RANDOM POLYNOMIALS OVER FINITE FIELDS

by

Andrew Sharkey

A thesis submitted to
the Faculty of Science
at the University of Glasgow
for the degree of
Doctor of Philosophy

June 22, 1999

Questa tesi è dedicata ad Alessandro Paluello (1975-1997).

Preface

This thesis is submitted in accordance with the regulations for the degree of Doctor of Philosophy in the University of Glasgow. It is the record of research carried out at Glasgow University between October 1995 and September 1998. No part of it has previously been submitted by me for a degree at any university.

I should like to express my gratitude to my supervisor, Prof. R.W.K.Odoni, for his advice and encouragement throughout the period of research. I would also like to thank Dr. S.D.Cohen for helping out with my supervision, and the E.P.S.R.C. for funding my research. Finally, I would like to thank my family for their support, and 'the boys', Brightwell, Ratter, Iain, Rob and Moh, without whom this period of study would not have been nearly so enjoyable.

A.Sh.

Contents

Abstract

The idea of this thesis is to take some questions about polynomials over finite fields and 'answer' them using probability theory; that is, we give the average behaviour of certain properties of polynomials. We tend to deal with multivariate polynomials, so questions about factorisation are not considered. Questions which are considered are ones concerning images and pre-images under a random polynomial mapping, and more generic questions which lead to results on the distributions of certain character sums over finite fields.

The methods used are based on those used by Odoni (details in Chapter 2). The probability space from which our random polynomial is chosen is essentially the set of all polynomials up to a given degree d, and we define a random variable associated with this space (for example, the number of zeros of a random polynomial). Once we have enough information about the random variable in question, we obtain asymptotic results about the distribution of this variable by letting both d and the size of the field, q, tend to infinty.

The results in this work tend to rely on comparisons between random polynomials (of degree up to d) and random mappings. We therefore do a certain amount of work with random mappings, exploiting nice combinatorial properties which they exhibit, and also using some non-trivial results from the classical theory of random maps. The resulting theorems for random polynomials, when interpreted number-theoretically, are often what one would expect, but every once in a while they cough up a surprise.

Introduction

Finite Fields

The theory of finite fields is a meeting point for several branches of mathematical science, including number theory, combinatorics, computing science, coding theory and cryptography. With their roots firmly embedded in classical number theory, finite fields traditionally been thought of as playing a small part in modern pure mathematics. However, since the dawn of the computer age in the nineteen-seventies, they have enjoyed a massive resurgence in popularity and are now the focus of both pure and applied mathematicians, with applications throughout the computer and telecommunications industries.

The study of finite fields can be traced to two brilliant mathematicians - Carl Friedrich Gauss (1777-1855), whose work on the arithmetic of congruences laid down the foundations; and Evariste Galois (1811-1832) who formulated the abstract notions required to construct these objects. In fact, many authors refer to finite fields as 'Galois Fields', hence attributing their invention to the Frenchman rather than the German.

Galois was interested in the problem of finding roots of polynomials over arbitrary fields. By extending the field in question and studying symmetric properties which these roots displayed, he was able to invent the theory of field extensions and their automorphisms, known today as Galois Theory. A spin-off from this work was the invention of group theory, as his discoveries later led mathematicians to formulate the abstract concept of a group.

Since the construction of finite fields arises from polynomials, the study of these two mathematical objects cannot be separated. Number theorists have, since the time of Galois, been examining properties of polynomials and related objects in order to gain insight into the structure of finite fields and their related objects. One important such example is Weil's study of function fields in one variable, algebraic objects closely related to algebraic curves over finite fields (see [41]). In 1973, Deligne proved the celebrated Weil conjectures on function fields, including the Riemann Hypothesis (see [22]). The latter is an analogue of the classical Riemann Hypothesis, which still remains unproved to this day. Deligne's results are among the most important in the classical theory of finite fields, having many applications, including some to coding theory (see [19]).

Probability Theory

For hundreds of years, probability has fascinated both mathematicians and bad gamblers. Throughout the seventeen and eighteen-hundreds, the basic theory was developed by the likes of Bernoulli, Gauss, Poisson, Chebyshev and many others, while the twentieth century saw an introduction of a more rigorous analytical, measure-theoretic approach. The use of probability theory is now widespread in discrete mathematics. For example, in number theory, primality testing is probabilistic while, in game theory, one often wants to know the 'chances of winning' a particular game. In combinatorics, there is a well-documented theory of random graphs, invented by Erdös et al. (see [6]) and probability theory is used more and more in the theory of finite groups, too.

Like number theory, probability theory is a beautiful area of classical mathematics. It can also turn out to be a useful tool when not enough is known about a mathematical system, giving a rigorous argument to a heuristic idea and shedding light on the general behaviour of the system. The methods involved are usually a combination of clever counting arguments and classical convergence theorems (essentially to do with measure theory). The arguments, therefore, are often a mixture of discrete and continuous mathematics, which giving them a certain charm.

This Thesis

In this thesis, we apply probabilistic methods to questions about polynomials over finite fields. The main reasons for this is that some interesting questions can be answered and, remarkably, few authors have thought of using probabilistic methods in this area.

The idea of applying probability theory to polynomials goes back to Offord and Littlewood [31], who did this over the real numbers. Over finite fields, we have authors such as Birch and Swinnerton-Dyer [5], Carlitz [9], Cohen [12], Odoni[36] and Knopfmacher [24] who have all, in some way, used probability in finite fields. However, there appears to be scope for the development of a more coherent and comprehensive theory of random polynomials over finite fields, akin to the existing theory of random graphs. Hopefully, this theory would eventually have applications in modern areas of research, too, for example in information technology.

In this work, we make a start towards this goal by gathering together and expanding on existing results in the literature. We begin by answering questions about image sizes of polynomial maps and stating some asymptotic results. Ultimately, we prove some results on the average behaviour of character sums which involve random polynomials, giving examples of different types of sum which can arise.

Chapter 1 is a general introduction to the results which we will need from classical probability theory. This is followed in Chapter 2 by a discussion of some of the basic concepts associated with random polynomials over a finite field. In Chapter 3, we then go on to give asymptotic distributions for the sizes of direct and inverse images of sets under random polynomial maps. Finally, Chapters 4&5 are concerned with constructing and finding the behaviour of a general complex-valued random variable which mimics character sums over a finite field. We conclude the thesis with a short sixth chapter discussing several directions in which the theory could go, showing that finite fields (and related mathematical structures) are objects which are particularly amenable to probabilistic techniques.

Chapter 1

Background: Probability Theory

1.1 Probability Spaces and Measures

In this chapter we gather together the various pieces of probability theory which will be used as 'tools' in the rest of the thesis. Some of this theory is elementary and some not so elementary. We first introduce formally the notion of a probability space.

Def 1.1.1. Let S be a set. A *σ-algebra* of sets in S is a collection $\mathcal{A}$ of subsets of S satisfying the following properties:

1. $\emptyset \in \mathcal{A}$

2. $T \in \mathcal{A} \Rightarrow S \backslash T \in \mathcal{A}$

3. If $\{T_i \ : \ i \in \mathbb{N}\}$ is a countable collection of elements of $\mathcal{A}$, then

$$\bigcup_{i \in \mathbb{N}} T_i \in \mathcal{A} \quad \text{and} \quad \bigcap_{i \in \mathbb{N}} T_i \in \mathcal{A}$$

That is, $\mathcal{A}$ contains the empty set, and is closed under the actions of taking complements and taking countable unions and intersections.

Def 1.1.2. A *probability measure* P on a σ-algebra $\mathcal{A}$ is a function $P : \mathcal{A} \rightarrow [0, 1]$ such that

1. $P(S) = 1$

2. If $\{T_i \ : \ i \in \mathbb{N}\}$ is a collection of disjoint sets in $\mathcal{A}$, then

$$P\left(\bigcup_{i \in N} T_i\right) = \sum_{i \in N} P(T_i)$$

Def 1.1.3. A *probability space* is a triple $(S, \mathcal{A}, P)$, where S is a set, $\mathcal{A}$ is a σ-field of sets in S, and P is a probability measure on $\mathcal{A}$.
S is called the *sample space*, while the sets in $\mathcal{A}$ are called events.

