ABSTRACT

RANDOM WALKS ON FINITE FIELDS AND HEISENBERG GROUPS

by

Yi Zhu

April, 2011

Chair: Dr. Chal Benson

Major Department: Mathematics

Let H be a finite group and μ a probability measure on H. This data determines an invariant random walk on H beginning from the identity element. The probability distribution for the state of the random walk after n steps is given by the n'th convolution power of the probability measure μ. The random walk and measure μ are said to be ergodic if the support of this distribution is the entire group for n sufficiently large. In this case a specialization of the Markov Ergodic Theorem ensures that the distribution after n steps converges point-wise to the uniform distribution. One employs the total variation distance on probability measures to analyze the rate of convergence to equilibrium. Suppose now that a finite group K acts on H by automorphisms. We say that the action pair $K : H$ is ergodic when the K-invariant probability measure μ supported on some K-orbit is ergodic. We call, moreover, $K : H$ a Gelfand action pair when the convolution algebra of K-invariant functions on H is commutative. Specializing the theory of spherical functions to the context of Gelfand action pairs we obtain a version of the Diaconis-Shahshahani Upper Bound Lemma, controlling the total variation distance to equilibrium for the random walk determined by μ.

The main results in this thesis concern invariant random walks on finite fields and three dimensional Heisenberg groups over finite fields. Let F be a finite field of odd characteristic and K a subgroup of the multiplicative group for F with even order. We obtain a necessary and sufficient condition for ergodicity of the action pair $K : F$ and an explicit summation formula for the upper bound on total variation distance to equilibrium guaranteed by the Upper Bound Lemma. Let $\widetilde{F}$ be a quadratic extension field for F and U denote the kernel of the norm mapping from $\widetilde{F}$ to F. An application of our field theoretic criterion for ergodicity shows that $U : \widetilde{F}$ is an ergodic action pair and we specialize our upper bound result to this context. Forming the three dimensional Heisenberg group $H = \widetilde{F} \times F$ over F the action of U on $\widetilde{F}$ induces an action of U on H by automorphisms. Benson and Ratcliff have shown that $U : H$ is a Gelfand action pair and determined the associated spherical functions. We prove that the pair $U : H$ is ergodic and make explicit the bound given by the Upper Bound Lemma.

RANDOM WALKS ON FINITE FIELDS AND HEISENBERG GROUPS

A Thesis

Presented to

The Faculty of the Department of Mathematics

East Carolina University

In Partial Fulfillment

of the Requirements for the Degree

Master of Arts in Mathematics

by

Yi Zhu

April, 2011

RANDOM WALKS ON FINITE FIELDS AND HEISENBERG GROUPS

by

Yi Zhu

APPROVED BY:

DIRECTOR OF THESIS:

Dr. Chal Benson

COMMITTEE MEMBER:

Dr. Johannes Hattingh

COMMITTEE MEMBER:

Dr. Gail Ratcliff

COMMITTEE MEMBER:

Dr. Kevin O'Brien

CHAIR OF THE DEPARTMENT
OF MATHEMATICS:

Dr. Johannes Hattingh

DEAN OF THE
GRADUATE SCHOOL:

Dr. Paul Gemperline

ACKNOWLEDGEMENTS

First of all, I would like to express my deepest gratitude to my thesis advisor, Dr.Chal Benson, for without his constant expertise and patience this would not have been possible.

I would also like to give special acknowledgement to Dr.Gail Ratcliff, for all of her guidance which gave me a deeper understanding on mathematics.

Finally, I am extremely grateful for my family for everything they have done for me. This thesis is dedicated to them.

TABLE OF CONTENTS

CHAPTER 1: PRELIMINARIES: ANALYSIS ON FINITE GROUPS

In this chapter we summarize background material concerning harmonic analysis with finite groups needed for our subsequent study of certain discrete random walks with invariance properties. General references for this material are [4] and [9].

1.1 The space $L(X)$

For a finite non-empty set X let

$$L(X) := \{f : X \to \mathbb{C}\}.$$

denote the set of all complex-valued functions on X. This is a vector space over the field $\mathbb{C}$ in the usual way. We adopt the following notation and terminology.

- For a subset $A \subset X$ write $\delta_A \in L(X)$ for the characteristic function

$$\delta_A(x) = \begin{cases} 1 & \text{if } x \in A \\ 0 & \text{if } x \notin A \end{cases}. \tag{1.1}$$

- For points $x_0 \in X$, we usually write δ_{x_0} in place of $\delta_{\{x_0\}}$. Thus

$$\delta_{x_0}(x) = \delta_{x_0,x} = \begin{cases} 1 & \text{if } x = x_0 \\ 0 & \text{if } x \neq x_0 \end{cases}.$$

- The *support* of a function $f \in L(X)$ is $Supp(f) := \{x \in X : f(x) \neq 0\}$. So for example, $Supp(\delta_A) = A$, $Supp(\delta_{x_\circ}) = \{x_\circ\}$.

Observe that the point masses $\{\delta_{x_0} : x_0 \in X\}$ form a basis for the vector space $L(X)$. Indeed each function $f \in L(X)$ can be written uniquely as a linear combination of the point masses via

$$f = \sum_{x_\circ \in X} f(x_\circ)\delta_{x_\circ}.$$

In particular, $L(X)$ is finite dimensional with $dim(L(X)) = |X|$.

1.1.1 Inner product and norms

We define a Hermitian inner product on the complex vector space $L(X)$ via

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

for $f, g \in L(X)$. As

$$\langle \delta_{x_1}, \delta_{x_2}\rangle = \sum_{y\in X} \delta_{x_1}(y)\overline{\delta_{x_2}(y)} = \delta_{x_1,x_2}$$

for points $x_1, x_2 \in X$ we see that $\{\delta_{x_\circ} : x_\circ \in X\}$ is an *orthonormal* basis for $L(X)$ with respect to the inner product $\langle\cdot,\cdot\rangle$.

The l^1- and l^2-norms on the complex vector space $L(X)$ are given by

$$\|f\|_1 := \sum_{x\in X} \mid f(x) \mid, \qquad \|f\|_2 := \sqrt{\langle f, f\rangle} = \left(\sum_{x\in X} \mid f(x) \mid^2\right)^{\frac{1}{2}}.$$

The *Cauchy-Schwarz inequality* asserts that

$$\mid \langle f, g\rangle \mid \leqslant \|f\|_2\|g\|_2$$

(see, for example, [7, Page 333] for a proof) and the two norms are related by the following inequalities:

Lemma 1.1.

$$\frac{1}{\sqrt{\mid X \mid}}\|f\|_1 \leq \|f\|_2 \leq \|f\|_1 \leq \sqrt{\mid X \mid}\,\|f\|_2.$$

Proof. To establish the inequality $\|f\|_1 \leq \sqrt{\mid X \mid}\|f\|_2$ observe that by Cauchy-

Schwarz

$$\|f\|_1 = |\langle |f|, 1\rangle| \leq \|f\|_2 \|1\|_2 = \sqrt{|X|}\,\|f\|_2.$$

To see that $\|f\|_2 \leq \|f\|_1$ first observe that this holds for $f \equiv 0$. So suppose now that $f \not\equiv 0$ and let $c := \|f\|_1 (> 0)$ and $g := (1/c)f$. We have $\|g\|_1 = \|f\|_1/c = 1$ and $|g(x)| \leq (\|g\|_1 = 1)$ for all $x \in X$. Thus also $|g(x)|^2 \leq |g(x)|$ for each $x \in X$ and now

$$\|g\|_2 = \left[\sum_{x\in X} |g(x)|^2\right]^{1/2} \leq \left[\sum_{x\in X} |g(x)|\right]^{1/2} = (\|g\|_1)^{1/2} = 1.$$

As $\|g\|_2 = \|(1/c)f\|_2 = (1/c)\|f\|_2 = \|f\|_2/\|f\|_1$ this gives $\|f\|_2 \leq \|f\|_1$ as claimed. □

1.1.2 Convergence in $L(X)$

Given a sequence $(f_m)_{m=1}^{\infty}$ in $L(X)$ and $f \in L(X)$ we say (f_m) *converges* (pointwise) to f and write $f_m \to f$ when $\lim_{m\to\infty} f_m(x) = f(x)$ for all $x \in X$. But we can also use any of our norms to characterize convergence. That is,

$$f_m \to f \iff \lim_{m\to\infty} \|f_m - f\|_1 = 0 \iff \lim_{m\to\infty} \|f_m - f\|_2 = 0. \tag{1.2}$$

Indeed, as the sums in the formulas for $\|f_m - f\|_1$ and $\|f_m - f\|_2$ are finite it is clear that these norms converge to zero as m tends to infinity if and only if $|f_m(x) - f(x)|$ converges to zero for every $x \in X$.

1.2 Probability measures

Definition 1.2. A function $\mu \in L(X)$ is a *probability measure* (or *distribution*) if μ is real valued with $0 \leq \mu(x) \leq 1$, for all $x \in X$ and $\sum_{x\in X} \mu(x) = 1$.

The *uniform distribution* on X is the probability measure

$$\mathfrak{u}(x) = \frac{1}{|X|}.$$

More generally for nonempty subsets $A \subset X$ we let μ_A denote the probability measure

$$\mu_A = \frac{1}{|A|}\delta_A. \tag{1.3}$$

Thus $\mathfrak{u} = \mu_X$.

Let (μ_m) be a sequence of probability measures and suppose that $\mu_m \to \nu$ in $L(X)$. Then ν is also a probability measure. Note that the set

$$\mathcal{M}(X) := \{\mu \in L(X) : \mu \text{ a probability measure}\}$$

is not a vector subspace of $L(X)$. Nonetheless it makes sense to consider the norm distance $\|\mu - \nu\|$ (using $\|\cdot\|_1$ or $\|\cdot\|_2$) between pairs of probability measures and to use such distance when considering convergence of sequences of probability measures. In fact, however, it is standard to employ another notation of distance between probability measures-the *total variation distance*:

$$\|\mu - \nu\|_{TV} := \max_{A \subset X} |\mu(A) - \nu(A)| \tag{1.4}$$

where here for $\pi \in \mathcal{M}(X)$ and $A \subset X$, $\pi(A) := \sum_{a \in A} \pi(a)$, so

$$\|\mu - \nu\|_{TV} = \max_{A \subset X} \left| \sum_{a \in A} (\mu(a) - \nu(a)) \right|.$$

Note that $0 \leq \|\mu - \nu\|_{TV} \leq 1$. In fact the total variation and l^1-metrics are related

as follows.

Lemma 1.3. *For all $\mu, \nu \in \mathcal{M}(X)$ one has*

$$||\mu - \nu||_{TV} = \frac{1}{2}||\mu - \nu||_1.$$

Proof. (See [4, Page 37].) Given $\mu, \nu \in \mathcal{M}(X)$ let

$$X^+ := \{x \in X : \mu(x) > \nu(x)\}$$

$$X^\circ := \{x \in X : \mu(x) = \nu(x)\}$$

$$X^- := \{x \in X : \mu(x) < \nu(x)\}$$

so that $X = X^+ \coprod X^\circ \coprod X^-$. Now for any subset $A \subset X$ we have

$$\begin{aligned}
|\mu(A) - \nu(A)| &= \left| \sum_{a \in A \cap X^+} (\mu(a) - \nu(a)) + \sum_{a \in A \cap X^-} (\mu(a) - \nu(a)) \right| \\
&\leq \max\left(\left| \sum_{a \in A \cap X^+} (\mu(a) - \nu(a)) \right|, \left| \sum_{a \in A \cap X^-} (\mu(a) - \nu(a)) \right| \right) \\
&= \max\left(\left| \mu(A \cap X^+) - \nu(A \cap X^+) \right|, \left| \mu(A \cap X^-) - \nu(A \cap X^-) \right| \right) \\
&= \max\left(\mu(X^+) - \nu(X^+),\, \nu(X^-) - \mu(X^-) \right).
\end{aligned}$$

But

$$\mu(X^+) + \mu(X^-) + \mu(X^\circ) = \mu(X) = 1 = \nu(X) = \nu(X^+) + \nu(X^-) + \nu(X^\circ)$$

and $\mu(X^\circ) = \nu(X^\circ)$ hence $\mu(X^+) - \nu(X^+) = \nu(X^-) - \mu(X^-)$. So we conclude that

$$\|\mu - \nu\|_{TV} = \max_{A \subset X} |\mu(A) - \nu(A)|$$

$$= \mu(X^+) - \nu(X^+) = \nu(X^-) - \mu(X^-).$$

Finally we observe that

$$\begin{aligned}
\|\mu - \nu\|_1 &= \sum_{x \in X} |\mu(x) - \nu(x)| \\
&= \sum_{x \in X^+} (\mu(x) - \nu(x)) + \sum_{x \in X^-} (\nu(x) - \mu(x)) \\
&= \big(\mu(X^+) - \nu(X^+)\big) + \big(\nu(X^-) - \mu(X^-)\big) \\
&= 2\|\mu - \nu\|_{TV}.
\end{aligned}$$

□

Now given a sequence of probability measures $(\mu_m)_{m=1}^{\infty}$, and $\nu \in \mathcal{M}(X)$, we have, in view of Lemma 1.3 and (1.2),

$$\mu_m \to \nu \iff \|\mu_m - \nu\|_{TV} \to 0.$$

The following example, from [5, Page 25], explains why "total variation" is the *right* distance to use on the space $\mathcal{M}(X)$.

Example 1.4. Let $X = \{1, 2, \cdots, 2m\}$, $A = \{1, 2, \cdots, m\}$, and consider the probability measures $\mathfrak{u}$, $\mu_A \in \mathcal{M}(X)$. We have $\mathfrak{u}(x) = \frac{1}{2m}$ for all x and

$$\mu_A(x) = \frac{1}{|A|}\delta_A(x) = \begin{cases} \frac{1}{m} & \text{if } 1 \le x \le m \\ 0 & \text{if } x > m \end{cases}.$$

Thus

$$\|\mu_A - \mathfrak{u}\|_{TV} = \frac{1}{2}\|\mu_A - \mathfrak{u}\|_1 = \frac{1}{2}\sum_{j=1}^{2m} |\mu_A(j) - \mathfrak{u}(j)| = \frac{1}{2}\left(\sum_{j=1}^{m} |\frac{1}{m} - \frac{1}{2m}| + \sum_{j=m+1}^{2m} |0 - \frac{1}{2m}|\right)$$

$$= \frac{1}{2}\left(m \cdot \frac{1}{2m} + m \cdot \frac{1}{2m}\right) = \frac{1}{2}.$$

On the other hand

$$\|\mu_A - \mathfrak{u}\|_2 = \left(\sum_{j=1}^{2m} |\mu_A(j) - \mathfrak{u}(j)|^2\right)^{\frac{1}{2}} = \left(\sum_{j=1}^{m} (\frac{1}{2m})^2 + \sum_{j=m+1}^{2m} (-\frac{1}{2m})^2\right)^{\frac{1}{2}}$$
$$= \left(m \cdot \frac{1}{4m^2} + m \cdot \frac{1}{4m^2}\right)^{\frac{1}{2}} = \left(\frac{1}{2m}\right)^{\frac{1}{2}} = \frac{1}{\sqrt{2m}}.$$

Therefore,

$$\|\mu_A - \mathfrak{u}\|_{TV} = \frac{1}{2} \quad \text{whereas} \quad \|\mu_A - \mathfrak{u}\|_2 = \frac{1}{\sqrt{2m}}.$$

Intuitively μ_A and $\mathfrak{u}$ are not close and the distance should not converges to 0 as m increases.

1.3 Convolution

Next let H be a finite group. We use multiplicative notation for the group operation $(H \times H \to H, (x, y) \to xy)$ and $e \in H$ is the identity element. For subsets $A, B \subset H$

$$AB = \{ab : a \in A, b \in B\} = \{h \in H : h = ab \text{ for some elements } a \in A, b \in B\}$$

is the set of all products of elements from A with elements from B. Also $A^2 = AA$,

$$A^m = \underbrace{AA \cdots A}_{m \text{ times}} = \{a_1 \cdots a_m : a_1, \cdots, a_m \in A\}.$$

Definition 1.5. For functions $f, g \in L(H)$ the convolution product $f \star g \in L(H)$ is

$$(f \star g)(x) = \sum_{y \in H} f(y) g(y^{-1}x) \tag{1.5}$$

The space $L(H)$ equipped with the convolution product is called the *group algebra*.

The convolution product can also be written as $(f \star g)(x) = \sum_{z\in H} f(xz^{-1})g(z) = \sum_{z\in H} f(z^{-1})g(zx)$ and

$$(f \star g)(x) = \sum_{ab=x} f(a)g(b) \tag{1.6}$$

where the sum is over all pairs $(a, b) \in H \times H$ with $ab = x$.

We compute that $\delta_{x_\circ} \star \delta_{x_1} = \delta_{x_\circ x_1}$ for any $x_\circ, x_1 \in H$. From this we conclude that the group algebra is commutative (i.e $f \star g = g \star f$) if and only if H is an abelian group. One has $\delta_e \star f = f = f \star \delta_e$, for all $f \in L(H)$, so the point mass δ_e is an identity element for the convolution product.

Convolution has the following properties. For all $f, g, h \in L(H)$, $c \in \mathbb{C}$:

$$\begin{gathered}(f + g) \star h = f \star h + g \star h,\\ f \star (g + h) = f \star g + f \star h,\\ (cf) \star g = c(f \star g) = f \star (cg),\\ (f \star g) \star h = f \star (g \star h).\end{gathered}$$

It is straightforward to check the first three of these identities. As regards associativity we compute

$$\begin{aligned}((f \star g) \star h)(x) &= \sum_{y\in H} (f \star g)(y)h(y^{-1}x)\\ &= \sum_{y\in H}\sum_{z\in H} f(z)g(z^{-1}y)h(y^{-1}x) \quad \text{let } y = zt,\ z^{-1}y = t,\ t^{-1}z^{-1}x = y^{-1}x\\ &= \sum_{z\in H}\sum_{t\in H} f(z)g(t)h(t^{-1}z^{-1}x)\\ &= \sum_{z\in H} f(z)(g \star h)(z^{-1}x)\\ &= (f \star (g \star h))(x),\end{aligned}$$

and hence $(f \star g) \star h = f \star (g \star h)$.

For the uniform distribution $\mathfrak{u}$, we compute

$$\mathfrak{u} \star f = s(f)\mathfrak{u} = f \star \mathfrak{u},$$

for any $f \in L(H)$ where $s(f) = \sum_{x \in H} f(x)$. In particular for any probability measure $\nu \in \mathcal{M}(H)$, we have $\mathfrak{u} \star \nu = \mathfrak{u} = \nu \star \mathfrak{u}$.

Lemma 1.6. *For $f, g \in L(H)$, one has $Supp(f \star g) \subset Supp(f)Supp(g)$. Moreover if f, g are non-negative real valued functions, then $Supp(f \star g) = Supp(f)Supp(g)$. In particular $Supp(\mu \star \nu) = Supp(\mu)Supp(\nu)$ holds for for all $\mu, \nu \in \mathcal{M}(H)$.*

Proof. This follows immediately from (1.6). □

Now writing

$$\mu^{\star m} := \underbrace{\mu \star \cdots \star \mu}_{m \text{ times}}$$

for probability measures $\mu \in \mathcal{M}(H)$ one has

$$Supp(\mu^{\star m}) = Supp(\mu)^m. \tag{1.7}$$

1.4 Characters

Definition 1.7. Let H be a finite group. A *character* on H is a group homomorphism

$$\chi : H \to \mathbb{C}^\times$$

where $\mathbb{C}^\times$ denotes the multiplicative group of non-zero complex numbers.

Observe that if $|H| = m$, then for all $x \in H$ we have $1 = \chi(e) = \chi(x^m) = \chi(x)^m$. So $\chi(x)$ is an m'th root of unity, $\chi(x) \in \{exp(\frac{2\pi i}{m}k) : 0 \leq k \leq m-1\}$. In particular

$|\chi(x)| = 1$. That is, χ takes value on the unit circle

$$\mathbb{T} = \{z \in \mathbb{C} : |z| = 1\} = \{e^{i\theta} : \theta \in \mathbb{R}\}.$$

The *trivial character* is the homomorphism $\chi_0 \equiv 1$. One has $\sum_{x\in H} \chi_0(x) = |H|$.

Lemma 1.8. *If $\chi : H \to \mathbb{C}^\times$ is a non-trivial character, then $\sum_{x\in H} \chi(x) = 0$.*

Proof. (See [8, Page 188].) As $\chi \neq \chi_\circ$ there is some $x_\circ \in H$ with $\chi(x_\circ) \neq 1$. Now

$$\chi(x_\circ) \sum_{x\in H} \chi(x) = \sum_{x\in H} \chi(x_\circ)\chi(x) = \sum_{x\in H} \chi(x_\circ x) = \sum_{y\in H} \chi(y)$$

and hence

$$(\chi(x_\circ) - 1) \sum_{x\in H} \chi(x) = 0$$

As $\chi(x_\circ) - 1 \neq 0$ we must have $\sum_{x\in H} \chi(x) = 0$. □

Corollary 1.9. *(Orthogonality Relations) For any pair of characters χ, χ' on H one has*

$$\langle \chi, \chi' \rangle = \begin{cases} 0 & \text{if } \chi \neq \chi' \\ |H| & \text{if } \chi = \chi' \end{cases}.$$

Proof. We have $\langle \chi, \chi' \rangle = \sum_{x\in H} \chi(x)\overline{\chi'(x)} = \sum_{x\in H} \chi''(x)$ where $\chi''(x) = \chi(x)\overline{\chi'(x)}$. It's easy to check that $\chi'' : H \to \mathbb{C}^\times$ is a character and that χ'' is non-trivial when $\chi \neq \chi'$. So the previous lemma shows that $\langle \chi, \chi' \rangle = 0$ when $\chi \neq \chi'$. On the other hand when $\chi = \chi'$ we have $\chi''(x) = |\chi(x)|^2 = 1$ for each x and hence $\langle \chi, \chi' \rangle = \sum_{x\in H} 1 = |H|$. □

Let $\mathcal{C}(H)$ denote the set of all characters on H. Each character is a non-zero function and the corollary shows that distinct characters are pair-wise orthogonal. In

particular $\mathcal{C}(H)$ is a linearly independent set in $L(H)$. As $dim(L(H)) = |H|$ it follows that $|\mathcal{C}(H)| \leq |H|$. So there are at most $|H| = m$ distinct characters on H.

