Fourier analysis on finite Abelian groups: some graphical applications

Andrew Goodall
Department of Mathematics
University of Bristol, Bristol BS8 1TW, United Kingdom
a.goodall@bristol.ac.uk

Abstract

A survey of basic techniques of Fourier analysis on a finite Abelian group Q with subsequent applications in graph theory. In particular, evaluations of the Tutte polynomial of a graph G in terms of cosets of the Q-flows (or dually Q-tensions) of G. Other applications to spanning trees of Cayley graphs and group-valued models on phylogenetic trees are also used to illustrate methods.

1 Introduction

Fourier analysis on finite vector spaces over $\mathbb{F}_2$, where the Fourier transform is also known as the Walsh-Hadamard transform, has been a rich source for applications in combinatorics. This is due to the fact that many combinatorial problems involve subsets of a finite set, and a subset of an m-set can be represented by its incidence vector in $\mathbb{F}_2^m$. See for example [13, 14] for the relationship of the Fourier transform to matroid duality and more generally to duals of functions defined on subsets of a finite set. As observed elsewhere [2, 24], Fourier analysis on other finite Abelian groups has not been so widespread in combinatorics, even though many of the methods used for (vector spaces over) $\mathbb{F}_2$ extend *mutatis mutandis* to (modules over) arbitrary Abelian groups. In [8, 9] Biggs uses the Fourier transform on an Abelian group Q of order q to exhibit the duality between interaction models defined in terms of vertex q-colourings of a graph (such as the q-state Potts model) and models defined in terms of Q-flows of a graph (such as the ice model on 4-regular graphs when $Q = \mathbb{F}_3$). The purpose of this article is to give further advertisement to how the elementary techniques of Fourier analysis on finite Abelian groups may be used to derive theorems in graph theory.

Properties of the Fourier transform relevant to our aims are outlined in Section 2. No avail is made of the Bonami-Gross-Beckner hypercontractive inequality for functions on finite vector spaces over $\mathbb{F}_2$ that has been of such utility in studying the phenomena of influences of Boolean variables and thresholds of

monotone Boolean functions: see for example the recent survey article [21] and references therein. A graphical application of this inequality in its version for functions on finite modules over finite cyclic groups can be found in [2].

Our graphical applications appear in Section 3, although in Section 2 there is a further illustrative example involving spanning trees of a Cayley graph drawing upon a known result from [27]. Cayley graphs are a natural object for study in graph theory by Fourier methods, since the Fourier transform diagonalises the adjacency matrix of a Cayley graph on an Abelian group. The most substantial application explored in Section 3 is in the search for new combinatorial interpretations for evaluations of the Tutte polynomial, an activity that has occupied the author elsewhere [16], and in this regard much influenced as a doctoral student by the perspicacious and patient supervision of Dominic Welsh.

2 Preliminaries

For background in the theory of Abelian groups and modules see for example [18]. An accessible introduction to Fourier analysis on finite groups and its applications can be found in [31]. A recent and exhaustive source for coding theory is [20]. Proofs can be found in these books for many of the facts quoted in this section, although gradually proofs will be included for results that are not so readily located in the literature.

For a consideration of analogues of the MacWillams extension and duality theorems in coding theory over finite (possibly non-commutative) rings see [36], in which there is an illuminating account of the role of generating characters. For a generalisation of matroid duality as defined over vector spaces to analogous structures defined over modules over finite rings see [32].

Abelian groups, rings, modules

Let R be an Abelian group written additively. R admits the structure of a commutative ring. Indeed, by the structure theorem for finite Abelian groups, if ℓ is the exponent of R (the least common multiple of the orders of its elements) then there is a unique sequence of positive integers $q_1, q_2, \ldots, q_m$ such that $2 \leq q_1 \mid q_2 \mid \cdots \mid q_m = \ell$ and

$$R \cong \mathbb{Z}_{q_1} \oplus \mathbb{Z}_{q_2} \oplus \cdots \oplus \mathbb{Z}_{q_m}.$$

This direct sum of additive groups can be interpreted as a direct sum of rings (the cyclic group $\mathbb{Z}_q$ is the additive group of the ring of integers modulo q). Multiplication in this case is componentwise: the product of $x = (x_1, x_2, \ldots, x_m)$, $y = (y_1, y_2, \ldots, y_m) \in \mathbb{Z}_{q_1} \oplus \mathbb{Z}_{q_2} \oplus \cdots \oplus \mathbb{Z}_{q_m}$ is $xy = (x_1y_1, x_2y_2, \ldots, x_my_m)$.

However, alternative multiplicative structures on R are possible. When the additive group R is the m-fold direct sum $\mathbb{Z}_p \oplus \mathbb{Z}_p \oplus \cdots \oplus \mathbb{Z}_p$ for prime p, R may be endowed with the structure of the finite field $\mathbb{F}_q$ for $q = p^m$.

Assume now that R is a commutative ring. If further $R = Q^m$ is the m-fold direct sum of another commutative ring Q, then R has the additional structure of a module over Q. In the sequel this will be the form R takes: m will be the number of edges in a graph G and elements of Q^m will be vectors indexed by edges of G. When $Q = \mathbb{F}_q$ is a finite field on q elements Q^m is a vector space over Q.

Characters

For an additive Abelian group R, a *character* of R is a homomorphism $\chi : R \to \mathbb{C}$ from the additive group R to the multiplicative group of $\mathbb{C}$. If the exponent of R is ℓ then characters map R to the multiplicative subgroup of ℓth roots of unity in $\mathbb{C}$. The set of characters form a group under pointwise multiplication. When R is a finite Abelian group $\widehat{R}$ is isomorphic to R. If $R = R_1 \oplus R_2$ then $\widehat{R} \cong \widehat{R_1} \times \widehat{R_2}$. In particular, if $R = Q^m$ is the m-fold direct sum of an Abelian group Q, then $\widehat{Q^m} \cong \widehat{Q}^m$.

For each $x \in R$, write χ_x for the image of x under a fixed isomorphism of R with $\widehat{R}$. In particular, the principal (trivial) character χ_0 is defined by $\chi_0(y) = 1$ for all $y \in R$, and $\chi_{-x}(y) = \overline{\chi_x(y)}$ for all $x, y \in R$, where the bar denotes complex conjugation.

Now introduce a multiplicative structure on R to make it into a commutative ring. A character $\chi \in \widehat{R}$ is a *generating character* for R if $\chi_x(y) = \chi(xy)$ for each character $\chi_x \in \widehat{R}$. If the commutative ring R has a generating character for its additive group then $\chi_x(y) = \chi_y(x)$.

Examples of rings with generating characters are the ring of integers $\mathbb{Z}_q$ modulo q, which has generating character $\chi(x) = e^{2\pi i x/q}$, and the finite field $\mathbb{F}_q$ for prime power $q = p^m$, which has generating character $\chi(x) = e^{2\pi i \mathrm{Tr}(x)/p}$, where $\mathrm{Tr}(x) = x + x^p + \cdots + x^{p^{m-1}}$ is the trace of x. Direct sums of rings with generating characters also have generating characters, so that given an Abelian group R there is always a ring with the additive structure of R which has a generating character. The ring $\mathbb{F}_2[X, Y]/(X^2, XY, Y^2)$ with additive group isomorphic to $\mathbb{Z}_2 \oplus \mathbb{Z}_2 \oplus \mathbb{Z}_2$ does not have a generating character; see [36] for the reason why.

If $R = Q^m$ is the m-fold direct sum of a commutative ring Q and ψ a generating character for Q, then χ defined by $\chi(x_1, \ldots, x_m) := \psi(x_1) \cdots \psi(x_m)$ for $(x_1, \ldots, x_m) \in Q^m$ is a generating character for Q^m. The *Euclidean inner product* (dot product) on Q^m is defined for $x = (x_1, \ldots, x_m), y = (y_1, \ldots, y_m) \in Q^m$ by $x \cdot y = x_1 y_1 + \cdots + x_m y_m$. Since $\psi(x_1) \cdots \psi(x_m) = \psi(x_1 + \cdots x_m)$, it follows that ψ has the property that $\chi_x(y) = \chi(xy) = \psi(x \cdot y)$ for $x, y \in Q^m$.

The algebra $\mathbb{C}^R$

Denote by $\mathbb{C}^R$ the vector space over $\mathbb{C}$ of all functions from R to $\mathbb{C}$. This is an inner product space with Hermitian inner product $\langle\ ,\ \rangle$ defined for $f, g \in \mathbb{C}^R$ by

$$\langle f, g\rangle = \sum_{x\in R} f(x)\overline{g(x)}.$$

Associated with this inner product is the Euclidean norm $\|f\|_2$ of f, defined by

$$\|f\|_2^2 = \langle f, f\rangle = \sum_{x\in R} |f(x)|^2.$$

The vector space $\mathbb{C}^R$ has the additional structure of an algebra under either of the following two definitions of multiplication:

(i) the *pointwise product* $f \cdot g$ of $f, g \in \mathbb{C}^R$, defined for $x \in R$ by $f \cdot g(x) = f(x)g(x)$,

(ii) the *convolution* $f * g$ of $f, g \in \mathbb{C}^R$, defined for $x \in R$ by

$$f * g(x) = \sum_{y\in R} f(y)g(x-y)$$

The r-fold convolution $f*f*\cdots*f$ is abbreviated to f^{*r}, the r-fold pointwise product $f \cdot f \cdots f$ to f^r. Note that for functions $f_1, \ldots, f_r \in \mathbb{C}^R$,

$$f_1 * f_2 * \cdots * f_r(x) = \sum_{\substack{x_1,\ldots,x_r\in R\\ x_1+\cdots+x_r=x}} f_1(x_1)f_2(x_2)\cdots f_r(x_r).$$

The set $\{1_x : x \in R\}$ of indicator functions defined by

$$1_x(y) = \begin{cases} 1 & x = y,\\ 0 & x \neq y,\end{cases}$$

form an orthonormal basis for $\mathbb{C}^R$, with $\langle 1_x, 1_y\rangle = 1_x(y)$. The indicator function notation is extended to subsets S of R by setting $1_S = \sum_{x\in S} 1_x$.

The characters of R are also orthogonal in this inner product space,

$$\langle \chi_x, \chi_y\rangle = \begin{cases} |R| & x = y,\\ 0 & x \neq y,\end{cases} \tag{1}$$

and form an orthogonal basis for $\mathbb{C}^R$.