Examples

1. If S is a finite set, the natural way of making S into a probability space is to take $\mathcal{A}$ to be the power set of S and to define
$$P(T) = \frac{\sharp T}{\sharp S} \qquad \forall T \in \mathcal{A}$$
2. Let S be a compact subset of $\mathbb{R}^n$. Then the natural way of making S into a probability space is to take $\mathcal{A}$ to be the σ-field of Borel subsets of S and to define P to be the Lebesgue measure on S, normalised so that $P(S) = 1$.
3. Let K be a finite extension of the p-adic numbers $\mathbb{Q}_p$ with ring of integers $\mathcal{O}$. If π is a local parameter for $\mathcal{O}$ and the residue field $\frac{\mathcal{O}}{\pi\mathcal{O}}$ has degree u over $\mathbb{F}_p$ then Haar measure is defined by:
$$\mu_{haar}(a + \pi^n\mathcal{O}) = \frac{1}{p^{nu}} \qquad \forall a \in \mathcal{O}$$
This is the p-adic analogue of Lebesgue measure in $\mathbb{R}$. Thus, to make $\mathcal{O}$ into a probability space we define $\mathcal{A}$ to be the smallest σ-field containing all the cosets $a+\pi^n\mathcal{O} \quad (a \in \mathcal{O},\ n \in \mathbb{Z}_{\geq 0})$ and put $P = \mu_{haar}$.

A probability space is in reality a mathematical model for the process of randomisation. Given a probability space, we will talk about "choosing an element at random" from the sample space. The probability measure then gives us an idea of the likelihood that the element chosen lies in a particular subset of the sample space.

Since we will be dealing with finite fields, most of our probability spaces will be finite. Thus, Example 1 is the one which will occur throughout this work, although example 2 will also crop up.

Notation Let $(S, \mathcal{A}, P)$ be a probability space and let $T \in \mathcal{A}$. For a random element x of S we shall often write $Prob(x \in T)$ to mean $P(T)$, as the former is more intuitive.

1.2 Random Variables and Independence

Def 1.2.1. Let $(S, \mathcal{A}, P)$ be a probability space. A *real-valued random variable* is a function $X : \mathcal{A} \to \mathbb{R}$ such that, for each $t \in \mathbb{R}$ the set $S_t = \{y \in S : X(y) \leq t\}$ is in $\mathcal{A}$.

This means that each S_t can be assigned a probability $P(S_t)$, which we shall write as $Prob(X \leq t)$. The function $F_X : \mathbb{R} \to [0, 1]$ given by $F_X(t) = Prob(X \leq t)$ is called the *probability distribution function* of X.

When dealing with two or more random variables on the same sample space, it is often necessary to know whether or not there is any relation between them.

Def 1.2.2. Let $A, B \in \mathcal{A}$ be two events of the sample space S. We say that A and B are *independent* if

$$P(A \cap B) = P(A)P(B)$$

Now let X and Y be two random variables on S and define $A_s = \{a \in S \ : \ X(a) \leq s\}$ and $B_t = \{b \in S \ : \ Y(b) \leq t\} \quad (s, t \in \mathbb{R})$. We say that X and Y are *independent random variables* if A_s and B_t are independent events for all $(s,t) \in \mathbb{R}^2$.
We say $\{X_i \ : \ i \in \mathcal{I}\}$ is a collection of independent random variables if the X_i are *pair-wise* independent.

Often we shall refer of the pair $\mathbf{X} = (X, Y)$ as a *random variable in* $\mathbb{R}^2$. By the same token, we define a complex-valued random variable Z by

$$Z = X + iY$$

where (X, Y) is a random variable in $\mathbb{R}^2$.

The *joint distribution function* of a random variable $\mathbf{X} \in \mathbb{R}^2$ is defined to be

$$F_{X,Y}(s,t) = Prob(X \leq s \text{ and } Y \leq t)$$

Thus, if X and Y are independent then $F_{X,Y}(s,t) = F_X(s)F_Y(t)$.

The notion of a random variable in $\mathbb{R}^2$ generalises suitably to $\mathbb{R}^n$, but in this work we shall only be concerned with real and complex variables.

1.3 Expectation

The *expectation* or *mean* of a random variable X is defined in its most general form via the Lebesgue-Stieltjes integral

$$\mathbb{E}(X) = \int_{t \in \mathbb{R}} t \, dF_X(t)$$

If the distribution function F_X is piecewise differentiable with derivative f_X we say that X is a *continuous* random variable with *density (function)* f_X. In this case, the expectation is then given by

$$\mathbb{E}(X) = \int_{t \in \mathbb{R}} t \, f_X(t) \, dt$$

If, on the other hand, X takes values in a countable subset N of $\mathbb{R}$ (or $\mathbb{C}$) then we say that X is a *discrete* random variable and the expectation is then given by

$$\mathbb{E}(X) = \sum_{t \in N} t \, Prob(X = t)$$

In this case we say that the density function is $f_X(t) = Prob(X = t)$ $(\forall t \in N)$.
Note that the integrals/sums above need not converge, in which case the expectation is said to be infinite. In our work, however, it will always be finite.

The definition of expectation extends naturally to two variables:

$$\mathbb{E}((X,Y)) = (\mathbb{E}(X), \mathbb{E}(Y)) \quad \text{and} \quad \mathbb{E}(X + iY) = \mathbb{E}(X) + i\mathbb{E}(Y)$$

It is often necessary to know the expectation of a function of a random variable, but we first need to know when this is well-defined. The function $g : \mathbb{R} \to \mathbb{R}$ is called *$\mathcal{A}$-measurable* if the composition $g \circ X$ is itself a random variable. Similarly, $h : \mathbb{R}^2 \to \mathbb{R}$ is $\mathcal{A}$-measurable if $h \circ (X,Y)$ is a random variable.
In these cases, we have

$$\mathbb{E}(g(X)) = \int_{t\in\mathbb{R}} g(t)\; dF_X(t)$$

$$\text{and} \qquad \mathbb{E}(h(X,Y)) = \int\int_{\mathbf{t}\in\mathbb{R}^2} h(t_1,t_2)\; dF_{X,Y}(t_1,t_2)$$

1.4 Moments, Variance and Covariance

Def 1.4.1. Let X be a real-valued random variable on a probability space $(S, \mathcal{A}, P)$. Then, for each $k \in \mathbb{N}$, X^k is also a real-valued random variable. The *kth moment* of X is defined by

$$m_k = \mathbb{E}(X^k)$$

The moment sequence $\{m_k\}_{k\in\mathbb{N}}$ of a real-valued random variable is very useful in that knowledge of it can yield information about the distribution function. The simplest example of this is Chebyshev's inequality (below) while the most important example is the *method of moments*, explained in § 1.7.

The *variance* of a real-valued random variable is defined by

$$Var(X) = \mathbb{E}\left((X - \mathbb{E}(X))^2\right)$$

This is always positive and simplifies to

$$Var(X) = m_2 - m_1^2$$

The *standard deviation* of X, usually denoted σ_X, is defined to be the positive square root of $V(X)$ and is a measurement of the 'spread' of values of X around the mean. $\mathbb{E}(X)$ and $Var(X)$ will often be denoted μ_X and σ_X^2 in our work.

Theorem 1.4.2. *(Chebyshev's inequality) Let X be a real-valued random variable with mean μ and variance σ^2. Then, for $t > 0$*

$$Prob(\ |X-\mu| \geq t\sigma\ ) \leq \frac{1}{t^2}$$

Proof. See [16], p151 . □

Next, consider a random-variable (X, Y) in $\mathbb{R}^2$. We define the *covariance* of (X, Y) by

$$Cov(X,Y) = \mathbb{E}((X-\mathbb{E}(X))(Y-\mathbb{E}(Y)))$$

This simplifies to

$$Cov(X,Y) = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y) = Cov(Y,X)$$

X and Y are called *uncorrelated* if $Cov(X, Y) = 0$. It is easily shown that independent variables are uncorrelated but that the converse is false.

The correlation coefficient of (X, Y) is defined by

$$\rho_{X,Y} = \frac{Cov(X,Y)}{\sigma_X \sigma_Y}$$

and it lies in the range $[-1, 1]$. Noting that

$$Var(X) = Cov(X,X) \quad \text{and} \quad Var(Y) = Cov(Y,Y)$$

we define the *covariance matrix* of (X, Y) by

$$\begin{bmatrix} Cov(X,X) & Cov(X,Y) \\ Cov(X,Y) & Cov(Y,Y) \end{bmatrix}$$

This matrix, which is positive semi-definite, contains all the 'second order' moment information about (X, Y).

The idea of covariance extends naturally to n variables (see [15]).

