 2, the maximum bit operations are required in the last step of the recursion. We know that the bit complexity of module $\boldsymbol{TSV(\frac{n}{p_r^{e_r}}, p_r^{e_r}, \mathrm{SV}(\frac{n}{p_r^{e_r}}), \mathrm{SV}(p_r^{e_r}))}$ is:

$$\begin{aligned}
& O^{\sim}\left(n^{2+\epsilon}\mu_h\Big(\mathrm{SV}\,(\frac{n}{p_r^{e_r}}), \mathrm{SV}(p_r^{e_r})\Big) + n^{3+\epsilon}\right) && \text{using (4.4)} \\
&= O^{\sim}\left(n^{2+\epsilon}\mu\Big(\frac{n}{p_r^{e_r}}, p_r^{e_r}\Big) + n^{3+\epsilon}\right) && \text{by Lemma 4.9} \\
&= O^{\sim}(n^{3+\epsilon}) && (4.5)
\end{aligned}$$

Since $r = O(\log(n))$, the bit complexity of the module $\boldsymbol{SV(n)}$ is $O(n^{3+\epsilon})$ for any $\epsilon > 0$.

4.4.1. Space Complexity

Note that the bit length of an integer that appears while executing the module $\boldsymbol{SV(n)}$ is $O(n^{1+\epsilon})$ for any $\epsilon > 0$. The maximum space is needed to store a Smith vector for n. Since a Smith vector has $O(n^2)$ terms, the space complexity of $\mathrm{SV}(n)$ is $O(n^{3+\epsilon})$ for any $\epsilon > 0$.

We summarise the above discussion in the following theorem:

THEOREM 4.10. *Given a positive integer n, the algorithm $\boldsymbol{SV(n)}$ gives a Smith vector for n. The bit complexity and space complexity of this algorithm are both $O(n^{3+\epsilon})$ for any $\epsilon > 0$.*

We may compare this theorem with the results in the literature in this direction. It appears to us that the best known algorithm for determining Smith normal form $S(A)$ of an integer matrix A and the unimodular transforming matrices U and V are due to Arne Storjohann in his PhD thesis [16]. Let $O(n^\theta)$ be the algebraic complexity involved in multiplying two $n \times n$ matrices with integer entries; best known algorithms give $2 < \theta \leqslant 3$ (for example, V. Vassilevska Williams gives an algorithm with $\theta = 2.373$ in [18]). He proves:

THEOREM ([16, Proposition 7.20]). *For a $n \times m$ matrix $A = (a_{ij})$ of rank r with integer entries, the Smith normal form $S(A)$ and the unimodular transforming matrices U and V may be obtained in $O^{\sim}(nmr^{\theta-1}\log\|A\| + nm\mu(r\log\|A\|))$ bit operations where $\|A\| = \max_{i,j}|a_{ij}|$.*

Specialising to our case, it is seen that Storjohann's algorithm would require $O(n^{2+\theta+\epsilon})$ bit operations. Here we substituted $\log\|A\| = O(n^{1+\epsilon})$ due to Lemma 4.8 and Lemma 4.5. Thus, our algorithm is an improvement to this best known algorithm in the special case we are interested in.

A. Determinant of A_n

We now calculate the determinant of the matrix A_n in terms of resultants of cyclotomic polynomials. The advantage of this new approach is the fact that this generalises to any monic polynomial f over a unique factorisation domain (UFD) and any of its factorisations into pairwise relatively prime polynomials. Moreover, this approach determines the sign of $\det(A_n)$ unambigously.

To calculate the determinant of A_n, we propose the following simplification:

A.1. Simplification

Let Ω_n be the cyclic group of all nth roots of unity. The cyclotomic polynomials $\{\Phi_d\}_{d|n}$ factorise over the ring $\mathbf{Z}[\Omega_n]$ of cyclotomic integers. We will see that calculating $\det(A_n)$ becomes very simple when we work over the ring of cyclotomic integers. Now, we shall argue that passing to $\mathbf{Z}[\Omega_n]$ does not affect the computation.