The Fourier transform

Fix an isomorphism $x \mapsto \chi_x$ of R with $\widehat{R}$ and let χ be a generating character for R such that $\chi_x(y) = \chi(xy)$.

For $f \in \mathbb{C}^R$ the *Fourier transform* $\widehat{f} \in \mathbb{C}^R$ is defined for $y \in R$ by

$$\widehat{f}(y) = \langle f, \chi_y \rangle = \sum_{x \in R} f(x)\chi_y(-x).$$

The definition of the Fourier transform depends on the choice of isomorphism $R \to \widehat{R}$, $x \mapsto \chi_x$ but is independent of this choice up to an automorphism of R.

The Fourier transform maps the basis of indicator functions to the basis of characters: $\widehat{1_y} = \chi_{-y}$. The *Fourier inversion formula* $\widehat{\widehat{f}}(x) = |R|f(-x)$, gives the inverse transform

$$f(x) = \frac{1}{|R|}\langle \widehat{f}, \chi_{-x} \rangle = \frac{1}{|R|} \sum_{y \in R} \widehat{f}(y)\chi_x(y). \tag{2}$$

Note that $\widehat{1_R} = |R|1_0$ and $\widehat{1_0} = 1_R$, since $\langle 1_R, \chi_y \rangle = \langle \chi_0, \chi_y \rangle = |R|1_0(y)$.

Suppose g is a translation of f, i.e. $g(x) = f(x+z)$ for fixed z and all $x \in R$. Then $\widehat{g}(x) = \widehat{f} \cdot \chi_{-z}(x)$ is a modulation of $\widehat{f}(x)$. Now suppose g is a dilation of f by an invertible element of R, i.e. $g(x) = f(ux)$ for fixed unit u and all $x \in R$. Then $\widehat{g}(x) = \widehat{f}(u^{-1}x)$ is a dilation of $\widehat{f}$ by u^{-1}.

The orthogonality of characters (1) yields *Plancherel's identity*

$$\langle f, g \rangle = \frac{1}{|R|}\langle \widehat{f}, \widehat{g} \rangle, \tag{3}$$

a special case of which is *Parseval's identity*

$$\|f\|_2^2 = \frac{1}{|R|}\|\widehat{f}\|_2^2. \tag{4}$$

Identities (2), (3) and (4) depend only on the fact (1) that the characters $\{\chi_x : x \in R\}$ form an orthogonal basis for $\mathbb{C}^R$. That each χ_x is a homomorphism of R into the multiplicative group of $\mathbb{C}$ leads to the following key property. The Fourier transform gives an isomorphism of the algebra $\mathbb{C}^R$ with multiplication pointwise product with the algebra $\mathbb{C}^R$ with multiplication convolution: for $y \in R$

$$\widehat{f * g}(y) = \widehat{f} \cdot \widehat{g}(y), \tag{5}$$

$$\widehat{f \cdot g}(y) = \frac{1}{|R|}\widehat{f} * \widehat{g}(y). \tag{6}$$

Subgroups, submodules, annihilators, orthogonals

For a subgroup C of the additive group R, the *annihilator* $C^\sharp$ of C is defined

$$C^\sharp := \{x \in R : \forall_{y \in C}\ \chi_x(y) = 1\},$$

the set of x for which the kernel of χ_x contains C. The annihilator $C^\sharp$ is a subgroup of R isomorphic to R/C.

A second key property of the Fourier transform is that it takes indicators of subgroups to (scalar multiples of) indicators of their annihilators:

$$\widehat{1_C}(y) = \sum_{x \in C} \chi_x(y) = |C| 1_{C^\sharp}(y). \tag{7}$$

By (2), (5) and (7) there follows the *Poisson summation formula*

$$\sum_{x \in C} f(x+z) = \frac{1}{|C^\sharp|} \sum_{x \in C^\sharp} \widehat{f}(x) \chi_z(x).$$

Suppose now that $R = Q^m$ for some commutative ring Q.

The *orthogonal* to C (with respect to the Euclidean inner product) is defined by

$$C^\perp = \{y \in Q^m : \forall_{x \in C}\ x \cdot y = 0\}.$$

The orthogonal to C and annihilator of C coincide provided certain conditions are met:

Lemma 2.1 *Let Q be a commutative ring possessing a generating character and C a Q-submodule of Q^m. Then $C^\sharp = C^\perp$.*

Proof. If ψ is a generating character for Q, then Q^m has generating character χ, for which $\chi_x(y) = \chi(xy) = \psi(x \cdot y)$ for all $x, y \in Q^m$. Hence $C^\perp \subseteq C^\sharp$, since $\psi(x \cdot y) = \psi(0) = 1$ for all $y \in C^\perp$ and $x \in Q^m$. Suppose $\chi_x(y) = \psi(x \cdot y) = 1$ for all $y \in C$. Then for all $a \in Q$ we have $1 = \psi(x \cdot ay) = \psi(a(x \cdot y))$. This forces $x \cdot y = 0$ since this equation holds for all $a \in Q$ and ψ is a generating character for Q. □

For Q-submodule C of Q^m, denote the coset $\{x + z : x \in C\}$ of C in the additive group Q^m by $C + z$, an element of the quotient module Q^m/C. Abbreviating sums of the form $\sum_{x \in C+z} g(x) = g * 1_C(z)$ by the notation $g(C+z)$, the Poisson summation formula says that

$$f(C+z) = \frac{1}{|C^\perp|} \widehat{f} \cdot \chi_z(C^\perp) \tag{8}$$

whenever C is a Q-submodule of a ring Q^m that has a generating character.

In later applications the submodule C will be the image or kernel of a linear transformation defined by a matrix indexed by the vertices and edges of a graph, or, in the last application, the paths and edges of a rooted tree. In making use of (8) it will be helpful to record the following facts.

Consider a linear transformation $T : Q^m \to Q^n$ of Q-modules. The transpose linear transformation $T^t : Q^n \to Q^m$ is adjoint to T (relative to the Euclidean inner product) in that $Tx \cdot y = x \cdot T^t y$ for all $x \in Q^m$, $y \in Q^n$.

Lemma 2.2 *Let Q be a commutative ring with a generating character, $T : Q^m \to Q^n$ a linear transformation and $T^t : Q^n \to Q^m$ its transpose. Then $(\operatorname{im} T^t)^\perp = \ker T$ and $(\ker T)^\perp = \operatorname{im} T^t$.*

Proof. If $y \in \ker T$ then, for all $x \in Q^n$, $T^t x \cdot y = x \cdot Ty = x \cdot 0 = 0$. If $y \notin \ker T$ then $Ty \neq 0$ and $T^t x \cdot y = x \cdot Ty$ which is non-zero for some $x \in Q^n$ (for example x a unit vector with non-zero position coinciding with a non-zero position of Ty).

Although $(\ker T)^\perp = (\operatorname{im} T^t)^{\perp\perp} \supseteq \operatorname{im} T^t$ holds for any ring Q, that Q has a generating character is needed to show equality. Consider that, by (7) and Lemma 2.1, $\widehat{1_{\operatorname{im} T^t}} = |\operatorname{im} T^t| 1_{\ker T}$, and use the Fourier inversion formula (2) to deduce that $q^m 1_{\operatorname{im} T^t} = |\operatorname{im} T^t| \widehat{1_{\ker T}}$. With $\ker T$ a Q-submodule of Q^m,

$$|\ker T| 1_{(\ker T)^\perp} = \widehat{1_{\ker T}} = \frac{q^m}{|\operatorname{im} T^t|} 1_{\operatorname{im} T^t}.$$

This implies that $(\ker T)^\perp = \operatorname{im} T^t$, $q^m = |\operatorname{im} T^t| \cdot |\ker T|$. □

Hamming weight, Krawtchouck polynomials

For each $a \in Q$ and $x \in Q^m$, define $n_a(x) := \#\{j : x_j = a\}$, the number of occurences of a in x. The *Hamming weight* $|x| := m - n_0(x)$ is the number of non-zero entries of x. The sets $S_j := \{x \in Q^m : |x| = j\}$ are often called shells (or levels), and $\cup_{0 \leq j \leq i} S_j$ spheres.

The *Krawtchouck polynomial* $K_j(k; m, q)$ *of degree* j is defined for $0 \leq j, k \leq m$ by

$$K_j(k; m, q) = [z^j]\ (1+(q-1)z)^{m-k}(1-z)^k = \sum_{0 \leq i \leq j} (-1)^i (q-1)^{j-i} \binom{k}{i} \binom{m-k}{j-i}.$$

In particular, $K_0(k; m, q) = 1$, $K_1(k; m, q) = m(q-1) - qk$, and $K_m(k; m, q) = (-1)^k (q-1)^{m-k}$.

The Fourier transform of indicator functions of shells are given (see for example [20]) by

$$\widehat{1_{S_j}}(x) = K_j(|x|; m, q). \tag{9}$$

For C a Q-submodule of Q^m, by the Poisson summation formula (8),

$$|C \cap S_j| = \#\{x \in C : |x| = j\} = \frac{1}{|C^\perp|} \sum_{x \in C^\perp} K_j(|x|; m, q).$$

Spanning trees of Cayley graphs

Our first application of the Fourier transform to graph theory is adapted from [27, chapter 5] to which the reader is referred for further details.

Given an additive group R with subset $S \subseteq R$ satisfying $-S = S$, the *Cayley graph* $\operatorname{Cayley}(R; S)$ is defined to have vertices elements of R and edges joining x and y whenever $x - y \in S$. The characters χ_x of the Abelian group R form eigenvectors of the adjacency matrix of $\operatorname{Cayley}(R; S)$, with eigenvalues given by the Fourier transform $\widehat{1_S}(x)$.

A *Hamming graph* is a Cayley graph on Q^m with $S = S_j$ for some $1 \leq j \leq m$. The m-dimensional hypercube is the case $Q = \mathbb{F}_2$ and $S = S_1$, where two vertices x, y are adjacent if and only if $x, y \in \mathbb{F}_2^m$ differ in exactly one place.

