The discrete Fourier uncertainty principle

by

Marcel K. Goh

6 October 2022

1. Introduction

Let Z be a finite abelian group. A *character* on Z is a homomorphism from Z to the multiplicative group $\mathbf{C} \setminus \{0\}$. It is easily seen that $|\chi(x)|$ must equal 1 for all $x \in Z$. The set of characters forms a group, which we shall denote by $\widehat{Z}$. This is the Pontryagin dual of Z. Letting $\mathbf{Z}_n$ be the n-element cyclic group, if $Z = \mathbf{Z}_{n_1} \times \mathbf{Z}_{n_2} \times \cdots \times \mathbf{Z}_{n_r}$, then for every $u = (u_1, \ldots, u_r) \in Z$ the function $\chi_u : Z \to \mathbf{C}$ given by

$$\chi_u(x_1, \ldots, x_r) = \prod_{i=1}^{r} \exp\biggl(\frac{2\pi i u_i x_i}{n_i}\biggr)$$

is a character, and in fact the map $u \mapsto \chi_u$ gives an isomorphism of groups from Z to $\widehat{Z}$.

The space of functions from Z to $\mathbf{C}$ can be made into an inner product space by setting

$$\langle f, g \rangle = \mathbf{E}_{x\in Z}\, f(x)\overline{g(x)},$$

where $\mathbf{E}_{x\in Z}\, F(x) = |Z|^{-1} \sum_{x\in Z} F(x)$, and likewise we define an inner product on the space of functions from $\widehat{Z}$ to $\mathbf{C}$ by putting

$$\langle \widehat{f}, \widehat{g} \rangle = \sum_{\chi \in \widehat{Z}} \widehat{f}(\chi)\overline{\widehat{g}(\chi)}.$$

For $f : Z \to \mathbf{C}$, the *Fourier transform* of f is the function $\widehat{f} : \widehat{Z} \to \mathbf{C}$ given by

$$\widehat{f}(\chi) = \langle f, \chi \rangle = \mathbf{E}_{x\in Z}\, f(x)\overline{\chi(x)}.$$

Of course, we can associate to any $\alpha \in Z$ the character $\chi_\alpha \in \widehat{Z}$, so we may write $\widehat{f}(\alpha)$ to mean $\widehat{f}(\chi_\alpha)$, and this is called the *Fourier coefficient of f at α.*

It is not difficult to prove that any two distinct characters are orthogonal in the space of functions from Z to $\mathbf{C}$. Furthermore, for any $x \in Z$ we can define a function $F_x : \widehat{Z} \to \mathbf{C}$ by $F_x(\chi) = \chi(x)$, and one can similarly show that if $x \neq y$, then $\langle F_x, F_y \rangle = 0$. So since both of the vector spaces $\mathbf{C}^Z$ and $\mathbf{C}^{\widehat{Z}}$ have dimension n, we have found orthogonal bases for these spaces, namely $\{\chi_u : u \in Z\}$ and $\{F_x : x \in Z\}$ respectively.

We have the following important formulas, whose proofs can be found in any book on Fourier analysis.

Theorem P (*Parseval–Plancherel identity*). *Let Z be a finite abelian group and let $f, g : Z \to \mathbf{C}$. If $\widehat{f}$ and $\widehat{g}$ are the Fourier transforms of f and g respectively, then $\langle f, g\rangle = \langle \widehat{f}, \widehat{g}\rangle$.* ∎

Theorem I (*Fourier inversion formula*). *Let Z be a finite abelian group and let $f : Z \to \mathbf{C}$. Then*

$$f(x) = \sum_{\chi \in \widehat{Z}} \widehat{f}(\chi)\chi(x). \quad ∎$$

Recall also the *Cauchy–Schwarz inequality*, which wears many disguises but in our context says that

$$\Big(\sum_{x\in Z} |f(x)| \cdot |g(x)|\Big)^2 \le \Big(\sum_{x\in Z} |f(x)|^2\Big)\Big(\sum_{x\in Z} |g(x)|^2\Big)$$

for all $f, g : Z \to \mathbf{C}$.

2. The uncertainty principle

The *support* of a function $f : Z \to \mathbf{C}$ is the set $\{x \in Z : f(x) \neq 0\}$. We will write $\|f\|_0$ for the size $|\operatorname{supp}(f)|$ of the support, and it is also convenient to write $\|f\|_\infty$ for the quantity $\max_{x\in Z} |f(x)|$. (These are defined analogously for functions on $\widehat{Z}$.) The uncertainty principle states that the support of $f : Z \to \mathbf{C}$ and the support of its Fourier transform $\widehat{f} : \widehat{Z} \to \mathbf{C}$ cannot both be small. We will make this fact quantitative very soon. First off, let us prove a lemma.