Def 1.4.3. Let $Z = X + iY$ be a complex-valued random variable and let $k, l \in \mathbb{N}$. Then the (k, l)th moment of Z can be defined by

$$m_{k,l} = \mathbb{E}\left(X^k Y^l\right)$$

or by

$$M_{k,l} = \mathbb{E}\left(Z^k \overline{Z}^l\right)$$

However, since $X = \frac{1}{2}(Z + \overline{Z})$ and $Y = \frac{1}{2}(Z - \overline{Z})$, knowing all the moments $m_{k,l}$ is equivalent to knowing all the moments $M_{k,l}$ $(k, l \in \mathbb{N})$. In our work, the $M_{k,l}$ will prove to be easier to calculate.

1.5 Characteristic Functions

Def 1.5.1. The characteristic function of a real-valued random variable X is defined by

$$C_X(t) = \mathbb{E}(e^{itX}) = \int_{u\in\mathbb{R}} e^{itu}\, dF_X(u)$$

Due to their relation to complex integration theory and, in particular, to Fourier transforms, characteristic functions exhibit some very nice properties (see [17], § 5.7) These make them a powerful tool in determining the convergence of random variables to certain distributions.

If $Z = X + iY$ is a complex-valued random variable, then its characteristic function is defined via

$$C_Z(\mathbf{t}) = \mathbb{E}(e^{i\mathbf{t}.Z}) = \int\int_{\mathbf{u}\in\mathbb{R}^2} e^{i\mathbf{t}.\mathbf{u}}\, dF_{X,Y}(u_1, u_2)$$

where $\mathbf{t} = (t_1, t_2)$ and $\mathbf{t}.Z = t_1X + t_2Y$

The following result is useful when manipulating characteristic functions:

Proposition 1.5.2. *If* $S = X_1 + \ldots + X_k$ *is a sum of independent random variables. then*

$$C_S(\mathbf{t}) = \prod_{j=1}^{k} C_{X_j}(\mathbf{t}) \qquad (*)$$

Conversely, if $(*)$ *holds, then* S *is identical in distribution to the sum of the independent variables* $X_1, \ldots, X_k$.

Proof. See [17], p164 □

1.6 Some Probability Distributions

What follows is a description of each of the different types of random variable which occur in this thesis.

Bernoulli p-trials

Let $0 \le p \le 1$. A *Bernoulli p-trial* is an experiment which has two results: success and failure. This is modelled by the discrete random variable X which takes values in $\{0, 1\}$ and satisfies

$$Prob(X = 1) = p \qquad \text{and} \qquad Prob(X = 0) = 1 - p$$

This variable has mean p and variance $p(1 - p)$.

Binomial Distribution

A discrete random variable X is said to have the *binomial distribution* with parameters N, p if

$$Prob(X = k) = \binom{N}{k} p^k (1-p)^{N-k}$$

This variable occurs naturally as the sum of N independent Bernoulli p-trials. For example,

$$X = Y_1 + \ldots + Y_N$$

where the Y_i are pairwise independent Bernoulli p-variables.
X has mean Np and variance $Np(1-p)$.

Poisson Variables

A discrete random variable X which takes values on the non-negative integers is said to have a *Poisson distribution* with parameter λ $(\lambda \in \mathbb{R}, \lambda > 0)$ if

$$Prob(X = k) \; = e^{-\lambda} \frac{\lambda^k}{k!} \qquad (\forall k \geq 0)$$

This variable has mean and variance both equal to λ.

Gaussian Variables

A real-valued random variable X is said to have a *Gaussian (Normal) distribution* if

$$F_X(t) \; = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{t} e^{-\frac{1}{2}u^2} \, du$$

This variable has mean 0 and variance 1.

Bivariate Gaussian Distributions

Let $\mathbf{H}$ be a positive-definite real 2 symmetric matrix. A random variable (X, Y) in $\mathbb{R}^2$ is said to have a *bivariate Gaussian distribution* with covariance $\mathbf{H}$ if

$$F_{X,Y}(x_1, x_2) \; = \frac{1}{2\pi (\det \mathbf{H})^{\frac{1}{2}}} \int_{-\infty}^{x_1} \int_{-\infty}^{x_2} \exp\left[-\frac{1}{2} \mathbf{t}^t \mathbf{H}^{-1} \mathbf{t} \right] \, dt_1 dt_2$$

This is the same as

$$F_{X,Y}(x_1, x_2) \; = \frac{1}{2\pi \sigma_X \sigma_Y \sqrt{1-\rho^2}} \int_{-\infty}^{x_1} \int_{-\infty}^{x_2} \exp\left[-\frac{1}{2(1-\rho^2)} q(t_1, t_2) \right] \, dt_1 dt_2$$

where $\rho = \rho_{X,Y}$ is the correlation coefficient of X and Y and

$$q(x, y) = \left(\frac{x}{\sigma_X}\right)^2 - 2\rho \left(\frac{x}{\sigma_X}\right) \left(\frac{y}{\sigma_Y}\right) + \left(\frac{y}{\sigma_Y}\right)^2$$

is a positive-definite quadratic form in two variables.

This variable has mean zero and we shall say that (X,Y) has parameters σ_X, σ_Y and ρ. In the case that $\mathbf{H} = \mathbf{I}$ we say that (X,Y) has an *isotropic* $\mathbb{R}^2$-Gaussian distribution.

Compound Poisson Variables

Let $\mathcal{Z} = \{Z_i \ : \ i \in \mathbb{N}\}$ be a set of independent random variables with a common distribution F and let Y be a Poisson variable with parameter λ. We define the *compound Poisson variable X with parameters λ and F* (or λ and $\mathcal{Z}$) by

$$X = \sum_{i=1}^{Y} Z_i$$

Clearly if F is concentrated at 1 then X reduces to an ordinary Poisson variable.

Weibull Distributions

A real-valued random variable X on $[0,\infty)$ has a *Weibull distribution with parameters $a, b > 0$* if

$$F_X(t) = 1 - e^{-at^b}$$

This variable will occur with parameters $-\frac{1}{2\sigma^2}, 2$ when looking at the distribution of the modulus of a two-dimensional Gaussian variable with covariance matrix $\sigma\mathbf{I}$.

Bessel Distributions

Here we use terminology which differs from that used in standard texts (*e.g.* [40]; see Appendix E for clarification). We define the Bessel function of order $v > -1$, as in [16], by

$$I_v(\lambda) = \sum_{k=0}^{\infty} \frac{1}{k!\,\Gamma(k+v+1)} \left(\frac{\lambda}{2}\right)^{2k+v} \qquad (\forall t \in \mathbb{R})$$

where $\Gamma(s)$ is the Gamma function

$$\Gamma(s) = \int_0^{\infty} e^{-u} x^{s-1} du \qquad (\forall s \in \mathbb{C},\ \Re(s) > 0)$$

If v is an integer, then we get

$$I_v(\lambda) = \sum_{k=0}^{\infty} \frac{1}{k!(k+v)!} \left(\frac{\lambda}{2}\right)^{2k+v} \qquad (\forall t \in \mathbb{R})$$

In this case, we further define $I_v(\lambda) = I_{-v}(\lambda)$ for $v \in \mathbb{Z}, v \leq -1$. A $\mathbb{Z}$-valued random variable X with density function $f(v) = e^{-\lambda} I_v(\lambda)$ will be said to have a *Bessel distribution with parameter λ*. For more on Bessel functions, see appendix E.

1.7 Convergence Theorems

Def 1.7.1. Let X be a real-valued random variable and let $\{X_j\}$ $(j \in \mathbb{N})$ be a sequence of real-valued random variables. We say that the X_j *converge in distribution* to X if, for each compact interval of continuity $I \subseteq \mathbb{R}$ of F,

$$F_j(x) \to F(x) \qquad (\forall x \in I) \quad \text{as } j \to \infty$$

where F and F_j denote the distribution functions of X and X_j respectively.

For random variables in $\mathbb{R}^2$, the definition is analogous to the previous one:

Def 1.7.2. Let (X, Y) be a random variable in $\mathbb{R}^2$ and let $\{(X_j, Y_j)\}$ $(j \in \mathbb{N})$ be a sequence of random variables in $\mathbb{R}^2$. We say that the (X_j, Y_j) *converge in distribution* to (X, Y) if, for each compact interval of continuity $I \subseteq \mathbb{R}^2$ of F,

$$F_j(x, y) \to F(x, y) \qquad (\forall (x, y) \in I) \quad \text{as } j \to \infty$$

where F and F_j denote the joint distribution functions of (X, Y) and (X_j, Y_j) respectively.

As always we shall identify a complex-valued random variable $X+iY$ with its $\mathbb{R}^2$ counterpart (X, Y) so that the above definition can be applied to sequences of complex-valued variables.