Let $f \in \mathbb{C}^R$ have support $S = \{x \in R : f(x) \neq 0\}$ and the property that $f(-x) = f(x)$. The function f may be thought of as assigning non-zero $\mathbb{C}$-valued weights to the edges of Cayley$(R; S)$. Define the linear transformation $M : \mathbb{C}^R \to \mathbb{C}^R$ by

$$Mg(x) = \sum_{y \in R} f(y)g(x + y) = f * g(x).$$

For a given vertex x of Cayley$(R; S)$, $Mg(x)$ is the 'f-weighted average' of the values of g at the vertices adjacent to x. With

$$M\chi_z(x) = \sum_{y \in R} f(y)\chi_z(y)\chi_z(x) = \widehat{f}(z)\chi_z(x),$$

χ_z is an eigenvector of M for each $z \in R$. By expressing the Laplacian of Cayley$(R; S)$ in terms of M and then appealing to Kirchoff's Matrix Tree Theorem, the result of [27, exercise 5.68] is that

$$\sum_T \prod_{e \in T} f(e) = \frac{1}{|R|} \prod_{0 \neq y \in R} \left(\widehat{f}(0) - \widehat{f}(y)\right),$$

where the left-hand summation is over the edge-sets of all spanning trees T of Cayley$(R; S)$. In particular,

$$\#\{\text{spanning trees of Cayley}(R; S)\} = \frac{1}{|R|} \prod_{0 \neq y \in R} (|S| - \widehat{1_S}(y)). \tag{10}$$

Taking $R = Q^m$ for Q of size q and $f = 1_{S_1}$, where $S_1 = \{x \in Q^m : |x| = 1\}$,

$$\#\{\text{spanning trees of Cayley}(Q^m; S_1)\} = q^{q^m - m - 1} \prod_{1 \leq k \leq m} k^{\binom{m}{k}(q-1)^k}, \tag{11}$$

since $\widehat{1_{S_1}}(y) = K_1(|y|; m, q) = m(q-1) - q|y|$ and $\#\{y \in Q^m : |y| = k\} = |S_k| = \binom{m}{k}(q-1)^k$. Stanley [27, example 5.6.10] gives as an example (11) for the case $Q = \mathbb{F}_2$, remarking that a direct combinatorial proof is not known.

By (10) and (9) the number of spanning trees of Cayley$(Q^m; S_j)$ is given by

$$q^{-m} \prod_{1 \leq k \leq m} (K_j(0; m, q) - K_j(k; m, q))^{\binom{m}{k}(q-1)^k}.$$

Weight enumerators

Before defining weight enumerators a technical lemma is recorded.

Lemma 2.3 *Let R be an Abelian group and $R = R_1 \oplus R_2$. Suppose that $f \in \mathbb{C}^R$ is defined for $x \in R$ by $f(x) = g(x_1)h(x_2)$ for functions $g \in \mathbb{C}^{R_1}$ and $h \in \mathbb{C}^{R_2}$. Then $\widehat{f}(x) = \widehat{g}(x_1)\widehat{h}(x_2)$, with Fourier transforms defined on the appropriate spaces.*

In particular, if $R = Q^m$ for an Abelian group Q and $f \in \mathbb{C}^R$ is defined for each $x = (x_1, \ldots, x_m) \in Q^m$ by

$$f(x) = f_1(x_1)f_2(x_2)\cdots f_m(x_m),$$

where $f_1, \ldots, f_m \in \mathbb{C}^Q$, then

$$\widehat{f}(x) = \widehat{f_1}(x_1)\widehat{f}(x_2)\cdots \widehat{f_m}(x_m),$$

where the Fourier transforms on the right-hand side are on $\mathbb{C}^Q$.

Let Q be a commutative ring with a generating character, $h : Q \to \mathbb{C}$ and C a submodule of Q^m. For each $z \in Q^m$ the *complete weight enumerator* of the coset $C + z$ of C is defined by

$$\mathsf{cwe}(C + z; h) := \sum_{x \in C+z} \prod_{a \in Q} h(a)^{n_a(x)}. \tag{12}$$

(It is usual to take the values $h(a)$ as indeterminates t_a so that the complete weight enumerator is a q-variable polynomial in $\{t_a : a \in Q\}$, but we shall mainly be interested in evaluations of the complete weight enumerator in $\mathbb{C}$.) The MacWilliams duality theorem for complete weight enumerators states that

$$\mathsf{cwe}(C; h) = \frac{1}{|C^\perp|}\mathsf{cwe}(C^\perp; \widehat{h}), \tag{13}$$

where $\widehat{h}$ is the Fourier transform of h, $\widehat{h}(b) = \sum_{a \in Q} h(a)\chi_a(-b)$.

The identity (13) is proved by defining $f : Q^m \to \mathbb{C}$ for $x = (x_1, \ldots, x_m)$ by

$$f(x) = h(x_1)h(x_2)\cdots h(x_m),$$

so that the complete weight enumerator is equal to $f(C)$, and using Lemma 2.3 and the Poisson summation formula (8) with $z = 0$.

The *Hamming weight enumerator* of $C + z$ is the specialisation of (12) obtained by setting $h(0) = t$ for indeterminate t and $h(a) = 1$ for $a \neq 0$:

$$\mathsf{hwe}(C + z; t) := \sum_{x \in C+z} t^{n_0(x)} = \sum_{x \in C+z} t^{m-|x|}.$$

Specialising (13) to the Hamming weight enumerator, the MacWilliams duality theorem states that

$$\mathsf{hwe}(C; t) = \frac{(t-1)^m}{|C^\perp|}\mathsf{hwe}(C^\perp, \frac{t-1+q}{t-1}), \tag{14}$$

where $q = |Q|$.

In the following, recall that for a function $g : Q^m \to \mathbb{C}$ the notation $g(C+z)$ stands for $\sum_{x \in C} g(x + z)$.

Lemma 2.4 *Let Q^m be a commutative ring with a generating character. For submodule C of Q^m and $r \geq 1$ functions $f_1, \ldots, f_r : Q^m \to \mathbb{C}$*

$$\sum_{C+z \in Q^m/C} f_1(C+z) \cdots f_r(C+z) = \frac{1}{|C^\perp|^{r-1}} \sum_{\substack{x_1,\ldots,x_r \in C^\perp \\ x_1+\cdots+x_r=0}} \widehat{f}_1(x_1) \cdots \widehat{f}_r(x_r). \tag{15}$$

Also, for $r \geq 1$ and $2r$ functions $f_1, \ldots, f_{2r} : Q^m \to \mathbb{C}$,

$$\sum_{C+z \in Q^m/C} f_1(C+z) \cdots f_r(C+z) \overline{f_{r+1}(C+z)} \cdots \overline{f_{2r}(C+z)}$$

$$= \frac{1}{|C^\perp|^{2r-1}} \sum_{\substack{x_1,\ldots,x_{2r} \in C^\perp \\ x_1+\cdots+x_r=x_{r+1}+\cdots+x_{2r}}} \widehat{f}_1(x_1) \cdots \widehat{f}_r(x_r) \overline{\widehat{f_{r+1}}(x_{r+1})} \cdots \overline{\widehat{f_{2r}}(x_{2r})}. \tag{16}$$

Proof. By (2), (5), (6), (7) and Lemma 2.1, the left-hand side of (15) is

$$\frac{1}{|C|} \sum_{z \in Q^m} f_1 * 1_C \cdot f_2 * 1_C \cdots f_r * 1_C(z) = \frac{1}{|C|} \cdot \frac{|C|^r}{q^{mr}} \cdot q^m \cdot \widehat{f}_1 \cdot 1_{C^\perp} * \cdots * \widehat{f}_r \cdot 1_{C^\perp}(0)$$

$$= \frac{1}{|C^\perp|^{r-1}} \sum_{x_1+\cdots+x_r=0} \widehat{f}_1(x_1) 1_{C^\perp}(x_1) \cdots \widehat{f}_r(x_r) 1_{C^\perp}(x_r),$$

which is the right-hand side of (15). The second statement (16) can be deduced from (15) by using the fact that if $g : Q^m \to \mathbb{C}$ then $\widehat{\widehat{g}}(y) = \overline{\widehat{g}(-y)}$. □

On setting $f_1 = \cdots = f_r = f_{r+1} = \cdots = f_{2r}$ in Lemma 2.4 we obtain the following.

Corollary 2.5 *Let Q^m be a commutative ring with a generating character. For submodule C of Q^m, function $f : Q^m \to \mathbb{C}$ and integer $r \geq 2$,*

$$\sum_{C+z \in Q^m/C} f(C+z)^r = \frac{1}{|C^\perp|^{r-1}} \sum_{x_1,\ldots,x_{r-1} \in C^\perp} \widehat{f}(x_1) \cdots \widehat{f}(x_{r-1}) \widehat{f}(-x_1 - x_2 - \cdots - x_{r-1}).$$

For integer $r \geq 1$,

$$\sum_{C+z \in Q^m/C} |f(C+z)|^{2r}$$

$$= \frac{1}{|C^\perp|^{2r-1}} \sum_{\substack{x_1,\ldots,x_{2r} \in C^\perp \\ x_1+\cdots+x_r=x_{r+1}+\cdots+x_{2r}}} \widehat{f}(x_1) \cdots \widehat{f}(x_r) \overline{\widehat{f}(x_{r+1})} \cdots \overline{\widehat{f}(x_{2r})}.$$

Another special case of Lemma 2.4 in conjunction with Lemma 2.3 will be useful later:

Corollary 2.6 *Let Q^m be a commutative ring with a generating character. For submodule C of Q^m and functions $h, k : Q \to \mathbb{C}$*

$$\sum_{C+z \in Q^m/C} \mathsf{cwe}(C+z;h)\overline{\mathsf{cwe}(C+z;k)} = \frac{1}{|C^\perp|}\mathsf{cwe}(C^\perp;\widehat{h}\cdot\overline{\widehat{k}}).$$

Proof. Set $r = 1$ in identity (16) of Lemma 2.4 and define $f_1, f_2 : Q^m \to \mathbb{C}$ by $f_1(x) = h(x_1)h(x_2)\cdots h(x_m)$, $f_2(x) = k(x_1)k(x_2)\cdots k(x_m)$ for $x = (x_1, x_2, \ldots, x_m)$. Lemma 2.3 and the definition (12) of the complete weight enumerator now give the result.□

3 Graphical applications

Let $G = (V, E)$ be a graph on n vertices and m edges with an arbitrary orientation assigned to its edges, and Q a commutative ring on q elements possessing a generating character. The number of components of G is denoted by $k(G)$. The *rank* $r(E)$ of G is $n - k(G)$.