Lemma 1. *Let f be a function from an abelian group Z to $\mathbf{C}$ and let $\widehat{f}$ be its Fourier transform. Then*

$$\|\widehat{f}\|_\infty \le \mathbf{E}_{x\in Z}\, |f(x)|.$$

Proof. Let $\chi \in \widehat{Z}$ be given. We have, by the definition of Fourier transform and the triangle inequality,

$$|\widehat{f}(\chi)| = \Big|\mathbf{E}_{x\in Z}\, f(x)\overline{\chi(x)}\Big| \le \mathbf{E}_{x\in Z}\big|f(x)\overline{\chi(x)}\big|,$$

but since $|\chi(x)| = 1$ for all x, this is exactly the right-hand side of the lemma statement and we are done since χ was arbitrary. ∎

We now state and prove the Fourier uncertainty principle.

Theorem 2 (*Fourier uncertainty principle*). *Let Z be a finite abelian group and $\widehat{Z}$ be its dual. If $f : Z \to \mathbf{C}$ is not identically zero and $\widehat{f} : \widehat{Z} \to \mathbf{C}$ is its Fourier transform, then*

$$\|f\|_0 \cdot \|\widehat{f}\|_0 \ge |Z|.$$

Proof. By the previous lemma and the definition of the support,

$$\|\widehat{f}\|_\infty \le \mathbf{E}_{x\in Z}\, |f(x)| = \frac{1}{|Z|}\sum_{x\in Z} |f(x)| = \frac{1}{|Z|}\sum_{x\in \operatorname{supp}(f)} |f(x)|.$$

We then use the Cauchy–Schwarz inequality to obtain

$$\sum_{x\in\operatorname{supp}(f)} |f(x)| \le \sqrt{\sum_{x\in\operatorname{supp}(f)} |f(x)|}\sqrt{\sum_{x\in\operatorname{supp}(f)} 1^2} = \sqrt{\|f\|_0 \sum_{x\in Z} |f(x)|^2},$$

and so far we have shown that

$$\|\widehat{f}\|_\infty \le \frac{1}{|Z|}\sqrt{\|f\|_0 \sum_{x\in Z} |f(x)|^2}.$$

But by the Parseval–Plancherel identity, we have

$$\sum_{x\in Z} |f(x)|^2 = |Z| \sum_{\chi\in\widehat{Z}} |\widehat{f}(\chi)|^2 \le |Z|\cdot\|\widehat{f}\|_0\cdot\|\widehat{f}\|_\infty^2,$$

and plugging this in above, we have

$$\|\widehat{f}\|_\infty \le \|\widehat{f}\|_\infty \sqrt{\frac{\|f\|_0\cdot\|\widehat{f}\|_0}{|Z|}}.$$

Since f is not the zero function, we can divide both sides by $\|\widehat{f}\|_\infty$, square the inequality, then rearrange to get the theorem statement. ▮

It can be shown that we have equality above if and only if f is (some multiple of) the characteristic function of a coset of a subgroup of Z.

So far so good, but for $Z = \mathbf{Z}_p$ a much stronger uncertainty principle holds, and the rest of these notes will be dedicated to establishing the algebraic machinery needed to prove it.

3. Cyclotomic polynomials

Let n be a positve integer. An *nth root of unity* is any complex number ω such that $\omega^n = 1$. Note that if d divides n, then any ω with $\omega^d = 1$ also satisfies $\omega^n = 1$, so in some sense this number should be associated to d and not n. An nth root of unity is called *primitive* if it is not an mth root of unity for any $1 \le m < n$. (Thus any nth root of unity is a primitive dth root of unity for exactly one d dividing n.) The *nth cyclotomic polynomial*, which we shall denote by Φ_n, is given by

$$\Phi_n(z) = \prod_\omega (z-\omega),$$

where in the product, ω runs over the primitive nth roots of unity. As some small examples, we have $\Phi_1(z) = z-1$, $\Phi_2(z) = z+1$, $\Phi_3(z) = z^2+z+1$, and $\Phi_4(z) = z^2+1$. Observe that so far, all the coefficients have been integers, a fact which is not obvious from the definition but can be shown by induction (and indeed we shall).

In the proof of the next lemma we will also require the *von Mangoldt function* $\Lambda(n)$, which is defined on positive integers by the rule

$$\Lambda(n) = \begin{cases} \log p, & \text{if } n = p^k \text{ for some prime } p \text{ and some integer } k \geq 1; \\ 0, & \text{otherwise.} \end{cases}$$

By the fundamental theorem of arithmetic, any integer n can be factored into $n = {p_1}^{e_1}{p_2}^{e_2} \cdots {p_s}^{e_s}$, and taking logarithms of both sides we see that

$$\log n = \sum_{i=1}^{s} e_i \log p_i = \sum_{d \backslash n} \Lambda(d).$$

Lemma 3. *Let $n \geq 1$. The nth cyclotomic polynomial Φ_n is monic with integer coefficients and we have*

$$\Phi_n(1) = \begin{cases} 0, & \text{if } n = 1; \\ p, & \text{if } n = p^k \text{ for some integer } k \geq 1; \\ 1, & \text{otherwise.} \end{cases}$$