Def 1.7.3. Let $\{X_j\}$ $(j \in \mathbb{N})$ be a sequence of real-valued random variables. We say that the X_j *converge in probability to zero* if, $\forall \epsilon > 0$,

$$Prob(|X_j| > \epsilon) \to 0 \qquad \text{as } j \to \infty$$

Lemma 1.7.4. *Let $\{X_j\}$ be as above and suppose that for each $j \in \mathbb{N}$ the first and second moments $m_{j,1}, m_{j,2}$ exist. If $m_{j,1}$, $m_{j,2} \to 0$ as $j \to \infty$ then the X_j converge in probability to zero.*

Proof. This is a simple application of Chebyshev's inequality (Theorem 1.4.2). □

The comparison of one distribution to another will be a recurring theme in our work. Indeed, we shall often be trying to prove that a given random variable converges to a well-known distribution as certain parameters vary. There are two ways in which this will be done: one using moments and the other using characteristic functions.

The Method of Moments

The following results essentially say that the convergence of a sequence of distributions is to a great extent governed by the convergence of the moments.

Theorem 1.7.5. *Let $\{X_j\}$ $(j \in \mathbb{N})$ and X be as above and suppose that the moments m_k of X and $m_{j,k}$ of X_j all exist $(k \in \mathbb{N})$. Suppose further that no other distribution has the same moment sequence as X. If, for each $k \in \mathbb{N}$,*

$$m_{j,k} \to m_k \qquad \textit{as } j \to \infty$$

then the X_j converge in distribution to X.

Proof. See [16], p269. □

We also have the analogous result for random variables in $\mathbb{R}^2$:

Theorem 1.7.6. *Let $\{(X_j, Y_j)\}$ $(j \in \mathbb{N})$ and (X, Y) be as above and suppose that the moments $m_{k,l}$ of X and $m_{j,k,l}$ of (X_j, Y_j) all exist $(k, l \in \mathbb{N})$. Suppose further that no other distribution has the same moment sequence as (X, Y). If, for each $k, l \in \mathbb{N}$*

$$m_{j,k,l} \to m_{k,l} \qquad \textit{as } j \to \infty$$

then the (X_j, Y_j) converge in distribution to (X, Y).

Proof. This is a relatively straight forward generalisation of 1.7.5. □

Note 1.7.7. The problem of two different distributions having the same moment sequence will not occur in our work, as all of our standard probability distributions have moment sequences which are peculiar to them.

The Continuity Theorem for Characteristic Functions

Unlike moment sequences, characteristic functions have the property that two different distributions have different characteristic functions. This is a consequence of the Fourier inversion theorem ([17] p170). The following theorem is therefore more powerful than 1.7.5:

Theorem 1.7.8. *(Continuity Theorem) The sequence $\{X_j\}$ of random variables (in $\mathbb{R}$ or $\mathbb{R}^2$) converges in distribution to the variable X if and only if the sequence $\{C_j(t)\}$ of characteristic functions converges point-wise to $C(t)$, the characteristic function of X.*

Proof. A slightly stronger version is proved in [16], p508. □

1.8 Central Limit Theorems

The *central limit theorem* plays an important role in many areas of probability theory and will be used often in this work, in various forms. We begin by stating the most basic version:

Theorem 1.8.1. *Let $X_1, \dots, X_N$ be mutually independent random variables with the same distribution F, such that $\mathbb{E}(X_j) = 0$ and $Var(X_j) = 1$. Define*

$$S_N^* = \frac{1}{\sqrt{N}} \sum_{j=1}^{N} X_j$$

Then, as $N \to \infty$, S_N^* converges in distribution to a Gaussian distribution.

Proof. See [16], p259. □

The next version is concerned with the convergence of the binomial distribution. Note that cases (ii) and (iii) are in a sense degenerate cases since 'freak' conditions are required for them to occur.

Theorem 1.8.2. *Let $X_1, \dots, X_N$ be mutually independent Bernoulli p-variables, where $p = p(N)$ and put*

$$S_N = \sum_{j=1}^{N} X_j \qquad \text{and} \qquad S_N^* = \frac{S_N \; - Np}{\sqrt{Np(1-p)}}$$

Then, as $N \to \infty$ we have the following:

1. *If $Np \to \infty$ then S_N^* converges in distribution to a Gaussian variable.*

2. *If $Np \to \lambda > 0$ then S_N converges in distribution to a Poisson variable with parameter λ.*

3. *If $Np \to 0$ then S_N converges in probability to zero.*

Proof. 1. See [17], p175.

2. We calculate the characteristic function of S_N:

$$\begin{aligned} C(t) &= \mathbb{E}(\exp(itS_N)) \\ &= \prod_{j=1}^{N} \mathbb{E}(\exp(itX_j)) \qquad \text{(by 1.5.2)} \\ &= \mathbb{E}(\exp(itX_1))^N \\ &= \left(1 + p(e^{it} - 1)\right)^N \qquad \text{(by Appendix F)} \\ &\sim \left(1 + \frac{\lambda(e^{it} - 1)}{N}\right)^N \\ &\to \exp \lambda(e^{it} - 1) \end{aligned}$$

Since the latter is the characteristic function of a Poisson-λ variable (see Appendix F), the result follows from the continuity theorem 1.7.8.

3. Follows from 1.7.4.

□

We shall also need a two-dimensional version of the central limit theorem:

Theorem 1.8.3. *Let* $\{\mathbf{X}_j\} = \{(X_j, Y_j)\}$ *be a sequence of mutually independent two-dimensional random variables with a common joint distribution* F*. Suppose that, for all* $j \in \mathbb{N}$,

$$\mathbb{E}(\mathbf{X}_j) = 0, \quad Var(X_j) = \sigma_1^2, \quad Var(Y_j) = \sigma_2^2, \quad Cov(X_j, Y_j) = \rho\sigma_1\sigma_2$$

Further, let us define

$$S_N^* = \frac{\mathbf{X}_1 + \ldots + \mathbf{X}_N}{\sqrt{N}}$$

Then as $N \to \infty$, S_N *converges in distribution to a bivariate Gaussian distribution with parameters* σ_1, σ_2, *and* ρ*; that is, a Gaussian distribution with covariance matrix*

$$\mathbf{H} = \begin{bmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix}$$

Proof. See [16], p 260. □

Note 1.8.4. In the event that all the $\mathbf{X}_j$ are real (*i.e.* $\sigma_2 = 0$) we get the degenerate case of a one-dimensional Gaussian distribution. More precisely, $\frac{S_N^*}{\sigma_1}$ converges in distribution to a Gaussian variable as $N \to \infty$.

1.9 A Note on Moments

In our work, often the moments $\mathbb{E}(X^k)$ of a discrete random variable X will be difficult to calculate directly. This is why we will revert to the 'trick' of calculating the expectation of the binomial coefficient:

$$M_k = \mathbb{E}\left(\binom{X}{k}\right)$$

Since $X^k \mapsto \binom{X}{k}$ is a non-singular $\mathbb{R}$-linear map on the space $\mathbb{R}[X]$ it follows that knowledge of the standard moment sequence $\{m_k\}$ is equivalent to knowledge of the binomial moment sequence $\{M_k\}$. Hence in Theorem 1.7.5 we may replace m_k with M_k.

We also note that we can retrieve the variance of X from M_2 and M_1 via the formula

$$Var(X) = 2\left(\mathbb{E}\left(\binom{X}{2}\right) - \binom{\mathbb{E}(X)}{2}\right) \tag{1.1}$$

Chapter 2

Random Polynomials

2.1 A Simple Model

Let q be a prime power, let $\mathbb{F}_q$ be the finite field with q elements and let X denote an indeterminate over $\mathbb{F}_q$. To begin with, we wish to pick a polynomial at random from $\mathbb{F}_q[X]$ and look at a certain property of it. The obvious way to do this is to take a positive integer d and define the probability space

$$\mathcal{F} = \mathcal{F}(q,d) = \{h \in \mathbb{F}_q[X] : \deg h \leq d\}$$

in the natural way (see §1.1). Clearly, $\mathcal{F}$ has q^{d+1} elements.

In general, we do some enumeration in $\mathcal{F}$ with respect to the property we are looking at (*e.g.* calculation of mean/variance and higher order moments). Once we have have some information about the distribution we then let d tend to infinity, thus giving the impression of choosing a polynomial at random from the whole of $\mathbb{F}_q[X]$. Hopefully, the probabilistic properties found will show some asymptotic behaviour as d tends to infinity.

As a simple example, consider the question: "How many zeros, on average, does a polynomial in $\mathbb{F}_q[X]$ have?" (Note that fuller treatments of this problem can be found in [36, 24].)