1.4.1 Characters on abelian groups

Now suppose that H is a *finite abelian group*. In this context it is standard to write $\widehat{H}$ in place of $\mathcal{C}(H)$ for the set of all characters.

Lemma 1.10. *For* finite abelian group H*, one has* $|\widehat{H}| = |H|$.

Proof. 1. First suppose that H is cyclic of order m with generator a,

$$H = \langle a \rangle = \{e, a, a^2, \cdots, a^{m-1}\} \quad (a^m = e).$$

For each $j \in \mathbb{Z}$ one checks that the map $\chi_j : H \to \mathbb{C}^\times$, $\chi_j(a^k) = exp(\frac{2\pi i}{m} jk)$ is a character and that $\chi_\circ, \chi_1, \cdots, \chi_{m-1}$ are all different. As we have shown that $|\widehat{H}| \leq m$, it follows that $\widehat{H} = \{\chi_\circ, \chi_1, \cdots, \chi_{m-1}\}$ and in particular, that $|\widehat{H}| = m = |H|$.

2. Suppose that H is a product $H = H_1 \times H_2 \times \cdots \times H_\ell$ of finite abelian groups $H_1, H_2, \cdots, H_\ell$. Given any characters $\psi_1, \cdots, \psi_\ell$ on $H_1, H_2, \cdots, H_\ell$ one checks that $\chi(x_1, \cdots, x_\ell) := \psi_1(x_1)\psi_2(x_2)\cdots\psi_\ell(x_\ell)$ is a character on H. Moreover distinct ℓ-tuples $(\psi_1, \cdots, \psi_\ell) \in \widehat{H_1} \times \cdots \times \widehat{H_\ell}$ produce distinct characters $\chi \in \widehat{H}$. Thus $|\widehat{H}| \geq |\widehat{H_1}| \cdot |\widehat{H_2}| \cdots |\widehat{H_\ell}|$.

3. The *Classification Theorem for Finite Abelian Groups* [1, Page 472] shows that any finite abelian group is isomorphic to a product of (one or more) cyclic groups. The result now follows from 1 and 2. □

Let H be finite abelian with $|H| = m$ elements. We have shown that the set $|\widehat{H}|$ of characters forms a set of m non-zero pair-wise orthogonal function on H. So $|\widehat{H}|$

is an orthogonal basis for the space $L(H)$. Moreover we have $\langle \chi, \chi \rangle = m$ for each $\chi \in \widehat{H}$. From linear algebra the expression for given $f \in L(H)$ in the basis $|\widehat{H}|$ is as follows.

Proposition 1.11. *(Fourier Inversion Formula) For any $f \in L(H)$ one has*

$$f = \frac{1}{m} \sum_{\chi \in \widehat{H}} \langle f, \chi \rangle \chi. \qquad (m = |H|)$$

Remark 1.12. Sometimes people write $\widehat{f}(\chi) := \langle f, \chi \rangle$ and call the map $\widehat{f} : \widehat{H} \to \mathbb{C}$ the *discrete fourier transform (DFT)* for $f : H \to \mathbb{C}$. The proposition shows how to recover a function f from its DFT.

1.5 Action pairs

In this section we consider pairs of finite groups, K, H.

Definition 1.13. An *action of K on H by automorphisms* is a map

$$K \times H \to H, \quad (k, h) \to k \cdot h$$

satisfying the following properties:

- For each $k \in K$ the map $\varphi_k : H \to H$, $\varphi_k(h) = k \cdot h$ is an automorphism of the group H (so in particular $k \cdot e_H = e_H$ for all k),
- $(k_1 k_2) \cdot h = k_1 \cdot (k_2 \cdot h)$, for all $k_1, k_2 \in K$, $h \in H$,
- $e_K \cdot h = h$, for all $h \in H$ (where e_K is the identity element in K).

We write $K : H$ to indicate that group K acts on H via automorphisms and refer to $K : H$ as an *action pair*. (This notation and terminology are non-standard.)

Definition 1.14. Let $K : H$ be an action pair and $h \in H$. The *K-orbit* of h is

$$K \cdot h = \{k \cdot h : k \in K\}.$$

Note that $K \cdot e_H = \{e_H\}$ and that $h \in K \cdot h$. For non-trivial action some elements $h \neq e_H$ will have $K \cdot h \supsetneq \{h\}$. The various orbits partition the group H. That is, H is the *disjoint* union of the distinct K-orbits. This follows from:

Lemma 1.15. *For elements $h_1, h_2 \in H$ if $(K \cdot h_1) \cap (K \cdot h_2) \neq \emptyset$, then $K \cdot h_1 = K \cdot h_2$. Thus distinct K-orbits are necessarily disjoint.*

Proof. Suppose $x_\circ \in (K \cdot h_1) \cap (K \cdot h_2)$, so $x_\circ = k_1 \cdot h_1 = k_2 \cdot h_2$ for some $k_1, k_2 \in K$. Let $x \in K \cdot h_1$, so $x = k_3 \cdot h_1$. But $k_1 \cdot h_1 = k_2 \cdot h_2$ implies that $h_1 = k_1^{-1} \cdot (k_2 \cdot h_2) = (k_1^{-1} k_2) \cdot h_2$, so $x = k_3 \cdot h_1 = k_3 \cdot ((k_1^{-1} k_2) \cdot h_2) = (k_3 k_2^{-1} k_1) \cdot h_2 \in K \cdot h_2$. Thus $K \cdot h_1 \subset K \cdot h_2$. Similarly, one shows $K \cdot h_2 \subset K \cdot h_1$. So $K \cdot h_2 = K \cdot h_1$. □

It also follows that $K \cdot h_1 = K \cdot h_2 \Leftrightarrow h_2 \in K \cdot h_1$ and that the distinct K-orbits partition H. That is, for some elements $x_1 = e_H, x_2, x_3, \cdots, x_m \in H$,

$$H = (K \cdot x_1) \coprod (K \cdot x_2) \coprod \cdots \coprod (K \cdot x_m).$$

Definition 1.16. Let $K : H$ be an action pair. We say that a function $f \in L(H)$ is *K-invariant* if $f(k \cdot x) = f(x)$ for all $k \in K$ $x \in H$ and let

$$L_K(H) = \{f \in L(H) : f \text{ is } K\text{-invariant}\}$$

be the set of all K-invariant $\mathbb{C}$-valued functions on the group H.

Lemma 1.17. *$L_K(H)$ is a* sub-algebra *of the group algebra $L(H)$. That is,*

- $L_K(H)$ *is a vector subspace of* $L(H)$ *and*
- $f \star g \in L_K(H)$ *for all* $f, g \in L_K(H)$.

Proof. It is clear that $L_K(H)$ is a subspace of $L(H)$. Let $f, g \in L_K(H)$, $x \in H, k \in K$. Now

$$\begin{aligned}(f \star g)(k \cdot x) &= \sum_{y \in H} f(y) g(y^{-1}(k \cdot x)) \quad \text{let } y = k \cdot z,\ z = k^{-1} \cdot y \\ &= \sum_{z \in H} f(k \cdot z) g((k \cdot z)^{-1}(k \cdot x)) = \sum_{z \in H} f(k \cdot z) g((k \cdot z^{-1})(k \cdot x)) \\ &= \sum_{z \in H} f(k \cdot z) g(k \cdot (z^{-1} x)) = \sum_{z \in H} f(z) g(z^{-1} x) = (f \star g)(x). \qquad \square\end{aligned}$$

Lemma 1.18. *The characteristic functions* $\{\delta_{K \cdot x_1}, \delta_{K \cdot x_2}, \cdots, \delta_{K \cdot x_r}\}$ *for the distinct* K*-orbits in* H *form a basis for* $L_K(H)$*. Moreover they are pairwise orthogonal with* $\langle \delta_{k \cdot x_j}, \delta_{k \cdot x_j} \rangle = |K \cdot x_j|$*. In particular we have* $dim(L_K(H)) = r$*, the number of* K*-orbits.*

Proof. For $f \in L_K(H)$ we can write $f = f(x_1)\delta_{K \cdot x_1} + f(x_2)\delta_{K \cdot x_2} + \cdots + f(x_r)\delta_{K \cdot x_r}$, and hence $\{\delta_{K \cdot x_1}, \cdots, \delta_{K \cdot x_r}\}$ spans $L_K(H)$. Also $\{\delta_{k \cdot x_1}, \cdots, \delta_{k \cdot x_r}\}$ are linearly independent. Indeed, if $c_1\delta_{K \cdot x_1} + \cdots + c_r\delta_{K \cdot x_r} = 0$, then evaluation at x_1 yields $c_1 + 0 + \cdots + 0 = 0$ because $x_1 \in K \cdot x_1$ and $K \cdot x_j \cap K \cdot x_1 = \emptyset$ for $j > 1$. Hence $c_1 = 0$. Similarly, $c_2 = 0, \cdots, c_r = 0$.

Next we compute

$$\langle \delta_{K \cdot x_j}, \delta_{K \cdot x_i} \rangle = \sum_{y \in H} \delta_{K \cdot x_i}(y) \overline{\delta_{K \cdot x_j}(y)} = \sum_{y \in H} \delta_{K \cdot x_i}(y) \delta_{K \cdot x_j}(y).$$

But as $\delta_A(y)\delta_B(y) = \delta_{A\cap B}(y)$ we have $\langle \delta_{K\cdot x_i}, \delta_{K\cdot x_j}\rangle = \sum_{y\in H} \delta_{(K\cdot x_i)\cap(K\cdot x_j)}(y)$. As

$$(K\cdot x_i)\cap(K\cdot x_j) = \begin{cases} \emptyset & \text{if} \\ \\ K\cdot x_j & \text{if } i=j \end{cases}.$$

we conclude that $\langle \delta_{K\cdot x_i}, \delta_{K\cdot x_j}\rangle = 0$ if $i \neq j$ and $\langle \delta_{K\cdot x_i}, \delta_{K\cdot x_i}\rangle = \sum_{y\in H} \delta_{K\cdot x_i}(y) = |K\cdot x_i|$. □

1.6 Gelfand action pairs and spherical functions

Definition 1.19. $K : H$ is a *Gelfand action pair* when $L_K(H)$ is a commutative algebra under convolution. That is, when $f \star g = g \star f$ holds for all K-invariant functions $f, g \in L_K(H)$.

Definition 1.20. Let $K : H$ be a Gelfand action pair. A K-invariant function $\phi \in L_K(H)$ is said to be *K-spherical* if

- $f \star \phi$ is a scalar multiple of ϕ for every $f \in L_K(H)$; and
- $\phi(e) = 1$

We let $\Delta(K : H)$ denote the set of all K-spherical function $\phi : H \to \mathbb{C}$.

Note that the constant function $\phi_\circ \equiv 1$ is K-spherical. In particular we observe that $f \star \phi_\circ = s(f)\phi_\circ$ where $s(f) := \sum_{x\in H} f(x)$. We call $\phi_\circ$ the *trivial* K-spherical function.

Theorem 1.21. *Let $K : H$ be a Gelfand action pair and r be the number of distinct K-orbits in H. Then $\Delta(K : H)$ is an orthogonal basis for $L_K(H)$. In particular, in view of Lemma 1.18, there are precisely r distinct K-spherical functions and one has*

$$f = \sum_{\phi\in\Delta(K:H)} \frac{\langle f, \phi\rangle}{\langle \phi, \phi\rangle}\phi$$

for any $f \in L_K(H)$. Moreover the spherical functions have the following properties for all $f \in L_K(H)$ and $\phi, \phi' \in \Delta(K:H)$:

1. $f \star \phi = \langle f, \phi \rangle \phi$,
2. $\langle \phi, \phi' \rangle = 0$ *for* $\phi \neq \phi'$,
3. $\phi \star \phi' \equiv 0$ *for* $\phi \neq \phi'$.

Together properties 1 and 2 show

$$\phi \star \phi' = \delta_{\phi,\phi'} \langle \phi, \phi \rangle \phi$$

for all spherical functions $\phi, \phi' \in \Delta(K:H)$. Our terminology and context (Gelfand action pairs) is not standard. Elsewhere the reader will find a well developed theory of *Gelfand pairs* and associated spherical functions. A (finite) Gelfand pair (G, K) consists of a finite group G and subgroup $K \leq G$ for which the convolution algebra

$$L(G/K) := \{f \in L(G) \ : \ f(k_1 g k_2) = f(g) \text{ for all } k_1, k_2 \in K \text{ and } g \in G\}$$

of K-*bi-invariant* functions on G is commutative. Given a Gelfand action pair $K:H$ one obtains a Gelfand pair (G, K) by taking $G := H \rtimes K$, the *semi-direct product* of H with K. The reader may find the general theory of spherical functions for finite Gelfand pairs presented in [4, §4.5] and [9, Pages 341-350]. Here we have specialized this more general theory to the action pairs context. In particular, Theorem 1.21 follows immediately from corresponding results in the Gelfand pairs context.

We remark that property 3 in Theorem 1.21 follows immediately from the defining properties of spherical functions. Indeed, given a pair of spherical functions ϕ, ϕ' we have both $\phi \star \phi' = \lambda' \phi'$ and $\phi \star \phi' = \phi' \star \phi = \lambda \phi$ for some scalars $\lambda, \lambda' \in \mathbb{C}$. Thus if

$\phi \star \phi' \not\equiv 0$ we must have that ϕ' is a multiple of ϕ. As $\phi(e) = 1 = \phi'(e)$ this implies that $\phi' = \phi$. Moreover orthogonality property 2 in the theorem now follows from properties 1 and 3. Indeed for spherical functions $\phi \neq \phi'$ we have

$$\langle \phi, \phi' \rangle = \langle \phi, \phi' \rangle \phi'(e) = (\phi \star \phi')(e) = 0.$$

The following result is a specialization of the *functional equation* for spherical functions associated with Gelfand pairs. (See [4, Theorem 4.5.3].)

Proposition 1.22. *For $\phi \in \Delta(K : H)$ one has*

$$\frac{1}{|K|} \sum_{k \in K} \phi\big(x(k \cdot y)\big) = \phi(x)\phi(y)$$

for all $x, y \in H$.

The proof for Proposition 1.22 requires the following lemma.

Lemma 1.23. *For $\phi \in \Delta(K : H)$ one has*

$$\phi(x^{-1}) = \overline{\phi(x)}$$

for all $x \in H$.

Proof. For fixed $x \in H$ one has

$$\langle \delta_{K \cdot x}, \phi \rangle = \sum_{y \in H} \delta_{K \cdot x}(y) \overline{\phi(y)} = |K \cdot x| \ \overline{\phi(x)},$$

by K-invariance of ϕ. But using the spherical function properties from Theorem 1.21

we also can write

$$\begin{aligned}\langle \delta_{K\cdot x}, \phi\rangle &= \langle \delta_{K\cdot x}, \phi\rangle \phi(e) \\ &= (\delta_{K\cdot x} \star \phi)(e) \\ &= \sum_{y\in H} \delta_{K\cdot x}(y)\phi(y^{-1}e) \\ &= |K\cdot x|\phi(x^{-1}).\end{aligned}$$

In the last step we used the fact that the function $y \mapsto \phi(y^{-1})$ is K-invariant. This is easily seen. So now

$$|K\cdot x|\overline{\phi(x)} = |K\cdot x|\phi(x^{-1})$$

and hence $\phi(x^{-1}) = \overline{\phi(x)}$ as stated. □

Proof of Proposition 1.22. Fix $x \in H$ and consider the function $F_x : H \to \mathbb{C}$ defined as

$$F_x(y) = \frac{1}{K}\sum_{k\in K} \phi(x(k\cdot y)).$$

This is clearly K-invariant and so, by Theorem 1.21, has spherical function expansion

$$F_x = \sum_{\phi'\in\Delta(K:H)} \frac{\langle F_x, \phi'\rangle}{\langle \phi', \phi'\rangle}\phi'.$$

But for $\phi' \in \Delta(K:H)$ we compute

$$\begin{aligned}\langle F_x, \phi'\rangle &= \sum_{y\in H} F_x(y)\overline{\phi'(y)} \\ &= \sum_{y\in H} F_x(y)\phi'(y^{-1}) \quad \text{by Lemma 1.23} \\ &= \sum_{y\in H} \frac{1}{|K|}\sum_{k\in K} \phi(x(k\cdot y))\phi'(y^{-1})\end{aligned}$$

$$
\begin{aligned}
&= \frac{1}{|K|} \sum_{k \in K} \sum_{y \in H} \phi(x \underbrace{(k \cdot y)}_{z}) \phi'(y^{-1}) \\
&= \frac{1}{|K|} \sum_{k \in K} \sum_{y \in H} \phi(xz) \underbrace{\phi'(k^{-1} \cdot z^{-1})}_{\phi'(z^{-1})} \\
&= \frac{|K|}{|K|} \sum_{y \in H} \phi(xz) \phi'(z^{-1}) \\
&= (\phi \star \phi')(x) \\
&= \delta_{\phi,\phi'} \langle \phi, \phi \rangle \phi(x) \quad \text{using Theorem 1.21.}
\end{aligned}
$$

So now

$$F_x = \sum_{\phi' \in \Delta(K:H)} \delta_{\phi,\phi'} \frac{\langle \phi, \phi \rangle}{\langle \phi', \phi' \rangle} \phi(x) \phi' = \phi(x) \phi.$$

That is, $F_x(y) = \phi(x)\phi(y)$ as desired. □

Note that if H is an finite *abelian* group then any action pair $K : H$ is automatically a Gelfand action pair. Indeed as $L(H)$ commutes under convolution so does the subalgebra $L_K(H)$. In the final section of this chapter we identify the spherical functions for Gelfand action pair's $K : H$ with H abelian and give a complete proof for Theorem 1.21 in this context. In particular, if H is abelian and $K = \{e_K\}$ trivial we will see that spherical functions are just the characters $\chi \in \widehat{H}$. This is consistent with the functional equation, which reduces to

$$\phi(xy) = \phi(x)\phi(y)$$

in this very special situation. The spherical function expansion for $f \in L_{\{e_K\}}(H) = L(H)$ given in Theorem 1.21 here reduces to the Fourier Inversion Formula.(Proposition 1.11)

1.7 K-averaged characters

Definition 1.24. Let $K : H$ be an action pair. For functions $f \in L(H)$ we define the *K-average* by

$$^K f(x) := \frac{1}{|K|} \sum_{k \in K} f(k \cdot x)$$

K-averaging has the following properties.

Lemma 1.25. *(a) $f \mapsto {}^K f$ is a linear operator on $L(H)$,*

(b) ${}^K f \in L_K(H)$ for each $f \in L(H)$,

(c) ${}^K f = f$ when $f \in L_K(H)$,

(d) $\langle {}^K f, g \rangle = \langle f, {}^K g \rangle$ for all $f, g \in L(H)$.

Proof for part d.

$$\begin{aligned}
\langle {}^K f, g \rangle &= \sum_{x \in H} \left(\frac{1}{|K|} \sum_{k \in K} f(k \cdot x) \right) \overline{g(x)} = \frac{1}{|K|} \sum_{k \in K} \sum_{x \in H} f(k \cdot x) \overline{g(x)} \\
&= \frac{1}{|K|} \sum_{k \in K} \sum_{y \in H} f(y) \overline{g(k^{-1} \cdot y)} = \sum_{y \in H} f(y) \left(\frac{1}{|K|} \sum_{k \in K} \overline{g(k^{-1} \cdot y)} \right) \\
&= \sum_{y \in H} f(y) \overline{\left(\frac{1}{|K|} \sum_{k' \in K} g(k' \cdot y) \right)} = \langle f, {}^K g \rangle.
\end{aligned}$$

□

Theorem 1.26. *Let $K : H$ be a* Gelfand action pair *and $\chi : H \to \mathbb{C}^\times$ a character on H. Then the K-average ${}^K\chi : H \to \mathbb{C}$ is a K-*spherical function*. Moreover if H is abelian then every K-spherical function has this form. That is, $\Delta(K : H) = \{ {}^K\chi : \chi \in \widehat{H} \}$ when H is abelian.*

Proof. Let $K : H$ be a Gelfand action pair, $\chi : H \to \mathbb{C}^\times$ a character and set $\phi := {}^K\chi$. Now $\phi \in L_K(H)$ and $\phi(e) = \frac{1}{|K|} \sum_{k \in K} \chi(k \cdot e) = \frac{|K|}{|K|} \chi(e) = \chi(e) = 1$ as χ is a

homomorphism. To show that ϕ is K-spherical it remains to check that $f \star \phi$ is a scalar multiple of ϕ for each $f \in L_K(H)$. Indeed we compute

$$\begin{aligned}
f \star \phi(x) &= \sum_{y\in H} f(y)\,{}^K\chi(y^{-1}x) = \sum_{y\in H} f(y)\frac{1}{|K|}\sum_{k\in K}\chi\left((k\cdot y)^{-1}(k\cdot x)\right)\\
&= \frac{1}{|K|}\sum_{k\in K}\sum_{y\in H} f(y)\chi\Big((\underbrace{k\cdot y}_{z})^{-1}(k\cdot x)\Big)\\
&= \frac{1}{|K|}\sum_{k\in K}\sum_{z\in H}\underbrace{f(k^{-1}\cdot z)}_{\mathrm{f(z)}}\underbrace{\chi\left(z^{-1}(k\cdot x)\right)}_{\chi(z^{-1})\chi(k\cdot x)}\\
&= \left(\sum_{z\in H} f(z)\underbrace{\chi(z^{-1})}_{\overline{\chi(z)}}\right)\left(\frac{1}{|K|}\sum_{k\in K}\chi(k\cdot x)\right)\\
&= \langle f,\chi\rangle\phi(x)\\
&= \langle {}^K f,\chi\rangle\phi(x)\\
&= \langle f, {}^K\chi\rangle\phi(x) \quad \text{using Lemma 1.25(d)}\\
&= \langle f,\phi\rangle\phi(x).
\end{aligned}$$

So $f \star \phi$ is a multiple of ϕ and, moreover,

$$f \star \phi = \langle f,\phi\rangle\phi \tag{1.8}$$

as in Theorem 1.21.