Consider the linear transformation $T : Q^E \to Q^V$ defined by the $n \times m$ incidence matrix $T = (T_{v,e})$ of G with (v, e) entry

$$T_{v,e} = \begin{cases} 1 & v \text{ the positive end of } e, \\ -1 & v \text{ the negative end of } e, \\ 0 & v \text{ not incident with } e. \end{cases}$$

The *boundary* operator T takes an edge colouring $x \in Q^E$ and colours a vertex with the net flow of colours into it:

$$(Tx)_v = \sum_{e \in E} T_{v,e}x_e.$$

The *coboundary* operator $T^t : Q^V \to Q^E$,

$$(T^t y)_e = \sum_{v \in V} T_{v,e}y_v,$$

takes a vertex colouring $y \in Q^V$ and colours a directed edge $e = (u, v)$ with the difference $y_v - y_u$ of colours on its endpoints. A loop (v, v) always has coboundary 0.

For a graph $G = (V, E)$ with boundary operator $T : Q^E \to Q^V$, $\ker T$ is the submodule of Q-*flows* of G and $\operatorname{im} T^t$ is the submodule of Q-*tensions* of G. By Lemma 2.2, $(\ker T)^\perp = \operatorname{im} T^t$ provided Q has a generating character, and always $(\operatorname{im} T^t)^\perp = \ker T$. There are $q^{n-k(G)}$ Q-tensions of G and $q^{m-n+k(G)}$ Q-flows of G. To each Q-tension $y \in \operatorname{im} T^t$ there are $q^{k(G)}$ vertex Q-colourings, all of whose coboundaries are equal to y.

The *boundary polynomial* of G is the Hamming weight enumerator of the submodule of Q-flows of G,

$$\mathsf{hwe}(\ker T; s) = \sum_{x \in \ker T} s^{m-|x|},$$

and the *coboundary polynomial* the Hamming weight enumerator of the submodule of Q-tensions of G,

$$\mathsf{hwe}(\operatorname{im} T^t; t) = \sum_{y \in \operatorname{im} T^t} t^{m-|y|} = q^{-k(G)} \sum_{z \in Q^V} t^{m-|T^t z|}.$$

The boundary and coboundary polynomials are related by the MacWilliams duality identity (14). The exponent $m - |T^t z|$ in the second expression for the coboundary polynomial is the number of monochrome edges in the vertex Q-colouring z, i.e. edges whose endpoints have the same colour. This relates the coboundary polynomial more transparently to the partition function of the q-state Potts model of statistical physics. See for example [33, 35].

For $F \subseteq E$ the rank $r(F)$ of F is defined to be the rank of the subgraph (V, F) obtained from G by deleting the edges in $E \backslash F$. The *Tutte polynomial* of G is defined by

$$T(G; s, t) = \sum_{F \subseteq E} (s-1)^{r(E)-r(F)} (t-1)^{|F|-r(F)}. \tag{17}$$

For more information about the Tutte polynomial see for example Dominic's book [33] or survey paper [34]. Many evaluations of the Tutte polynomial of a graph have interpretations in terms of combinatorial properties of the graph, some more surprising than others. So for example that $T(G; 1, 1)$ counts spanning trees of a connected graph G is clear upon observing that $r(F) = |F|$ if and only if (V, F) is acyclic and $r(F) = r(E)$ if and only if (V, F) is spanning. But that $T(G; -1, -1) = (-1)^m (-2)^d$ where d is the dimension of bicycle space of G might appear less so [25]. For boundary map $T : \mathbb{F}_2^m \to \mathbb{F}_2^n$, the bicycle space of G is $\ker T \cap \operatorname{im} T^t$, the subspace of elements that are both flows (cycles) and tensions (cocycles). Also not obvious from the definition (17) is that the coefficients of $T(G; s, t)$ are non-negative integers: they count spanning trees of G according to their internal activity and external activities; see for example [11, 10] for an explanation of these terms. Below we shall obtain evaluations of the Tutte polynomial of the form

$$s_2^{n-k(G)} t_2^{m-n+k(G)} T(G; \frac{s_1}{s_2}, \frac{t_1}{t_2}). \tag{18}$$

If $s_2 = 0$ then (18) is well-defined, despite appearances, with value $t_2^{m-n+k(G)-\ell(G)} s_1^{n-k(G)} t_1^{\ell(G)}$, where $\ell(G)$ is the number of loops in G; similarly, if $t_2 = 0$ then (18) is equal to $s_2^{n-k(G)-b(G)} s_1^{b(G)} t_1^{m-n+k(G)}$, where $b(G)$ is the number of bridges (coloops) in G.

In order to obtain our evaluations of the Tutte polynomial we use the fact (see for example [35]) that the boundary and coboundary polynomials are specialisations of Tutte polynomial on the hyperbola $H_q := \{(s, t) : (s-1)(t-1) = q\}$. If $T : Q^m \to Q^n$ is the boundary operator on G and $|Q| = q$ then

$$\mathsf{hwe}(\ker T; s) = (s-1)^{m-n+k(G)} T(G; s, \frac{s-1+q}{s-1}), \tag{19}$$

and

$$\mathsf{hwe}(\operatorname{im} T^t; t) = (t-1)^{n-k(G)} T(G; \frac{t-1+q}{t-1}, t). \tag{20}$$

The Ashkin-Teller model and the 4-state Potts model

Let G be a graph on n vertices and m edges, C the subspace of $\mathbb{F}_2$-flows of G, and $\widetilde{C}$ the subspace of $\mathbb{F}_4$-flows of G. (So $C^\perp$ is the subspace of $\mathbb{F}_2$-tensions and $\widetilde{C}^\perp$ the subspace of $\mathbb{F}_4$-tensions.) The Cartesian product $C \times C$ is isomorphic to $\widetilde{C}$ and $(C \times C)^\perp = C^\perp \times C^\perp$ is isomorphic to $\widetilde{C}^\perp$.

Let $\{0, 1, \omega, \overline{\omega} = 1 + \omega\}$ be the set of elements of $\mathbb{F}_4$. Identify the subspaces $C \times C$ and $C^\perp \times C^\perp$ of $\mathbb{F}_2^m \times \mathbb{F}_2^m$ with their images $\widetilde{C}$ and $\widetilde{C}^\perp$ in $\mathbb{F}_4^m$ under the isomorphism $\mathbb{F}_2 \times \mathbb{F}_2 \to \mathbb{F}_4$ defined by $(0,0) \mapsto 0$, $(1,1) \mapsto 1$, $(0,1) \mapsto \omega$, $(1,0) \mapsto \overline{\omega}$.

The partition function of the Ashkin-Teller model [5] is an evaluation of the complete weight enumerator of the subspace of $\mathbb{F}_4$-tensions

$$\mathsf{cwe}(\widetilde{C}^\perp; t_0, t_1, t_\omega, t_{\overline{\omega}}) = \sum_{z \in \widetilde{C}^\perp} t_0^{n_0(z)} t_1^{n_1(z)} t_\omega^{n_\omega(z)} t_{\overline{\omega}}^{n_{\overline{\omega}}(z)},$$

with the specialisation $t_0 = t_1 t_\omega t_{\overline{\omega}}$ and an assignment of positive real values to $t_1, t_\omega, t_{\overline{\omega}}$ (corresponding to interaction energies). The symmetric Ashkin-Teller model takes the further specialisation $t_\omega = t_{\overline{\omega}}$. The 4-state Potts model (Hamming weight enumerator) is the specialisation $t_1 = t_\omega = t_{\overline{\omega}}$.

Identifying $z \in \mathbb{F}_4$ with $(x, y) \in \mathbb{F}_2 \times \mathbb{F}_2$, the Ashkin-Teller specialisation of the complete weight enumerator of $\widetilde{C}^\perp$ is alternatively given by

$$\sum_{z \in \widetilde{C}^\perp} (t_1 t_\omega t_{\overline{\omega}})^{n_0(z)} t_\omega^{n_\omega(z)} t_{\overline{\omega}}^{n_{\overline{\omega}}(z)} t_1^{n_1(z)} = \sum_{x,y \in C^\perp} t_\omega^{n_0(x)} t_{\overline{\omega}}^{n_0(y)} t_1^{n_0(x+y)}. \tag{21}$$

When $t_1 = 1$ this reduces to $\mathrm{hwe}(C^\perp; t_\omega)\mathrm{hwe}(C^\perp; t_{\overline{\omega}})$, and the Ashkin-Teller model reduces to two independent Ising models (2-state Potts models).

From Corollary 2.5 the sum of the cubes of coset weight enumerators of the subspace C of $\mathbb{F}_2$-flows is, with the help of (21), given by

$$\sum_{C + z \in \mathbb{F}_2^m / C} \mathsf{hwe}(C + z; t)^3 = \frac{(t-1)^{3m}}{|C^\perp|^2} \sum_{x,y \in C^\perp} \left(\frac{t+1}{t-1}\right)^{n_0(x)+n_0(y)+n_0(x+y)}$$

$$= \frac{(t-1)^{3m}}{|C^\perp|^2} \cdot \left(\frac{t+1}{t-1}\right)^m \sum_{z \in C^\perp \times C^\perp} \left(\frac{t+1}{t-1}\right)^{2n_0(z)}$$

$$= \frac{(t-1)^m (t^2-1)^m}{|C^\perp|^2} \mathrm{hwe}(\widetilde{C}^\perp; \left(\frac{t+1}{t-1}\right)^2).$$

Thus we have the following theorem, by using (20) to replace the Hamming weight enumerator in the above expression by the Tutte polynomial on H_4.