Proof. Let Ω_n be the set of all nth roots of unity, primitive or not. Then the polynomial $z^n - 1$ factors as

$$z^n - 1 = \prod_{\omega \in \Omega_n} (z - \omega).$$

Now since every nth root of unity is a primitive dth root of unity for exactly one d dividing n, we can group roots together and write

$$z^n - 1 = \prod_{d \backslash n,} \Phi_d(z).$$

Let us prove the formula for $\Phi_n(1)$ first. Of course, $\Phi_1(1) = 1 - 1 = 0$. Then for $n > 1$,

$$\frac{z^n - 1}{\Phi_1(z)} = \lim_{z \to 1} \frac{z^n - 1}{z - 1} = \lim_{z \to 1} \frac{nz^{n-1}}{1} = n,$$

giving us the formula

$$n = \prod_{d \backslash n,\, d > 1} \Phi_d(1).$$

Taking logarithms of both sides, we have

$$\log n = \sum_{d \backslash n,\, d > 1} \log \Phi_d(1),$$

and by the formula above for the von Mangoldt function Λ, as well as the fact that $\Lambda(1) = 0$, we have

$$\sum_{d \backslash n,\, d > 1} \Lambda(d) = \sum_{d \backslash n,\, d > 1} \log \Phi_d(1).$$

The claim is that these two sums are actually equal term-by-term. When n is prime, the statement above already shows that $\log \Phi_p(1) = \Lambda(p) = \log p$, and supposing the claim proven for all $m < n$, we cancel all smaller terms in the formula to conclude that $\Lambda(n) = \log \Phi_n(1)$, which is what we needed to show.

Now we prove that Φ_n has integer coefficients. Again, the proof starts with the decomposition of $z^n - 1$ into linear factors, which this time we write as

$$z^n - 1 = \Phi_n(z) \prod_{d\backslash n,\, d<n} \Phi_d(z).$$

With the base case $\Phi_1(z) = z - 1$, strong induction would prove the claim if we can show that in a factorisation

$$z^n - 1 = (a_0 + a_1 z + \cdots + a_r z^r)(b_0 + b_1 z + \cdots + b_s z^s),$$

the hypotheses $b_s = 1$ and b_j being integer for all $1 \le j < s$ implies that the coefficients a_i are all integer for $1 \le i \le r$ and that this polynomial is monic as well. The fact that $a_r = 1$ is obvious. Then since b_0 is an integer and $a_0 b_0 = -1$, both a_0 and b_0 must be ± 1. Now assume that for some $t \ge 0$, a_i is integral for all $1 \le i \le t$, and consider the coefficient of z^{t+1} of the left-hand side. Call this coefficient c_{t+1} and note that it is an integer (in fact, it is either 0 or 1, but that is unimportant). We expand

$$c_{t+1} = a_{t+1} b_0 + a_t b_1 + \cdots + a_0 b_{t+1},$$

and rearrange to obtain

$$a_{t+1} = \frac{c_{t+1} - a_t b_1 - a_{t-1} b_2 - \cdots - a_0 b_{t+1}}{b_0},$$

from which we conclude by induction on t that

$$a_{t+1} = \pm(c_{t+1} - a_t b_1 - a_{t-1} b_2 - \cdots - a_0 b_{t+1})$$

is an integer. This also completes the induction on n, so we have shown that Φ_n is a monic polynomial with integer coefficients for all n. ▮

4. Irreducibility of cyclotomic polynomials

A polynomial $f(z)$ with integer coefficients is said to be *irreducible over* $\mathbf{Z}$ if it cannot be expressed as a product of two nonconstant polynomials in $\mathbf{Z}[z]$. This section will be devoted to proving that the cyclotomic polynomials Φ_n are irreducible over $\mathbf{Z}$. First we need a lemma in the ring of formal polynomials $\mathbf{F}_p[z]$.

Lemma 4. *Let $f(z)$ be a polynomial with coefficients in $\mathbf{F}_p$. Then $f(z^p) = f(z)^p$ in $\mathbf{F}_p[z]$.*

Proof. Let $f(z) = a_0 + a_1 z + \cdots + a_m z^m$. We have

$$f(z)^p = (a_0 + a_1 z + \cdots + a_m z^m)^p = \sum_{k_1+\cdots+k_m=p} \binom{p}{k_1, \ldots, k_m} \prod_{j=1}^{m} (a_j z^j)^{k_j}$$

by the multinomial theorem. But unless some $k_i = p$ and all the others are 0, there is a p in the numerator of the multinomial coefficient that does not appear in the denominator. Hence in $\mathbf{F}_p$ we have

$$\binom{p}{k_1, \ldots, k_m} = \begin{cases} 1, & \text{if } k_i = p \text{ for some } i; \\ 0, & \text{otherwise.} \end{cases}$$

Applying Fermat's little theorem, which states that $a^p = a$ in $\mathbf{F}_p$, we have

$$f(z)^p = \sum_{i=1}^{m} (a_i z^i)^p = \sum_{i=1}^{m} a_i (z^i)^p = f(z^p),$$

which is what we wanted to show. ▮

Theorem 5. *The nth cyclotomic polynomial is irreducible over $\mathbf{Z}$.*

Proof. Suppose, towards a contradiction, that $\Phi_n = fg$ for nonconstant f and g in $\mathbf{Z}[z]$. Then we can partition the primitive n roots of unity into two disjoint nonempty classes A and B such that

$$f(z) = \prod_{\omega \in A} (z - \omega) \qquad \text{and} \qquad g(z) = \prod_{\omega \in B} (z - \omega).$$