If $\zeta = \zeta(q,d)$ is the number of zeros of the random polynomial $f \in \mathcal{F}$ then we can calculate the expectation:

$$\begin{aligned}\mathbb{E}(\zeta) &= \sum_{h \in \mathcal{F}} \sharp\{a \in \mathbb{F}_q : h(a) = 0\} q^{-(d+1)} \\ &= \frac{1}{q^{d+1}} \sum_{a \in \mathbb{F}_q} \sharp\{h \in \mathcal{F} : h(a) = 0\}\end{aligned}$$

Now, by the factor theorem

$$h(a) = 0 \iff h(X) = (X-a)h_1(X) \qquad \text{for some } h_1 \in \mathbb{F}_q[X]$$

From this we see that the number of h vanishing at each $a \in \mathbb{F}_q$ is q^d. Hence

$$\begin{aligned} \mathbb{E}(\zeta) &= \frac{1}{q^{d+1}} \sum_{a \in \mathbb{F}_q} q^d \\ i.e.\ \mathbb{E}(\zeta) &= 1 \end{aligned}$$

Similarly we can calculate the second moment of ζ:

$$\begin{aligned} \mathbb{E}\left(\binom{\zeta}{2}\right) &= \sum_{h \in \mathcal{F}} \sharp\{\{a,b\} \subseteq \mathbb{F}_q : h(a) = h(b) = 0\} q^{-(d+1)} \\ &= \frac{1}{q^{d+1}} \sum_{\{a,b\} \subseteq \mathbb{F}_q} \sharp\{h \in \mathcal{F} : h(a) = h(b) = 0\} \end{aligned}$$

Again, by the factor theorem, for $a \neq b$,

$$h(a) = h(b) = 0 \iff h(X) = (X-a)(X-b)h_1(X) \qquad \textit{provided} \ \deg h \geq 2$$

From this we get that the number of h in $\mathcal{F}$ vanishing on $\{a,b\}$ must be q^{d-1}, provided $d \geq 2$. Hence

$$\begin{aligned} \mathbb{E}\left(\binom{\zeta}{2}\right) &= \frac{1}{q^{d+1}} \sum_{\{a,b\} \subseteq \mathbb{F}_q} q^{d-1} \\ i.e.\ \mathbb{E}\left(\binom{\zeta}{2}\right) &= \frac{1}{q^2}\binom{q}{2} \end{aligned}$$

We can now find the variance, using the formula in §1.9:

$$Var(\zeta) = 1 - \frac{1}{q}$$

We conclude (trivially) that as $d \to \infty$,

$$\mathbb{E}(\zeta) \sim 1 \qquad \text{and} \qquad Var(\zeta) \sim 1 - \frac{1}{q}$$

There are some points to note from this example:

1. In both the first and second moment calculations, the key step which allows us to do the enumeration is the 'reversal' of the sum. This will occur again and again in our work.
2. To calculate the first moment we required $d \geq 1$ while the second moment required $d \geq 2$. We shall see later that knowledge of the kth moment in this type of problem requires $d \geq k$.
3. We can also let $q \to \infty$ to get that $Var(\zeta) \sim 1$. We shall see later that, for questions of distribution, we we will always have to let q tend to infinity as well as d.

2.2 Several Variables - a Key Lemma

Let $r \in \mathbb{N}$ and let $X_1, \ldots, X_r$ be independent indeterminates over $\mathbb{F}_q$. Because of the existence of a Euclidean algorithm in $\mathbb{F}_q[X]$, polynomials in one variable lend themselves to properties which multivariate polynomials do not have. In $\mathbb{F}_q[X_1, \ldots, X_r]$ we therefore turn to some ideal theory to overcome this.

Firstly, we define our probability space via

$$\mathcal{F} = \{h \in \mathbb{F}_q[X_1, \ldots, X_r] : \deg h \leq d\}$$

where the degree of h is defined to be the maximum degree of all monomials $X_1^{i_1} X_2^{i_2} \ldots X_r^{i_r}$, the degree of such a monomial being $i_1 + \ldots + i_r$. Hence $\mathcal{F}$ is a finite set and is made into a probability space in the natural way.

The first thing to note about polynomials in several variables is that we have an extension to the factor theorem (see [2]):

$$\begin{aligned} h(a_1, \ldots, a_r) = 0 \quad &\Longleftrightarrow \quad h \in (X_1 - a_1, \ldots, X_r - a_r) \\ &\Longleftrightarrow \quad h = (X_1 - a_1)h_1 + \ldots + (X_r - a_r)h_r \end{aligned}$$

for some $h_1, \ldots, h_r \in \mathbb{F}_q[X_1, \ldots, X_r]$. We shall abbreviate $(a_1, \ldots, a_r)$ to $\mathbf{a}$, $(X_1, \ldots, X_r)$ to $\mathbf{X}$ and $(X_1 - a_1, \ldots, X_r - a_r)$ to $(\mathbf{X} - \mathbf{a})$.

In line with the example of §2.1, we would like to be able to count the number of h in $\mathcal{F}$ which vanish at a given point $\mathbf{a} \in \mathbb{F}_q^r$. For a fixed q and variable i, define

$$\mathcal{R}_i = \{h \in \mathbb{F}_q[X_1, \ldots, X_r] : \deg h \leq i\}$$

Then $\mathcal{R}_1 \subseteq \mathcal{R}_2 \subseteq \ldots$ is an increasing chain of $\mathbb{F}_q$-subspaces of the $\mathbb{F}_q$-algebra $\mathcal{R} = \mathbb{F}_q[X_1, \ldots X_r]$. Note that the $\mathcal{R}_i$ have finite codimension in $\mathcal{R}$ $(i \in \mathbb{N})$. Also, an ideal $I \lhd \mathcal{R}$ is an $\mathbb{F}_q$-subspace and so we have the increasing chain

$$I \cap \mathcal{R}_1 \subseteq I \cap \mathcal{R}_2 \subseteq \ldots$$

Recall that $\deg I$ is defined to be the codimension of I in $\mathcal{R}$. Assuming $\deg I$ to be finite, we have that the $\mathcal{R}_i \cap I$ $(i \in \mathbb{N})$ are of finite codimension in $\mathcal{R}$. Define λ_i to be the $\mathbb{F}_q$-dimension of $\frac{\mathcal{R}_i}{\mathcal{R}_i \cap I}$. Since

$$\frac{\mathcal{R}_i}{\mathcal{R}_i \cap I} \cong \frac{\mathcal{R}_i + I}{I}$$

we have that

$$\lambda_1 \leq \lambda_2 \leq \ldots$$

is a non-decreasing sequence of positive integers, tending to $\deg I$. The next lemma gives us a useful bound on how long it takes for the sequence to reach $\deg I$.

Lemma 2.2.1. *With the notation above,*

$$\lambda_i = \deg I \qquad \textit{provided } i \geq \deg I$$

Proof. The idea is to show that if $\lambda_i = \lambda_{i+1}$ for some $i \in \mathbb{N}$, then $\lambda_i = \lambda_{i+2}$. It would then follow that the sequence $\lambda_1, \lambda_2, \ldots$ is strictly increasing up to a critical index, i_0, and $\lambda_i = \deg I$ for all $i \geq i_0$. This would imply that the smallest possible value of i_0 is $\deg I$, occurring in the case that all the 'jumps' are of size 1. The result would then follow.

Now, the space $\mathcal{R}_i$ is spanned by the set of monomials of degree at most i. Suppose that we have found an i satisfying $\lambda_i = \lambda_{i+1}$. If M is a monomial of degree $i+2$, then $M = X_j M'$ for some $1 \leq j \leq r$ and some monomial M' of degree $i+1$. Since $\lambda_{i+1} = \lambda_i$ there exists an $h \in \mathcal{R}_i$ with $M' \equiv h \bmod I$. Now, $X_j h \in \mathcal{R}_{i+1}$ and so $X_j h \equiv h' \bmod I$, for some $h' \in \mathcal{R}_i$. Hence $M \equiv h' \bmod I$ and this means that $\lambda_{i+2} = \lambda_i$. □

Note 2.2.2. This observation, due to P.Vamos, appears in [36] with *strict* inequality. This is why we have reproved it here.