Theorem 3.1 *Let $G = (V, E)$ be a graph on n vertices, m edges and $k(G)$ components, and let C be the subspace of $\mathbb{F}_2$-flows of G. Then, for $t \in \mathbb{C}$,*

$$\sum_{C+z\in\mathbb{F}_2^m/C} \mathsf{hwe}(C+z;t)^3 = (t+1)^m t^{n-k(G)}(t-1)^{2(m-n+k(G))}T(G;\frac{t^2-t+1}{t},\left(\frac{t+1}{t-1}\right)^2).$$

In particular, when $t = \sqrt{-1}$,

$$\sum_{C+z\in\mathbb{F}_2^m/C} \mathsf{hwe}(C+z;\sqrt{-1})^3 = (\sqrt{-1}-1)^m(-2)^{m-n+k(G)}T(G;-1,-1). \quad (22)$$

Also from Corollary 2.5 we have the identity

$$\sum_{C+z\in\mathbb{F}_2^m/C} \mathsf{hwe}(C+z;t)^2 = (2t)^{n-k(G)}(t-1)^{2(m-n+k(G))}T(G;\frac{t^2+1}{2t},\left(\frac{t+1}{t-1}\right)^2), \quad (23)$$

this time giving a specialisation of the Tutte polynomial to H_2, and which for $t = \sqrt{-1}$ evaluates to $(2\sqrt{-1})^m$ if G is Eulerian and 0 otherwise. Similarly

$$\sum_{C+z\in\mathbb{F}_2^m/C} |\mathsf{hwe}(C+z;t)|^2 = (t+\bar{t})^{n-k(G)}|t-1|^{2(m-n+k(G))}T(G;\frac{|t|^2+1}{t+\bar{t}},\left|\frac{t+1}{t-1}\right|^2), \quad (24)$$

which for $t = \sqrt{-1}$ evaluates to 2^m. Evaluations such as (22), (23) and (24) yield information about the distribution of Hamming weights modulo 4 in cosets of $\mathbb{F}_2$-flows: writing $N_j = N_j(C+z) := \#\{x \in C+z : n_0(x) \equiv j \pmod 4\}$ for $j = 0, 1, 2, -1$, we have $\mathsf{hwe}(C+z;\sqrt{-1}) = N_0 - N_2 + \sqrt{-1}(N_1 - N_{-1})$.

Evaluations of the Tutte polynomial

In all of this section Q will be an Abelian group of order q which will be assumed to have been given the further structure of a ring with a generating character so as to be able to apply the results of Section 2. Let $G = (V, E)$ be a graph on n vertices, m edges and with $k(G)$ components, C the Q-submodule of Q-flows of G, and $C^\perp$ the Q-submodule of Q-tensions of G.

Let $A \subseteq Q$ and $B = A^m \subseteq Q^m$. Supposing $C + z \in Q^m/C$ is chosen uniformly at random, what can be said about the probability distribution of $|C + z \cap B|$? When $A = Q$ this distribution is clearly uniform, but when for example $A = Q \setminus 0$, so that $B = \{x \in Q^m : |x| = m\}$, the answer is not so clear. In particular, the size of $|C \cap B|$ in this case is $(-1)^{m-n+k(G)}T(G; 0, 1-q)$, the number of nowhere-zero Q-flows of G.

Corollary 2.5 with $f = 1_B$ yields the rth moments of $|C + z \cap B|$ as specialisations of $\mathsf{cwe}(\widetilde{C}^\perp)$ where $\widetilde{C}$ is the $(r-1)$-fold Cartesian product $C \times C \times \cdots C$, equal to the set of Q^{r-1}-flows of G, and $\widetilde{C}^\perp$ the set of Q^{r-1}-tensions of G (using a similar argument to above for $Q = \mathbb{F}_2, r = 3$).

For $r = 2$, in order to have an evaluation of the Tutte polynomial on H_q the evaluation of the complete weight enumerator of $C^\perp$ provided by Corollary 2.6,

$$\sum_{C+z \in Q^m/C} |C + z \cap B|^2 = \frac{1}{|C^\perp|} \mathsf{cwe}(C^\perp; |\widehat{1_A}|^2), \tag{25}$$

needs to be an evaluation of Hamming weight enumerator of $C^\perp$; i.e. A must have the property that $|\widehat{1_A}|^2$ is constant on $Q \setminus 0$. Identity (20) tells us that

$$\frac{1}{|C^\perp|} \mathsf{hwe}(C^\perp; t) = \left(\frac{t-1}{q}\right)^{n-k(G)} T(G; \frac{t-1+q}{t-1}, t), \tag{26}$$

so that if $|\widehat{1_A}|^2$ is constant on $Q \setminus 0$ then (26) yields an evaluation of the Tutte polynomial on H_q interpreted in terms of the left-hand side of (25), and hence the second moment of $|C + z \cap B|$ when a coset $C + z$ is chosen uniformly at random from Q^m/C.

It will be convenient to use a variation on convolution of functions defined on a ring R, namely the *cross-correlation* $f \star g$ of two functions $f, g \in \mathbb{C}^R$, defined by

$$f \star g(x) = \sum_{y \in R} \overline{f(y)} g(x + y).$$

This has the property that

$$\widehat{f \star g}(y) = \overline{\widehat{f}} \cdot \widehat{g}(y), \tag{27}$$

and in particular $\widehat{f \star f} = |\widehat{f}|^2$. Most pertinently for this section, for each $a \in Q$,

$$1_A \star 1_A(a) = \#\{(a_1, a_2) \in A \times A : a_1 - a_2 = a\},$$

and since a function $g : Q \to \mathbb{C}$ is constant on $Q \setminus 0$ if and only if $\widehat{g}$ is constant on $Q \setminus 0$ it follows that $|\widehat{1_A}|^2$ has this property if and only if $1_A \star 1_A$ does.

For $2 \le s \le q$, a (q, s, ℓ)-*difference set* in Q is a subset A of s elements of Q with the property that for each $0 \ne a \in Q$ there are precisely ℓ pairs $(a_1, a_2) \in A \times A$ such that $a_1 - a_2 = a$, i.e. $1_A \star 1_A$ is constant on $Q \setminus 0$. (There are s pairs $(a, a) \in A \times A$ with difference equal to 0.) The parameters of a (q, s, ℓ)-difference set must satisfy $s(s-1) = (q-1)\ell$. For any Abelian group Q on $q \ge 2$ elements the set $A = Q$ forms a (q, q, q)-difference set and the set of nonzero elements $A = Q \setminus 0$ form a $(q, q-1, q-2)$-difference set in A. A (q, s, ℓ_0, ℓ_1)-*partial difference set* in Q is a subset A of size s with the property that for each $0 \ne a \in A$ there are precisely ℓ_0 pairs $(a_1, a_2) \in A \times A$ such that $a_1 - a_2 = a$ and for each $0 \ne a \notin A$ there are precisely ℓ_1 pairs $(a_1, a_2) \in A \times A$ such that $a_1 - a_2 = a$, i.e. $1_A \star 1_A$ is constant on A and constant on $Q \setminus (A \cup 0)$. If $Q = \mathbb{F}_q$ for prime power $q \equiv 1 \pmod 4$ then the subset of non-zero squares is a $(q, \frac{q-1}{2}, \frac{q-5}{4}, \frac{q-1}{4})$-partial difference set. For more on difference sets see for example [7, 17, 22].

Equation (25) yields the following evaluation of the Tutte polynomial on H_q.

Theorem 3.2 [16] *Let $G = (V, E)$ be a graph on n vertices, m edges and with $k(G)$ components. Let Q be an Abelian group of order q, C the set of Q-flows of G, A a (q, s, ℓ)-difference set in Q and $B = A^m$. Then*

$$\sum_{C+z\in Q^m/C} |C + z \cap B|^2 = \ell^{n-k(G)}(s-\ell)^{m-n+k(G)}T(G; \frac{s}{\ell}, \frac{s^2}{s-\ell}). \qquad (28)$$

Furthermore, the left-hand side of (28) *is an evaluation of the Tutte polynomial on H_q only if A is a difference set in Q.*

Taking $A = Q \setminus 0$ in Theorem 3.2 gives

$$\sum_{C+z\in Q^m/C} |C + z \cap (Q \setminus 0)^m|^2 = (q-2)^{n-k(G)}T(G; \frac{q-1}{q-2}, (q-1)^2).$$

When $Q = \mathbb{F}_q$ for prime power $q \equiv -1 \pmod 4$, the set A of non-zero squares in $\mathbb{F}_q$ forms a $(q, \frac{q-1}{2}, \frac{q-3}{4})$-difference set. The set $A \cup 0$ of squares then forms a $(q, \frac{q+1}{2}, \frac{q+1}{4})$-difference set, since $\mathbb{F}_q = A \cup (-A) \cup 0$. Thus by Theorem 3.2, when B is the set of vectors in $\mathbb{F}_q^m$ all of whose components are squares in $\mathbb{F}_q$ and C the subspace of $\mathbb{F}_q$-flows,

$$\sum_{C+z\in \mathbb{F}_q^m/C} |C + z \cap B|^2 = \left(\frac{q+1}{4}\right)^m T(G; 2, q+1).$$

Note that A may not be a difference set in Q and yet give an evaluation of the Tutte polynomial in Theorem 3.2, only not on H_q but a different hyperbola. For example, if P is a subgroup of Q of size p and $A = Q \setminus P$, then $\widehat{1_A} = q1_0 - p1_{P^\sharp}$ which leads to an evaluation of the Tutte polynomial on $H_{q/p}$:

$$\sum_{C+z\in Q^m/C} |C + z \cap (Q \setminus P)^m|^2 = \left(\frac{q-2p}{p}\right)^{n-k(G)} T(G; \frac{q-p}{q-2p}, \left(\frac{q-p}{p}\right)^2).$$

Henceforth we shall just seek evaluations on H_q when Q has order q so as to avoid obscuring arguments by extra complications.

For subset A of Q, let $n_A(x) := \sum_{a\in A} n_a(x)$ denote the number of entries in $x \in Q^m$ belonging to A. Suppose now that A is partitioned into two sets A_0 and A_1 and that

$$B_0 = \{x \in A^m : n_{A_1}(x) \equiv 0 \pmod 2\},\ B_1 = \{x \in A^m : n_{A_1}(x) \equiv 1 \pmod 2\}.$$

Note that $|B_0| = |B_1|$ when $|A_0| = |A_1|$. The set $C + z \cap B$ is partitioned into $C + z \cap B_0$ and $C + z \cap B_1$. How are the differences $|C + z \cap B_0| - |C + z \cap B_1|$ distributed when $C + z$ ranges uniformly over Q^m/C? The expected difference $|C + z \cap B_0| - |C + z \cap B_1|$ is $(|B_0| - |B_1|)/|C^\perp|$, which is equal to zero if $|A_0| = |A_1|$. Turning to the second moment, by Corollary 2.6

$$\sum_{C+z\in Q^m/C} (|C + z \cap B_0| - |C + z \cap B_1|)^2 = \frac{1}{|C^\perp|}\mathsf{cwe}(C^\perp; |\widehat{1_{A_0}} - \widehat{1_{A_1}}|^2). \quad (29)$$

By (27), equation (29) yields an evaluation of the Tutte polynomial precisely when $1_{A_0} \star 1_{A_0} + 1_{A_1} \star 1_{A_1} - 1_{A_0} \star 1_{A_1} - 1_{A_1} \star 1_{A_0}$ is constant on $Q \setminus 0$.