To see this, consider the following diagram of maps:

$$\begin{array}{ccccc}
\frac{\mathbf{Z}[X]}{(X^n-1)} & \xhookrightarrow{\iota_1} & \frac{\mathbf{Z}[\Omega_n][X]}{(X^n-1)} & & \\
\Big\downarrow{\scriptstyle \Psi_n} & & \Big\downarrow{\scriptstyle \widetilde{\Psi}_n} & \searrow^{\rho_2} & \\
\bigoplus_{d|n} \frac{\mathbf{Z}[X]}{\Phi_d(X)} & \xhookrightarrow{\iota_2} & \bigoplus_{d|n} \frac{\mathbf{Z}[\Omega_n][X]}{(\Phi_d(X))} & \xrightarrow{\rho_1} & \bigoplus_{\omega\in\Omega_n} \frac{\mathbf{Z}[\Omega_n][X]}{(X-\omega)}
\end{array}$$

where $\widetilde{\Psi}_n$ is the $\mathbf{Z}[\Omega_n]$-linear extension of Ψ_n, ρ_1 is the canonical quotient map, ρ_2 is the composition $\rho_1 \circ \widetilde{\Psi}_n$ and ι_1, ι_2 are canonical inclusions. The matrix of $\widetilde{\Psi}_n$ with respect to $(1, \overline{X}, \ldots, \overline{X}^{n-1})$ as a basis for $\mathbf{Z}[\Omega_n][X]/(X^n-1)$ and $(1, \overline{X}, \ldots, \overline{X}^{\phi(d)-1})$ as a basis for $\mathbf{Z}[\Omega_n][X]/\Phi_d(X)$ is equal to A_n. So we have:

$$\det(A_n) = \det(\widetilde{\Psi}_n) = \frac{\det(\rho_2)}{\det(\rho_1)}. \tag{A.1}$$

However, to write the matrix of the $\mathbf{Z}[\Omega_n]$-linear maps ρ_1 and ρ_2, we need an ordered basis for the codomain of ρ_1 which is the same as that of ρ_2. So, we may fix any total order $\prec$ on Ω_n so that if d_1 and d_2 are two divisors of n and $d_1 < d_2$, then, all the primitive d_1th roots of unity precede all the primitive d_2th roots of unity in the order $\prec$. Say,

$$\Omega_n = \coprod_{d|n} \{\omega_{d,1}, \ldots, \omega_{d,\phi(d)}\} = \{\omega_{d,j} : d \mid n, 1 \leqslant j \leqslant \phi(d)\} \tag{A.2}$$

where $\omega_{d,j}$ is a primitive dth root of unity ($1 \leqslant j \leqslant \phi(d)$) and Ω_n is ordered lexicographically.

The rest of the calculation will determine $\det(\rho_1)$ and $\det(\rho_2)$. The matrix of ρ_1 with respect to the chosen basis is a block matrix $\operatorname{diag}(\mathcal{D}_1, \ldots, \mathcal{D}_d, \ldots, \mathcal{D}_n)_{d|n}$

where $\mathcal{D}_d$ is a $\phi(d) \times \phi(d)$ matrix of the following form:

$$\mathcal{D}_d = \begin{pmatrix} 1 & \omega_{d,1} & \omega_{d,1}^2 & \cdots & \omega_{d,1}^{\phi(d)-1} \\ 1 & \omega_{d,2} & \omega_{d,2}^2 & \cdots & \omega_{d,2}^{\phi(d)-1} \\ 1 & \omega_{d,3} & \omega_{d,3}^2 & \cdots & \omega_{d,3}^{\phi(d)-1} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & \omega_{d,\phi(d)} & \omega_{d,\phi(d)}^2 & \cdots & \omega_{d,\phi(d)}^{\phi(d)-1} \end{pmatrix} \tag{A.3}$$

where $\omega_{d,1} \prec \cdots \prec \omega_{d,\phi(d)}$ are primitive dth roots of unity.

The matrix $[\rho_2]$ of ρ_2 with respect to the chosen basis is a block column matrix

$$(R_1, \ldots, R_d, \ldots, R_n)^t_{d|n}$$

where R_d is the following matrix:

$$R_d = \begin{pmatrix} 1 & \omega_{d,1} & \ldots & \omega_{d,1}^{\ell} & \ldots & \omega_{d,1}^{n-1} \\ \vdots & \vdots & & \vdots & & \vdots \\ 1 & \omega_{d,j} & \ldots & \omega_{d,j}^{\ell} & \ldots & \omega_{d,j}^{n-1} \\ \vdots & \vdots & & \vdots & & \vdots \\ 1 & \omega_{d,\phi(d)} & \ldots & \omega_{d,\phi(d)}^{\ell} & \ldots & \omega_{d,\phi(d)}^{n-1} \end{pmatrix} \tag{A.4}$$