Since any two primitive roots are powers of one another, there exists $\omega \in A$ and an integer $m > 1$ such that $\omega^m \in B$. Factor m into primes $m = p_1 p_1 \cdots p_k$. Let $\omega_0 = \omega$ and for $1 \leq i \leq k$ let $\omega_i = \omega^{p_1 p_2 \cdots p_i}$. Let j be the smallest integer such that $\omega_j \in B$. (Since $\omega_0 \in A$ and $\omega_k = \omega^m \in B$, such a j must exist.) Now letting $\omega' = \omega^{p_1 \cdots p_{j-1}}$ and setting $p = p_j$, we have found some $\omega' \in A$ and some prime p such that $\omega^p \in B$.

This means that ω' is a root of both $f(z)$ and $g(z^p)$. Let $h(z)$ be the greatest common divisor of $f(z)$ and $g(z^p)$. By the Euclidean algorithm there exist polynomials $r(z)$ and $s(z)$ such that

$$h(z) = f(z) r(z) + g(z^p) s(z),$$

showing that $h(z)$ has ω' as a root and, in particular, is not constant. Now we consider everything as polynomials in $\mathbf{F}_p[z]$, which is a unique factorisation domain. Applying the previous lemma twice, we have $h(z^p) = h(z)^p$ and $z^{np} - 1 = (z^n - 1)^p$ in this ring. Now since $\Phi_n(z^p) = f(z^p) g(z^p) = f(z)^p g(z^p)$, we find that in $\mathbf{F}_p[z]$, the polynomial $h(z)^{p+1}$ divides $\Phi_n(z^p)$, and because $\Phi_n(z^p)$ divides $z^{np}-1=(z^n-1)^p$, we see that $h(z)^{p+1}$ divides $(z^n-1)^p$ as well. This means that $h(z)^2$ divides z^n-1. Putting $p(z)=z^n-1$, this means that there is some polynomial q such that $p=h^2q$. Then we find that $nz^{n-1}=p'=2hh'q+h^2q'$ is divisible by h, and thus z^n-1 and nz^{n-1} have a nonconstant common factor.

On the other hand, letting n^{-1} be the multiplicative inverse of n in $\mathbf{F}_p$, we can run the Euclidean algorithm on z^n-1 and nz^{n-1}:

$$\begin{aligned} z^n-1 &= (n^{-1}z)(nz^{n-1})+(-1)\\ nz^{n-1} &= (-1)(-nz^{n-1})+0, \end{aligned}$$

discovering that the greatest common divisor of these two polynomials is 1. This contradiction shows that $\Phi_n(z)$ is irreducible over $\mathbf{Z}$. ▮

5. Vandermonde determinants

In our journey towards proving a stronger uncertainty principle over $\mathbf{F}_p$, we will require special polynomials called *Vandermonde determinants.* These are indexed by $n\ge 1$ and defined by

$$\Delta_n(z_1,\ldots,z_n)=\prod_{i=1}^{n}\prod_{j=i+1}^{n}(z_j-z_i).$$

The next lemma justifies the name “determinant”.

Lemma 6. *Let $z_1,\ldots,z_n$ be indeterminates. Letting*

$$V=\begin{pmatrix} 1 & z_1 & {z_1}^2 & \cdots & {z_1}^{n-1}\\ 1 & z_2 & {z_2}^2 & \cdots & {z_2}^{n-1}\\ \vdots & \vdots & \vdots & \ddots & \vdots\\ 1 & z_n & {z_n}^2 & \cdots & {z_n}^{n-1} \end{pmatrix},$$

we have $\Delta_n(z_1,\ldots,z_n)=\det V$.

Proof. By the Leibniz formula for determinants, we have

$$\det V=\sum_{\pi\in\mathfrak{S}_n}\operatorname{sgn}(\pi)\prod_{i=1}^{n}{z_i}^{\pi(i)-1},$$

where $\mathfrak{S}_n$ is the symmetric group of all permutations on n letters. (The factor $\operatorname{sgn}(\pi)$ is 1 if the permutation π factors as a product of an even number of transpositions, and -1 if it factors as an odd number of transpositions.) Since the ith column of V contains only monomials of degree $i-1$, every term of $\det V$ is a monomial in which the sum of the degrees over all coefficients is $0+1+\cdots n-1=n(n-1)/2$.