2.3 The General Model

Let $r, s \in \mathbb{N}$ and consider the set $\mathbb{F}_q[X_1, \ldots, X_r]^s$. This consists of all polynomial vectors of the form $\mathrm{f} = (f_1, \ldots, f_s)$ $(f_1, \ldots, f_s \in \mathbb{F}_q[X_1, \ldots, X_r])$ and each of these vectors induces a map from $\mathbb{F}_q^r$ to F_q^s via evaluation

$$\mathbf{a} \mapsto \mathbf{f}(\mathbf{a}) = (f_1(\mathbf{a}), \ldots, f_s(\mathbf{a}))$$

We will be looking at certain properties of this map by using random polynomials. Our probability space is defined as follows:

Let $g_1, \ldots, g_s \in \mathcal{F}_q[X_1, \ldots X_r]$ be fixed polynomials and let $d_1, \ldots, d_s \in \mathbb{N}$. For $1 \leq j \leq s$, we define

$$\mathcal{F}_j = g_j + \mathcal{R}_{d_j} = \{g_j + h : \deg h \leq d_j\}$$

and

$$\mathcal{F} = \mathcal{F}_1 \times \cdots \times \mathcal{F}_s \subseteq \mathbb{F}_q[X_1, \ldots, X_r]^s$$

Thus $\mathcal{F}$, as before, is a finite set and is made into a probability space in the natural way. $\mathcal{F}$ depends on parameters $q, d_1, \ldots, d_s$ and on $g_1, \ldots, g_s$. A random element of $\mathcal{F}$ will be referred to as a *'random polynomial (vector) of type* (r, s)*'*. We now give two results, essentially corollaries of 2.2.1, which will allow us to handle polynomial vectors in $\mathcal{F}$.

Corollary 2.3.1. *Let* $\mathbf{a}_1, \dots, \mathbf{a}_u \in \mathbb{F}_q^r$ *be distinct and let* $\mathbf{b}_1, \dots, \mathbf{b}_u \in \mathbb{F}_q^s$*. Then*

$$\sharp\{\mathbf{h} \in \mathcal{F} \;:\; \mathbf{h}(\mathbf{a}_i) = \mathbf{b}_i \;\; \forall i\} = q^{-su}\sharp\mathcal{F}$$

provided $d_j \geq u \qquad (1 \leq j \leq s)$

Proof. Let $\mathbf{h} \in \mathcal{F}$ and, for each $1 \leq i \leq r$, write $\mathbf{b}_i = (b_{i1}, \dots, b_{is})$. Then

$$\begin{aligned} & \mathbf{h}(\mathbf{a}_i) = \mathbf{b}_i \quad (\forall i) \\ \iff \quad & (h_1(\mathbf{a}_i), \dots, h_s(\mathbf{a}_i)) = (b_{i1}, \dots, b_{is}) \quad (\forall i) \\ \iff \quad & h_j(\mathbf{a}_i) = b_{ij} \quad (\forall i, j) \\ \iff \quad & h_j \equiv b_{ij} \mod (\mathbf{X} - \mathbf{a}_i) \quad (\forall i, j) \end{aligned}$$

For each j, the Chinese remainder theorem gives us a unique solution modulo I, where

$$I = \prod_{i=1}^{u} (\mathbf{X} - \mathbf{a}_i)$$

so that *deg* $I = u$. Hence, for each j, the number of solutions in $\mathcal{F}_j$ is $\sharp(\mathcal{F}_j \cap I) = \sharp(R_{d_j} \cap I)$. By Lemma 2.2.1, this number is $q^{-u}\sharp\mathcal{F}_j$ in each case, provided $d_j \geq u$.

In this case, the total number of solutions is just the product over all j, namely

$$\prod_{j=1}^{s} q^{-u}\sharp\mathcal{F}_j = q^{-su}\sharp\mathcal{F}$$

as required. □

Def 2.3.2. In the light of this result, we define $d = \min\{d_1, \dots, d_s\}$.

Corollary 2.3.3. *Let* $U \subseteq \mathbb{F}_q^r$, $V \subseteq \mathbb{F}_q^s$ *be non-empty sets of sizes u and v respectively. Then*

$$\sharp\{\mathbf{h} \in \mathcal{F} \;:\; \mathbf{h}(U) \subseteq V\} = \left(\frac{v}{q^s}\right)^u \sharp\mathcal{F}$$

provided $d \geq u$

Proof. Let $U = \{\mathbf{a}_1, \dots, \mathbf{a}_u\}$. Then

$$\begin{aligned} & \mathbf{h}(U) \subseteq V \\ \iff \quad & \mathbf{h}(\mathbf{a}_i) \in V \quad (\forall i) \\ \iff \quad & \mathbf{h}(\mathbf{a}_i) = \mathbf{b}_i \quad (\forall i) \text{ for some } \mathbf{b}_1, \dots, \mathbf{b}_u \in V \end{aligned}$$

By corollary 2.3.1, the above has $q^{-su}\sharp\mathcal{F}$ solutions (provided $d_j \geq u \;\forall j$), for each choice of $(\mathbf{b}_1, \dots, \mathbf{b}_u)$ in V^u. Since there are v^u such choices, the result follows. □

Note 2.3.4. Later, we shall use the fact that if $d_j < u$ for some j, then

$$\sharp\{\mathbf{h} \in \mathcal{F} \;:\; \mathbf{h}(U) \subseteq V\} \leq \left(\frac{v}{q^s}\right)^u \sharp\mathcal{F}$$

2.4 Random Polynomials versus Random Maps

Let A and B be arbitrary finite sets of sizes n and N respectively. The set B^A of all maps from A to B is also finite set, of cardinality N^n, and so we make this into a probability space in the natural way. A random element of this space will be called a *'random map of type (n, N)'*.

The basic idea of this thesis is to compare the random polynomial of type (r, s) with the random map of type (q^r, q^s). That is, we ask a question about a random polynomial vector, try to answer it for the random map between $\mathbb{F}_q^r$ and $\mathbb{F}_q^s$, and then find out under which conditions the polynomial vector has a similar behaviour. It is a well-known property of finite fields that, given any map from $\mathbb{F}_q$ to itself, there exists a polynomial in one variable of degree less than q which represents that map (see [30]). There is also an extension of this to r variables:

Proposition 2.4.1. *There is a one-to-one correspondence between mappings $\phi \in (\mathbb{F}_q)^{\mathbb{F}_q^r}$ and polynomials h in $\mathbb{F}_q[X_1, \dots, X_r]$ which satisfy $\deg_j h \leq q-1 \quad (\forall j)$. (Here, $\deg_j h$ means the degree of h as a polynomial in X_j.)*

Proof. Two polynomials g and h give rise to the same map

$$\begin{aligned}
&\iff \quad g(\mathbf{a}) = h(\mathbf{a}) \quad \forall \mathbf{a} \in \mathbb{F}_q^r \\
&\iff \quad g \equiv h \quad \bmod (X_1 - a_1, \dots, X_r - a_r) \quad \forall \mathbf{a} \in \mathbb{F}_q^r \\
&\iff \quad g \equiv h \quad \bmod \prod_{\mathbf{a} \in \mathbb{F}_q^r} (X_1 - a_1, \dots, X_r - a_r) \\
&\iff \quad g \equiv h \quad \bmod (X_1^q - X_1, \dots, X_r^q - X_r)
\end{aligned}$$

Since S is a complete set of residues modulo $(X_1^q - X_1, \dots, X_r^q - X_r)$, two distinct elements of S will give rise to different maps. Finally, note that $\sharp S = q^{q^r} = \sharp(\mathbb{F}_q)^{\mathbb{F}_q^r}$ and the result follows. □

Corollary 2.4.2. *There is a one-to-one correspondence between mappings $\phi \in (\mathbb{F}_q^s)^{\mathbb{F}_q^r}$ and polynomial vectors $\mathbf{h} = (h_1, \dots, h_s)$ which satisfy $deg_j h_i \leq q-1 \quad (\forall i, j)$.*

Proof. Immediate. □

These results tell us, that if the degrees involved are large enough, then polynomials and maps are interchangeable. This is echoed by our enumeration results, 2.3.1 & 2.3.3, if one looks at them carefully. For this reason our problems will usually be formulated with a large q and a smaller (but still large) d in mind.

2.5 Independence

After we have looked at the basic image-size problems in Chapter 3, we will go on to explore a way of applying random polynomials to character sums in Chapters 4&5. This will involve some tedious moment calculations which will be greatly simplified by means of the next proposition.

Let $A = \mathbb{F}_q^r$, $B = \mathbb{F}_q^s$ so that $n = q^r$, $N = q^s$ and let $\psi : B \to \mathbb{C}$ be any complex-valued function. If f is a random polynomial vector of type (r, s) and ϕ is a random map of type (n, N) then, for each $a \in A$, $\psi(\mathbf{f}(a)))$ and $\psi(\phi(a))$ are complex-valued random variables.