Suppose first that $A = Q$. For a partition of Q into two subsets A_0 and A_1, $\widehat{1_{A_0}} - \widehat{1_{A_1}} = 2\widehat{1_{A_0}} - \widehat{1_Q} = 2\widehat{1_{A_0}} - q1_0$. Hence $|\widehat{1_{A_0}} - \widehat{1_{A_1}}|^2$ is constant on $Q \setminus 0$ if and only if $|\widehat{1_{A_0}}|^2$ is constant on $Q \setminus 0$, i.e. A_0 is a (q, s, ℓ)-difference set in Q, $|\widehat{1_{A_0}} - \widehat{1_{A_1}}|^2 = (2s-q)^2 1_0 + 4(s-\ell)1_{Q\setminus 0}$, and (29) is an evaluation of the Tutte polynomial in this case.

Theorem 3.3 *Let $G = (V, E)$ be a graph on n vertices, m edges and with $k(G)$ components. Let Q be an Abelian group of order q and C the set of Q-flows of G. If A_0 is a (q, s, ℓ)-difference set in Q, $A_1 = Q \setminus A_0$, $B_0 = \{x \in Q^m : n_{A_1}(x) \equiv 0 \pmod 2\}$ and $B_1 = Q^m \setminus B_0$, then*

$$\sum_{C+z \in Q^m/C} (|C + z \cap B_0| - |C + z \cap B_1|)^2$$

$$= [q - 4(s-\ell)]^{n-k(G)} [4(s-\ell)]^{m-n+k(G)} T(G; \frac{q}{q-4(s-\ell)}, \frac{(2s-q)^2}{4(s-\ell)}) \quad (30)$$

Furthermore, the left-hand side of (30) *is an evaluation of the Tutte polynomial on H_q only if A_0 is a difference set in Q.*

For example, when $A_0 = Q \setminus 0$, $A_1 = \{0\}$, the set B_0 comprises elements with an even number of zero entries and B_1 those with an odd number. Here equation (30) is

$$\sum_{C+z \in Q^m/C} \mathsf{hwe}(C + z; -1)^2 = (q-4)^{n-k(G)} 4^{m-n+k(G)} T(G; \frac{q}{q-4}, (\frac{q}{2} - 1)^2). \quad (31)$$

(When $q = 2$ equation (31) is (23) with $t = -1$. When $q = 4$ the right-hand side of equation (31), equal to 4^m, depends only on the number of edges of G.)

Take now $A = Q \setminus 0$, partitioned into subsets A_0 and $A_1 = Q \setminus (A_0 \cup 0)$. The condition now for (29) to be an evaluation of the Tutte polynomial is that $|2\widehat{1_{A_0}} - q1_0 + 1_Q|^2$ is constant on $Q \setminus 0$. It is not difficult to show that this equivalent to requiring that $1_{A_0} \star 1_{A_0} + 1_{A_0 \cup 0} \star 1_{A_0 \cup 0}$ is constant on $Q \setminus 0$ and in turn that this is the case if and only if, for some constants $s(= |A_0|), \ell_0, \ell, \ell_1$,

$$1_{A_0} \star 1_{A_0} = s1_0 + \ell_0 1_{A_0 \cap (-A_0)} + \ell 1_{A_0 \Delta (-A_0)} + \ell_1 1_{A \setminus (A_0 \cup (-A_0))}. \quad (32)$$

Moreover, the constants necessarily satisfy $\ell_1 = \ell_0 + 1$ if $A_0 \cap (-A_0) \neq \emptyset$ and if $A_0 \cap (-A_0) = \emptyset$ it must be the case that $A_0 \cup (-A_0) = A = Q \setminus 0$ (for otherwise the impossible condition $2\ell + 1 = 2\ell_1$ arises).

If A_0 is a (q, s, ℓ_0, ℓ_1)-partial difference set with $\ell_0 \neq \ell_1$ then $A_0 = -A_0$ (since $1_{A_0} \star 1_{A_0}(a) = 1_{A_0} \star 1_{A_0}(-a)$ so that $a \in A_0$ if and only if $-a \in A_0$) and A_0 satisfies (32). If A_0 is a (q, s, ℓ)-difference set then it also satisfies (32) with $\ell_0 = \ell_1 = \ell$. The conditions $\ell_1 = \ell_0 + 1$ if $A_0 \cap (-A_0) \neq \emptyset$ and $A_0 \cup (-A_0) = Q \setminus 0$ if $A_0 \cap (-A_0) = \emptyset$ both force $s = |A_0| = \frac{q-1}{2}$ and the following theorem results.

Theorem 3.4 *Let $G = (V, E)$ be a graph on n vertices, m edges and with $k(G)$ components. Let Q be an Abelian group of odd order $q \geq 3$ and C the set of Q-flows of G. Suppose that $Q \setminus 0$ is partitioned into two subsets A_0 and A_1 each of size $\frac{q-1}{2}$. Let $B_0, B_1 \subseteq (Q \setminus 0)^m$ be defined by*

$$B_0 = \{x \in (Q\backslash 0)^m : n_{A_1}(x) \equiv 0 \,(\mathrm{mod}\, 2)\},\ B_1 = \{x \in (Q\backslash 0)^m : n_{A_1}(x) \equiv 1 \,(\mathrm{mod}\, 2)\}.$$

Then

$$\sum_{C+z \in Q^m/C} \left(|C + z \cap B_0| - |C + z \cap B_1|\right)^2 = (-1)^{n-k(G)} q^{m-n+k(G)} T(G; 1-q, 0) \tag{33}$$

precisely when A_0 (and A_1) is a

$$\begin{cases} (q, \frac{q-1}{2}, \frac{q-3}{4})\text{-difference set in } Q & q \equiv -1 \,(\mathrm{mod}\, 4), \\ (q, \frac{q-1}{2}, \frac{q-5}{4}, \frac{q-1}{4})\text{-partial difference set in } Q & q \equiv 1 \,(\mathrm{mod}\, 4). \end{cases}$$

Furthermore, the only partition of $Q \setminus 0$ into two sets A_0 and A_1 for which the left-hand side of (33) is an evaluation of the Tutte polynomial on H_q is when $|A_0| = |A_1|$ and A_0 is a (partial) difference set, with parameters as above.

There are (see for example [17]) difference sets and partial difference sets not equivalent to the non-zero squares in a finite field but that have the Paley parameters $(q, \frac{q-1}{2}, \frac{q-3}{4})$ or $(q, \frac{q-1}{2}, \frac{q-5}{4}, \frac{q-1}{4})$. Theorem 3.4 with $q = 3$, $A_1 = \{-1\}$ and G the line graph of a cubic graph is equivalent to [23, Theorem 1.1], one of Matiyasevich's restatements of the Four Colour Theorem.[1]

Theorem 3.4 shows that if $A = Q \setminus 0$ is to be partitioned into subsets A_0 and A_1 in such a way that (29) is an evaluation of the Tutte polynomial, then $|Q|$ is odd and $|A_0| = |A_1|$. Consider more generally a partition of $A \subseteq Q$ size s into $r \geq 2$ subsets A_i each of size s/r and indexed by an additive Abelian group I of order r. Define

$$B_i = \bigcup_{\substack{i_1,\ldots,i_m \in I \\ i_1+\cdots+i_m=i}} A_{i_1} \times \cdots \times A_{i_m}, \tag{34}$$

and $B = \cup_{i \in I} B_i = A^m$. Let $\beta : B \to I$ be the function defined by $\beta(x) = i$ if $x \in B_i$.

Corresponding to the case $r = 2$ of $(|C+z\cap B_0| - |C+z\cap B_1|)^2$, the quantity

$$r\left(\sum_{i \in I} |C + z \cap B_i|^2\right) - |C + z \cap B|^2,$$

[1] A chapter of the author's thesis [15] explores Matiyasevich's use of Petersen's 'graph polynomial' in obtaining his results in greater detail. For a cyclic group $Q = \mathbb{Z}_q$ the vector space $\mathbb{C}^{Q^m}$ is isomorphic to the quotient of the vector space $\mathbb{C}[t_1, \ldots, t_m]$ of polynomials in m indeterminates by the ideal generated by the polynomials $t_j^q - 1$, $1 \leq j \leq m$. This links the method of Fourier analysis on Q^m with Alon and Tarsi's more general polynomial method, for which see in particular [3, 4, 1].

always non-negative by the Cauchy-Schwarz inequality, is equal to zero if and only if $|C + z \cap B_i|$ is independent of i. Moreover, since the A_i, and hence the B_i, are all the same size, if $x, y \in B$ are chosen uniformly at random then

$$s^{-2m} \sum_{C+z \in Q^m/C} \left(r \sum_{i \in I} |C + z \cap B_i|^2 - |C + z \cap B|^2 \right)$$

$$= \mathbb{P}(x - y \in C \mid \beta(x) = \beta(y)) - \mathbb{P}(x - y \in C)$$

is the correlation between the event that x, y lie in the same coset of C (i.e. $Tx = Ty$ where $T : Q^m \to Q^n$ is the boundary map, whose kernel is C) and the event that x, y belong to the same set B_i (i.e. $\beta(x) = \beta(y)$).

Theorem 3.5 *Let $G = (V, E)$ be a graph on n vertices, m edges and with $k(G)$ components, Q an Abelian group of order $q \geq 3$ and C the set of Q-flows of G. Suppose that $Q \setminus 0$ is partitioned into r subsets A_i each of size $\frac{q-1}{r}$ and indexed by an additive Abelian group I of order r. Let $B_i \subseteq (Q \setminus 0)^m$ be defined by (34) above and $\beta : (Q \setminus 0)^m \to I$ be defined by $\beta(x) = i$ if $x \in B_i$.*

If x, y are chosen independently uniformly at random from $(Q \setminus 0)^m$, then

$$\mathbb{P}(x - y \in C \mid \beta(x) = \beta(y)) - \mathbb{P}(x - y \in C)$$

$$= (-1)^{n-k(G)}(r-1)(q-1)^{-2m}q^{m-n+k(G)}T(G; 1-q, 0) \tag{35}$$

if and only if $\sum_{i \in I} 1_{A_{i+j}} \star 1_{A_i}$ is constant on $Q \setminus 0$ for all $j \in I$.