Now let χ, χ' be characters on H and set $\phi = {}^K\chi$, $\phi' = {}^K\chi'$. Using (1.8) it follows that $\langle \phi, \phi'\rangle = 0$ unless $\phi \equiv \phi'$, as explained in the preceding section. Thus the set

$$S := \{{}^K\chi : \chi \textit{ is a character on } H\}$$

of K-averaged characters is a set of pair-wise orthogonal spherical functions. In

particular, S is linearly independent.

Next suppose that H is abelian. We will show that $S = \Delta(K : H)$. First observe that S spans $L_K(H)$ and hence forms an orthogonal basis for this space. Indeed given $f \in L_K(H)$ we may apply the Fourier Inversion Formula to write

$$f = \frac{1}{|H|} \sum_{\chi \in \widehat{H}} \langle f, \chi \rangle \chi.$$

K-averaging both sides yields

$$f = \frac{1}{|H|} \sum_{\chi \in \widehat{H}} \langle f, \chi \rangle {}^K\chi$$

since ${}^K f = f$. So $f \in Span(S)$. Now suppose that $\alpha : H \to \mathbb{C}$ were a spherical function with $\alpha \notin S$. We must have $\alpha \star \phi = 0$ for each $\phi \in S$ and (1.8) shows that also $\langle \alpha, \phi \rangle = 0$ for each $\phi \in S$. As S in an orthogonal basis for $L_K(H)$ this implies $\alpha \equiv 0$. As $\alpha(e) = 1$ we have obtained a contradiction. In conclusion, $S = \Delta(K : H)$ as desired. □

Proof of Theorem 1.21 for H abelian. By Theorem 1.26 we have

$$\Delta(K : H) = \{ {}^K\chi : \chi \in \widehat{H} \}.$$

Moreover, the proof showed the set of K-averaged characters forms an orthogonal basis for $L_K(H)$ and property 1 in Theorem 1.21 is established above as (1.8). □

Remark 1.27. For characters χ, χ' we may have ${}^K\chi = {}^K\chi'$ although $\chi \neq \chi'$. For a non-trivial action $K : H$ with H abelian this must be the case for certain characters. For here $\Delta(K : H) = \{ {}^K\chi : \chi \in \widehat{H} \}$ and we know $|\widehat{H}| = m = |H|$ whereas $|\Delta(K : H)| = r < m$, the number of distinct K-orbits in H.

CHAPTER 2: DISCRETE RANDOM WALKS

2.1 Random walks on a finite set

Let X be a finite non-empty set.

Definition 2.1. A *stochastic matrix* P on X is a function $P : X \times X \to \mathbb{R}$ with

(a) $P(x, y) \geq 0$ *for all* x, y,

(b) $\sum_{y \in X} P(x, y) = 1$ *for all* x.

That is, $P_{x_0}(y) = P(x_0, y)$ is a probability measure on X for each fixed $x_0 \in X$. If $|X| = n$ then P amounts to an $n \times n$ real matrix indexed by the set X, each row of which is a probability measure. Such a P determines a random walk on X. Elements $x_0 \in X$ represent possible positions of a random walk. At each discrete time, $t = 0, 1, 2, \cdots$, our walker is at some position $x \in X$. If we are at position x at time t, then $P(x, y)$ gives the probability that we move to position y at time $t+1$. Note that the probabilities $P_x(y) = P(x, y)$ only depend on the current position x, not on time t or the past position of the walker at prior times. We say that the stochastic matrix P is *the transition matrix* for the random walk.

Suppose the walker is at position x_0 at time $t = 0$. Then

$$P^{(2)}_{x_0}(y) = P^{(2)}(x_0, y) = \sum_{x_1 \in X} P(x_0, x_1)P(x_1, y)$$

gives the probability the walker is at position y at time $t = 2$. Likewise

$$P^{(3)}_{x_0}(y) = P^{(3)}(x_0, y) = \sum_{x_2 \in X} P^{(2)}(x_0, x_2)P(x_2, y) = \sum_{x_1, x_2 \in X} P(x_0, x_1)P(x_1, x_2)P(x_2, y)$$

is the probability distribution for the walker's position at time $t = 3$. In general

$$P_{x_0}^{(m)}(y) = P^{(m)}(x_0, y) = \sum_{x_{m-1} \in X} P^{(m-1)}(x_0, x_{m-1}) P(x_{m-1}, y)$$
$$= \sum_{x_1, x_2, \cdots, x_{m-1} \in X} P(x_0, x_1) P(x_1, x_2) \cdots P(x_{m-2}, x_{m-1}) P(x_{m-1}, y)$$

is the probability distribution after m random steps.

Note that: each of $P^{(2)}, P^{(3)}, \cdots, P^{(m)}, \cdots$ is itself a stochastic matrix. Regarding P as an $n \times n$ matrix $P^{(2)}, P^{(3)}, \cdots$ are the powers of P in the usual sense.

Definition 2.2. A stochastic matrix P on X is *ergodic* if for some $m_0 \geq 0$, we have $P^{(m_0)}(x, y) > 0$ for all $x, y \in X$.

The main result in this field is the following. The reader may find a proof in [4, Pages 18-20].

Theorem 2.3 (The Markov Ergodic Theorem)**.** *P is ergodic if and only if there is a probability measure $\pi \in \mathcal{M}(X)$ with $Supp(\pi) = X$ and*

$$\lim_{m \to \infty} P^{(m)}(x, y) = \pi(y).$$

That is, $P_{x_0}^{(m)} \to \pi$ in $\mathcal{M}(x)$ independent of x_0.

2.2 Random walks on finite groups

Definition 2.4. Let H be a finite group. A stochastic matrix $P : H \times H \to \mathbb{R}$ on H is said to be *left-H-invariant* if $P(ax, ay) = P(x, y)$ for all $a, x, y \in H$.

Lemma 2.5. *Given $\mu \in \mathcal{M}(H)$, $P_\mu(x,y) := \mu(x^{-1}y)$ is a left-H-invariant stochastic matrix and, moreover, every left-H-invariant stochastic matrix P is of the form $P = P_\mu$ for some probability measure μ, namely $\mu(z) := P(e,z)$.*

Proof. Let $\mu \in \mathcal{M}(H)$ be given and set $P_\mu(x,y) := \mu(x^{-1}y)$. Now P_μ is a left-H-invariant stochastic matrix because:

1. $P_\mu(x,y) = \mu(x^{-1}y) \geq 0$ as $\mu \in \mathcal{M}(H)$,
2. $\sum_{y\in H} P_\mu(x,y) = \sum_{y\in H} \mu(x^{-1}y) = \sum_{z\in H} \mu(z) = 1$ as $\mu \in \mathcal{M}(H)$,
3. $P_\mu(ax,ay) = \mu((ax)^{-1}(ay)) = \mu(x^{-1}a^{-1}ay) = \mu(x^{-1}y) = P_\mu(x,y)$.

Conversely, given a left-H-invariant stochastic matrix P let $\mu(x) := P(e,x) = P_e(x)$. Now

$$P(x,y) = P(x^{-1}x, x^{-1}y) = P(e, x^{-1}y) = \mu(x^{-1}y) = P_\mu(x,y).$$

So $P = P_\mu$ as stated. □

Given $\mu \in \mathcal{M}(H)$ one computes that $P_\mu^{(2)} = P_{\mu\star\mu}$. Indeed

$$\begin{aligned}
P_\mu^{(2)}(x,y) &= \sum_{z\in H} P_\mu(x,z)P_\mu(z,y) \\
&= \sum_{z\in H} \mu(\underbrace{x^{-1}z}_{w})\mu(z^{-1}y) \quad (\text{let } w = x^{-1}z \text{ so } z = xw \text{ and } z^{-1} = w^{-1}x^{-1}) \\
&= \sum_{w\in H} \mu(w)\mu(w^{-1}x^{-1}y) \\
&= (\mu \star \mu)(x^{-1}y) \\
&= P_{\mu\star\mu}(x,y).
\end{aligned}$$

Iterating we obtain

$$P_\mu^{(m)} = P_{\mu^{\star m}}$$

where

$$\mu^{\star m} := \underbrace{\mu \star \cdots \star \mu}_{m \text{ times}}.$$

So, in particular, $\mu^{\star m}$ is the probability for the state of the random walk after m steps using transition matrix P_μ beginning from $e \in H$.

One executes such a walk as follows.

> Choose a sequence $x_1, x_2, x_3, \ldots$ of points in H independently distributed according to μ. After m steps our walker is at state $x_1 x_2 \cdots x_m$ (product in H).

Definition 2.6. We will say that a probability measure $\mu \in \mathcal{M}(H)$ on an finite group H is *ergodic* when the associated left-H-invariant stochastic matrix P_μ is ergodic in the sense of Definition 2.2.

Proposition 2.7. *The following are equivalent for $\mu \in \mathcal{M}(H)$.*

(a) μ is ergodic,

(b) $Supp(\mu)^{m_\circ} = H$ for some $m_\circ \geq 1$,

(c) $\mu^{\star m} \to \mathfrak{u}$.

In particular ergodicity for μ only depends on $Supp(\mu)$ and the equilibrium distribution from the Markov Ergodic Theorem is necessarily the uniform distribution $\mathfrak{u}$.

Proof. $\underline{(a) \iff (b)}$: By definition μ is ergodic if and only if $P_\mu^{(m_\circ)}(x, y) = \mu^{\star m_\circ}(x^{-1}y) > 0$ for some $m_\circ$ and all $x, y \in H$. Equivalently, $Supp(\mu^{\star m_\circ}) = H$ for some $m_\circ \geq 1$. But

as μ is non-negative real valued we have $Supp(\mu^{\star m_\circ}) = Supp(\mu)^{m_\circ}$ (see (1.7)). So μ is ergodic if and only if $Supp(\mu)^{m_\circ} = H$ for some $m_\circ \geq 1$.

$\underline{(b) \implies (c)}$: We adapt the argument given in [4, Pages 18-20] to establish the The Markov Ergodic Theorem. For each $m = 1, 2, \cdots$, let

$$\alpha_m := \min_{x \in H} \mu^{\star m}(x), \quad \beta_m := \max_{x \in H} \mu^{\star m}(x)$$

so that $0 \leq \alpha_m \leq \beta_m \leq 1$ for each m. Now for every $x \in H$ we have

$$\mu^{\star(m+1)}(x) = \sum_{y \in H} \mu^{\star m}(xy^{-1})\mu(y) \geq \sum_{y \in H} \alpha_m \mu(y) = \alpha_m 1 = \alpha_m$$

and hence $\big(\alpha_{m+1} = \min_{x \in H} \mu^{\star(m+1)}(x)\big) \geq \alpha_m$. Likewise one has $\beta_{m+1} \leq \beta_m$ for all m. Thus $(\alpha_m)_m$ and $(\beta_m)_m$ are bounded monotone sequences hence convergent. Let

$$\alpha := \lim_m \alpha_m, \quad \beta := \lim_m \beta_m$$

as $\alpha_m \leq \beta_m$ we have $\alpha \leq \beta$. Using (b) we will show that in fact

$$\alpha = \beta \tag{2.1}$$

holds. Assuming this as $\alpha_m \leq \mu^{\star m}(x) \leq \beta_m$ it follows also that $\lim_m \mu^{\star m}(x) = \alpha$ for each $x \in H$. As each $\mu^{\star m}$ is a probability measure this gives

$$|H|\alpha = \sum_{x \in H} \Big(\lim_m \mu^{\star m}(x)\Big) = \lim_m \underbrace{\left(\sum_{x \in H} \mu^{\star m}(x)\right)}_{1} = 1$$

So $\alpha = \frac{1}{|H|}$. Thus $\mu^{\star m}(x) \xrightarrow[m]{} \frac{1}{|H|}$ for every $x \in H$. That is, $\mu^{\star m} \to \mathfrak{u}$ as claimed.

To complete the proof that $(b) \implies (c)$ it remains to show (2.1). Suppose that $Supp(\mu^{\star m_0}) = H$ and let $\varepsilon := \min_{x\in H} \mu^{\star m_0}(x) > 0$. As $\mu^{\star m_0}$ is a probability measure and $Supp(\mu^{\star m_0}) = H$ we have $0 < \varepsilon < 1$ here. Now for any $k \geq 0,\ x \in H$

$$\begin{aligned}\mu^{\star(m_0+k)}(x) &= \sum_{y\in H} \mu^{\star m_0}(y)\mu^{\star k}(y^{-1}x) \\ &= \sum_{y\in H} \left(\mu^{\star m_0}(y) - \varepsilon\mu^{\star k}(x^{-1}y)\right)\mu^{\star k}(y^{-1}x) + \varepsilon\sum_{y\in H}\mu^{\star k}(x^{-1}y)\mu^{\star k}(y^{-1}x).\end{aligned}$$

But for each $y \in H$ we have

$$\begin{aligned}\mu^{\star m_0}(y) - \varepsilon\mu^{\star k}(x^{-1}y) &\geq \mu^{\star m_0}(y) - \mu^{\star m_0}(y)\mu^{\star k}(x^{-1}y) \\ &= \mu^{\star m_0}(y)(1 - \mu^{\star k}(x^{-1}y)) \geq 0\end{aligned}$$

and hence

$$\left(\mu^{\star m_0}(y) - \varepsilon\mu^{\star k}(x^{-1}y)\right)\mu^{\star k}(y^{-1}x) \geq \left(\mu^{\star m_0}(y) - \varepsilon\mu^{\star k}(x^{-1}y)\right)\alpha_k.$$

So now

$$\begin{aligned}\sum_{y\in H}\left(\mu^{\star m_0}(y) - \varepsilon\mu^{\star k}(x^{-1}y)\right)\mu^{\star k}(y^{-1}x) &\geq \alpha_k\sum_{y\in H}\left(\mu^{\star m_0}(y) - \varepsilon\mu^{\star k}(x^{-1}y)\right) \\ &= \alpha_k(1-\epsilon).\end{aligned}$$

Also

$$\sum_{y\in H}\mu^{\star k}(x^{-1}y)\mu^{\star k}(y^{-1}x) = \sum_{z\in H}\mu^{\star k}(z)\mu^{\star k}(z^{-1}) = \mu^{\star 2k}(e).$$

So for each $x \in H$ we have

$$\mu^{\star(m_0+k)}(x) \geq (1-\varepsilon)\alpha_k + \varepsilon\mu^{\star 2k}(e)$$

and hence

$$\alpha_{m_0+k} \geq (1-\varepsilon)\alpha_k + \varepsilon\mu^{\star 2k}(e).$$

Likewise one shows

$$\beta_{m_0+k} \leq (1-\varepsilon)\beta_k + \varepsilon\mu^{\star 2k}(e)$$

and hence

$$\beta_{m_0+k} - \alpha_{m_0+k} \leq (1-\varepsilon)(\beta_k - \alpha_k).$$

Iterating now given $\beta_{jm_0+k} - \alpha_{jm_0+k} \leq (1-\varepsilon)^j(\beta_k - \alpha_k)$ for each $k, j \geq 1$. In particular taking $k = m_0$ we see

$$0 \leq \beta_{(j+1)m_0} - \alpha_{(j+1)m_0} \leq (1-\varepsilon)^j(\beta_{m_0} - \alpha_{m_0}).$$

As $0 < 1-\varepsilon < 1$ we have $(1-\varepsilon)^j \to 0$ as $j \to \infty$ so $\beta - \alpha = \lim_{j\to\infty}\left(\beta_{(j+1)m_0} - \alpha_{(j+1)m_0}\right) = 0$. That is, $\alpha = \beta$ holds as claimed.

$\underline{(c) \implies (b)}$: Suppose that (b) fails. Then for each $m \geq 1$ there is some point $x_m \in H$ with $\mu^{\star m}(x_m) = 0$. But now

$$\|\mu^{\star m} - \mathfrak{u}\|_1 = \sum_{x\in H}\left|\mu^{\star m}(x) - \frac{1}{|H|}\right| \geq \left|\mu^{\star m}(x_m) - \frac{1}{|H|}\right| = \frac{1}{|H|}$$

for every m. Hence $(\|\mu^{\star m} - \mathfrak{u}\|_1)_m$ does not converge to zero and $(\mu^{\star m})_m$ does not converge to $\mathfrak{u}$. So if $\mu^{\star m} \to \mathfrak{u}$ then (b) must hold. $\square$

2.3 Random walks with action pairs

Let H and K be finite groups with K acting by automorphisms on H to form an action pair $K : H$. We let $\mathcal{M}_K(H) = \mathcal{M}(H) \cap L_K(H)$ be the set of K-invariant probability measures on H.

Lemma 2.8. *For $\mu \in \mathcal{M}_K(H)$ the stochastic matrix P_μ is K-invariant, i.e.*

$$P_\mu(k \cdot x, k \cdot y) = P_\mu(x, y).$$

Proof. For any $x, y \in H$ and $k \in K$ we have

$$\begin{aligned} P_\mu(k \cdot x, k \cdot y) &= \mu((k \cdot x)^{-1}(k \cdot y)) \\ &= \mu((k \cdot x^{-1})(k \cdot y)) \\ &= \mu(k \cdot (x^{-1}y)) = \mu(x^{-1}y) \\ &= P_\mu(x, y). \end{aligned}$$

□

The following definition is non-standard. Recall that $K \cdot x_\circ$ denotes the K-orbit through a point $x_\circ \in H$ and $\mu_{K \cdot x_0}$ is the uniform probability measure supported on $K \cdot x_\circ$ as in Equation 1.3.

Definition 2.9. Let $K : H$ be an action pair. We say that

(a) $K : H$ is *ergodic* if $\mu_{K \cdot x_0}$ is ergodic for some $x_0 \in H$,

(b) $K : H$ is *strongly ergodic* if $\mu_{K \cdot x_0}$ is ergodic for every $x_0 \neq e$.

The measures $\mu_{K \cdot x_0}$ are K-invariant with smallest possible support, a single K-orbit. So an ergodic action pair is one for which there is an ergodic invariant walk with minimal support. A strongly ergodic action pair is one for which all invariant

walks with minimal support are ergodic. Here we specify "$x_0 \neq e$" in the definition because $K \cdot e = \{e\}$ and $\mu_{\{e\}} = \delta_e$ is not ergodic (unless H is the trivial group).

Now consider $(K : H)$-invariant random walks where the action pair $K : H$ is a *Gelfand action pair*. So let $K : H$ be a Gelfand action pair and $\mu \in \mathcal{M}_K(H)$. We apply the spherical function expansion to μ:

$$\mu = \sum_{\phi \in \Delta(K:H)} \frac{\langle \mu, \phi \rangle}{\langle \phi, \phi \rangle} \phi$$

where $\Delta(K : H)$ is the set of K-spherical functions on H. As we know $\phi \star \phi' = \delta_{\phi,\phi'} \langle \phi, \phi \rangle \phi$ for $\phi, \phi' \in \Delta(K : H)$. We obtain

$$\mu \star \mu = \sum_{\phi \in \Delta(K:H)} \frac{\langle \mu, \phi \rangle}{\langle \phi, \phi \rangle}^2 \phi$$

Proof. Indeed

$$\begin{aligned}
\mu \star \mu &= \left(\sum_{\phi} \frac{\langle \mu, \phi \rangle}{\langle \phi, \phi \rangle} \phi \right) \star \left(\sum_{\phi'} \frac{\langle \mu, \phi' \rangle}{\langle \phi', \phi' \rangle} \phi \right) \\
&= \sum_{\phi} \sum_{\phi'} \frac{\langle \mu, \phi \rangle \langle \mu, \phi' \rangle}{\langle \phi, \phi \rangle \langle \phi', \phi' \rangle} \underbrace{\phi \star \phi'}_{\delta_{\phi,\phi} \langle \phi, \phi' \rangle \phi} \\
&= \sum_{\phi} \frac{\langle \mu, \phi \rangle \langle \mu, \phi \rangle}{\langle \phi, \phi \rangle \langle \phi, \phi \rangle} \langle \phi, \phi \rangle \phi \\
&= \sum_{\phi} \frac{\langle \mu, \phi \rangle}{\langle \phi, \phi \rangle}^2 \phi.
\end{aligned}$$

□

Iterating gives:

$$\mu^{\star m} = \sum_{\phi \in \Delta(K:H)} \frac{\langle \mu, \phi \rangle}{\langle \phi, \phi \rangle}^m \phi.$$

The term in this summation involving the trivial spherical function $\phi_\circ$ is

$$\frac{\langle \mu, \phi_\circ \rangle}{\langle \phi_\circ, \phi_\circ \rangle}^m \phi_\circ = \mathfrak{u}.$$

Indeed $\langle \phi_\circ, \phi_\circ \rangle = \sum_{x \in H} |\phi_\circ(x)|^2 = |H|$ and

$$\langle \mu, \phi_\circ \rangle = \sum_{x \in H} \mu(x) \overline{\phi_\circ(x)} = \sum_{x \in H} \mu(x) = 1$$

Thus for each x:

$$\frac{\langle \mu, \phi_\circ \rangle}{\langle \phi_\circ, \phi_\circ \rangle}^m \phi_\circ(x) = \frac{1}{|H|} = \mathfrak{u}(x).$$

Thus we have established the following.