Since the matrices $\mathcal{D}_d$ and $[\rho_2]$ are Vandermonde matrices, we have at once that:

$$\begin{aligned} \det(\rho_1) &= \prod_{d|n} \det(\mathcal{D}_d) \\ &= \prod_{d|n} \prod_{1 \leqslant i < j \leqslant \phi(d)} (\omega_{d,j} - \omega_{d,i}) \qquad \text{(A.5)} \\ \det(\rho_2) &= \prod_{\substack{d_1, d_2 | n \\ 1 \leqslant d_1 < d_2 \leqslant n}} \prod_{\substack{1 \leqslant i \leqslant \phi(d_1) \\ 1 \leqslant j \leqslant \phi(d_2)}} (\omega_{d_2,j} - \omega_{d_1,i}) \prod_{d|n} \prod_{1 \leqslant i < j \leqslant \phi(d)} (\omega_{d,j} - \omega_{d,i}) \qquad \text{(A.6)} \end{aligned}$$

Now, from (A.1), we see that:

$$\begin{aligned} \det(A_n) &= \prod_{\substack{d_1, d_2 | n \\ 1 \leqslant d_1 < d_2 \leqslant n}} \left(\prod_{\substack{1 \leqslant i \leqslant \phi(d_1) \\ 1 \leqslant j \leqslant \phi(d_2)}} (\omega_{d_2,j} - \omega_{d_1,i}) \right) \qquad \text{(A.7)} \\ &= \prod_{\substack{d_1, d_2 | n \\ 1 \leqslant d_1 < d_2 \leqslant n}} \mathcal{R}(\Phi_{d_2}, \Phi_{d_1}) \qquad \text{(A.8)} \end{aligned}$$

where $\mathcal{R}(f, g)$ is the resultant of the polynomials f and g (for definition and basic properties of resultants of polynomials, see [14, Chapter 1, Section 3]). The resultant of pairs of cyclotomic polynomials first appears in print in the

work of Diederichsen [5, §3, Hilfssatz 2] on integral representations of cyclic groups. We also refer to Apostol [1] and Dresden [6] for alternative proofs. The following result will be used to finish off the computation:

THEOREM A.1 (Diederichsen). *Let m and n be positive integers.*

1. $\mathcal{R}(\Phi_m, \Phi_n) = 0$ *if and only if* $m = n$.

Assume now that $m > n$.

2. *If* $n = 1$, *then,*

$$\mathcal{R}(\Phi_m, \Phi_n) = \begin{cases} (-1)^{\phi(m)} p & \text{if } m = p^\alpha \text{ for some } \alpha > 0 \\ (-1)^{\phi(m)} & \text{otherwise} \end{cases} \tag{A.9}$$

3. *If* $n > 1$ *and* $\gcd(m, n) = 1$, *then,* $\mathcal{R}(\Phi_m, \Phi_n) = 1$.

4. *If* $n > 1$ *and* $\gcd(m, n) > 1$, *then,*

$$\mathcal{R}(\Phi_m, \Phi_n) = \begin{cases} p^{\phi(m)} & \text{if } \frac{m}{n} = p^\alpha \text{ for some } \alpha > 0 \\ 1 & \text{otherwise} \end{cases} \tag{A.10}$$

We now start from (A.8) and use Theorem A.1 to get a closed form expression for $\det(A_n)$ in terms of the prime factorisation of n, say, $n = p_1^{\alpha_1} \cdots p_r^{\alpha_r}$.

We first rewrite (A.8) as follows:

$$\prod_{\substack{1 \leqslant d_1 < d_2 \leqslant n \\ d_1, d_2 \mid n}} \mathcal{R}(\Phi_{d_2}, \Phi_{d_1}) = \prod_{\substack{d \mid n \\ d \neq 1}} \mathcal{R}(\Phi_d, X - 1) \prod_{\substack{1 < d_1 < d_2 \leqslant n \\ d_1, d_2 \mid n}} \mathcal{R}(\Phi_{d_2}, \Phi_{d_1}) \tag{A.11}$$