Moreover, if any subset A of Q is partitioned into subsets A_i of equal size such that the correlation on the left-hand side of (35) is an evaluation of the Tutte polynomial on H_q, then $|A| = q - 1$ and the evaluation is equal to the right-hand side of (35).

Proof. Suppose $A \subseteq Q$ size s is partitioned into $r \geq 2$ subsets A_i size s/r indexed by I and B_i is defined as above (34). Define for each $j \in I$ the function $h_j : Q \to \mathbb{Z}$ by

$$h_j = \sum_{i \in I} 1_{A_{i+j}} \star 1_{A_i},$$

i.e. $h_j(a)$ counts the total multiplicity of a in the differences $A_{i+j} - A_i$ for $i \in I$. Note that $h_j(0) = s$ for $j = 0$ and $h_j(0) = 0$ for $j \neq 0$

Let $i \mapsto \pi_i$ be an isomorphism of I with the group of characters $\widehat{I}$. Define for each $k \in I$

$$f_k = \sum_{i \in I} \pi_k(i) 1_{A_i}.$$

Note that $f_0 = 1_A$. By orthogonality of the characters of I and by Corollary 2.6,

$$\sum_{C+z \in Q^m/C} r \left(\sum_{i \in I} |C + z \cap B_i|^2 \right) - |C + z \cap B|^2$$

$$= \sum_{0\neq k\in I} \sum_{C+z\in Q^m/C} |\mathrm{cwe}(C+z;f_k)|^2 = \frac{1}{|C^\perp|}\sum_{0\neq k\in I} \mathsf{cwe}(C^\perp;|\widehat{f_k}|^2).$$

With

$$|\widehat{f_k}|^2 = \left|\sum_{i\in I}\pi_k(i)\widehat{1_{A_i}}\right|^2 = \sum_{j\in I}\pi_k(j)\sum_{i\in I}\widehat{1_{A_{i+j}}}\,\overline{\widehat{1_{A_i}}}$$

by (27) and Fourier inversion (2), for each $b\in Q$,

$$\frac{1}{q}\widehat{|\widehat{f_k}|^2}(-b) = \sum_{j\in I}\pi_k(j)\sum_{i\in I} 1_{A_{i+j}}\star 1_{A_i}(b)$$

$$= \sum_{j\in I}\pi_k(j)\sum_{a\in Q} h_j(a)1_a(b) = \sum_{a\in Q}1_a(b)\sum_{j\in J}h_j(a)\pi_k(j).$$

Define for each $a\in Q$ the function $g_a: I\to\mathbb{Z}$ by $g_a(j)=h_j(a)$ for each $j\in I$. Then

$$\frac{1}{q}\widehat{|\widehat{f_k}|^2}(-b) = \sum_a \widehat{g_a}(k)1_a(b),$$

where the Fourier transform on the right-hand side is on $\mathbb{C}^I$, those on the left-hand side on $\mathbb{C}^Q$. Thus $|\widehat{f_k}|^2$ is constant on $Q\setminus 0 \Leftrightarrow \widehat{g_a}=\widehat{g_b}$ for all $a,b\in Q\setminus 0$ $\Leftrightarrow g_a=g_b$ for all $a,b\in Q\setminus 0 \Leftrightarrow h_j(a)=h_j(b)$ for all $a,b\in Q\setminus 0$ and $j\in I$, i.e. h_j is a constant ℓ_j on $Q\setminus 0$ for all $j\in I$. If $h_0(a)=\ell_0$ for $a\neq 0$ then $r\frac{s}{r}(\frac{s}{r}-1)=(q-1)\ell_0$; similarly for $j\neq 0$, if $h_j(a)=\ell_j$ for $a\neq 0$ then $r(\frac{s}{r})^2=(q-1)\ell_j$. Hence, for $|\widehat{f_k}|^2$ to be constant on $Q\setminus 0$ it is necessary and sufficient that

$$h_j = \sum_{i\in I}1_{A_{i+j}}\star 1_{A_i} = \begin{cases} s1_0+\frac{s(s-r)}{(q-1)r}1_{Q\setminus 0} & j=0,\\ \frac{s^2}{(q-1)r}1_{Q\setminus 0} & j\neq 0.\end{cases}$$

Note then that since $\frac{s}{(q-1)r}(s-(s-r))=\frac{s}{q-1}\in\mathbb{Z}$, the integer s must be a multiple of $q-1$. Since $s\leq q$ and $q\geq 3$ it follows that $s=q-1$ and up to translation A can be assumed to be equal to $Q\setminus 0$, partitioned into $r\mid q-1$ subsets A_i. For $a\neq 0$,

$$\widehat{g_a}(k) = \sum_{j\in I}\pi_k(j)h_j(a) = \frac{s(s-r)}{(q-1)r} - \frac{s^2}{(q-1)r} = -\frac{s}{q-1} = -1$$

and $\widehat{g_0}(k)=\sum_{j\in I}h_j(0)=s=q-1$. Hence

$$\frac{1}{q}\widehat{|\widehat{f_k}|^2} = \sum_{a\in Q}1_a\widehat{g_a}(k) = (q-1)1_0 - 1_{Q\setminus 0}$$

and it follows that, for each $0\neq k\in I$, $|\widehat{f_k}|^2 = q1_{Q\setminus 0}$. This establishes the theorem.□

To illustrate Theorem 3.5 with a concrete example, take $Q = \mathbb{F}_q$, $A = \mathbb{F}_q \setminus 0 = \mathbb{F}_q^\times$ and $I = \mathbb{Z}_r$. Let c be a multiplicative generator of $\mathbb{F}_q^\times$ and define $A_i = \{c^j \in \mathbb{F}_q^\times : j \equiv i \,(\mathrm{mod}\, r)\}$, so that A_0 is the set of rth powers in $\mathbb{F}_q^\times$. Let τ be an order r character of the multiplicative group $\mathbb{F}_q^\times$, so that $\ker \tau = A_0$, and set $\tau(0) = 0$. For $k \in \mathbb{Z}_r$ define τ^k to be the kth power of τ (i.e. regarding the exponent k as an integer). Note that $\tau^0 = 1_{\mathbb{F}_q \setminus 0}$. The Fourier transform of $\tau^k : \mathbb{F}_q \to \mathbb{C}$ is the *Gauss sum*

$$\widehat{\tau^k}(b) = \sum_{a \in Q} \tau^k(a)\chi_b(a),$$

where $b \mapsto \chi_b$ is an isomorphism $\mathbb{F}_q \to \widehat{\mathbb{F}_q}$. It is a well known result (see for example [6]) that $|\widehat{\tau^k}|^2 = q1_{\mathbb{F}_q \setminus 0}$ for $k \neq 0$.

Theorem 3.6 [16] *Let $G = (V, E)$ be a graph on n vertices, m edges, $k(G)$ components, and $T : \mathbb{F}_q^m \to \mathbb{F}_q^n$ the boundary operator. Let $x = (x_1, \ldots, x_m)$ and $y = (y_1, \ldots, y_m)$ be chosen uniformly at random from $(\mathbb{F}_q^\times)^m$. Suppose τ is a multiplicative character of order r on $\mathbb{F}_q^\times$. Then*

$$\begin{aligned}&\mathbb{P}(Tx = Ty \mid \tau(x_1 \cdots x_m) = \tau(y_1 \cdots y_m)) - \mathbb{P}(Tx = Ty)\\ &= (-1)^{n-k(G)}(r-1)(q-1)^{-2m}q^{m-n+k(G)}T(G; 1-q, 0).\end{aligned}$$

Phylogenetic trees

In our final graphical application a different linear transformation $T : Q^m \to Q^n$ is considered with which to use the machinery set up in Section 2.

Rooted trees have been widely used in phylogenetics to model evolution; see for example [26, 12, 30, 28, 29] to add further detail and context to the following account. Let $\mathcal{T}$ be a rooted tree with n leaves and m edges. The tree $\mathcal{T}$ represents the evolution of a set of n taxa: the leaves of the tree are the observed taxa, the interior nodes representing hypothetical ancestors to subsets of the taxa, and the root a hypothetical common ancestor.

Label the leaves $\mathcal{T}$ arbitrarily by $[n] := \{1, 2, \ldots, n\}$ and label the edges by $[m]$ in such as way that an edge labelled $i \in [n]$ has as an endpoint the leaf labelled i (and other edges are labelled arbitrarily by $[m] \setminus [n]$). Define for each leaf i the *path* $P(i)$ to be the set of edges on the unique path in $\mathcal{T}$ from the root to the leaf i. Define for each edge $j \in [m]$ the *cluster* $C(j)$ by $C(j) = \{i \in [n] : j \in P(i)\}$. Deleting the edge j splits the tree into two components, one containing the root, the other containing the leaves in $C(j)$. The *initial* vertex of an edge j is the endpoint nearer the root, and its *final* vertex the endpoint nearer the leaves in $C(j)$.

Let Q be an additive Abelian group of order q. A Markov process on vertex Q-colourings of $\mathcal{T}$ is defined by assigning to each edge j a transition probability $p_j : Q \times Q \to [0, 1]$, described by a square matrix indexed by Q, with (a, b) entry $p_j(a, b)$ the conditional probability that the final vertex of j is coloured b given

that the initial vertex of j is coloured a. When $p_j(a,b) =: f_j(b-a)$ depends only on the colour difference $c = b-a$, a simpler desciption of this model results. For each edge $j \in [m]$ and $c \in Q$ there is associated a probability $f_j(c)$ that the colour change along edge j is c: the probability that the final vertex of j is coloured $a+c$ given that the initial vertex of j is coloured a.

In order to use the results of Section 2, assume that Q has the further structure of a commutative ring with a generating character. Define the linear transformation $T : Q^m \to Q^n$ for each $i \in [n]$ and $x \in Q^m$ by

$$(Tx)_i = \sum_{j \in P(i)} x_j,$$

colouring the leaf i with the sum of the colours on the path from the root to i.

The transpose $T^t : Q^n \to Q^m$ is the linear transformation

$$(T^t y)_j = \sum_{i \in C(j)} y_i,$$

colouring the edge j with the sum of the colours on the leaves that it separates from the root. Let $C = \ker T$, for which $|C| = q^{m-n}$. By Lemma 2.2, $C^\perp = \operatorname{im} T^t$ and $|C^\perp| = q^n$.