Proposition 2.10. *If $K : H$ is a Gelfand action pair and $\mu \in \mathcal{M}_K(H)$ then*

$$\mu^{\star m} - \mathfrak{u} = \sum_{\phi \neq \phi_\circ} \frac{\langle \mu, \phi \rangle}{\langle \phi, \phi \rangle}^m \phi$$

where the sum is over all non-trivial K-spherical functions $\phi \in \Delta(K : H)$.

Corollary 2.11. *$\mu \in \mathcal{M}_K(H)$ is ergodic if and only if $|\langle \mu, \phi \rangle| < 1$ for all non-trivial spherical functions $\phi \neq \phi_\circ$.*

Proof. By Proposition 2.7 μ is ergodic if and only if $\mu^{\star m} \xrightarrow[m]{} \mathfrak{u}$. Let $\phi \in \Delta(K : H)$ with $\phi \neq \phi_\circ$. Then

$$\begin{aligned}
\langle \mu^{\star m} - \mathfrak{u}, \phi \rangle &= \left\langle \sum_{\phi' \neq \phi_\circ} \frac{\langle \mu, \phi' \rangle}{\langle \phi', \phi' \rangle}^m \phi', \phi \right\rangle \\
&= \sum_{\phi' \neq \phi_\circ} \frac{\langle \mu, \phi' \rangle}{\langle \phi', \phi' \rangle}^m \underbrace{\langle \phi', \phi \rangle}_{\delta_{\phi', \phi} \langle \phi, \phi \rangle} \\
&= \frac{\langle \mu, \phi \rangle}{\langle \phi, \phi \rangle}^m \langle \phi, \phi \rangle = \langle \mu, \phi \rangle^m
\end{aligned}$$

If μ is ergodic, then $\mu^{\star m} - \mathfrak{u} \to 0$ and hence $\langle \mu^{\star m} - \mathfrak{u}, \phi \rangle = \langle \mu, \phi \rangle^m \to 0$. This implies $|\langle \mu, \phi \rangle| < 1$.

Conversely if $|\langle \mu, \phi \rangle| < 1$ for all spherical functions $\phi \neq \phi_\circ$ we have

$$\mu^{\star m} - \mathfrak{u} = \sum_{\phi \neq \phi_\circ} \frac{\langle \mu, \phi \rangle}{\langle \phi, \phi \rangle}^m \phi \xrightarrow[m \to \infty]{} 0$$

and thus μ is ergodic. □

Taking $\mu = \mu_{K \cdot x_0}$, the probability measure supported on a single K-orbit, we observe that for spherical functions ϕ,

$$\begin{aligned} \langle \mu_{K \cdot x_0}, \phi \rangle &= \frac{1}{|K \cdot x_0|} \sum_{x \in H} \delta_{K \cdot x}(x) \overline{\phi(x)} \\ &= \frac{1}{|K \cdot x_0|} \sum_{x \in K \cdot x_0} \overline{\phi(x)} \\ &= \frac{1}{|K \cdot x_0|} |K \cdot x_0| \overline{\phi(x_0)} = \overline{\phi(x_0)} \end{aligned}$$

since ϕ is K-invariant. So Proposition 2.10 yields

$$\mu_{k \cdot x_0}^{\star m} - \mathfrak{u} = \sum_{\phi \neq \phi_\circ} \frac{(\overline{\phi(x_0)})}{\langle \phi, \phi \rangle}^m \phi$$

and Corollary 2.11 gives:

Corollary 2.12. *A Gelfand action pair $K : H$ is*

(a) ergodic if and only if for some $x_\circ \in H$ one has $|\phi(x_\circ)| < 1$ for all $\phi \neq \phi_\circ$,

(b) strongly ergodic if and only if $|\phi(x_\circ)| < 1$ for all $\phi \neq \phi_\circ$ and all $x_\circ \neq e$.

Lemma 2.13 (Upper Bound Lemma)**.** *Given a Gelfand action pair $K : H$ and a*

probability measure $\mu \in \mathcal{M}_K(H)$ *one has*

$$\|\mu^{\star m} - \mathfrak{u}\|_{TV}^2 \leq \frac{|H|}{4} \sum_{\phi \neq \phi_\circ} \frac{|\langle \mu, \phi \rangle|^{2m}}{\langle \phi, \phi \rangle}.$$

In particular, for points $x_\circ \in H$

$$\|\mu_{K \cdot x_\circ}^{\star m} - \mathfrak{u}\|_{TV}^2 \leq \frac{|H|}{4} \sum_{\phi \neq \phi_\circ} \frac{|\phi(x_\circ)|^{2m}}{\langle \phi, \phi \rangle}.$$

Here $\| \cdot \|_{TV}$ *denotes the total variation distance given by Equation 1.4.*

Proof. Recall that the total variation distance between a pair of probability measures is one half the l^1-distance (Lemma 1.3). Thus we have

$$\|\mu^{\star m} - \mathfrak{u}\|_{TV}^2 = \frac{1}{4}\|\mu^{\star m} - \mathfrak{u}\|_1^2 \leq \frac{|H|}{4}\|\mu^{\star m} - \mathfrak{u}\|_2^2$$

because, by Lemma 1.1, $\|f\|_1 \leq \sqrt{|H|}\|f\|_2$ for $f \in L(H)$. But now

$$\begin{aligned}
\|\mu^{\star m} - \mathfrak{u}\|_2^2 &= \langle \mu^{\star m} - \mathfrak{u}, \mu^{\star m} - \mathfrak{u} \rangle \\
&= \left\langle \sum_{\phi \neq \phi_\circ} \frac{\langle \mu, \phi \rangle}{\langle \phi, \phi \rangle}^m \phi, \sum_{\phi' \neq \phi_\circ} \frac{\langle \mu, \phi' \rangle}{\langle \phi', \phi' \rangle}^m \phi' \right\rangle \\
&= \sum_{\phi \neq \phi_\circ \neq \phi'} \frac{\langle \mu, \phi \rangle^m \overline{\langle \mu, \phi' \rangle^m}}{\langle \phi, \phi \rangle \langle \phi', \phi' \rangle} \langle \phi, \phi' \rangle \\
&= \sum_{\phi \neq \phi_\circ} \frac{|\langle \mu, \phi \rangle|^{2m}}{\langle \phi, \phi \rangle}
\end{aligned}$$

since $\langle \phi, \phi' \rangle = 0$ for $\phi \neq \phi'$. So

$$\|\mu^{\star m} - \mathfrak{u}\|_{TV}^2 \leq \frac{|H|}{4} \sum_{\phi \neq \phi_\circ} \frac{|\langle \mu, \phi \rangle|^{2m}}{\langle \phi, \phi \rangle}. \quad \square$$

Diaconis and Shahshahani pioneered the application of harmonic analysis on finite groups to the study of random walks in their ground breaking 1981 paper on random permutations [6]. This included a version of the Upper Bound Lemma. A version in the Gelfand pairs context can be found in [5, Page 55] and [4, Page 144]. Lemma 2.13 is a specialization of this result to the action pairs setting.

CHAPTER 3: RANDOM WALKS ON FINITE FIELDS

3.1 The context

In this chapter we study invariant random walks arising in connection with finite fields. We refer the reader to [1, Chapter 13] or [8, Chapter 2] for background on the structure of finite fields. Let F denote a finite field. It is well known that the order of F is a prime power,

$$|F| = q = p^s,$$

say for some prime p and exponent $s \geq 1$. Moreover for each prime p and exponent s there is exactly one field of order $q = p^s$, up to isomorphism, and we sometimes write $F \cong \mathbb{F}_q$. The field F is an extension of its prime field (smallest subfield) $\mathbb{Z}_p = \mathbb{Z}/p\mathbb{Z}$. We think of $\mathbb{Z}_p$ concretely as $\mathbb{Z}_p = \{0, 1, \ldots, p-1\}$ with arithmetic performed modulo p. We assume throughout that

p is an odd prime.

This restriction is needed for Theorem 3.3 below and in Section 3.4.

We let F^+ denote the additive group for F and $F^\times$ the multiplicative group of non-zero field elements. The group $F^\times$ acts via multiplication on F^+ to yield an action pair $F^\times : F^+$. This is an immediate consequence of the field axioms. We will study action pairs of the form $K : F^+$ where $K \leq F^\times$ denotes a subgroup of $F^\times$. As F^+ is abelian these are Gelfand action pairs. It is known that the multiplicative group $F^\times$ is cyclic, and a generator is called a *primitive element*. As subgroups of cyclic groups are cyclic, K is a cyclic group whose order divides $|F^\times| = q - 1$. Moreover

the additive group F^+ is isomorphic to the product

$$\underbrace{\mathbb{Z}_p^+ \times \cdots \times \mathbb{Z}_p^+}_{s \text{ times}}$$

of s copies of the additive cyclic group $\mathbb{Z}_p^+$.

3.2 Ergodicity for action pairs $K : F^+$

Proposition 3.1. *Let $K \leq F^\times$. Then for any $x_\circ \neq 0$ and $m \geq 1$*

$$\|\mu_{K\cdot x_\circ}^{\star m} - \mathfrak{u}\|_{TV} = \|\mu_K^{\star m} - \mathfrak{u}\|_{TV}$$

Proof. Let $x_\circ \neq 0$ be fixed. We have for $x \in F$,

$$\begin{aligned} x \in K \cdot x_\circ &\iff x = k \cdot x_\circ \ for\ some\ k \in K \\ &\iff x_\circ^{-1} x = k \ for\ some\ k \in K \\ &\iff x_\circ^{-1} x \in K. \end{aligned}$$

Also note that $|K \cdot x_\circ| = |K|$ because the map $K \to K \cdot x_\circ$, $k \mapsto k x_\circ$ is a bijection. So for any $x \in F$,

$$\begin{aligned} \mu_{K\cdot x_\circ}(x) &= \begin{cases} \frac{1}{|K\cdot x_\circ|} & \text{if } x \in K \cdot x_\circ \\ 0 & \text{if } x \notin K \cdot x_\circ \end{cases} \\ &= \begin{cases} \frac{1}{|K|} & \text{if } x_\circ^{-1} x \in K \\ 0 & \text{if } x_\circ^{-1} x \notin K \end{cases} \\ &= \mu_K(x_\circ^{-1} x) \end{aligned}$$

Now we compute that

$$\begin{aligned}
(\mu_{K\cdot x_\circ} \star \mu_{K\cdot x_\circ})(x) &= \sum_{y\in F} \mu_{K\cdot x_\circ}(-y)\mu_{K\cdot x_\circ}(y+x) \\
&= \sum_{y\in F} \mu_K(x_\circ^{-1}(-y))\mu_K(x_\circ^{-1}(y+x)) \\
&= \sum_{y\in F} \mu_K(-(x_\circ^{-1}y))\mu_K(x_\circ^{-1}y + x_\circ^{-1}x) \quad (\text{let } z = x_\circ^{-1}y) \\
&= \sum_{z\in F} \mu_K(-z)\mu_K(z + x_\circ^{-1}x) \\
&= (\mu_K \star \mu_K)(x_\circ^{-1}x).
\end{aligned}$$

Iterating gives

$$\mu_{K\cdot x_\circ}^{\star m}(x) = \mu_K^{\star m}(x_\circ^{-1}x)$$

for all $x \in F$, $m \geq 1$.

Finally now

$$\begin{aligned}
\|\mu_{K\cdot x_\circ}^{\star m} - \mathfrak{u}\|_{TV} &= \frac{1}{2}\|\mu_{K\cdot x_\circ}^{\star m} - \mathfrak{u}\|_1 \\
&= \frac{1}{2}\sum_{x\in F} |\mu_{K\cdot x_\circ}^{\star m}(x) - \mathfrak{u}(x)| \\
&= \frac{1}{2}\sum_{x\in F} \left|\mu_K^{\star m}(x_\circ^{-1}x) - \frac{1}{q}\right| \quad (\text{let } y = x_\circ^{-1}x) \\
&= \frac{1}{2}\sum_{y\in F} \left|\mu_K^{\star m}(y) - \frac{1}{q}\right| \\
&= \frac{1}{2}\|\mu_K^{\star m} - \mathfrak{u}\|_1 \\
&= \|\mu_K^{\star m} - \mathfrak{u}\|_{TV}
\end{aligned}$$

as claimed. □

Corollary 3.2. *Let $K \leq F^\times$. Then the following are equivalent:*

1. *$K : F^+$ is ergodic.*

2. *$K : F^+$ is strongly ergodic.*

3. *μ_K is an ergodic probability measure.*

Proof. Recall that $K : F^+$ is ergodic when $\mu_{K\cdot x_\circ}$ is an ergodic measure for at least one point $x_\circ \neq 0$. This means that $\|\mu_{K\cdot x_\circ}^{\star m} - \mathfrak{u}\|_{TV} \underset{m}{\longrightarrow} 0$ for some $x_\circ \neq 0$. Moreover $K : F^+$ is strongly ergodic when this happens for every $x_\circ \neq 0$. Proposition 3.1 shows that $\|\mu_{K\cdot x_\circ}^{\star m} - \mathfrak{u}\|_{TV} \underset{m}{\longrightarrow} 0$ for all $x_\circ \neq 0$ if and only if $\|\mu_K^{\star m} - \mathfrak{u}\|_{TV} \underset{m}{\longrightarrow} 0$. So $K : F^+$ is strongly ergodic if and only if it is ergodic if and only if μ_K is an ergodic measure. □

Below we will apply Proposition 2.7 to obtain a criterion for erodicity of μ_K. We know that $F^\times$ is a cyclic group (see [1, Page 46]) and hence so is any subgroup $K \leq F^\times$. Let α denote a generator for K. Thus $K = \langle\alpha\rangle = \{\alpha^j : j \in \mathbb{Z}\} = \{1, \alpha, \alpha^2, \cdots, \alpha^{d-1}\}$ where $d = ord_{F^\times}(\alpha)$, the order of α in $F^\times$. For $m \in \mathbb{N}$, we now see that the m'th "power" of $K = K \cdot 1$ in the additive group F^+ is

$$\begin{aligned} mK &:= \underbrace{K + K + \cdots + K}_{\text{m-times}} \\ &= \left\{a_\circ + a_1\alpha + a_2\alpha^2 + \cdots + a_{d-1}\alpha^{d-1} | a_j \in \mathbb{N} \cup \{0\},\ a_\circ + a_1 + \cdots + a_{d-1} = m\right\} \\ &= \left\{\overline{a}_\circ + \overline{a}_1\alpha + \overline{a}_2\alpha^2 + \cdots + \overline{a}_{d-1}\alpha^{d-1} | a_j \in \mathbb{N} \cup \{0\},\ a_\circ + a_1 + \cdots + a_{d-1} = m\right\} \end{aligned}$$

where $\overline{a}_j \in \mathbb{Z}_p$ denotes the residue of the non-negative integer a_j modulo p.

Recall that the subfield $\mathbb{Z}_p(\alpha)$ of F generated by $\alpha \in F^\times$ is

$$\mathbb{Z}_p(\alpha) = \{f(\alpha) : f \in \mathbb{Z}_p[x]\}$$

and we have field extensions

$$\mathbb{Z}_p \subset \mathbb{Z}_p(\alpha) \subset F.$$

We see that all elements of mK belong to the intermediate field $\mathbb{Z}_p(\alpha)$, that is

$$mK \subset \mathbb{Z}_p(\alpha)$$

for all m.

Theorem 3.3. *Let α be a generator for K.*

(a) If $K : F^+$ is (strongly) ergodic then $\mathbb{Z}_p(\alpha) = F$. Equivalently, the field element $\alpha \in F$ has degree s over the prime field $\mathbb{Z}_p$, where $|F| = p^s$.

(b) Conversely, if $\mathbb{Z}_p(\alpha) = F$ and $-1 \in K$ then $K : F^+$ is (strongly) ergodic.

Proof. Suppose $K : F^+$ is ergodic. Corollary 3.2 implies that μ_K is an ergodic measure and by Proposition 2.7 we must have

$$\left(Supp(\mu_K^{\star m_\circ}) = m_\circ K\right) = F$$

for some $m_\circ \geq 1$. As $m_\circ K \subset \mathbb{Z}_p(\alpha) \subset \mathbb{F}$, this gives $\mathbb{Z}_p(\alpha) = F$.

Next assume that $\mathbb{Z}_p(\alpha) = F$ and $-1 \in K$. We have $K = \{1, \alpha, \alpha^2, \ldots, \alpha^{d-1}\}$ as above and now

$$\begin{aligned}\mathbb{Z}_p(\alpha) &= \left\{\overline{a}_\circ + \overline{a}_1\alpha + \overline{a}_2\alpha^2 + \cdots + \overline{a}_{d-1}\alpha^{d-1} | a_j \in \mathbb{N} \cup \{0\}\right\} \\ &= \left(\bigcup_{m=1}^{\infty} mK\right) \cup \{0\}.\end{aligned}$$

In fact as $\pm 1 \in K$ we have $0 \in K + K = 2K$ and can write simply $\mathbb{Z}_p(\alpha) = \cup_{m=1}^{\infty} mK$. Now say $F = \{\beta_1, \beta_2, \cdots, \beta_q\}$ and for each j let $m_j \geq 1$ be such that $\beta_j \in m_j K$. In

fact we can assume m_j is odd because p is odd and we can write

$$\beta_j = \beta_j + \underbrace{\alpha + \alpha + \cdots + \alpha}_{\text{p-times}},$$

so $m_j K \subset (m_j + p)K$ and if m_j is even then $m_j + p$ is odd. As $0 \in 2K$ we also have $K \subset 3K \subset 5K \subset \cdots$. Now let $m_\circ = \max(m_1, m_2, \cdots, m_q)$. We have

$$F \subset \bigcup_{j=1}^{q} m_j K \subset m_\circ K \subset F.$$

So $m_\circ K = F$ and hence $K : F^+$ is ergodic by Proposition 2.7. □

Lemma 3.4. $-1 \in K \iff |K|$ *is even.*

Proof. Here $K = \langle \alpha \rangle = \{1, \alpha, \alpha^2, \cdots, \alpha^{d-1}\}$. Suppose that $-1 \in K$. Thus $-1 = \alpha^j$ for some exponent $1 \le j \le d-1$. Now $\alpha^{2j} = 1$ with $2 \le 2j \le 2d-2$. But the only exponent i with $2 \le i \le 2d-1$ with $\alpha^i = 1$ is $i = d$. Thus $|K| = d = 2j$ is even.

Conversely, suppose that d is even, $d = 2j$ say. Then $\alpha^j \in K$ has $(\alpha^j)^2 = \alpha^d = 1$ and $(\alpha^j \neq 1)$. As the only roots for $x^2 - 1$ in F are 1 and -1 we must have $\alpha^j = -1$. So $-1 \in K$. □

3.3 Upper bound on $\|\mu_K^{\star m} - \mathfrak{u}\|_{TV}^2$

Let $\widehat{F}^+$ denote the set of all additive characters on F. These are the maps $\psi : F \to \mathbb{C}^\times$ for which $\psi(a+b) = \psi(a)\psi(b)$ for all $a, b \in F$. It is known that

$$\widehat{F}^+ = \{\psi_a \ : \ a \in F\} \quad \text{where} \quad \psi_a(x) := \exp\left(\frac{2\pi i}{p} tr(ax)\right) \tag{3.1}$$

and $tr : F \to \mathbb{Z}_p$ is the absolute trace mapping for the field extension $\mathbb{Z}_p \subset (F = \mathbb{F}_q = \mathbb{F}_{p^s})$, namely

$$tr(x) = x + x^p + x^{p^2} + \cdots + x^{p^{s-1}}. \tag{3.2}$$

(See [8, Theorem 5.7 and Definition 2.22].)

Let $K \leq F^\times$ be a subgroup of $F^\times$. As F^+ is abelian $K : F^+$ is a Gelfand action pair and Theorem 1.26 ensures that the K-spherical functions are K-averaged characters. Thus

$$\Delta(K : F^+) = \{\phi_a \ : \ a \in F\} \quad \text{where} \quad \phi_a := {}^K\psi_a.$$

That is,

$$\phi_a(x) = \frac{1}{|K|}\sum_{k\in K}\psi_a(kx) = \frac{1}{|K|}\sum_{k\in K}\psi_1(kax) = \frac{1}{|K|}\sum_{k\in K}\psi_{ka}(x).$$

Note that $\phi_a(x) = \phi_1(ax)$ and that $\phi_\circ \equiv 1$ is the trivial spherical function. The spherical functions ϕ_a are not all distinct. In fact

$$\begin{aligned}\left|\Delta(K : F^+)\right| &= \text{the number of } K \text{ orbits in } F \\ &= 1 + \left|F^\times/K\right| \\ &= 1 + \frac{|F^\times|}{|K|}.\end{aligned}$$

So there are exactly $1 + |F^\times|/|K|$ distinct K-spherical functions in all.