In the formula

$$\Delta_n(z_1,\dots,z_n) = \prod_{i=1}^{n}\prod_{j=i+1}^{n}(z_j - z_i),$$

note that since the product runs over $\binom{n}{2} = n(n-1)/2$ linear factors, every term in this sum is a monomial of total degree $n(n-1)/2$ as well. Because $\det V$ is equal to zero if any two of the z_i are equal, $\det V$ is divisible by the linear polynomial $z_j - z_i$ for all $i < j$. Repeating this for all such i and j, we conclude that $\Delta_n(z_1,\dots,z_n)$ divides $\det V$; that is, $\det V = \Delta_n f$ for some polynomial f. But since both of them consist purely of monomials of total degree n, f must be a constant polynomial. To find out what this constant factor is, note that the term corresponding to the identity permutation in the Leibniz formula is ${z_2}{z_3}^2{z_4}^3\cdots{z_n}^{n-1}$, and expanding

$$\Delta_n(z_1,\dots,z_n) = (z_2 - z_1)(z_3 - z_2)(z_3 - z_1)(z_4 - z_3)\cdots(z_n - z_1),$$

a moment's scrutiny reveals that this term is also ${z_2}{z_3}^2{z_4}^3\cdots{z_n}^{n-1}$, and hence f must equal 1, which is what we needed. ∎

Lemma 7. *Let $n_1,\dots,n_k$ be positive integers and let $P \in \mathbf{Z}[z_1,\dots,z_k]$ be the polynomial given by*

$$P(z_1,\dots,z_n) = \sum_{\pi\in\mathfrak{S}_k}\operatorname{sgn}(\pi)\prod_{i=1}^{k}{z_i}^{n_{\pi(i)}}.$$

Then we may factor $P = \Delta_k Q$, where $Q \in \mathbf{Z}[z_1,\dots,z_k]$ is such that

$$Q(1,1,\dots,1) = \Delta_k(n_1,\dots,n_k)/\Delta_k(1,\dots,k).$$

Proof. By the Leibniz formula, $P(z_1,\dots,z_n)$ is the determinant of the $k\times k$ matrix whose entry in the ith row and jth column is $z_j^{n_i}$. As in the previous proof, P is divisible by $z_j - z_i$ for all $i < j$ and dividing out these linear factors, we obtain a polynomial Q such that $P = \Delta_k Q$.

It remains to compute $Q(1,1,\dots,1)$. To do so, we make use of the normalised differentiation operators $D_i = z_i(\partial/\partial z_i)$. It is easy to see that these operators obey the product rule $D_i(fg) = fD_ig + D_ifg$. Since

$$D_i({z_1}^{n_1}\cdots{z_k}^{n_k}) = n_i({z_1}^{n_1}\cdots{z_k}^{n_k}),$$

this monomial is an eigenfunction of D_i with eigenvalue n_i. Now consider the polynomial

$$(D_2{D_3}^2{D_4}^3\cdots{D_k}^{k-1})P = (D_2{D_3}^2{D_4}^3\cdots{D_k}^{k-1})\Big(Q\cdot\prod_{i<j}(z_j - z_i)\Big),$$

evaluated at 1. There are $k(k-1)/2$ differentiation operators to be applied and the same number of linear factors on the right-hand side. Repeatedly applying the product rule, we obtain $2^{k(k-1)/2}$ terms, but the only terms that survive when evaluated at $(1,\ldots,1)$ are the ones in which each linear factor $z_j - z_i$ is acted upon by either D_j or D_i, yielding z_j or $-z_i$ respectively.

Note that there are $k-1$ instances of the operator D_k to be applied, and there are only $k-1$ factors with the variable z_k appearing, namely the factors of the form $z_k - z_i$ for some $i < k$. So all those operators must hit those factors (yielding z_k), and there are $(k-1)!$ ways for this to happen. With those out of the way, there are now $k-2$ instances of D_{k-1} to be applied, and the only undifferentiated factors with the variable z_{k-1} appearing are the factors of the form $z_{k-1} - z_i$, of which there are $k-2$. So there are $(k-2)!$ ways for this to happen. Continuing in this manner, we see that

$$(D_2 {D_3}^2 {D_4}^3 \cdots {D_k}^{k-1} P)(1,\ldots,1) = 0!1!2!\cdots(k-1)!Q(1,\ldots,1).$$

But note that

$$\Delta_k(1,\ldots,k) = \prod_{i=1}^{k} \prod_{j=i+1}^{k} (j-i) = \prod_{i=1}^{k}(i-1)! = 0!1!2!\cdots(k-1)!,$$

so

$$(D_2 {D_3}^2 {D_4}^3 \cdots {D_k}^{k-1} P)(1,\ldots,1) = \Delta_k(1,\ldots,k)Q(1,\ldots,1).$$

But from the definition of P and the observation that $z_i^{n_{\pi(i)}}$ is an eigenfunction of the operator D_i with eigenvalue $n_{\pi(i)}$, we directly compute

$$(D_2 {D_3}^2 {D_4}^3 \cdots {D_k}^{k-1} P)(z_1,\ldots,z_k) = \sum_{\pi\in\mathfrak{S}_n} \operatorname{sgn}(\pi) \prod_{i=1}^{k} {n_{\pi(i)}}^{i-1} z_i^{n_{\pi(i)}}.$$

Evaluating at $z_1 = \cdots = z_k = 1$ and noting that the sum over all π in $\mathfrak{S}_n$ also runs over all π^{-1} in $\mathfrak{S}_n$ (π and π^{-1} have the same sign), we see that

$$(D_2 {D_3}^2 {D_4}^3 \cdots {D_k}^{k-1} P)(1,\ldots,1) = \Delta_k(n_1,\ldots,n_k).$$

Combining this with our earlier computation gives the conclusion

$$Q(1,\ldots,1) = \Delta_k(n_1,\ldots,n_k)/\Delta_k(1,\ldots,k),$$

which is what we wanted to prove. ▮

6. Chebotarëv's lemma

Armed with the irreducibility of cyclotomic polynomials and the computation from the last section, we can now prove a useful lemma concerning matrices of qth roots of unity, where q is a prime power. First we prove a criterion for the nonvanishing of polynomials on qth roots of unity.