The Jukes-Cantor model for evolution of purine-pyrimidine sequences takes $Q = \mathbb{Z}_2$, and the Kimura model takes $Q = \mathbb{Z}_2 \oplus \mathbb{Z}_2$, the additive group of $\mathbb{F}_4 = \{0, 1, \omega, \overline{\omega}\}$, for encoding nucleotide sequences —the purines, adenine and guanine, and the pyrimidines, cystosine and thymine. *Transitions* are substitutions within a family, *transversions* substitutions between families. If we take (adenine, guanine, cystosine, thymine) $= (0, 1, \omega, \overline{\omega})$ then a transition corresponds to adding an element of $\{0, 1\}$, a transversion to adding an element of $\{\omega, \overline{\omega}\}$.

The Kimura 3-parameter model takes the transition probability at an edge j to be of the form

$$f_j = p_0 1_0 + p_1 1_1 + p_\omega 1_\omega + p_{\overline{\omega}} 1_{\overline{\omega}},$$

where $p_0 + p_1 + p_\omega + p_{\overline{\omega}} = 1$, for which

$$\widehat{f_j} = 1_0 + [1 - 2(p_\omega + p_{\overline{\omega}})]1_1 + [1 - 2(p_1 + p_\omega)]1_\omega + [1 - 2(p_1 + p_{\overline{\omega}})]1_{\overline{\omega}}.$$

The Kimura 2-parameter model distinguishes only between transitions and transversions, i.e. $p_\omega = p_{\overline{\omega}}$,

$$f_j = p_0 1_0 + p_1 1_1 + p_\omega 1_{\{\omega, \overline{\omega}\}},$$

where $p_0 + p_1 + 2p_\omega = 1$, and

$$\widehat{f_j} = 1_0 + [1 - 4p_\omega]1_1 + [1 - 2(p_1 + p_\omega)]1_{\{\omega, \overline{\omega}\}}.$$

The Jukes-Cantor model does not distinguish between the three types of substitution, i.e. $p_1 = p_\omega = p_\omega := p/3$,

$$f_j = (1-p)1_0 + \frac{p}{3} 1_{\{1, \omega, \overline{\omega}\}},$$

$$\widehat{f_j} = 1_0 + (1 - \frac{4p}{3})1_{\{1,\omega,\overline{\omega}\}}.$$

The Kimura models are analogous to the Ashkin-Teller model in non-symmetric and symmetric versions, the Jukes-Cantor model to the 4-state Potts model.

Let $f : Q^m \to [0,1]$ be a probability distribution on the edge Q-colourings of $\mathcal{T}$, edge j receiving colour c with probability $f_j(c)$, $f_j = \sum_{c\in Q} f_j(c)1_c$, and $f(x) = f_1(x_1)\cdots f_m(x_m)$ for $x = (x_1, \ldots, x_m) \in Q^m$. Then for random variable x on Q^m with probability distribution f, the leaf-colouring $Tx = y$ is a random variable on Q^n with probability distribution

$$f_T(y) := \sum_{x\in Q^m,\, Tx=y} f(x),$$

given by

$$f_T(y) = f * 1_{\ker T}(y') = f(C + y'),$$

where $y' \in Q^m$ is any edge colouring such that $Ty' = y$. A convenient choice for y' is to set $y'_j = y_j$ for $j \in [n]$ and $y'_j = 0$ for $j \in [m] \setminus [n]$. Hence by the Poisson summation formula (8), for $y \in Q^n$,

$$f_T(y) = q^{-n} \sum_{x\in \operatorname{im} T^t} \widehat{f}(x)\chi_x(y') = q^{-n} \sum_{x\in \operatorname{im} T^t} \widehat{f}(x)\chi_y(x'), \tag{36}$$

where $x' \in Q^n$ is the truncation of x to $[n]$, $x'_i = x_i$ for $i \in [n]$. Identity (36) is a restatement in different language of a known result: for $Q = \mathbb{Z}_2$ and $f_j = (1-p_j)1_0 + p_j 1_1$, $\widehat{f_j} = 1_0 + (1-2p_j)1_1$ it is a theorem of Hendy and Penny [19], which was subsequently generalised to elementary Abelian 2-groups by Székely et al. [28], and finally to any Abelian group Q by Székely et al. [30]. To (36) we can add the following.

Theorem 3.7 *Suppose $f : Q^m \to [0,1]$ defines a probability distribution on the edge Q-colourings of a phylogenetic tree $\mathcal{T}$ and that x_1, x_2 are independent random variables each with probability density function f. Then, the probability that the same leaf Q-colouring of $\mathcal{T}$ results from x_1 and x_2 is given by*

$$\mathbb{P}(Tx_1 = Tx_2) = q^{-n} \sum_{x\in \operatorname{im} T^t} |\widehat{f}(x)|^2.$$

In particular, if f is defined on each edge $j \in [m]$ by $f_j = (1-p)1_0 + \frac{p}{q-1}1_{Q\setminus 0}$, where $0 \le p \le 1$, and y_1, y_2 are independent random variables on Q^n with probability distribution f_T then

$$\mathbb{P}(y_1 = y_2) = q^{-n} \sum_{x\in \operatorname{im} T^t} \left(1 - \frac{pq}{q-1}\right)^{2|x|}.$$

The problem for phylogenetics is how far it is possible to determine the tree $\mathcal{T}$ from knowledge of (an estimate of) the distribution f_T alone.

References

[1] N. Alon. Combinatorial Nullstellensatz. *Combinatorics, Probability and Computing*, 8(1 & 2):7–29, 1999.

[2] N. Alon, I. Dinur, E. Friedgut, and B. Sudakov. Graph products, Fourier analysis and spectral techniques. *Geometric And Functional Analysis*, 14(5):913–940, 2004.

[3] N. Alon and M. Tarsi. Colorings and orientations of graphs. *Combinatorica*, 12:125–134, 1992.

[4] N. Alon and M. Tarsi. A note on graph colorings and graph polynomials. *Journal of Combinatorial Theory Series B*, 70:197–201, 1997.

[5] J. Ashkin and E. Teller. Statistics of two-dimensional lattices with four components. *Physical Review*, 64(5 & 6):178–184, 1943.

[6] B. Berndt, R. Evans, and K. Williams. *Gauss and Jacobi sums*. Wiley-Interscience, 1998.

[7] T. Beth, D. Jungnickel, and H. Lenz. *Design Theory*, volume 1. Cambridge University Press, Cambridge, 2nd edition, 1999.

[8] N. Biggs. On the duality of interaction models. *Mathematical Proceedings of the Cambridge Philosophical Society*, 80:429–436, 1976.

[9] N. Biggs. *Interaction Models*. Cambridge University Press, 1977.

[10] N. Biggs. *Algebraic Graph Theory*. Cambridge University Press, Cambridge, 2nd edition, 1993.

[11] B. Bollobás. *Modern Graph Theory*. Springer, New York, 1998.

[12] S. Evans. Fourier analysis and phylogenetic trees. *Modern Signal Processing*, 46:117–136, 2003.

[13] G. Farr. Tutte-Whitney polynomials: some history and generalisations. In this volume.

[14] G. Farr. A generalization of the Whitney rank generating function. *Mathematical Proceedings of the Cambridge Philosophical Society*, 113:267–280, 1993.

[15] A. Goodall. *Graph polynomials and the discrete Fourier transform*. PhD thesis, University of Oxford, 2004.

[16] A. Goodall. Some new evaluations of the Tutte polynomial. *Journal of Combinatorial Theory Series B*, 96:207–224, 2006.

[17] M. Hall. *Combinatorial Theory*. Wiley-Interscience, New York, 1986.

[18] B. Hartley and T. Hawkes. *Rings, Modules and Linear Algebra*. Chapman and Hall, London, 1970.

[19] M. Hendy and D. Penny. A framework for the quantative study of evolutionary trees. *Systematic Zoology*, 38(4):297–309, 1989.

[20] W. Huffman and V. Pless. *Fundamentals of Error-correcting Codes*. Cambridge University Press, Cambridge, 2003.

[21] G. Kalai and M. Safra. Threshold phenomena and influence. In A. Percus, G. Istrate, and C. Moore, editors, *Computational Complexity and Statistical Physics*. Oxford University Press, New York, 2005.

[22] S. Ma. A survey of partial difference sets. *Designs, Codes and Cryptography*, 4(4):221–261, 1994.

[23] Y. Matiyasevich. Some probabilistic restatements of the four color conjecture. *Journal of Graph Theory*, 46:167–179, 2004.

[24] A. Pott. Nonlinear functions in abelian groups and relative difference sets. *Discrete Applied Mathematics*, 138:177–193, 2004.

[25] R. C. Read and P. Rosenstiehl. On the principal edge tripartition of a graph. *Annals of Discrete Mathematics*, pages 195–226, 1978.

[26] C. Semple and M. Steel. *Phylogenetics*. Oxford University Press, Oxford, 2003.

[27] R. Stanley. *Enumerative Combinatorics*, volume 2. Cambridge University Press, 1999.

[28] L. Székely, P. Erdős, M. Steel, and D. Penny. A Fourier inversion formula for evolutionary trees. *Applied Mathematics Letters*, 6(2):13–16, 1993.

[29] L. Székely, M. Steel, and P. Erdős. Fourier calculus on finite sets and evolutionary trees. Technical report, 1991.

[30] L. Székely, M. Steel, and P. Erdős. Fourier calculus on evolutionary trees. *Advances in Applied Mathematics*, 14:200–216, 1993.

[31] A. Terras. *Fourier analysis on finite groups and applications*. Cambridge University Press, 1999.

[32] D. Vertigan. Latroids and their representation by codes over modules. *Transactions of the American Mathematical Society*, 356(10):3841–3868, 2003.

[33] D. Welsh. *Complexity: Knots, Colourings and Counting*. Cambridge University Press, 1993.

[34] D. Welsh. The Tutte polynomial. *Random Structures and Algorithms*, 15:210–228, 1999.

[35] D. Welsh and C. Merino. The Potts model and the Tutte polynomial. *Journal of Mathematical Physics*, 41(3):1127–1152, 2000.

[36] J. Wood. Duality for modules over finite rings and applications to coding theory. *American Journal of Mathematics*, 121(3):555–575, 1999.