Lemma 3.5. *$\langle\phi_a, \phi_a\rangle = \frac{q}{|K|}$ for $a \neq 0$ (and of course $\langle\phi_\circ, \phi_\circ\rangle = q$).*

Proof. If $a = 0$ then $\phi_a = \phi_\circ \equiv 1$ and $\langle\phi_a, \phi_a\rangle = \langle\phi_\circ, \phi_\circ\rangle = \sum_{x\in F} 1 = q$. If $a \neq 0$, we

get, using Lemma 1.25,

$$\begin{aligned}\langle \phi_a, \phi_a \rangle &= \langle {}^K\psi_a, {}^K\psi_a \rangle = \langle {}^K({}^K\psi_a), \psi_a \rangle = \langle {}^K\psi_a, \psi_a \rangle \\ &= \frac{1}{|K|} \sum_{x \in F} \sum_{k \in K} \psi_a(kx) \overline{\psi_a(x)} \\ &= \frac{1}{|K|} \sum_{k \in K} \sum_{x \in F} exp\left(\frac{2\pi i}{p} \underbrace{(tr(kax) - tr(ax))}_{tr((k-1)ax)} \right) \\ &= \frac{1}{|K|} \sum_{k \in K} \sum_{x \in F} \psi_{(k-1)a}(x)\end{aligned}$$

But Lemma 1.8 shows

$$\begin{aligned}\sum_{x \in F} \psi_{(k-1)a}(x) &= \begin{cases} q & \text{if } (k-1)a = 0 \\ 0 & \text{if } otherwise \end{cases} \\ &= q\delta_{k,1}\end{aligned}$$

So $\langle \phi_a, \phi_a \rangle = \frac{1}{|K|} \sum_{k \in K} q\delta_{k,1} = \frac{q}{|K|}$. □

Lemma 3.6. *For $a, b \in F$ one has*

$$\phi_a = \phi_b \iff K \cdot a = K \cdot b.$$

Proof. Say $Ka = Kb$, equivalently $a \in Kb$. So $a = k_\circ b$ for some $k_\circ \in K$. Now for any $x \in F$:

$$\begin{aligned}\phi_a(x) &= \frac{1}{|K|} \sum_{k \in K} \psi_a(kx) = \frac{1}{|K|} \sum_{k \in K} \psi_1(kax) \\ &= \frac{1}{|K|} \sum_{k \in K} \psi_1(kk_\circ bx) = \frac{1}{|K|} \sum_{k_1 \in K} \psi_1(k_1 bx)\end{aligned}$$

$$= \frac{1}{|K|} \sum_{k_1 \in K} \psi_b(k_1 x) = \phi_b(x).$$

Conversely suppose $\phi_a = \phi_b$. Then $\langle \phi_a, \phi_b \rangle = \langle \phi_a, \phi_a \rangle \neq 0$, by the preceding lemma. But

$$\begin{aligned}
\langle \phi_a, \phi_b \rangle = \langle {}^K\psi_a, {}^K\psi_b \rangle &= \langle {}^K({}^K\psi_a), \psi_b \rangle = \langle {}^K\psi_a, \psi_b \rangle \\
&= \sum_{x \in F} {}^K\psi_a(x)\overline{\psi_b(x)} \\
&= \sum_{x \in F} \frac{1}{|K|} \sum_{k \in K} \psi_a(kx)\overline{\psi_b(x)} \\
&= \frac{1}{|K|} \sum_{k \in K} \sum_{x \in F} exp\left(\frac{2\pi i}{p} \underbrace{tr(akx - bx)}_{tr((ka-b)x)} \right) \\
&= \frac{1}{|K|} \sum_{k \in K} \left(\sum_{x \in F} \psi_{ka-b}(x) \right).
\end{aligned}$$

But if $ka - b \neq 0$ we have $\sum_{x \in F} \psi_{ka-b}(x) = 0$ by Lemma 1.8. As $\langle \phi_a, \phi_b \rangle \neq 0$ we must have $ka - b = 0$ for at least one $k \in K$. That is $b = ka$ for some $k \in K$. But this gives $b \in Ka$ and hence $Ka = Kb$ as desired. □

Lemma 3.7. $||\mu_K^{\star m} - \mathfrak{u}||_{TV}^2 \leq \frac{1}{4} \sum_{a \in F^\times} |\phi_a(1)|^{2m}$.

Proof. By the Upper Bound Lemma (Lemma 2.13) we have

$$\|\mu_K^{\star m} - \mathfrak{u}\|_{TV}^2 = \|\mu_{K \cdot 1}^{\star m} - \mathfrak{u}\|_{TV}^2 \leq \frac{|F|}{4} \sum_{\phi \neq \phi_0} \frac{|\phi(1)|}{\langle \phi, \phi \rangle}^{2m}$$

Using Lemma 3.6 we may now write

$$\|\mu_K^{\star m} - \mathfrak{u}\|_{TV}^2 \leq \frac{q}{4|K|} \sum_{a \in F^\times} \frac{|\phi_a(1)|}{\langle \phi_a, \phi_a \rangle}^{2m}.$$

Indeed for $a \neq 0$ we have $|Ka| = |K|$ since the map $K \mapsto Ka$, $k \mapsto ka$ is a bijection. (That is, in the previous summation each distinct spherical function ϕ_a appears $|K|$ times.) Finally as Lemma 3.5 shows $\langle \phi_a, \phi_a \rangle = \frac{q}{|K|}$ for each $a \neq 0$ we obtain the result as stated. □

Recall that both $F^\times$ and K are finite cyclic groups. Theorem 3.3 gives a sufficient condition for ergodicity of the action pair $K : F^+$ in terms of a generator for K. In this context we can make explicit the upper bound on $||\mu_K^{\star m} - \mathfrak{u}||_{TV}^2$ given by Lemma 3.7 as follows.

Theorem 3.8. *Let $g \in F^\times$ be a generator for $F^\times$ (primitive element). Suppose that $\mathbb{Z}_p(\alpha) = F$ and set $K := \langle \alpha \rangle$, $d := |K|$. Assume $d = 2d'$ is even. Then $K : F^+$ is (strongly) ergodic and*

$$||\mu_K^{\star m} - \mathfrak{u}||_{TV}^2 \leq \frac{1}{4(d')^{2m}} \sum_{\ell=1}^{q-1} C_\ell^{2m}$$

where

$$C_\ell := \sum_{j=0}^{d'-1} \cos\left(\frac{2\pi}{p} tr(g^{\ell+cj})\right), \quad \textit{with} \quad \alpha := g^c.$$

Proof. We have

$$F^\times = \langle g \rangle = \{1, g, g^2, \cdots, g^{q-2}\} = \{g, g^2, \cdots, g^{q-1}\}$$

and as $\alpha := g^c$ generates $K \leq F^\times$,

$$K = \{1, \alpha, \alpha^2, \cdots, \alpha^{d-1}\} = \{\alpha, \alpha^2, \cdots, \alpha^d\}$$

where $d := |K|$. As we are assuming that $d = 2d'$ is even, $-1 = \alpha^{d'}$ belongs to K.

Theorem 3.3 ensures that $K : F^+$ is (strongly) ergodic. Moreover we have

$$\begin{aligned} K &= \{1, \alpha, \cdots, \alpha^{d'-1}, -1, -\alpha, \cdots, -\alpha^{d'}\} \\ &= \{\pm\alpha^j : 0 \le j \le d'-1\} = \{\pm g^{cj} : 0 \le j \le d'-1\} \end{aligned}$$

Now

$$\begin{aligned} \phi_a(1) &= \frac{1}{|K|}\sum_k \psi_a(kx) \\ &= \frac{1}{d}\sum_{j=0}^{d'-1}\{\psi_a(g^{cj}) + \psi_a(-g^{cj})\} \\ &= \frac{1}{d}\sum_{j=0}^{d'-1}\underbrace{\left\{\exp\left(\frac{2\pi i}{p}tr(ag^{cj})\right) + \exp\left(-\frac{2\pi i}{p}tr(ag^{cj}))\right)\right\}}_{2\cos(2\pi tr(ag^{cj})/p)} \\ &= \frac{2}{d}\sum_{j=0}^{d'-1}\cos\left(\frac{2\pi}{p}tr(ag^{cj})\right) \\ &= \frac{1}{d'}\sum_{j=0}^{d'-1}\cos\left(\frac{2\pi}{p}tr(ag^{cj})\right). \end{aligned}$$

So

$$\frac{1}{4}\sum_{a\in F^\times}|\phi_a(1)|^{2m} = \frac{1}{4}\sum_{l=1}^{q-1}|\phi_{g^l}(1)|^{2m} = \frac{1}{4(d')^{2m}}\sum_{l=1}^{q-1}\left[\sum_{j=0}^{d'-1}\cos\left(\frac{2\pi}{P}tr(g^{l+cj})\right)\right]^{2m}.$$

This completes the proof in view of Lemma 3.7. □

3.4 The action pair $U : \widetilde{F}$

Below we specialize our results from the previous sections to an interesting family of examples that arise in connection with quadratic extensions of finite fields. The ideas

here will also play a key role in the following chapter concerning random walks on finite Heisenberg groups. As before F is a finite field with $q = p^s$ elements where p is an odd prime, $F \cong \mathbb{F}_q$. Note that the squaring map

$$s : F^\times \to F^\times, \quad s(x) = x^2$$

is a homomorphism on the multiplicative group $F^\times$ with $Ker(s) = \{\pm 1\}$. So, by the *First Isomorphism Theorem* in Group Theory, $s(F^\times) \cong F^\times/\{\pm 1\}$ has order $(q-1)/2$. Thus half of the non-zero elements in F are squares and half are non-squares. Choose $\varepsilon \in F^\times$ a non-square and form the quadratic extension field

$$\widetilde{F} = F(\sqrt{\varepsilon}) = \left\{a + b\sqrt{\varepsilon} \ : \ a, b \in F\right\}$$

with field operations

$$\begin{aligned}(a + b\sqrt{\varepsilon}) + (c + d\sqrt{\varepsilon}) &:= (a + c) + (b + d)\sqrt{\varepsilon},\\ (a + b\sqrt{\varepsilon})(c + d\sqrt{\varepsilon}) &:= (ac + bd\varepsilon) + (ad + bc)\sqrt{\varepsilon}.\end{aligned}$$

This is a field of order $q^2 = p^{2s}$, so $\widetilde{F} \cong \mathbb{F}_{q^2}$. We may regard $\widetilde{F}$ as a finite analog for the field $\mathbb{C} = \mathbb{R}(\sqrt{-1})$ of complex numbers and adopt "complex notation" to $\widetilde{F}$ writing, for $z = x + y\sqrt{\varepsilon}$ in $\widetilde{F}$,

$$Re(z) = x, \quad Im(z) = y \quad \text{and} \quad \overline{z} = x - y\sqrt{\varepsilon}.$$

It is easy to check that the map $\sigma(z) = \overline{z}$ is an automorphism of the field $\widetilde{F}$ fixing the subfield F. In fact σ coincides with the q'th power mapping $z \mapsto z^q$, which is well known to be an automorphism [1, Proposition 6.20].

Lemma 3.9. *$\overline{z} = z^q$ for all $z \in \widetilde{F}$.*

Proof. For $x, y \in F$, we have:

$$(x + y\sqrt{\varepsilon})^q = x^q + (\sqrt{y\varepsilon})^q = x^q + (\sqrt{\varepsilon})^q y^q = x + (\sqrt{\varepsilon})^q y$$

as $a^q = a$ for all $a \in F = \mathbb{F}_q$ [1, Page 512]. So it just remains to show $(\sqrt{\varepsilon})^q = -\sqrt{\varepsilon}$. We have $(\sqrt{\varepsilon})^{2q} = \varepsilon^q = \varepsilon$ as $\varepsilon \in F$, so $((\sqrt{\varepsilon})^q)^2 = \varepsilon$, and thus $(\sqrt{\varepsilon})^q = \sqrt{\varepsilon}$ or $(\sqrt{\varepsilon})^q = -\sqrt{\varepsilon}$. But the polynomial $x^q - x$ has at most q roots in $\widetilde{F}$ and all q elements of F are roots. Since $\sqrt{\varepsilon} \notin F$ we must have $(\sqrt{\varepsilon})^q \neq \sqrt{\varepsilon}$. So now it follows that $(\sqrt{\varepsilon})^q = -\sqrt{\varepsilon}$. □

It is easy to check that the map

$$N : \widetilde{F}^\times \to F^\times \quad N(z) = z\overline{z} = z^{q+1} \left(= Re(z) - \varepsilon Im(z)^2\right)$$

is a group homomorphism. This is called the *norm mapping* for the field extension $\widetilde{F} \supset F$ [8, Page 57]. We define the *finite unitary group* $U = U(1, \widetilde{F})$ as

$$U = Ker(N : \widetilde{F}^\times \to F^\times) = \left\{z \in \widetilde{F} : z\overline{z} = 1\right\}. \tag{3.3}$$

This is the natural analog for the unit circle group $U(1) = \mathbb{T} = \{z \in \mathbb{C} \ : \ |z|^2 = z\overline{z} = 1\}$ in the complex plane.

Lemma 3.10. *U is a cyclic subgroup of order $q+1$ in $\widetilde{F}^\times$. If $g \in \widetilde{F}^\times$ is a primitive element then $u = g^{q-1}$ generates U.*

Proof. As $\widetilde{F}^\times$ is cyclic and U a subgroup of $\widetilde{F}^\times$, the group U is cyclic. We have

$$U = \{z \in F^\times : z^{q+1} = 1\} = \{z \in F^\times : z^{q+1} - 1 = 0\}$$

and hence $|U| \leq q+1$, as the polynomial $z^{q+1}-1$ can have at most $q+1$ roots. Note, moreover, that $z^{q+1}-1$ divides $z^{q^2-1}-1$, indeed

$$z^{q^2-1}-1 = (z^{q+1}-1)f(z) \text{ where } f(z) := \sum_{j=0}^{q-2}(z^{q+1})^j.$$

As $|\widetilde{F}| = q^2-1$ we have $z^{q^2-1}=1$ for each $z \in \widetilde{F}^\times$. That is, each $z \in \widetilde{F}^\times$ is a root of $z^{q^2-1}-1$ and hence a root of either $z^{q+1}-1$ or $f(z)$. As $f(z)$ has at most $deg(f) = (q^2-1)-(q+1)$ roots it follows that $z^{q+1}-1$ must have at least $q+1$ roots. So now $|U| = q+1$ as stated.

Next define a mapping $\eta : \widetilde{F}^\times \to \widetilde{F}^\times$ via

$$\eta(z) = \overline{z}/z = z^q/z = z^{q-1}.$$

As $\eta(zw) = (zw)^{q-1} = z^{q-1}w^{q-1} = \eta(z)\eta(w)$, the map η is a homomorphism. Observe that $\eta(\widetilde{F}^\times) \subset U(1,\widetilde{F})$ since

$$N(\eta(z)) = N(\overline{z}/z) = N(\overline{z})/N(z) = \frac{\overline{z}\overline{\overline{z}}}{z\overline{z}} = \frac{z\overline{z}}{z\overline{z}} = 1.$$

Also

$$Ker(\eta) = \{z \in \widetilde{F}^\times : \overline{z}/z = 1\} = \{z \in \widetilde{F}^\times : \overline{z} = z\} = F^\times$$

By the First Isomorphism Theorem $\eta(\widetilde{F}^\times) \cong \widetilde{F}^\times/Ker(\eta)$ and so

$$|\eta(\widetilde{F})| = |\widetilde{F}^\times|/|Ker(\eta)| = |\widetilde{F}^\times|/|F^\times| = \frac{q^2-1}{q-1} = q+1 = |U(1,\widetilde{F})|$$

As $\eta(\widetilde{F}) \subset U(1,\widetilde{F})$ and $|\eta(\widetilde{F}^\times)| = |U(1,\widetilde{F})|$ we must have $\eta(\widetilde{F}^\times) = U(1,\widetilde{F})$. Since $\eta : \widetilde{F}^\times \to U(1,\widetilde{F})$ is a surjective homomorphism one generater for $U(1,\widetilde{F})$ is given

by $\eta(g)$ where $g \in \widetilde{F}^{\times}$ generates $\widetilde{F}^{\times}$. Thus $\eta(g) = g^{q-1}$ generates the cyclic group $U(1, \widetilde{F})$. □

Proposition 3.11. *$U : \widetilde{F}^{+}$ is (strongly) ergodic.*

Proof. We will apply Theorem 3.3. First note that $-1 \in U$. So we need only check that $\mathbb{Z}_p(u) = \widetilde{F}$ where $u \in U$ denotes any generator for the cyclic group U. (In view of Lemma 3.10 we can use $u = g^{q-1}$ with g a primitive element in $\widetilde{F}$.) Let $s_1 := deg_{\mathbb{Z}_p}(u)$. We must show that $s_1 = 2s = deg(\widetilde{F}/\mathbb{Z}_p)$. Here

$$\mathbb{Z}_p \subset (\mathbb{Z}_p(u) \cong \mathbb{F}_{p^{s_1}}) \subset (\widetilde{F} \cong \mathbb{F}_{p^{2s}})$$

and from Field Theory we know $s_1|(2s)$. (This is the *Subfield Criterion* [8, Theorem 2.6].) But $U \subset \mathbb{Z}_p(u)$ and hence

$$(p^{s_1} = |\mathbb{Z}_p(u)|) \geq (|U| = q + 1 = p^s + 1)$$

so that $s_1 > s$ must hold. Now as $s_1|(2s)$ and $s_1 > s$ we conclude that $s_1 = 2s$ as claimed. □

3.5 An upper bound on $\|\mu_U^{\star m} - \mathfrak{u}\|_{TV}^2$

Let $\widetilde{tr} : \widetilde{F} \to \mathbb{Z}_p$ denote the absolute trace mapping for the finite field $\widetilde{F}$ and set

$$UB_F(m) := \frac{1}{4}\left(\frac{2}{q+1}\right)^{2m} \sum_{\ell=1}^{q^2-1} c_\ell^{2m} \tag{3.4}$$

where

$$c_\ell = \sum_{j=0}^{(q-1)/2} \cos\left(\frac{2\pi}{p}\widetilde{tr}(g^{\ell+(q-1)j})\right) \tag{3.5}$$

for $1 \leq \ell \leq q^2 - 1$. With this notation we have:

Theorem 3.12. $\|\mu_U^{\star m} - \mathfrak{u}\|_{TV}^2 \leq UB_F(m)$.

Proof. This is just a specialization of Theorem 3.8. Here we have $\widetilde{F}$ in place of F, q^2 in place of q, $\widetilde{tr}$ in place of tr, $q-1$ in place of c (see Lemma 3.10) and $(q+1)/2$ in place of d'. □

The trace mapping $\widetilde{tr} : \widetilde{F} \to \mathbb{Z}_p$ in 3.5 can be written as

$$\widetilde{tr}(z) = tr(2Re(z)) = 2tr(Re(z)) \tag{3.6}$$

where $tr : F \to \mathbb{Z}_p$ is as in 3.2. This follows from the transitivity of trace [8, Theorem 2.26]. Indeed the trace map for the extension $F \subset \widetilde{F}$ is $z \mapsto z + \overline{z} = 2Re(z)$.

CHAPTER 4: RANDOM WALKS ON FINITE HEISENBERG GROUPS

4.1 The context

The (polarized) three dimensional Heisenberg group over a given field $\mathcal{K}$ is the set $H_1(\mathcal{K}) = \mathcal{K} \times \mathcal{K} \times \mathcal{K}$ with product

$$(x, y, t)(u, v, s) = (x + u,\ y + v,\ t + s + xv - yu).$$

One verifies the associative law for this group operation as follows.

$$\begin{aligned}
\big((x_1, y_1, t_1)(x_2, y_2, t_2)\big)(x_3, y_3, t_3) &= (x_1 + x_2, y_1 + y_2, t_1 + t_2 + x_1y_2 - y_1x_2)(x_3, y_3, t_3) \\
&= \big((x_1 + x_2) + x_3, (y_1 + y_2) + y_3, \\
&\qquad t_1 + t_2 + x_1y_2 - y_1x_2 + t_3 + (x_1 + x_2)y_3 - (y_1 + y_2)x_3\big) \\
&= \big(x_1 + (x_2 + x_3), y_1 + (y_2 + y_3), \\
&\qquad t_1 + t_2 + x_1y_2 - y_1x_2 + t_3 + x_1y_3 + x_2y_3 - y_1x_3 - y_2x_3\big) \\
&= \big(x_1 + (x_2 + x_3), y_1 + (y_2 + y_3), \\
&\qquad t_1 + t_2 + t_3 + x_2y_3 - y_2x_3 + x_1(y_2 + y_3) - y_1(x_2 + x_3)\big) \\
&= (x_1, y_1, t_1)(x_2 + x_3, y_2 + y_3, t_2 + t_3 + x_2y_3 - y_2x_3) \\
&= (x_1, y_1, t_1)\big((x_2, y_2, t_2)(x_3, y_3, t_3)\big).
\end{aligned}$$

It is easy to check that the identity element in $H_1(\mathcal{K})$ is $e = (0, 0, 0)$ and inverses are given by

$$(x, y, t)^{-1} = (-x, -y, -t).$$

The group $H_1(\mathcal{K})$ is a non-abelian group provided $\mathcal{K}$ has characteristic $char(\mathcal{K}) \neq 2$. In this case the center of $H_1(\mathcal{K})$ is $Z\big(H_1(\mathcal{K})\big) = \{(0,0,t) : t \in \mathcal{K}\}$.