Lemma 8. *Let p be a prime and q an integer power of p. Let $Q \in \mathbf{Z}[z_1, \dots, z_k]$ be such that $Q(1, \dots, 1)$ is not divisible by p. Then for any k-tuple $(\omega_1, \dots, \omega_k)$ of qth roots of unity, $Q(\omega_1, \dots, \omega_k) \neq 0$.*

Proof. We proceed by contraposition. Suppose that $\omega_1, \dots, \omega_k$ exist such that $Q(\omega_1, \dots, \omega_k) = 0$. Letting ω be a primitive root of unity, there are integers $n_1, \dots, n_k$ such that for all i, $\omega_i = \omega^{n_i}$. Let $R \in \mathbf{Z}[z]$ be given by $R(z) = Q(z^{n_1}, \dots, z^{n_k})$. Then $R(\omega) = 0$. Thus $R(z)$ has a root in common with the cyclotomic polynomial $\Phi_q(z)$. But we showed earlier that this polynomial is irreducible, implying that $\Phi_q(z)$ divides $R(z)$. So $Q(1, \dots, 1) = R(1)$ is divisible by $\Phi_q(1) = p$. ∎

We are now able to prove the following useful result, which is named after N. Chebotarëv.

Lemma 9 (*Chebotarëv*, 1926). *Let q be a prime power, let $1 \leq k < p$, and let $\omega_1, \dots, \omega_k$ be distinct qth roots of unity. Let $n_1, \dots, n_k$ be integers that are all distinct modulo p. Then the $k \times k$ matrix whose entry in the ith row and jth column is ${\omega_i}^{n_j}$ has nonzero determinant.*

Proof. Let $P(z_1, \dots, z_k)$ be the determinant of the matrix $({z_i}^{n_j})_{1 \leq i,j \leq k}$. This is the polynomial from Lemma 7, and that lemma says that we can factor $P = \Delta_k Q$, where Q is a polynomial with integer coefficients such that

$$Q(1, \dots, 1) = \Delta_k(n_1, \dots, n_k)/\Delta_k(1, \dots, k).$$

We want to show that $P(\omega_1, \dots, \omega_k)$ is not zero. Since ω_i are all distinct, $\Delta_k(\omega_1, \dots, \omega_k)$ is a product of $\binom{k}{2}$ nonzero elements, and in particular is nonzero. So we need only show that $Q(\omega_1, \dots, \omega_k) \neq 0$ and by the previous lemma, it suffices to show that $Q(1, \dots, 1)$ is not divisible by p. But the numerator in the formula for $Q(1, \dots, 1)$, namely $\Delta_k(n_1, \dots, n_k)$ is a product of differences $n_j - n_i$ for all $1 \leq i < j \leq k$, and these differences were all assumed to be nonzero modulo p. Thus their product is nonzero modulo p; in other words, their product is not divisible by p. This completes the proof. ∎

7. Tao's improved uncertainty principle

Chebotarëv's lemma is all we need to prove Tao's improved Fourier uncertainty principle for functions $f : \mathbf{Z}_p \to \mathbf{C}$. First we state a corollary of that lemma.

Corollary 10. *Let p be a prime and let A and B be subsets of $\mathbf{Z}_p$ with $|A| = |B|$. The linear transformation $\mathbf{C}^A \to \mathbf{C}^B$ define by $Tf = \widehat{f}|_B$ (that is, we restrict the Fourier transform of f to B) is invertible. (We write, for instance, $\mathbf{C}^A$ to denote functions from A to $\mathbf{C}$, or in other words, functions $f : \mathbf{Z}_p \to \mathbf{C}$ such that $\operatorname{supp}(f) \subseteq A$.)*

Proof. Write $Z = \mathbf{Z}_p$. Recall that the sets $\{\chi_u : u \in Z\}$ and $\{F_x : x \in Z\}$, as defined in the introduction, are orthogonal bases for $\mathbf{C}^Z$ and $\mathbf{C}^{\widehat{Z}}$ respectively.