Proposition 2.5.1. *For any subset* $\{a_1, \dots, a_t\}$ *of* A $(1 \leq t \leq n)$, *we have*

1. *The random variables* $\psi(\phi(a_1)), \dots, \psi(\phi(a_t))$ *are independent.*
2. *The random variables* $\psi(\mathbf{f}(a_1)), \dots, \psi(\mathbf{f}(a_t))$ *are independent,* provided $d \geq t$.

Proof. 1. We require to show that, given any $z_1, \dots, z_t \in \mathbb{C}$,

$$P\left(\psi(\phi(a_i)) \leq z_i \ \forall i)\right) = \prod_{i=1}^{t} P\left(\psi(\phi(a_i)) \leq z_i)\right) \tag{2.1}$$

Note that here, '$\leq$' refers to the natural partial order on $\mathbb{C}$, namely

$$\omega_1 \leq \omega_2 \quad \Longleftrightarrow \quad \Re(\omega_1) \leq \Re(\omega_2) \text{ and } \Im(\omega_1) \leq \Im(\omega_2)$$

Now, in (2.1) we have

$$\begin{aligned} L.H.S. &= \frac{\sharp\{\phi \in B^A : \ \psi(\phi(a_i)) \leq z_i \ \forall i\}}{N^n} \\ &= \frac{1}{N^n} \sum_{\substack{\mathbf{b} \in B^t \\ \psi(b_i) \leq z_i \ \forall i}} \sharp\{\phi \in B^A : \ \phi(a_i) = b_i \ \forall i\} \\ &= \frac{1}{N^n} \sum_{\substack{\mathbf{b} \in B^t \\ \psi(b_i) \leq z_i \ \forall i}} N^{n-t} \\ &= \prod_{i=1}^{t} \left(\frac{1}{N} \sum_{\substack{b \in B \\ \phi(b) \leq z_i}} 1 \right) \end{aligned}$$

On the other hand,

$$\begin{aligned}
R.H.S. &= \prod_{i=1}^{t}\left(\frac{\sharp\{\phi\in B^A:\ \psi(\phi(a_i))\leq z_i\}}{N^n}\right)\\
&= \prod_{i=1}^{t}\left(\frac{1}{N^n}\sum_{\substack{b\in B\\ \phi(b)\leq z_i}}\sharp\{\phi\in B^A:\ \phi(a_i)=b_i\}\right)\\
&= \prod_{i=1}^{t}\left(\frac{1}{N^n}\sum_{\substack{b\in B\\ \phi(b)\leq z_i}}N^{n-1}\right)\\
&= \prod_{i=1}^{t}\left(\frac{1}{N}\sum_{\substack{b\in B\\ \phi(b)\leq z_i}}1\right)
\end{aligned}$$

and hence the result.

2. Similarly, in this case we require to show that, given any $z_1,\dots,z_t\in\mathbb{C}$,

$$P\left(\psi(\mathbf{f}(a_i))\leq z_i,\ \forall i)\right)=\prod_{i=1}^{t}P\left(\psi(\mathbf{f}(a_i))\leq z_i)\right) \tag{2.2}$$

Again, '$\leq$' refers to the natural partial order on $\mathbb{C}$.

Now, in (2.2) we have

$$\begin{aligned}
L.H.S. &= \frac{\sharp\{\mathbf{f}\in\mathcal{F}:\ \psi(\mathbf{f}(a_i))\leq z_i\ \forall i\}}{\sharp\mathcal{F}}\\
&= \frac{1}{\sharp\mathcal{F}}\sum_{\substack{\mathbf{b}\in B^t\\ \psi(b_i)\leq z_i\ \forall i}}\sharp\{\mathbf{f}\in\mathcal{F}:\ \mathbf{f}(a_i)=b_i\ \forall i\}\\
&= \frac{1}{\sharp\mathcal{F}}\sum_{\substack{\mathbf{b}\in B^t\\ \psi(b_i)\leq z_i\ \forall i}}q^{-st}\sharp\mathcal{F} \qquad \text{provided } d\geq t\text{, by 2.3.1}\\
&= \prod_{i=1}^{t}\left(\frac{1}{q^s}\sum_{\substack{b\in B\\ \mathbf{f}(b)\leq z_i}}1\right)
\end{aligned}$$

On the other hand,

$$\begin{aligned}
R.H.S. &= \prod_{i=1}^{t} \left(\frac{\sharp\{\mathbf{f} \in \mathcal{F}\ \psi(\mathbf{f}(a_i)) \leq z_i\}}{\sharp\mathcal{F}} \right) \\
&= \prod_{i=1}^{t} \left(\frac{1}{\mathcal{F}} \sum_{\substack{b \in B \\ \mathbf{f}(b) \leq z_i}} \sharp\{\mathbf{f} \in \mathcal{F} :\ \mathbf{f}(a_i) = b_i\} \right) \\
&= \prod_{i=1}^{t} \left(\frac{1}{\sharp\mathcal{F}} \sum_{\substack{b \in B \\ \mathbf{f}(b) \leq z_i}} q^{-s} \sharp\mathcal{F} \right) \qquad \text{provided } d \geq 1, \text{ by 2.3.1} \\
&= \prod_{i=1}^{t} \left(\frac{1}{q^s} \sum_{\substack{b \in B \\ \mathbf{f}(b) \leq z_i}} 1 \right)
\end{aligned}$$

and so the result is proved.

□

Corollary 2.5.2. *With the same notation as above,*

$$\mathbb{E}\left(\prod_{i=1}^{t} \psi(\phi(a_i)) \right) = \prod_{i=1}^{t} \mathbb{E}\left(\psi(\phi(a_i))\right)$$

$$\mathbb{E}\left(\prod_{i=1}^{t} \psi(\mathbf{f}(a_i)) \right) = \prod_{i=1}^{t} \mathbb{E}\left(\psi(\mathbf{f}(a_i))\right) \quad \textit{provided } d \geq t$$

Proof. Immediate, from §1.4 □

Chapter 3

Inverse and Direct Image Sizes

3.1 Definitions

Let $\mathbf{f}$ be a random polynomial vector of type (r, s). Thinking of $\mathbf{f}$ as a map from $\mathbb{F}_q^r$ to $\mathbb{F}_q^s$ we can ask the following two questions:

1. "Given a subset C of $\mathbb{F}_q^s$, what is the size of the set

$$\mathbf{f}^{-1}(C) = \{\mathbf{a} \in \mathbb{F}_q^r : \mathbf{f}(\mathbf{a}) \in C\}$$

i.e. the inverse-image of C under $\mathbf{f}$?"

2. "What is the size of the (direct) image of $\mathbb{F}_q^r$ under $\mathbf{f}$?"

The inverse-image question will turn out to be quite straight forward while the direct-image question is a little more involved. For a random map, the latter is known as the *'classical occupancy problem'* and is often stated as follows:

"Suppose that n balls are fired randomly into N boxes, the probability that any particular box is hit being $\frac{1}{N}$ for each shot. Find the distribution of the number of empty boxes as $n, N \to \infty$."

Note that the idea is to work with the *complement* of the image set. This problem has been studied in great detail by several authors: Weiss [42], Rényi [38], Erdős [14], Békéssy [3, 4], Ivchenko & Medvedev [21], Kolchin [26], and Sevastyanov [39]. The results are best summarised in [27, 28], and in §3.5 we will adopt the approach of [27].

The direct-image size of a random polynomial in *one variable* has been studied by Cohen [12] and Knopfmacher & Knopfmacher [25], although their approach is different from the one we use here. There is also a related paper by Birch & Swinnerton-Dyer [5], involving a different type of 'random polynomial'.

Notation Throughout this chapter and, in fact, the rest of the thesis, several pieces of notation will remain constant:

$\mathbf{f}$ is a random polynomial in $\mathcal{F}$, where is $\mathcal{F}$ is as in §2.3. A and B are sets of sizes n and N respectively while ϕ is a random map between them. Often we (tacitly) put $A = \mathbb{F}_q^r$, $B = \mathbb{F}_q^s$ so that $n = q^r$ and $N = q^s$, allowing us to compare $\mathbf{f}$ to ϕ. C is a subset of B or $\mathbb{F}_q^s$, of size c.

With the image-size problems above in mind we define

$$\zeta_C = \sharp \mathbf{f}^{-1}(C) \qquad \text{and} \qquad \zeta_C^* = \sharp \phi^{-1}(C)$$

$$\eta = \sharp \left(\mathbb{F}_q^s \backslash \mathbf{f}(\mathbb{F}_q^r)\right) \qquad \text{and} \qquad \eta^* = \sharp \left(B \backslash \phi(A)\right)$$

The idea is to find the behaviours of the random map variables ζ_C^* and η^*, then attempt to match the moments of each with its polynomial counterpart (ζ_C and η, respectively). We begin with the inverse-image problem.