In the classical setting one has $\mathcal{K} = \mathbb{R}$. Identifying $\mathbb{R}\times\mathbb{R}$ with $\mathbb{C}$ via $(x,y) \leftrightarrow x+iy$ one can write

$$H_1(\mathbb{R}) = \mathbb{C}\times\mathbb{R} \quad \text{with product} \quad (z,t)(w,s) = \big(z+w,\, t+s-Im(z\overline{w})\big).$$

Now the unit circle group $U(1) = \mathbb{T} = \{z \in \mathbb{C} \ : \ |z|^2 = z\overline{z} = 1\}$ acts by automorphisms on $H_1(\mathbb{R})$ via

$$k\cdot(z,t) = (kz,t).$$

Here we take $\mathcal{K} = F$ to be a finite field with characteristic $char(F) = p \neq 2$ and write simply

$$H = H_1(F).$$

As in Chapter 3, $F \cong \mathbb{F}_q$ for $q = p^s$ with p an odd prime. Letting $\varepsilon \in F$ denote a non-square we form the quadratic extension field $\widetilde{F} = F(\sqrt{\varepsilon})$ and recall that we have adapted "complex notation" to $\widetilde{F}$ writing for $(z = x+\sqrt{\varepsilon}y) \in \widetilde{F}$,

$$\overline{z} = x - \sqrt{\varepsilon}y, \quad Re(z) = x, \quad Im(z) = y.$$

Now just as in the classical situation we may write

$$H = \widetilde{F}\times F \quad \text{with product} \quad (z,t)(w,s) := \big(z+w,\, t+s-Im(z\overline{w})\big)$$

and the finite unitary group (see (3.3))

$$U = U(1,\widetilde{F}) = \{k \in \widetilde{F} : N(k) = k\overline{k} = k^{q+1} = 1\}$$

acts by automorphisms on H via

$$k \cdot (z,t) = (kz,t).$$

It is shown on [2] that $U : H$ is a Gelfand action pair. As H is non-abelian this fact is not obvious. For functions $f, g \in L(H)$ the convolution product $f \star g$ is explicitly

$$\begin{aligned}(f \star g)(z,t) &= \sum_{(w,s)\in H} f(w,s)g\big((w,s)^{-1}(z,t)\big) \\ &= \sum_{(w,s)\in H} f(w,s)g(z-w,\ t-s+Im(w\overline{z})).\end{aligned}$$

4.2 Spherical functions for $U : H$

The spherical functions for the Gelfand action pair $U : H$ are computed in [3]. Given a U-spherical function ϕ on the Heisenberg group H let $\phi^\circ : \widetilde{F} \to \mathbb{C}$ and $\psi : F \to \mathbb{C}$ be defined as

$$\phi^\circ(z) := \phi(z,0), \qquad \psi(t) := \phi(0,t).$$

Lemma 4.1. *For any $\phi \in \Delta(U : H)$ the function ψ is an additive character on the field F ($\psi \in \widehat{F}^+$) and $\phi(z,t) = \phi^\circ(z)\psi(t)$ holds for all $(z,t) \in H$.*

Proof. Applying the *functional equation* for spherical functions (Proposition 1.22) we see

$$\begin{aligned}\phi^\circ(z)\psi(t) &= \phi(z,0)\phi(0,t) \\ &= \frac{1}{|U|}\sum_{k\in U}\phi\big((z,0)(k\cdot(0,t))\big) \\ &= \frac{1}{q+1}\sum_{k\in U}\phi\big((z,0)(k\,0,t)\big)\end{aligned}$$

$$
\begin{aligned}
&= \frac{1}{q+1}\sum_{k\in U}\phi\big((z,0)(0,t)\big)\\
&= \frac{1}{q+1}\sum_{k\in U}\phi\big(z, t - Im(z\,0)\big)\\
&= \frac{1}{q+1}\sum_{k\in U}\phi(z,t)\\
&= \phi(z,t).
\end{aligned}
$$

Moreover for $t, s \in F$ we have

$$
\begin{aligned}
\psi(t)\psi(s) = \phi(0,t)\phi(0,s) &= \frac{1}{q+1}\sum_{k\in K}\phi\big((0,t)(k\,0,s)\big)\\
&= \frac{1}{q+1}\sum_{k\in K}\phi(0,t+s) = \phi(0,t+s)\\
&= \psi(t+s).
\end{aligned}
$$

So $\psi : F \to \mathbb{C}$ is an additive character. □

Definition 4.2. We say $\phi \in \Delta(U : H)$ is of *type 1* when the additive character $\psi(t) = \phi(0,t)$ is non-trivial ($\psi \not\equiv 1$). Otherwise we say ϕ is a spherical function of *type 2*. We let $\Delta_1(U : H)$ and $\Delta_2(U : H)$ denote the sets of spherical functions of types 1 and 2 respectively, so that

$$\Delta(U : H) = \Delta_1(U : H) \coprod \Delta_2(U : H).$$

Lemma 4.3. *For $\phi \in \Delta_2(U : H)$ the map $\phi^\circ(z) = \phi(z, 0)$ is a U-spherical function on $\widetilde{F}$. Conversely, if $\phi^\circ \in \Delta(U : \widetilde{F}^+)$ then the map $\phi(z,t) = \phi^\circ(z)$ is a type 2 spherical function on H.*

Proof. For given maps $f : H \to \mathbb{C}$, and elements $s_\circ \in F$, we write

$$f^{s_\circ} : \widetilde{F} \to \mathbb{C}, \quad f^{s_\circ}(z) := f(z, s_\circ).$$

(a) Let $\phi : H \to \mathbb{C}$ be a type 2 U-spherical function on H. We must show $\phi^\circ : \widetilde{F} \to \mathbb{C}$ is a U-spherical function on the additive group $\widetilde{F}^+$. Clearly ϕ° is U-invariant and satisfies $\phi^\circ(0) = \phi(0,0) = 1$. It remains to show that for any U-invariant function $g \in L_U(\widetilde{F})$ the function $g \star_{\widetilde{F}} \phi^\circ$ is a scalar multiple of ϕ°. (Here $\star_{\widetilde{F}}$ denotes convolution in $\widetilde{F}^+$.) For this we check that

$$g \star_{\widetilde{F}} \phi^\circ = (\widetilde{g} \star \phi)^\circ$$

where $\widetilde{g} \in L(H)$ is $\widetilde{g}(z,t) := g(z)\delta_\circ(t)$. The convolution on the right hand side is in the Heisenberg group H.

$$\begin{aligned}
(\widetilde{g} \star \phi)^\circ(z) &= (\widetilde{g} \star \phi)(z, 0) \\
&= \sum_{w,s} \widetilde{g}(w,s)\phi(z - w, 0 - s + Im(w\overline{z})) \\
&= \sum_{w,s} g(w)\delta_\circ(s)\phi(z - w, -s + Im(w\overline{z})) \\
&= \sum_{w} g(w)\phi(z - w, Im(w\overline{z})) \\
&= \sum_{w \in \widetilde{F}} g(w)\phi(z - w, 0) \quad (\text{since } \phi(z,t) = \phi(z,0) \text{ as } \phi \in \Delta_2(U : H)) \\
&= \sum_{w \in \widetilde{F}} g(w)\phi^\circ(z - w) \\
&= g \star_{\widetilde{F}} \phi^\circ(z).
\end{aligned}$$

Now as $\widetilde{g}$ is U-invariant and $\phi : H \to \mathbb{C}$ is a U-spherical function we have

$\widetilde{g} \star \phi = \lambda\phi$ for some scalar $\lambda = \lambda_{g,\phi}$. But now

$$g \star_{\widetilde{F}} \phi^\circ = (\lambda\phi)^\circ = \lambda\phi^\circ$$

is a multiple of ϕ° as required. So ϕ° belongs to $\Delta(U : \widetilde{F}^+)$.

(b) Next suppose that we are given a U-spherical function $\phi^\circ : \widetilde{F} \to \mathbb{C}$ on the additive group $\widetilde{F}^+$ and define a map $\phi : H \to \mathbb{C}$ via $\phi(z,t) = \phi^\circ(z)$. We claim that ϕ is a U-spherical function on the Heisenberg group H. Clearly ϕ is U-invariant and satisfies $\phi(0,0) = \phi^\circ(0) = 1$. It remains to show that for any $f \in L_U(H)$ the map $f \star \phi$ is a scalar multiple of ϕ. For this we check that

$$(f \star \phi)(z,t) = \sum_{s \in F} (f^s \star_{\widetilde{F}} \phi^\circ)(z);$$

$$\begin{aligned}(f \star \phi)(z,t) &= \sum_{w,s} f(w,s)\phi(z-w, t-s+Im(w\overline{z})) \\ &= \sum_s \sum_w f^s(w)\phi^\circ(z-w) \\ &= \sum_s (f^s \star_{\widetilde{F}} \phi^\circ)(z).\end{aligned}$$

Now as $f^s : \widetilde{F} \to \mathbb{C}$ is U-invariant and ϕ° is a U-spherical function on $\widetilde{F}^+$ we have that $f^s \star_{\widetilde{F}} \phi^\circ = \lambda_s \phi^\circ$ from some scalar $\lambda_s \in \mathbb{C}$ so now

$$(f \star \phi)(z,t) = \sum_{s \in F} \lambda_s \phi^\circ(z) = (\sum_{s \in F} \lambda_s)\phi(z,t),$$

which shows $f \star \phi$ is a scalar multiple of ϕ as required. □

The U-spherical functions on $\widetilde{F}$ were the subject of Chapter 3. These are just

U-averaged additive characters on $\widetilde{F}$. There are q such U-spherical functions on $\widetilde{F}$ and hence q spherical functions of type 2 on H. These are the spherical functions on H obtained by U-averaging characters. The spherical functions of type 1 are more subtle. These were given in [3].

Theorem 4.4. *The U-spherical functions of type 1 on H are indexed by pairs $(\psi, \widetilde{\chi})$ where*

- *$\psi \in \widehat{F^+}$ is a non-trivial additive character on F and*
- *$\widetilde{\chi} \in (\widetilde{F}^\times / F^\times)\widehat{\ }$ a non-trivial multiplicative character on $\widetilde{F}$ whose restriction to $F^\times$ is trivial.*

The spherical function $\phi = \phi_{\psi,\widetilde{\chi}}$ for the pair $(\psi, \widetilde{\chi})$ can be written as

$$\phi(z,t) = \phi^\circ(z)\psi(t)$$

where $\phi^\circ(0) = 1$ and

$$\phi^\circ(z) = \frac{-1}{q^2-1} \sum_{w \in \widetilde{F}-F} \widetilde{\chi}(w)\psi\left(-\frac{1}{4\varepsilon}\frac{Re(w)}{Im(w)}N(z)\right) \tag{4.1}$$

$$= \frac{-1}{q+1} \sum_{a \in F} \widetilde{\chi}\big(a+\sqrt{\varepsilon}\big)\psi\left(-\frac{a}{4\varepsilon}N(z)\right) \tag{4.2}$$

for $z \neq 0$. (Recall that $N : \widetilde{F} \to F$ is the norm mapping $N(z) = z\overline{z}$.) There are precisely $(q-1)q = q^2 - q$ such character pairs $(\psi, \widetilde{\chi})$ and the associated spherical functions are all distinct. So

$$\Delta_1(U:H) = \{\phi_{\psi,\widetilde{\chi}} \ : \ (\psi, \widetilde{\chi}) \text{ as above}\}$$

is the set of type 1 spherical functions listed without repetition.

Corollary 4.5. *For each $\phi \in \Delta_1(U:H)$ and every $(z_\circ, t_\circ) \in H$ with $z_\circ \neq 0$ one has*

$$|\phi(z_\circ, t_\circ)| \leq \frac{q}{q+1}.$$

Proof. For $z_\circ \neq 0$ we have

$$\begin{aligned} |\phi(z_\circ, t_\circ)| &= \left| -\frac{1}{q+1} \sum_{a \in F} \widetilde{\chi}(a+\sqrt{\varepsilon}) \psi(-\frac{a}{4\varepsilon} N(z_\circ)) \psi(t_\circ) \right| \\ &\leq \frac{1}{q+1} \sum_{a \in F} \left| \widetilde{\chi}(a+\sqrt{\varepsilon}) \right| \left| \psi(-\frac{a}{4\varepsilon} N(z_\circ)) \right| |\psi(t_\circ)| \\ &= \frac{q}{q+1}. \end{aligned}$$

□

Lemma 4.6. *For each $\phi \in \Delta_1(U:H)$ the function $\phi^\circ(z) = \phi(z,0)$ is real valued.*

Proof. Recall that for $z \neq 0$,

$$\phi^\circ(z) = -\frac{1}{q^2-1} \sum_{w \in \widetilde{F} - F} \widetilde{\chi}(w) \psi\left(-\frac{1}{4\varepsilon} \frac{Re(w)}{Im(w)} N(z) \right).$$

For given $w \in \widetilde{F} - F$ observe that

$$\widetilde{\chi}(w^{-1}) = \widetilde{\chi}(w)^{-1} = \overline{\widetilde{\chi}(w)}$$

since $\widetilde{\chi}$ is a multiplicative character. Also as

$$w^{-1} = \frac{\overline{w}}{N(w)} = \frac{Re(w)}{N(w)} - \sqrt{\varepsilon}\, \frac{Im(w)}{N(w)}$$

we have

$$Re(w^{-1}) = \frac{Re(w)}{N(w)}, \quad Im(w^{-1}) = -\frac{Im(w)}{N(w)},$$

and hence

$$\psi\left(-\frac{1}{4\varepsilon}\frac{Re(w^{-1})}{Im(w^{-1})}N(z)\right)=\psi\left(\frac{1}{4\varepsilon}\frac{Re(w)}{Im(w)}N(z)\right)=\overline{\psi\left(-\frac{1}{4\varepsilon}\frac{Re(w)}{Im(w)}N(z)\right)}$$

since $\psi : F \to \mathbb{C}^{\times}$ is an additive character. Note, moreover, that $w^{-1} \neq w$ since

$$w^{-1}=w \implies w^2=1 \implies w=\pm 1 \implies w \in F.$$

Pairing the terms for w and w^{-1} in the above summation formula we see that $\phi^{\circ}(z)$ is a sum of real numbers

$$\begin{aligned}&\widetilde{\chi}(w)\psi\left(-\frac{1}{4\varepsilon}\frac{Re(w)}{Im(w)}N(w)\right)+\overline{\widetilde{\chi}(w)\psi\left(-\frac{1}{4\varepsilon}\frac{Re(w)}{Im(w)}N(w)\right)}\\&=2Re\left(\widetilde{\chi}(w)\psi\left(-\frac{1}{4\varepsilon}\frac{Re(w)}{Im(w)}N(w)\right)\right).\end{aligned}$$

□

In order to apply the *Upper Bound Lemma* (Lemma 2.13) we require the values $\langle\phi,\phi\rangle$ for each $\phi \in \Delta(U:H)$.

Lemma 4.7. *For each type 1 spherical function $\phi \in \Delta_1(U:H)$ one has $\langle\phi,\phi\rangle = q^2$.*

Proof.

$$\begin{aligned}\langle\phi,\phi\rangle &= \sum_{(z,t)\in H}|\phi(z,t)|^2=\sum_{(z,t)\in H}|\phi^{\circ}(z)\psi(t)|^2\\&=\sum_{(z,t)\in H}|\phi^{\circ}(z)|^2|\psi(t)|^2=\sum_{(z,t)\in H}|\phi^{\circ}(z)|^2\\&=q\sum_{z\in\widetilde{F}}|\phi^{\circ}(z)|^2=q\left\{|\phi^{\circ}(0)|^2+\sum_{z\in\widetilde{F}^{\times}}|\phi^{\circ}(z)|^2\right\}\\&=q\left\{1+\sum_{z\in\widetilde{F}^{\times}}\left|-\frac{1}{q+1}\sum_{a\in F}\widetilde{\chi}(a+\sqrt{\varepsilon})\psi(-\frac{a}{4\varepsilon}N(z))\right|^2\right\}\end{aligned}$$

$$= q\left\{1 + (\frac{1}{q+1})^2 \sum_{z\in\widetilde{F}^\times}\left|\sum_{a\in F}\widetilde{\chi}(a+\sqrt{\varepsilon})\psi(-\frac{a}{4\varepsilon}N(z))\right|^2\right\}$$
$$= q\left\{1 + (\frac{1}{q+1})^2 \sum_{z\in\widetilde{F}^\times}\sum_{a,b\in F}\widetilde{\chi}(a+\sqrt{\varepsilon})\overline{\widetilde{\chi}(b+\sqrt{\varepsilon})}\psi\left(-\frac{a}{4\varepsilon}N(z)\right)\overline{\psi\left(-\frac{b}{4\varepsilon}N(z)\right)}\right\}$$

But

$$\overline{\psi\left(-\frac{b}{4\varepsilon}N(z)\right)} = \psi\left(-\frac{b}{4\varepsilon}N(z)\right)^{-1} = \psi\left(--\frac{b}{4\varepsilon}N(z)\right) = \psi\left(\frac{b}{4\varepsilon}N(z)\right)$$

and

$$\psi\left(-\frac{a}{4\varepsilon}N(z)\right)\overline{\psi\left(-\frac{b}{4\varepsilon}N(z)\right)} = \psi\left(\frac{b-a}{4\varepsilon}N(z)\right)$$

since $\psi : F \to \mathbb{C}^\times$ is an additive character. So now

$$\langle\phi,\phi\rangle = q\left\{1 + \left(\frac{1}{q+1}\right)^2 \sum_{z\in\widetilde{F}^\times}\sum_{a,b\in F}\widetilde{\chi}(a+\sqrt{\varepsilon})\overline{\widetilde{\chi}(b+\sqrt{\varepsilon})}\psi\left(\frac{b-a}{4\varepsilon}N(z)\right)\right\}.$$

But

$$\sum_{z\in\widetilde{F}^\times}\sum_{a,b\in F}\widetilde{\chi}(a+\sqrt{\varepsilon})\overline{\widetilde{\chi}(b+\sqrt{\varepsilon})}\psi\left(\frac{b-a}{4\varepsilon}N(z)\right)$$
$$= \sum_{a,b\in F}\widetilde{\chi}(a+\sqrt{\varepsilon})\overline{\widetilde{\chi}(b+\sqrt{\varepsilon})}\left(\sum_{z\in\widetilde{F}^\times}\psi\left(\frac{b-a}{4\varepsilon}N(z)\right)\right)$$
$$= \sum_{a\in F}\widetilde{\chi}(a+\sqrt{\varepsilon})\left\{\overline{\widetilde{\chi}(a+\sqrt{\varepsilon})}(q^2-1) + \sum_{b\neq a}\overline{\widetilde{\chi}(b+\sqrt{\varepsilon})}\sum_{z\in\widetilde{F}^\times}\psi\left(\frac{b-a}{4\varepsilon}N(z)\right)\right\}$$
$$= \sum_{a\in F}\left\{\underbrace{|\widetilde{\chi}(a+\sqrt{\varepsilon})|^2}_{1}(q^2-1) + \widetilde{\chi}(a+\sqrt{\varepsilon})\sum_{b\neq a}\overline{\widetilde{\chi}(b+\sqrt{\varepsilon})}\sum_{z\in\widetilde{F}^\times}\psi\left(\frac{b-a}{4\varepsilon}N(z)\right)\right\}.$$

Note that the norm mapping $N : \widetilde{F}^\times \to F^\times$ is surjective and $(q+1)$-to-1. Indeed one

has

$$\left|\widetilde{F}^{\times}/Ker(N)\right| = \left|\widetilde{F}^{\times}/U\right| = \frac{q^2-1}{q+1} = q-1 = \left|F^{\times}\right|.$$

So

$$\begin{aligned}\sum_{z\in\widetilde{F}^{\times}} \psi\left(\frac{b-a}{4\varepsilon}N(z)\right) &= (q+1)\sum_{c\in F^{\times}} \psi\left(\frac{b-a}{4\varepsilon}c\right) = (q+1)\sum_{d\in F^{\times}}\psi(d)\\ &= (q+1)\left(\underbrace{\sum_{d\in F}\psi(d)}_{0} - \psi(0)\right) = -(q+1),\end{aligned}$$

in view of Lemma 1.8, since $\psi \in \widehat{F}$ is non-trivial. Thus

$$\begin{aligned}&\sum_{a,b\in F} \widetilde{\chi}(a+\sqrt{\varepsilon})\overline{\widetilde{\chi}(b+\sqrt{\varepsilon})}\sum_{z\in\widetilde{F}^{\times}}\psi\left(\frac{b-a}{4\varepsilon}N(z)\right)\\ &= \sum_{a\in F}\left\{(q^2-1) + \widetilde{\chi}(a+\sqrt{\varepsilon})\sum_{b\neq a}\overline{\widetilde{\chi}(b+\sqrt{\varepsilon})}(-(q+1))\right\}\\ &= q(q^2-1) - (q+1)\sum_{a\in F}\widetilde{\chi}(a+\sqrt{\varepsilon})\sum_{b\neq a}\overline{\widetilde{\chi}(b+\sqrt{\varepsilon})}.\end{aligned}$$