As a consequence of (A.9), the first product on the right hand side of (A.11) evaluates to:

$$\prod_{\substack{d \mid n \\ d \neq 1}} \mathcal{R}(\Phi_d, X - 1) = \prod_{\substack{d \mid n \\ d \neq 1}} (-1)^{\phi(d)} \Phi_d(1) = (-1)^{n-1} n. \tag{A.12}$$

For evaluating the second product, by Theorem A.1, note that a pair (d_1, d_2) of divisors of n contributes to the product if and only if $d_1 \neq 1$ and the ratio $\frac{d_2}{d_1}$ is a prime power, say $p_i^{e_i}$ for some $1 \leqslant i \leqslant r$ and $1 \leqslant e_i \leqslant \alpha_i$; also each such pair contributes $p_i^{\phi(d_2)}$ to the product. For a fixed prime p_i $(1 \leqslant i \leqslant r)$ and exponent e_i $(1 \leqslant e_i \leqslant \alpha_i)$, every divisor d_1 of $\frac{n}{p_i^{e_i}}$ with $d_1 \neq 1$ determines a contributing pair (d_1, d_2) of divisors and conversely. Therefore:

$$\prod_{\substack{1 < d_1 < d_2 \leqslant n \\ d_1, d_2 \mid n}} \mathcal{R}(\Phi_{d_2}, \Phi_{d_1}) = \prod_{i=1}^{r} \prod_{e_i=1}^{\alpha_i} \prod_{\substack{d \mid n p_i^{-e_i} \\ d \neq 1}} p_i^{\phi(d)} \tag{A.13}$$

$$= \frac{1}{n} \prod_{i=1}^{r} p_i^{n \sum_{e_i=1}^{\alpha_i} p_i^{-e_i}} \tag{A.14}$$

This finishes the computation and we now have:

THEOREM A.2. *For a positive integer n, we have:*

$$\det(A_n) = \prod_{\substack{1 \leqslant d_1 < d_2 \leqslant n \\ d_1, d_2 | n}} \mathcal{R}(\Phi_{d_2}, \Phi_{d_1}) \tag{A.15}$$

$$= (-1)^{n-1} \prod_{i=1}^{r} p_i^{\frac{n(1-p_i^{-\alpha_i})}{(p_i-1)}}. \tag{A.16}$$

Following the computations done before Theorem A.1, we may prove:

THEOREM A.3. *Suppose that f is a monic polynomial over a UFD and*

$$f = \prod_{i=1}^{n} f_i$$

is a factorisation of f into pairwise relatively prime polynomials. Then, the determinant of the canonical map Ψ_f written with respect to the standard basis is:

$$\det(\Psi_f) = \prod_{1 \leqslant i < j \leqslant n} \mathcal{R}(f_j, f_i). \tag{A.17}$$

REMARK A.4. From the above theorem, the resultant $\mathcal{R}(f_2, f_1)$ is the determinant of the map $\Psi_{f_1 f_2}$ written with respect to the standard basis. Specialising to $f_1 = \Phi_m(X)$ and $f_2 = \Phi_n(X)$ with $m > n$, we have the following exact sequence:

$$0 \longrightarrow \frac{\mathbf{Z}[X]}{\Phi_m(X)\Phi_n(X)} \longrightarrow \frac{\mathbf{Z}[X]}{\Phi_m(X)} \oplus \frac{\mathbf{Z}[X]}{\Phi_n(X)} \longrightarrow G(\Phi_m(X)\Phi_n(X)) \longrightarrow 0$$

Now, we have $|G(f)| = |\mathcal{R}(\Phi_m, \Phi_n)|$. Also, the intersection of the ideal generated by Φ_m and Φ_n with $\mathbf{Z}$ is given by ([6, Theorem 2]):

$$\langle \Phi_m(X), \Phi_n(X) \rangle \cap \mathbf{Z} = \begin{cases} p\mathbf{Z} & \text{if } \frac{m}{n} = p^\alpha \text{ for some } \alpha > 0 \\ \mathbf{Z} & \text{otherwise} \end{cases} \tag{A.18}$$

Now, setting $r = \phi(m) + \phi(n)$, if $(e_0, \ldots, e_{r-1})$ are elementary divisors of $\Psi_{\Phi_m \Phi_n}$, then, $e_i \mid p$ and $\prod e_i = |\mathcal{R}(\Phi_m, \Phi_n)|$. If $\frac{m}{n}$ is not power of a prime, then, $e_i = 1$ for all i. If $\frac{m}{n}$ is power of a prime p, the elementary divisors are given by $e_i = 1$ for $0 \leqslant i \leqslant \phi(n) - 1$ and $e_j = p$ for $\phi(n) \leqslant j \leqslant \phi(m) + \phi(n) - 1$.