In the first basis, by the Fourier inversion formula, the function f is represented by the vector $\big(\widehat{f}(0), \ldots, \widehat{f}(p-1))\big)$, and in the second basis, by the definition of Fourier transform, the function $\widehat{f}$ is represented by the vector $p^{-1}\big(f(0), \ldots, f(p-1)\big)$. If we set $\omega = e^{2\pi i/p}$, then the Fourier inversion formula can be expressed as

$$\begin{pmatrix} f(0) \\ f(1) \\ \vdots \\ f(p-1) \end{pmatrix} = \begin{pmatrix} \chi_0(0) & \chi_1(0) & \cdots & \chi_{p-1}(0) \\ \chi_0(1) & \chi_1(1) & \cdots & \chi_{p-1}(1) \\ \vdots & \vdots & \ddots & \vdots \\ \chi_0(p-1) & \chi_1(p-1) & \cdots & \chi_{p-1}(p-1) \end{pmatrix} \begin{pmatrix} \widehat{f}(0) \\ \widehat{f}(1) \\ \vdots \\ \widehat{f}(p-1) \end{pmatrix}$$
$$= \begin{pmatrix} (\omega^0)^0 & \omega^0 & \cdots & (\omega^0)^{p-1} \\ (\omega^1)^0 & \omega^1 & \cdots & (\omega^1)^{p-1} \\ \vdots & \vdots & \ddots & \vdots \\ (\omega^{p-1})^0 & \omega^{p-1} & \cdots & (\omega^{p-1})^{p-1} \end{pmatrix} \begin{pmatrix} \widehat{f}(0) \\ \widehat{f}(1) \\ \vdots \\ \widehat{f}(p-1) \end{pmatrix}.$$

Hence, letting $k = |A| = |B|$, the matrix of T is some $k \times k$ minor of the matrix above. But now letting the z_i be the ith powers of ω for $i \in A$ and letting $n_j = j$ for all $j \in B$, we see that this matrix satisfies the hypotheses of Chebotarëv's lemma, and must therefore be invertible. ∎

Theorem 11 (*Tao,* 2005). *Let p be a prime number. If $f : \mathbf{Z}_p \to \mathbf{C}$ is a nonzero function, then*

$$\|f\|_0 + \|\widehat{f}\|_0 \geq p + 1.$$

Conversely, if A and B are two nonempty subsets of $\mathbf{Z}_p$ such that $|A|+|B| \geq p+1$, then there exists a function f such that $\operatorname{supp}(f) = A$ *and* $\operatorname{supp}(\widehat{f}) = B$.

Proof. Suppose, for a contradiction, that f is such that $\|f\|_0 + \|\widehat{f}\|_0 \leq p$. Letting $A = \operatorname{supp}(f)$, we can then find a set B with $|B| = |A|$ that is disjoint from $\operatorname{supp}(\widehat{f})$. Now the Fourier transform of f restricted to B must be zero. But applying the corollary with A and B, we see that T should have nonzero determinant, which gives a contradiction since we have just found that $Tf = 0$ for some $f \neq 0$.

Now we prove the converse. First let us handle the case where $|A| + |B| = p+1$. In this situation we may choose A' with $|A'| = |A|$ such that $A' \cup B = \mathbf{Z}_p$ and $|A' \cap B| = 1$; say $A' \cap B = \{x\}$. Now, apply the corollary with A and A' to find that $T : \mathbf{C}^A \to \mathbf{C}^{A'}$ is invertible. Letting $g \in \mathbf{C}^{A'}$ be a function with $g(x) \neq 0$ and $g(y) = 0$ for all $y \in A' \setminus \{x\}$, we can find $f \in \mathbf{C}^A$ such that $Tf = g$, that is, the restriction of $\widehat{f}$ to A' equals g. Now, since $\operatorname{supp}(\widehat{f}) \subseteq A'^c \cup \{x\} = B$, we have $\|\widehat{f}\|_0 \leq |B|$ and then since $\operatorname{supp}(f) \subseteq A$, we have $\|f\|_0 \leq |A|$. But in order not to contradict what we proved in the previous paragraph, we must have $\operatorname{supp}(f) = A$ and $\operatorname{supp}(\widehat{f}) = B$.

Now assume that $|A| + |B| > p + 1$. Consider the set

$$S = \{(A', B') : A' \subseteq A, B' \subseteq B, |A| + |B| = p + 1\}.$$

This set is finite, so let us index its elements $(A_1, B_1), \dots, (A_s, B_s)$. From the previous paragraph, there exist functions $f_1, \dots, f_s$ such that for all $1 \le i \le s$, $\mathrm{supp}(f_i) = A_i$ and $\mathrm{supp}(\widehat{f_i}) = B_i$. Now let

$$f = \lambda_1 f_1 + \cdots \lambda_s f_s$$

for some scalars $\lambda_i \in \mathbf{C}$ to be chosen later. It is clear that $\mathrm{supp}(f) \subseteq A$ and, since the Fourier transform is linear, we also have $\mathrm{supp}(\widehat{f}) \subseteq B$. This is true regardless of our choices for the λ_i. Now we must prove that we can pick the λ_i so that $A \subseteq \mathrm{supp}(f)$ and $B \subseteq \mathrm{supp}(\widehat{f})$. For $x \in A$, let

$$V_x = \Big\{(\lambda_1, \dots, \lambda_s) \in \mathbf{C}^s : \sum_{i=1}^{s} \lambda_i f_i(x) = 0\Big\}.$$

This is a subspace of codimension 1 in $\mathbf{C}^s$, since we have s degrees of freedom and one nontrivial linear constraint. Similarly, for all $x \in B$, let

$$W_x = \Big\{(\lambda_1, \dots, \lambda_s) \in \mathbf{C}^s : \sum_{i=1}^{s} \lambda_i \widehat{f_i}(x) = 0\Big\}.$$