But as $\widetilde{\chi}$ is trivial on $F^{\times}$, we can write

$$\begin{aligned}\sum_{b\in F}\widetilde{\chi}(b+\sqrt{\varepsilon}) &= \sum_{b\in F}\frac{1}{q-1}\sum_{c\in F^{\times}}\widetilde{\chi}(b+\sqrt{\varepsilon})\\ &= \frac{1}{q-1}\sum_{b\in F}\sum_{c\in F^{\times}}\widetilde{\chi}(c)\widetilde{\chi}(b+\sqrt{\varepsilon})\\ &= \frac{1}{q-1}\sum_{b\in F,\ c\in F^{\times}}\widetilde{\chi}(cb+c\sqrt{\varepsilon}) \quad (\text{as } \widetilde{\chi}:\widetilde{F}^{\times}\to\mathbb{C}^{\times} \text{ is multiplicative})\\ &= \frac{1}{q-1}\sum_{w\in\widetilde{F}-F}\widetilde{\chi}(w)\end{aligned}$$

$$= \frac{1}{q-1}\Big\{ \underbrace{\sum_{w\in\widetilde{F}^\times} \widetilde{\chi}(w)}_{0} - \underbrace{\sum_{w\in F^\times} \widetilde{\chi}(w)}_{q-1} \Big\}$$

$$= -1$$

by Lemma 1.8, since $\widetilde{\chi} : \widetilde{F}^\times \to \mathbb{C}^\times$ is a multiplicative character with $\widetilde{\chi}|_{F^\times} \equiv 1$. Thus

$$\begin{aligned}
&\sum_{a\in F} \widetilde{\chi}(a+\sqrt{\varepsilon}) \sum_{b\neq a} \overline{\widetilde{\chi}(b+\sqrt{\varepsilon})} \\
&= \sum_{a\in F} \widetilde{\chi}(a+\sqrt{\varepsilon}) \left(\overline{\underbrace{\sum_{b\in F} \widetilde{\chi}(b+\sqrt{\varepsilon})}_{-1}} - \overline{\widetilde{\chi}(a+\sqrt{\varepsilon})} \right) \\
&= \sum_{a\in F} \widetilde{\chi}(a+\sqrt{\varepsilon}) \left\{ -1 - \overline{\widetilde{\chi}(a+\sqrt{\varepsilon})} \right\} \\
&= -\underbrace{\sum_{a\in F} \widetilde{\chi}(a+\sqrt{\varepsilon})}_{-1} - \sum_{a\in F} \underbrace{\left|\widetilde{\chi}(a+\sqrt{\varepsilon})\right|^2}_{1} \\
&= 1-q.
\end{aligned}$$

We have

$$\sum_{z\in\widetilde{F}^\times} \sum_{a,b\in F} \widetilde{\chi}(a+\sqrt{\varepsilon})\overline{\widetilde{\chi}(b+\sqrt{\varepsilon})} \sum_{z\in\widetilde{F}^\times} \psi\left(\frac{b-a}{4\varepsilon} N(z)\right) = q(q^2-1) - (q+1)(1-q)$$

and now

$$\begin{aligned}
\langle \phi, \phi \rangle &= q\left\{ 1 + \left(\frac{1}{q+1}\right)^2 \big(q(q^2-1) + (q-1)(q+1)\big) \right\} \\
&= q\left\{ 1 + \left(\frac{1}{q+1}\right)^2 \big((q+1)(q^2-1)\big) \right\}
\end{aligned}$$

$$= q\{1 + q - 1\}$$
$$= q^2. \qquad \square$$

Lemma 4.8. *For each non-trivial type 2 spherical function $\phi \in \Delta_2(U : H)$ one has $\langle \phi, \phi \rangle = q^3/(q+1)$ (and of course $\langle \phi_\circ, \phi_\circ \rangle = q^3$).*

Proof. Lemma 4.3 shows that $\phi^\circ \in \Delta(U : \widetilde{F}^+)$ and Lemma 3.5 gives

$$\langle \phi^\circ, \phi^\circ \rangle_{\widetilde{F}} = \begin{cases} \frac{q^2}{q+1} & \text{if } \phi^\circ \not\equiv 1 \\ q^2 & \text{if } \phi^\circ \equiv 1 \end{cases}.$$

So for $\phi \in \Delta_2(U : H)$

$$\begin{aligned} \langle \phi, \phi \rangle_H &= \sum_{(z,t) \in H} |\phi(z,t)|^2 \\ &= q \sum_{z \in \widetilde{F}} |\phi^\circ(z)|^2 \\ &= q \langle \phi^\circ, \phi^\circ \rangle_{\widetilde{F}} \\ &= \begin{cases} \frac{q^3}{q+1} & \text{if } \phi^\circ \not\equiv 1 \\ q^3 & \text{if } \phi^\circ \equiv 1 \end{cases}. \end{aligned}$$

$\square$

4.3 Ergodicity of $U : H$

We will write $\nu_{z_\circ, t_\circ} := \mu_{U \cdot (z_\circ, t_\circ)}$. The action pair $U : H$ is not *strongly* ergodic because U-orbits $U \cdot (0, t_\circ)$ through central elements $(0, t_\circ)$ in H are single points, $U \cdot (0, t_\circ) = \{(0, t_\circ)\}$. (The measure $\nu_{0,t_\circ} = \delta_{(0,t_\circ)}$ is of course not ergodic.) However, we will show $U : H$ is ergodic.

Theorem 4.9. *The action pair $U : H$ is ergodic. In fact $\nu_{z_\circ,t_\circ}$ is an ergodic probability measure for each point $(z_\circ, t_\circ)$ with $z_\circ \neq 0$.*

Proof. As $U : H$ is a Gelfand action pair Corollary 2.12 shows that $\nu_{z_\circ,t_\circ} = \mu_{U\cdot(z_\circ,t_\circ)}$ is ergodic if and only if $|\phi(z_\circ, t_\circ)| < 1$ for all $\phi \neq \phi_\circ \in \Delta(U : H)$.

For $\phi \in \Delta_1(U : H)$ we have shown in Corollary 4.5 that

$$|\phi(z_\circ, t_\circ)| \leq \frac{q}{q+1},$$

so $|\phi(z_\circ, t_\circ)| < 1$ for $\phi \in \Delta_1(U : H)$.

Now consider $\phi \in \Delta_2(U : H)$, $\phi \neq \phi_\circ$. We have $\phi(z_\circ, t_\circ) = \phi^\circ(z_\circ) = \phi(z_\circ, 0)$ where $\phi^\circ : \widetilde{F} \to \mathbb{C}$ is a spherical function for $U : \widetilde{F}^+$ with $\phi^\circ \not\equiv 1$. As we proved that $U : \widetilde{F}^+$ is strongly ergodic (see Proposition 3.11) it now follows that $|\phi^\circ(z_\circ)| < 1$, because $z_\circ \neq 0$ and $\phi^\circ \in \Delta(U : \widetilde{F}^+)$ and $\phi^\circ \not\equiv 1$. □

4.4 Convergence to equilibrium

The following result parallels Proposition 3.1.

Lemma 4.10. *Given $(z_\circ, t_\circ) \in H$ with $z_\circ \neq 0$ let $t_1 = t_\circ/N(z_\circ)$. Now one has*

$$\|\nu^{\star m}_{z_\circ,t_\circ} - \mathfrak{u}\|_{TV} = \|\nu^{\star m}_{1,t_1} - \mathfrak{u}\|_{TV}$$

for all $m \geq 1$.

Proof. Recall that convolution of functions $f, g \in L(H)$ is given by

$$(f \star g)(z,t) = \sum_{(w,s)\in H} f(w,s)g(z-w, t-s+Im(w\overline{z})).$$

Define an action of $\widetilde{F}^\times$ on H by

$$z_\circ \cdot (z,t) = \big(z_\circ z, N(z_\circ)t\big)$$

and an action of $\widetilde{F}^\times$ on $L(H)$ by

$$(z_\circ \cdot f)(z,t) = f\big(z_\circ^{-1}z, N(z_\circ^{-1})t\big).$$

Thus for $f, g \in L(H)$ we have

$$(z_\circ \cdot f) \star (z_\circ \cdot g)(z,t) = \sum_{(w,s)\in H} f\big(z_\circ^{-1}w, N(z_\circ^{-1})s\big) g\big(z_\circ^{-1}z - z_\circ^{-1}w, \\ N(z_\circ^{-1})t - N(z_\circ^{-1})s + N(z_\circ^{-1})Im(w\overline{z})\big).$$

But now as $N(z_\circ^{-1}) \in F$ we have $N(z_\circ)^{-1}Im(\alpha) = Im(N(z_\circ^{-1})\alpha)$ for any $\alpha \in \widetilde{F}$. (Indeed writing $\alpha = x + y\sqrt{\varepsilon}$, one has $N(z_\circ^{-1})\alpha = N(z_\circ^{-1})x + N(z_\circ^{-1})y\sqrt{\varepsilon}$ so

$$Im(N(z_\circ^{-1})\alpha) = N(z_\circ^{-1})y = N(z_\circ^{-1})Im(\alpha).)$$

Thus

$$N(z_\circ^{-1})Im(w\overline{z}) = Im(N(z_\circ^{-1})w\overline{z}) = Im(z_\circ^{-1}\overline{z_\circ^{-1}}w\overline{z}) = Im(z_\circ^{-1}w\overline{z_\circ^{-1}z})$$

and

$$(z_\circ \cdot f) \star (z_\circ \cdot g)(z,t) = \sum_{(w,s)\in H} f(\underbrace{z_\circ^{-1}w}_{w'}, \underbrace{N(z_\circ^{-1})s}_{s'}) g(z_\circ^{-1}z - \underbrace{z_\circ^{-1}w}_{w'}, \\ N(z_\circ^{-1})t - \underbrace{N(z_\circ^{-1})s}_{s'} + Im(\underbrace{z_\circ^{-1}w}_{w'}\overline{z_\circ^{-1}z}))$$

$$= \sum_{(w',s')\in H} f(w',s')g\big(z_\circ^{-1}z - w', N(z_\circ^{-1})t - s' + Im(w'\overline{z_\circ^{-1}z})\big)$$
$$= (f \star g)(z_\circ^{-1}z, N(z_\circ^{-1})t)$$
$$= \big(z_\circ \cdot (f \star g)\big)(z,t).$$

Now we have defined $\nu_{z_\circ,t_\circ}$ as the uniform measure on

$$U \cdot (z_\circ, t_\circ) = \big(Uz_\circ\big) \times \{t_\circ\} = z_\circ \cdot \left(U \times \left\{\frac{t_\circ}{N(z_\circ)}\right\}\right)$$

and so

$$\nu_{z_\circ,t_\circ} = z_\circ \cdot \nu_{1,t_1}$$

where $t_1 := t_\circ/N(z_\circ)$. Thus now

$$\nu_{z_\circ,t_\circ} \star \nu_{z_\circ,t_\circ} = (z_\circ \cdot \nu_{1,t_1}) \star (z_\circ \cdot \nu_{1,t_1}) = z_\circ \cdot \big(\nu_{1,t_1} \star \nu_{1,t_1}\big)$$

and iterating gives

$$\nu_{z_\circ,t_\circ}^{\star m} = z_\circ \cdot \left(\nu_{1,t_1}^{\star m}\right). \tag{4.3}$$

By using equation (4.3) we have

$$\begin{aligned}
\|\nu_{z_\circ,t_\circ}^{\star m} - \mathfrak{u}\|_{TV} &= \left\|z_\circ \cdot \left(\nu_{1,t_1}^{\star m}\right) - \mathfrak{u}\right\|_{TV} \\
&= \frac{1}{2}\left\|z_\circ \cdot \left(\nu_{1,t_1}^{\star m}\right) - \mathfrak{u}\right\|_1 \\
&= \frac{1}{2}\sum_{z,t}\left|\left(z_\circ \cdot \left(\nu_{1,t_1}^{\star m}\right)\right)(z,t) - \frac{1}{q^3}\right| \\
&= \frac{1}{2}\sum_{z,t}\left|\nu_{1,t_1}^{\star m}(z_\circ^{-1}z, N(z_\circ^{-1})t) - \frac{1}{q^3}\right| \\
&= \frac{1}{2}\sum_{w,s}\left|\nu_{1,t_1}^{\star m}(w,s) - \frac{1}{q^3}\right|
\end{aligned}$$

$$= \frac{1}{2} \left\| \nu_{1,t_1}^{\star m} - \mathfrak{u} \right\|_1$$
$$= \left\| \nu_{1,t_1}^{\star m} - \mathfrak{u} \right\|_{TV}. \qquad \square$$

Now fix $(z_\circ, t_\circ) \in H$ with $z_\circ \neq 0$ and write $\nu_{z_\circ,t_\circ} = \mu_{U\cdot(z_\circ,t_\circ)}$ as above. This is an ergodic probability measure by Theorem 4.9. We have shown that

$$\|\nu_{z_\circ,t_\circ}^{\star m} - \mathfrak{u}\|_{TV} = \|\nu_{1,t_1}^{\star m} - \mathfrak{u}\|_{TV}$$

where $t_1 = \frac{t_\circ}{N(z_\circ)} \in F$. By the Upper Bound Lemma (Lemma 2.13) now

$$\|\nu_{z_\circ,t_\circ}^{\star m} - \mathfrak{u}\|_{TV}^2 \leq \frac{|H|}{4} \sum_{\phi \neq \phi_\circ} \frac{|\phi(1,t_1)|}{\langle \phi, \phi \rangle}^{2m}$$

where the sum is over all non-trivial U-spherical functions $\phi \in \Delta(U : H)$. But for each such ϕ we have $\phi(1,t_1) = \phi^\circ(1)\psi(t_1)$ where $\phi^\circ(1) = \phi(1,0)$ and $\psi(t_1) = \phi(0,t_1)$ with $\psi : F \to \mathbb{C}$ an additive character. So

$$|\phi(1,t_1)| = |\phi^\circ(1)| \underbrace{|\psi(t_1)|}_{1} = |\phi^\circ(1)|.$$

Corollary 4.11. *For any $(z_\circ, t_\circ) \in H$ with $z_\circ \neq 0$ we have*

$$\|\nu_{z_\circ,t_\circ}^{\star m} - \mathfrak{u}\|_{TV}^2 \leq \frac{q^3}{4} \sum_{\phi \neq \phi_\circ} \frac{|\phi^\circ(1)|}{\langle \phi, \phi \rangle}^{2m}.$$

Now let

$$UB_1(m) := \frac{q^3}{4} \sum_{\phi \in \Delta_1} \frac{|\phi^\circ(1)|}{\langle \phi, \phi \rangle}^{2m}, \qquad UB_2(m) := \frac{q^3}{4} \sum_{\substack{\phi \in \Delta_2 \\ \phi \neq \phi_\circ}} \frac{|\phi^\circ(1)|}{\langle \phi, \phi \rangle}^{2m},$$

where $\Delta_1 = \Delta_1(U : H)$, $\Delta_2 = \Delta_2(U : H)$ so that

$$||\nu_{z_\circ,t_\circ}^{\star m} - \mathfrak{u}||_{TV}^2 \leq UB_1(m) + UB_2(m)$$

for all $m \geq 1$ and any $(z_\circ, t_\circ) \in H$ with $z_\circ \neq 0$.

Our study of the action pair $U : \widetilde{F}^+$ yields an explicit formula for $UB_2(m)$. Recall that $\widetilde{tr} : \widetilde{F} \to \mathbb{Z}_p$ denotes the absolute trace mapping for the finite field $\widetilde{F}$ (see Equation 3.5) and $g \in \widetilde{F}^\times$ is a primitive element ($\langle g \rangle = \widetilde{F}^\times$).

Proposition 4.12.

$$UB_2(m) = \frac{1}{4}\left(\frac{2}{q+1}\right)^{2m} \sum_{\ell=1}^{q^2-1} c_\ell^{2m}$$

where

$$c_\ell = \sum_{j=0}^{(q-1)/2} \cos\left(\frac{2\pi}{p}\widetilde{tr}(g^{\ell+(q-1)j})\right)$$

for $1 \leq \ell \leq q^2 - 1$.

Proof. Lemma 4.8 shows $\langle \phi, \phi \rangle = q^3/(q+1)$ for each $\phi \in \Delta_2(U : H) - \{\phi_\circ\}$ and hence

$$UB_2(m) = \frac{q+1}{4} \sum_{\substack{\phi \in \Delta_2 \\ \phi \neq \phi_\circ}} |\phi^\circ(1)|^{2m}.$$

But Lemma 4.3 shows $\{\phi^\circ \ : \ \phi \in \Delta_2(U : H)\} = \Delta(U : \widetilde{F}^+)$ and now Theorem 3.12 gives $UB_2(m) = UB_F(m)$ where $UB_F(m)$ is as in Equation 3.4. □

Next we will obtain an explicit summation formula for $UB_1(m)$. Namely:

Proposition 4.13.

$$UB_1(m) = \frac{q}{4}\left(\frac{2}{q^2-1}\right)^{2m} \sum_{a \in F^\times} \sum_{\ell=1}^{q} b_{a,\ell}^{2m}$$

where

$$b_{a,\ell} := \sum_{k\in\mathcal{S}} \cos\left(2\pi\left[\frac{\ell k}{q+1} + \frac{1}{p}tr\left(a\frac{Re(g^k)}{Im(g^k)}\right)\right]\right)$$

with, as before, $g \in \widetilde{F}^\times$ *a primitive element in the field* $\widetilde{F}$*, the map* $tr : F \to \mathbb{Z}_p$ *the absolute trace mapping for the field* F *(see Equation 3.2), and now*

$$\mathcal{S} := \left\{k \ : \ 1 \le k \le \frac{q^2-1}{2} \text{ and } (q+1) \nmid k\right\}.$$

(*Here* $Re(g^k)/Im(g^k)$ *means division in the field* F*. For* $k \in \mathcal{S}$ *one has* $Im(g^k) \neq 0$.)

The proof for Proposition 4.13, given below, requires preliminary Lemma 4.14. Recall, from Theorem 4.4, that each type 1 spherical function ϕ has the form $\phi = \phi_{\psi,\widetilde{\chi}}$ where

- $\psi \in \widehat{F^+}$ is a non-trivial additive character on F and
- $\widetilde{\chi} \in (\widetilde{F}^\times/F^\times)\widehat{\ }$ a non-trivial multiplicative character on $\widetilde{F}$ whose restriction to $F^\times$ is trivial.