REMARK A.5. Let S be a UFD. Then, $R = S[x_1, \ldots, x_n]$ is a UFD and considering the polynomial $f(X) = \prod_i (X - x_i)$ in Theorem A.3, we see that Vandermonde determinant falls out as a special case. The elementary divisors of a Vandermonde matrix over a Dedekind domain has been calculated by M. Bhargava [4, Lemma 2].

Acknowledgements

The authors would like to thank the Institute of Mathematical Sciences, Chennai and Indian Statistical Institute, Bangalore, where various parts of this work was carried out, for their hospitality and support. The authors thank Amritanshu Prasad for suggesting numerous improvements. The authors are grateful to Vikram Sharma for guiding them with references for the write-up in Section 4.4.

References

[1] Tom M. Apostol. Resultants of cyclotomic polynomials. *Proceedings of the American Mathematical Society*, 24(3):457–462, March 1970.

[2] Raymond G. Ayoub and Christine Ayoub. On the group ring of a finite abelian group. *Bull. Austral. Math. Soc.*, 1:245–261, 1969.

[3] P. T. Bateman. Note on the coefficients of the cyclotomic polynomial. *Bull. Amer. Math. Soc.*, 55:1180–1181, 1949.

[4] Manjul Bhargava. Generalized factorials and fixed divisors over subsets of a dedekind domain. *Journal of Number Theory*, 72(1):62–75, 1998.

[5] Fritz-Erdmann Diederichsen. Über die Ausreduktion ganzzahliger Gruppendarstellungen bei arithmetischer Äquivalenz. *Abh. Math. Sem. Univ. Hamburg*, 13(1):357–412, 1939.

[6] Gregory Dresden. Resultants of cyclotomic polynomials. *Rocky Mountain J. Math.*, 42(5):1461–1469, 2012.

[7] Martin Fürer. Faster integer multiplication. In *Proceedings of the Thirty-ninth Annual ACM Symposium on Theory of Computing*, STOC '07, pages 57–66, New York, NY, USA, 2007. ACM.

[8] Godfrey Harold Hardy and Edward Maitland Wright. *An Introduction to the Theory of Numbers*. Oxford science publications, illustrated, reprint edition, 1979.

[9] Nathan Jacobson. *Basic Algebra*, volume 1. W. H. Freeman and Company, 2nd edition, 1985.

[10] Serge Lang. *Algebra*, volume 211 of *Graduate Texts in Mathematics*. Springer-Verlag, third edition, 2002.

[11] Andrew D. Loveless. The general GCD-product function. *Integers*, 6, July 2006. #A19.

[12] Kamalakshya Mahatab. A problem about determinant of sum of permutation matrices. MathOverflow. URL:http://mathoverflow.net/q/137180 (version: 2013-07-20).

[13] Victor Y. Pan and Elias P. Tsigaridas. Nearly optimal refinement of real roots of a univariate polynomial. Available from: http://hal.inria.fr/hal-00960896, February 2014.

[14] Victor V. Prasolov. *Polynomials*, volume 11 of *Algorithms and Computations in Mathematics*. Springer-Verlag, Berlin, 2004. Translated from the 2001 Russian second edition by Dimitry Leites.

[15] A Schöhage and V Strassen. Schnelle multiplikation großer zahlen. *Computing*, 7(3 - 4):281 – 292, 1971.

[16] Arne Storjohann. *Algorithms for Matrix Canonical Forms*. PhD thesis, Swiss Federal Institute of Technology, Zürich, Switzerland, 2000.

[17] Joachim Von Zur Gathen and Jürgen Gerhard. *Modern Computer Algebra*. Cambridge University Press, 1st edition, 1999.

[18] Virginia Vassilevska Williams. Multiplying matrices faster than coppersmith-winograd. In *Proceedings of the 44th Symposium on Theory of Computing*, STOC '12, pages 887–898, New York, NY, USA, 2012. ACM.

[19] Chee-Keng Yap. *Fundamental Problems in Algorithmic Algebra*. Oxford University Press, 2000.