Now

$$\bigcup_{x \in A} V_x \cup \bigcup_{x \in B} W_x$$

is a finite union of subspaces of codimension 1. Thus its complement is nonempty and we can choose $\lambda_1, \dots, \lambda_s$ such that the resulting f has $f(x) \neq 0$ for all $x \in A$ and $\widehat{f}(x) \neq 0$ for all $x \in B$. This completes the proof. ∎

8. The Cauchy–Davenport theorem

We now use Tao's uncertainty principle for cyclic groups of prime order to prove the Cauchy–Davenport theorem. First, we need to define the *convolution* of two functions $f, g : Z \to \mathbf{C}$. This is denoted $f * g$ and given by

$$(f * g)(x) = \mathbf{E}_{y+z=x}\, f(y) g(z).$$

The following easy theorem describes the nice behaviour of convolutions under Fourier transform.

Theorem C (*Convolution law*). *Let Z be a finite abelian group and $f, g : Z \to \mathbf{C}$. For all $\chi \in \widehat{Z}$, we have $\widehat{f * g}(\chi) = \widehat{f}(\chi)\widehat{g}(\chi)$.*

Proof. We expand

$$\widehat{f * g}(\chi) = \mathbf{E}_{x \in Z} (f * g)(x) \overline{\chi(x)} = \mathbf{E}_{x \in Z}\, \mathbf{E}_{y+z=x}\, f(y) g(z) \overline{\chi(y)\chi(z)}.$$

But note that x does not appear anywhere in the summand, so we can rewrite this as an expectation over *all* y and z (their sum must equal x for some $x \in Z$). So

$$\widehat{f * g}(\chi) = \mathbf{E}_{y\in Z}\, \mathbf{E}_{z\in Z}\, f(y)\overline{\chi(y)}g(z)\overline{\chi(z)} = \widehat{f}(\chi)\widehat{g}(\chi),$$

which is what we wanted. ∎

Now let Z be a finite abelian group and let A and B be subsets of Z. We define the *sumset* $A + B$ to be the set of all sums $a + b$ where $a \in A$ and $b \in B$. This is related to the convolution operation above, since if $\operatorname{supp}(f) = A$ and $\operatorname{supp}(g) = B$, then $\operatorname{supp}(f * g) \subseteq A + B$. Also notice that the convolution law gives $\operatorname{supp}\bigl(\widehat{f * g}\bigr) = \operatorname{supp}(\widehat{f}) \cap \operatorname{supp}(\widehat{g})$.

Theorem 12 (*Cauchy–Davenport theorem*). *Let A and B be nonempty subsets of $\mathbf{Z}_p$. Then we have*

$$|A + B| \ge \min\bigl(|A| + |B| - 1, p\bigr).$$

Proof. Since A is nonempty, we can build a nonempty set $X \subseteq \mathbf{Z}_p$ with $p+1-|A|$ elements, so that $\mathbf{Z}_p \setminus X$ has $|A| - 1$ elements. Then since B is nonempty, if $|X| + 1 > |B|$, then we can pick $Y' \subseteq X$ with $|X| + 1 - |B|$ elements, and let $Y = Y' \cup (\mathbf{Z}_p \setminus X)$, so that

$$|Y| = p + 1 - |A| + 1 - |B| + |A| - 1 = p + 1 - |B|.$$

If instead $|X| + 1 \le |B|$, then we let Y' be any singleton subset of X and since $p - |X| \ge p + 1 - |B|$ we can find a subset Y'' of $\mathbf{Z}_p \setminus X$ of size $p - |B|$. Letting $Y = Y' \cup Y''$, we also have $|Y| = p + 1 - |B|$ in this case, and in both cases we have

$$|X \cap Y| = \max\bigl(|X| + 1 - |B|, 1\bigr) = \max\bigl(|X| + |Y| - p, 1\bigr).$$

Now since $|A| + |X| = |B| + |Y| = p + 1$, by Tao's uncertainty principle we can find $f : \mathbf{Z}_p \to \mathbf{C}$ with $\operatorname{supp}(f) = A$ and $\operatorname{supp}(\widehat{f}) = X$, as well as $g : \mathbf{Z}_p \to \mathbf{C}$ with $\operatorname{supp}(g) = B$ and $\operatorname{supp}(\widehat{g}) = Y$. As we saw before, the convolution $f * g$ has support contained in $A + B$ and the support of its Fourier transform equals $X \cap Y$. By the other direction of Tao's uncertainty principle, we must have

$$|A + B| + |X \cap Y| \ge |\operatorname{supp}(f * g)| + \bigl|\operatorname{supp}\bigl(\widehat{f * g}\bigr)\bigr| \ge p + 1.$$

But since $|X| + |Y| - p = p + 2 - |A| - |B|$,

$$|A + B| \ge p + 1 - \max\bigl(|X| + |Y| - 1, 1\bigr) = \min(|A| + |B| - 1, p),$$

exactly what we wanted to show. ∎

References