Explicitly

$$\phi^\circ(1) = \phi^\circ_{\psi,\widetilde{\chi}}(1) = \frac{-1}{q^2-1}\sum_{w\in\widetilde{F}-F}\widetilde{\chi}(w)\psi\left(-\frac{1}{4\varepsilon}\frac{Re(w)}{Im(w)}\right).$$

From Equation 3.1 we know that the non-trivial additive characters on F are

$$\{\psi_a : a \in F^\times\} \quad \text{where} \quad \psi_a(x) := exp\left(\frac{2\pi i}{p}tr(ax)\right)$$

with $tr : F \to \mathbb{Z}_p$ the absolute trace map for the field F. The multiplicative characters $\widetilde{\chi} \in \widehat{(\widetilde{F}^\times)}$ can be described as follows. Let $g \in \widetilde{F}^\times$ be a primitive element in the extension field $\widetilde{F}$. As $\widetilde{F}^\times = \langle g\rangle$ is cyclic of order q^2-1, each character $\widetilde{\chi} : \widetilde{F}^\times \to \mathbb{C}^\times$ has the form

$$\widetilde{\chi}(g^k) = exp\left(\frac{2\pi i}{q^2-1}jk\right)$$

for some fixed j with $0 \leq j \leq q^2 - 2$. For $j = 0$ we have $\widetilde{\chi} \equiv 1$. So we require $j \geq 1$ as $\widetilde{\chi}$ must be a non-trivial character. But we also require that $\widetilde{\chi} \mid_{F^\times} \equiv 1$. It is easy to see that g^{q+1} is a primitive element in $F^\times$. That is, $\langle g^{q+1} \rangle = F^\times$. Indeed g^{q+1} has order $q-1$ and g^{q+1} belongs to F since $g^{q+1} = N(g)$. So

$$\widetilde{\chi} \mid_{F^\times} \equiv 1 \iff \widetilde{\chi}(g^{q+1}) = 1 \iff exp\left(\frac{2\pi i}{q-1} j\right) = 1 \iff (q-1) \mid j.$$

Thus the non-trivial characters $\widetilde{\chi} \in \widehat{(\widetilde{F}^\times)}$ for which $\widetilde{\chi} \mid_{F^\times} \equiv 1$ have

$$j \in \{q-1, 2(q-1), \ldots, q(q-1)\}.$$

Setting $j = l(q-1)$ (with $1 \leq l \leq q$) in the formula for $\widetilde{\chi}$ we now write

$$\widetilde{\chi}_l(g^k) = exp\left(\frac{2\pi i}{q+1} lk\right).$$

We have proved:

Lemma 4.14. *The non-trivial characters $\widetilde{\chi} \in \widehat{(\widetilde{F}^\times)}$ with $\widetilde{\chi} \mid_{F^\times} \equiv 1$ are precisely*

$$\widetilde{\chi}_1, \widetilde{\chi}_2, \ldots, \widetilde{\chi}_q.$$

Proof of Proposition 4.13. Now for $a \in F^\times$, $1 \leq l \leq q$ let

$$\phi_{a,\ell} = \phi_{\psi,\widetilde{\chi}}$$

as in Theorem 4.4 where

$$\psi = \psi_{(-4\varepsilon a)}, \qquad \widetilde{\chi} = \widetilde{\chi}_l.$$

We have shown that

$$\Delta_1(U:H) = \{\phi_{a,l} : a \in F^\times, 1 \le l \le q\}$$

lists $\Delta_1(U:H)$ without repetition. Explicitly

$$\phi^\circ_{a,l}(1) = -\frac{1}{q^2-1} \sum_{w \in \widetilde{F}-F} \widetilde{\chi}_l(w)\psi_1\left(a\frac{Re(w)}{Im(w)}\right).$$

We have $\widetilde{F}^\times = \{g, g^2, \ldots, g^{q^2-1} = 1\}$ and $F^\times = \langle g^{q+1}\rangle$. So

$$\widetilde{F} - F = \{g^k : 1 \le k \le q^2 - 1 \; and \; (q+1) \nmid k\}.$$

The proof of Lemma 4.6 shows that for each $w \in \widetilde{F} - F$ one has $w^{-1} \in \widetilde{F} - F$, $w^{-1} \neq w$ and the pair of terms for w and w^{-1} in the sum for $\phi^\circ_{a,l}(1)$ are complex conjugates of each other. So letting

$$\mathcal{S} := \left\{k \; : \; 1 \le k \le \frac{q^2-1}{2} \text{ and } (q+1) \nmid k\right\},$$

we have

$$\widetilde{F} - F = \{g^k : k \in \mathcal{S}\} \amalg \{(g^k)^{-1} = g^{q^2-1-k} : k \in \mathcal{S}\}$$

and so

$$\phi^\circ_{a,l}(1) = -\frac{1}{q^2-1} \sum_{k \in \mathcal{S}} 2Re\left(\widetilde{\chi}_l(g^k)\psi_1\left(a\frac{Re(g^k)}{Im(g^k)}\right)\right).$$

But

$$\widetilde{\chi}_l(g^k)\psi_1\left(a\frac{Re(g^k)}{Im(g^k)}\right) = exp\left(\frac{2\pi i}{q+1}lk + \frac{2\pi i}{p}tr\left(\frac{aRe(g^k)}{Im(g^k)}\right)\right)$$

and so

$$\phi^{\circ}_{a,l}(1) = -\frac{2}{q^2-1}\sum_{k\in\mathcal{S}}\cos\left(2\pi\left[\frac{lk}{q+1}+\frac{1}{p}tr\left(\frac{aRe(g^k)}{Im(g^k)}\right)\right]\right).$$

Now

$$UB_1(m) = \frac{q^3}{4}\sum_{\phi\in\Delta_1}\frac{|\phi^{\circ}(1)|^{2m}}{\langle\phi,\phi\rangle} = \frac{q}{4}\sum_{\phi\in\Delta_1}|\phi^{\circ}(1)|^{2m}$$

since $\langle\phi,\phi\rangle = q^2$ for $\phi\in\Delta_1(U:H)$ by Lemma 4.7. Thus

$$\begin{aligned} UB_1(m) &= \frac{q}{4}\sum_{a\in F^{\times}}\sum_{l=1}^{q}|\phi^{\circ}_{a,l}(1)|^{2m} \\ &= \frac{q}{4}\left(\frac{2}{q^2-1}\right)^{2m}\sum_{a\in F^{\times}}\sum_{l=1}^{q}b_{a,l}^{2m} \end{aligned}$$

where

$$b_{a,l} := \sum_{k\in\mathcal{S}}\cos\left(2\pi\left[\frac{lk}{q+1}+\frac{1}{p}tr\left(\frac{aRe(g^k)}{Im(g^k)}\right)\right]\right)$$

as stated. □

In summary we have proved the following theorem.

Theorem 4.15. *For each point $(z_\circ,t_\circ)\in H$ with $z_\circ\neq 0$ the probability measure $\nu_{z_\circ,t_\circ}$ is ergodic and one has*

$$||\nu^{\star m}_{z_\circ,t_\circ}-\mathfrak{u}||^2_{TV}\leq UB_1(m)+UB_2(m)$$

where functions UB_1 and UB_2 are given explicitly in Propositions 4.13 and 4.12 above. Note that these bounds do not depend on $(z_\circ,t_\circ)$.

4.5 Crude exponential estimates on UB_1 and UB_2

The explicit formulas given for $UB_1(m)$ and $UB_2(m)$ implement the upper bound on $\|\nu^{\star m}_{z_\circ,t_\circ} - \mathfrak{u}\|^2_{TV}$ (with $z_\circ \neq 0$) guaranteed by the Upper Bound Lemma. Working from these we can produce (much) weaker but (much) more tractable bounds.

Theorem 4.16. *For all $m \geq 1$ one has*

$$UB_1(m) \leq \frac{q^2(q-1)}{4}\left(\frac{q}{q+1}\right)^{2m} \tag{4.4}$$

and

$$UB_2(m) \leq \frac{q^2-1}{4}\left(\frac{q-1+2\alpha}{q+1}\right)^{2m} \tag{4.5}$$

where $\alpha := \cos(\pi/p)$.

As $0 < q/(q+1) < 1$ the right hand side of (4.4) approaches zero as m tends to infinity. Also $0 < \alpha < 1$ here and so $q-1 < q-1+2\alpha < q-1+2 = q+1$ and hence

$$0 < \frac{q-1+2\alpha}{q+1} < 1.$$

It follows that the right hand side of (4.5) approaches zero as m tends to infinity.

Proof. Corollary 4.5 shows that $|\phi^\circ(1)| \leq \frac{q}{q+1}$ for each $\phi \in \Delta_1(U:H)$. There are q distinct U-orbits in $\widetilde{F}$ and hence q^2 distinct U-orbits in $H = \widetilde{F} \times F$. So

$$|\Delta_1(U:H)| = q^2 - |\Delta_2(U:H)| = q^2 - |\Delta_2(U:\widetilde{F}^+)| = q^2 - q = q(q-1).$$

Thus

$$UB_1(m) = \frac{q}{4}\sum_{\phi\in\Delta_1}|\phi^\circ(1)|^{2m} \leq \frac{q}{4}q(q-1)\left(\frac{q}{q+1}\right)^{2m} = \frac{q^2(q-1)}{4}\left(\frac{q}{q+1}\right)^{2m},$$

establishing (4.4).

Recall that

$$UB_2(m) = \frac{q+1}{4} \sum_{\substack{\phi \in \Delta_2 \\ \phi \neq \phi_\circ}} |\phi^\circ(1)|^{2m} = \frac{1}{4} \left(\frac{2}{q+1} \right)^{2m} \sum_{l=1}^{q^2-1} c_l^{2m}$$

where

$$c_l = \sum_{j=0}^{(q-1)/2} \cos\left(\frac{2\pi}{p} \widetilde{tr}\big(g^{l+(q-1)j}\big) \right).$$

As $\nu_{1,0}$ is ergodic (Theorem 4.9) we know that $|\phi^\circ(1)| < 1$ for each $\phi \in \Delta_2(U : H) - \{\phi_\circ\}$ (Corollary 2.12) and hence that $\lim_{m\to\infty} UB_2(m) = 0$. Thus we must also have

$$\left| \frac{2}{q+1} c_l \right| < 1$$

for each $\ell = 1 \ldots q^2 - 1$. It follows that not all of the $\frac{q+1}{2}$ terms in the summation for c_l can be equal to 1. That is, writing

$$c_l = \sum_{j=0}^{(q-1)/2} \cos\left(\frac{2\pi}{p} k_j \right)$$

where $k_j := \widetilde{tr}(g^{l+(q-1)j}) \in \{0, 1, \cdots, p-1\}$ we must have $k_j \geq 1$ for at least one j. Now consider the inequality

$$|c_l| \leq \sum_{j=0}^{(q-1)/2} \left| \cos\left(\frac{2\pi}{p} k_j \right) \right|.$$

The largest possible value of $\left|\cos\left(\frac{2\pi}{p}k\right)\right|$ with $k \in \{1, 2, \cdots, p-1\}$ occurs for $k =$

$(p \pm 1)/2$. That is

$$\left|\cos\left(\frac{2\pi}{p}\left(\frac{p-1}{2}\right)\right)\right| = \left|\cos\left(\pi - \frac{\pi}{p}\right)\right| = \left|-\cos\left(\frac{\pi}{p}\right)\right| = \cos\left(\frac{\pi}{p}\right) = \alpha$$

Thus

$$|c_l| \leq \left(\frac{q-1}{2}\,1\right) + \alpha = \frac{q-1+2\alpha}{2}.$$

(That is, in the worst case we have $k_j = 0$ for $\frac{q-1}{2}$ terms in our summation and $k_j = \frac{p-1}{2}$ or $\frac{p+1}{2}$ for one term.) Now finally,

$$\begin{aligned} UB_2(m) &= \frac{1}{4}\left(\frac{2}{q+1}\right)^{2m} \sum_{l=1}^{q^2-1} c_l^{2m} \\ &\leq \frac{1}{4}\left(\frac{2}{q+1}\right)^{2m} (q^2-1)\left(\frac{q-1+2\alpha}{2}\right)^{2m} \\ &= \frac{q^2-1}{4}\left(\frac{q-1+2\alpha}{q+1}\right)^{2m}, \end{aligned}$$

establishing (4.5). □

As noted above, both $q/(q+1)$ and $(q-1-2\alpha)/(q+1)$ lie in the open interval $(0,1)$. So the the right-hand sides in these crude estimates do approach 0 as $m \to \infty$. As the total variation distance is at most 1, these estimates are, however, not useful until m becomes sufficiently large. For the estimate on $UB_1(m)$ we require

$$m \geq \frac{2\log(q) + \log(q-1) - \log(4)}{2(\log(q+1) - \log(q))}$$

for the right-hand side to be at most 1. Likewise the estimate on $UB_2(m)$ is not meaningful unless

$$m \geq \frac{\log(q^2-1) - \log(4)}{2(\log(q+1) - \log(q-1+2\alpha))}.$$

Indeed as regards the estimate on $UB_1(m)$,

$$\begin{aligned}
\frac{q^2(q-1)}{4}\left(\frac{q}{q+1}\right)^{2m} \leq 1 &\iff \left(\frac{q+1}{q}\right)^{2m} \geq \frac{q^2(q-1)}{4} \\
&\iff \log\left(\frac{q+1}{q}\right)^{2m} \geq \log\frac{q^2(q-1)}{4} \\
&\iff 2m\left(\log(q+1)-\log q\right) \geq 2\log q + \log(q-1) - \log 4 \\
&\iff m \geq \frac{2\log(q) + \log(q-1) - \log(4)}{2(\log(q+1)-\log(q))}.
\end{aligned}$$

Also as regards the estimate on $UB_2(m)$,

$$\begin{aligned}
\frac{q^2-1}{4}\left(\frac{q-1+2\alpha}{q+1}\right)^{2m} \leq 1 &\iff \left(\frac{q+1}{q-1+2\alpha}\right)^{2m} \geq \frac{q^2-1}{4} \\
&\iff 2m(\log(q+1)-\log(q-1+2\alpha)) \\
&\qquad\qquad \geq \log(q^2-1)-\log 4 \\
&\iff 2m \geq \frac{\log(q^2-1)-\log 4}{\log(q+1)-\log(q-1+2\alpha)} \\
&\iff m \geq \frac{\log(q^2-1)-\log(4)}{2(\log(q+1)-\log(q-1+2\alpha))}.
\end{aligned}$$

4.6 Numerical data

In this final section we take $F = \mathbb{Z}_p$, that is, $q = p^1 = p$ here. The computer algebra system Maple was used to program the formulas for $UB_1(m)$ and $UB_2(m)$. Table 4.1 gives a non-square $\varepsilon \in \mathbb{Z}_p$ and a primitive element $g \in \mathbb{Z}_p(\sqrt{\varepsilon})$ for the first 10 odd primes. Euler's criterion was used to find the non-squares ε. Namely, an element $a \in \mathbb{Z}_p$ is a non-square if and only if $a^{(p-1)/2}$ is not congruent to 1 modulo p.

p	ε	g
3	2	$1+\sqrt{2}$
5	2	$1+2\sqrt{2}$
7	3	$1+\sqrt{3}$
11	2	$1+5\sqrt{2}$
13	2	$1+2\sqrt{2}$
17	3	$1+2\sqrt{3}$
19	2	$1+9\sqrt{2}$
23	5	$1+\sqrt{5}$
29	2	$1+4\sqrt{2}$
31	3	$1+6\sqrt{3}$

Table 4.1: Nonsquares ε and primitive elements $g \in \mathbb{Z}_p(\sqrt{\varepsilon})$.

Table 4.2 lists the values of $UB_1(m)$ and $UB_2(m)$ for the first 15 steps of our random walk with $p = 7$. Note that the values of $UB_1(m)$ exceed 1 until $m = 3$ and that $UB_2(1) > 1$. (Recall that total variation distances cannot exceed 1.) After 7 steps we have $UB_1(7) \approx .004$, $UB_2(7) \approx .0006$ and after 15 steps we find $UB_1(7) \approx 4\times10^{-7}$, $UB_2(7) \approx 6\times10^{-8}$. The calculations used floating point arithmetic in Maple with Digits=10. The calculations were repeated using Digits=20. This revealed the presence of round off error effecting the final 3 digits in the values listed in Table 4.2. In Table 4.3 we show the values of $UB_1(m)$ and $UB_2(m)$ for the 10 odd primes after $m = 4, 8, 12$ steps.

Table 4.4 lists the actual total variation distances $\|\mu_U^{*m} - \mathfrak{u}\|_{TV}$ together with the

upper bound estimates $\sqrt{UB_1(m)+UB_2(m)}$ for the first 15 steps of our random walk with $p=7$. This confirms the fact that

$$\|\mu_U^{*m}-\mathfrak{u}\|_{TV} \leq \sqrt{UB_1(m)+UB_2(m)}$$

as expected. The forth column of Table 4.4 shows the percentage by which $\sqrt{UB_1(m)+UB_2(m)}$ exceeds $\|\mu_U^{*m}-\mathfrak{u}\|_{TV}$.

In outline the values for $\|\mu_U^{*m}-\mathfrak{u}\|_{TV}$ in Table 4.4 were produced as follows. An array $H[0..p-1,\ 0..p-1,\ 0..p-1]$ was initialized to store values for the function $(\mu_U-\mathfrak{u})(x,y,t)$ and a procedure created to replace the contents of array $H[x,y,t]$ with the result upon convolution with μ_U according to the following formula. For functions $f(x,y,t)$ on the Heisenberg group $\mathbb{Z}_p\times\mathbb{Z}_p\times\mathbb{Z}_p$ we have

$$(\mu_U\star f)(x,y,t)=\frac{1}{p+1}\sum_{(a+b\sqrt{\varepsilon})\in U} f\big((x-a)\ \bmod p,\ (y-b)\ \bmod p,\ (t+bx-ay)\ \bmod p\big).$$

In this way we implement the successive powers $\mu_U^{*m}-\mathfrak{u}$ for $m=1,2,\ldots$. Floating point arithmetic was used with Digits=10 and Digits=20 in Maple. This revealed that round off error can accumulate as m increases. For this reason the data in Table 4.4 was produced using Digits=20 but we display only the 4 most significant digits.

Finally Table 4.5 shows the values for the total variation distances and upper bound estimates for the first 10 odd primes after $m=4$,8 and 12 steps.

m	$UB_1(m)$	$UB_2(m)$
1	9.187500000	1.281250000
2	2.009765621	0.2524414060
3	0.5091705312	0.06776428220
4	0.1390218731	0.02031576628
5	0.03985889982	0.006307523680
6	0.01182060125	0.001979932239
7	0.003590037240	0.0006236913635
8	0.001109084200	0.0001966941122
9	0.0003469146640	0.00006205568390
10	0.000109517521	0.00001958074840
11	0.00003481602888	0.000006178696675
12	0.00001112826214	0.000001949716156
13	0.000003572240978	$6.152453465 \times 10^{-7}$
14	0.000001150710836	$1.941449700 \times 10^{-7}$
15	$3.717434478 \times 10^{-7}$	$6.126384535 \times 10^{-8}$

Table 4.2: $p = 7$ data.

p	$m = 4$	$m = 8$	$m = 12$
3	0.07690429695	0.003387629995	0.0001594503854
	0.003921508789	0.00001525902190	$5.960464835 \times 10^{-8}$
5	0.1662880086	0.007428867070	0.0003935643558
	0.01271629941	0.00007949466855	$5.540642855 \times 10^{-7}$
7	0.1390218731	0.001109084200	0.00001112826214
	0.02031576628	0.0001966941122	0.000001949716156
11	0.1588488083	0.0005492444028	0.000002535115472
	0.01114909553	0.00002279480482	$5.703840620 \times 10^{-8}$
13	0.1391618473	0.0002491009246	$6.345398572 \times 10^{-7}$
	0.008680763975	0.00001057266380	$1.665990296 \times 10^{-8}$
17	0.1294386848	0.0001013296581	$1.066803465 \times 10^{-7}$
	0.008683238815	0.00001011584548	$1.423437902 \times 10^{-8}$
19	0.1153507806	0.00005035984680	$3.144003770 \times 10^{-8}$
	0.005585814250	0.000001898919512	$7.538543360 \times 10^{-10}$
23	0.1088292357	0.00002931527008	$1.163053074 \times 10^{-8}$
	0.003561406515	$3.714536714 \times 10^{-7}$	$4.470172364 \times 10^{-11}$
29	0.09251515380	0.00001065535641	$1.740143474 \times 10^{-9}$
	0.002879422850	$2.081903350 \times 10^{-7}$	$1.729124278 \times 10^{-11}$
31	0.08419123298	0.000006337059128	$6.865275162 \times 10^{-10}$
	0.002833508630	$2.626949896 \times 10^{-7}$	$3.265465680 \times 10^{-11}$

Table 4.3: Some UB_1 and UB_2 values for the first 10 odd primes.

m	$\|\mu_U^{*m} - \mathfrak{u}\|_{TV}$	$\sqrt{UB_1(m) + UB_2(m)}$	% error
1	0.9767	3.236	231%
2	0.8338	1.504	80%
3	0.3775	0.7596	101%
4	0.2480	0.3992	61%
5	0.1307	0.2149	64%
6	0.07630	0.1175	54%
7	0.04217	0.06491	54%
8	0.02530	0.03614	43%
9	0.01372	0.02022	47%
10	0.008286	0.01136	37%
11	0.004489	0.006403	43%
12	0.002695	0.003616	34%
13	0.001470	0.002046	39%
14	0.0008736	0.001160	33%
15	0.0004804	0.0006580	37%

Table 4.4: Total variation distances and upper bound estimates for $p = 7$.

p	$m = 4$	$m = 8$	$m = 12$
3	0.2373 0.2843 20%	0.04843 0.05833 20%	0.01032 0.01263 22%
5	0.3147 0.4231 34%	0.06499 0.08665 33%	0.01503 0.01985 32%
7	0.2480 0.3992 61%	0.02530 0.03614 43%	0.002695 0.003616 34%
11	0.2196 0.4123 88%	0.01599 0.02392 50%	0.001160 0.001610 39%
13	0.1911 0.3845 101%	0.01056 0.01611 53%	0.0005837 0.0008070 38%
17	0.1692 0.3716 120%	0.006555 0.01056 61%	0.0002387 0.0003477 46%
19	0.1506 0.3478 131%	0.004512 0.007229 60%	0.0001285 0.0001794 40%
23	0.1385 0.3352 142%	0.003367 0.005449 62%	0.00007651 0.0001081 41%
29	0.1198 0.3089 158%	0.001904 0.003296 73%	0.00002771 0.00004192 51%
31	0.1119 0.2950 164%	0.001509 0.002569 70%	0.00001826 0.00002682 47%

Table 4.5: Some TV-distances and upper bound estimates for the first 10 odd primes.

REFERENCES

[1] Artin, M. (1991). *Algebra*. Prentice Hall Inc., Englewood Cliffs, NJ.

[2] Benson, C. and Ratcliff, G. (2008). Gelfand pairs associated with finite Heisenberg groups. In *Representations, wavelets, and frames*, Appl. Numer. Harmon. Anal., pages 13–31. Birkhäuser Boston, Boston, MA.

[3] Benson, C. and Ratcliff, G. (2009). Spherical functions for the action of a finite unitary group on a finite Heisenberg group. In *New developments in Lie theory and geometry*, volume 491 of *Contemp. Math.*, pages 151–166. Amer. Math. Soc., Providence, RI.

[4] Ceccherini-Silberstein, T., Scarabotti, F., and Tolli, F. (2008). *Harmonic analysis on finite groups*, volume 108 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge. Representation theory, Gelfand pairs and Markov chains.

[5] Diaconis, P. (1988). *Group representations in probability and statistics*. Institute of Mathematical Statistics Lecture Notes—Monograph Series, 11. Institute of Mathematical Statistics, Hayward, CA.

[6] Diaconis, P. and Shahshahani, M. (1981). Generating a random permutation with random transpositions. *Z. Wahrsch. Verw. Gebiete*, 57(2):159–179.

[7] Friedberg, S. H., Insel, A. J., and Spence, L. E. (1997). *Linear algebra*. Prentice Hall Inc., Upper Saddle River, NJ, fourth edition.

[8] Lidl, R. and Niederreiter, H. (1997). *Finite fields*, volume 20 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge, second edition. With a foreword by P. M. Cohn.

[9] Terras, A. (1999). *Fourier analysis on finite groups and applications*, volume 43 of *London Mathematical Society Student Texts*. Cambridge University Press, Cambridge.