3.2 The Distribution of ζ_C^*

The binomial moments of ζ_C^* are easy to calculate:

Lemma 3.2.1. *For $k \geq 0$,*

$$\mathbb{E}\left(\binom{\zeta_C^*}{k}\right) = \binom{n}{k}\left(\frac{c}{N}\right)^k$$

In particular, the moments of ζ_C^* depend only on $c = \sharp C$ and not on the actual choice of C.

Proof.

$$\begin{aligned}
\mathbb{E}\left(\binom{\zeta_C^*}{k}\right) &= \mathbb{E}\left(\sharp\{\text{k-subsets of } \phi^{-1}(C)\}\right) \\
&= N^{-n} \sum_{\phi \in B^A} \sharp\{\text{k-subsets of } \phi^{-1}(C)\} \\
&= N^{-n} \sum_{\substack{K \subseteq A \\ \sharp K = k}} \sharp\{\phi \in B^A \ : \ \phi(K) \subseteq C\} \\
&= N^{-n} \binom{n}{k} c^k N^{n-k} \\
&= \binom{n}{k}\left(\frac{c}{N}\right)^k
\end{aligned}$$

□

So from now on, we will write ζ_c^* instead of ζ_C^*. We are now in a position to determine the asymptotic behaviour of ζ_c^*:

Theorem 3.2.2. *Let $nc \to \infty$ and/or $N \to \infty$. Then*

1. *If $\frac{nc}{N} \to \lambda > 0$ then ζ_c^* converges in distribution to a Poisson variable with parameter λ.*

2. *If $\frac{nc}{N} \to \infty$ then $\frac{\zeta_c^* - \mu}{\sigma}$ converges in distribution to a Gaussian (normal) random variable where $\mu = \frac{nc}{N}$, $\sigma^2 = \frac{nc}{N}(1 - \frac{c}{N})$.*

3. *If $\frac{nc}{N} \to 0$ then ζ_c^* converges in probability to zero.*

Proof. Let $A = \{a_1, \ldots, a_n\}$. We can write $\zeta_c^* = Y_1 + \ldots + Y_n$ where each Y_i is a Bernouilli variable defined by

$$Y_i = \begin{cases} 1 & \phi(a_i) \in C \\ 0 & \phi(a_i) \notin C \end{cases}$$

So ζ_c^* is the sum of n independent Bernoulli p-trials, where $p = \frac{c}{N}$. As a consequence, ζ_c^* has a binomial distribution with parameters n and p and so has mean $\mu = np$ and variance $\sigma^2 = np(1-p)$.

Applying Theorem 1.8.2 gives the result. □

3.3 The Distribution of ζ_C

The key to determining the asymptotic behaviour of ζ_C as $q, d \to \infty$ is the method of moments (section 1.7). That is, we attempt to match the moments of ζ_C with the moments of ζ_c^*:

Lemma 3.3.1. *Let $A = \mathbb{F}_q^r$, $B = \mathbb{F}_q^s$ so that $n = q^r$, $N = q^s$; we may compare the variables ζ_C and ζ_c^*. Then, for each $k \in \mathbb{Z}$ with $0 \leq k \leq d$ we have*

$$\mathbb{E}\left(\binom{\zeta_C}{k}\right) = \mathbb{E}\left(\binom{\zeta_c^*}{k}\right)$$

Proof.

$$\begin{aligned}
\mathbb{E}\left(\binom{\zeta_C}{k}\right) &= \mathbb{E}\left(\sharp\{\text{k-subsets of } \mathbf{h}^{-1}(C)\}\right) \\
&= \frac{1}{\sharp\mathcal{F}} \sum_{\mathbf{h}\in\mathcal{F}} \sharp\{\text{k-subsets of } \mathbf{h}^{-1}(C)\} \\
&= \frac{1}{\sharp\mathcal{F}} \sum_{\substack{K\subseteq A \\ \sharp K = k}} \sharp\{\mathbf{h}\in\mathcal{F} \,:\, \mathbf{h}(K)\subseteq C\} \\
&= \frac{1}{\sharp\mathcal{F}} \sum_{\substack{K\subseteq A \\ \sharp K = k}} \left(\frac{c}{q^s}\right)^k \sharp\mathcal{F} \qquad \textit{provided } d \geq k, \text{ by Cor.2.3.3} \\
&= \binom{q^r}{k}\left(\frac{c}{q^s}\right)^k \qquad \textit{provided } k \leq d \\
&= \mathbb{E}\left(\binom{\zeta_c^*}{k}\right) \qquad \textit{provided } k \leq d
\end{aligned}$$

□

Now suppose that we let q tend to infinity. Then $n, N \to \infty$ and ζ_c^* will behave in accordance with Theorem 3.2.2. If we also let d tend to infinity, we have by 3.3.1 that the moment sequence of ζ_C tends to that of ζ_c^*. Hence, by the method of moments, the two variables have the same limiting behaviour. We therefore have the main result on inverse-images:

Theorem 3.3.2. *Let $q \to \infty$ and let $d \to \infty$ arbitrarily slowly with q. Then*

1. *If $cq^{r-s} \to \lambda > 0$ then ζ_C converges in distribution to a Poisson variable with parameter λ.*

2. *If $cq^{r-s} \to \infty$ then $\frac{\zeta_C - \mu}{\sigma}$ converges in distribution to a Gaussian variable, where $\mu = cq^{r-s}$ and $\sigma^2 = cq^{r-s}(1 - q^{-s})$.*

3. *If $cq^{r-s} \to 0$ then ζ_C converges in probability to zero.*

3.4 Applications

1. Number of Zeros of f

Let us put $C = \{\mathbf{0}\}$. Then ζ_C is the number of zeros of $\mathbf{f}$. In fact, this problem was solved in [36], the paper which originally introduced the probability space $\mathcal{F}$ which we use in this thesis.

Corollary 3.4.1. *(Odoni) Let $q \to \infty$ and let $d \to \infty$ arbitrarily slowly with q. Then*

1. *If $r = s$ then $\zeta_{\mathbf{0}}$ converges in distribution to a Poisson variable with parameter* 1.

2. *If $r > s$ then $\frac{\zeta_{\mathbf{0}} - \mu}{\sigma}$ converges in distribution to a Gaussian variable, where $\mu = q^{r-s}$ and $\sigma^2 = q^{r-s}(1 - q^{-s})$.*

3. If $r < s$ then ζ_0 converges in probability to zero.

ζ_0 can be thought of as the number of $\mathbb{F}_q$-rational points on the variety in $\mathbb{F}_q^r$ defined by $\mathbf{f} = \mathbf{0}$. This application relates to deep results from algebraic geometry, due to Weil [41] and Deligne [13] (see [36] for details).

2. Polynomials Taking Values at Primitive Roots

Let $P = \{\text{primitive roots of } \mathbb{F}_q\}^s$. Then ζ_P is the number of $\mathbf{a} \in \mathbb{F}_q^r$ which are mapped to s-tuples of primitive roots under the random polynomial vector $\mathbf{f}$. Since there are $\varphi(q-1)$ primitive roots in $\mathbb{F}_q$, we have $c = \varphi(q-1)^s$.

It is known (see [18], p.267) that

$$\limsup_{m\to\infty} \frac{\varphi(n)}{n} = 1 \tag{3.1}$$

while

$$\liminf_{m\to\infty} \frac{\varphi(n)\log\log n}{n} = K \qquad (\text{where } K > 0) \tag{3.2}$$

It follows that

$$cq^{r-s} = q^r \left(\frac{\varphi(q-1)}{q}\right)^s \to \infty \qquad \text{as } q \to \infty \tag{3.3}$$

This gives us the following result:

Corollary 3.4.2. *Let $q \to \infty$ and let $d \to \infty$ arbitrarily slowly with q. Then $\frac{\zeta_P - \mu}{\sigma}$ converges in distribution to a Gaussian variable, where*

$$\mu = q^r \left(\frac{\varphi(q-1)}{q}\right)^s \quad \textit{and} \quad \sigma^2 = q^r \left(\frac{\varphi(q-1)}{q}\right)^s (1 - q^{-s})$$

Let $r = s = 1$ so that ζ_P is asymptotically Gaussian with mean $\varphi(q-1)$ and variance $(1-q^{-1})\varphi(q-1)$. In [32], Madden proved that, given a fixed $d \in \mathbb{N}$, then for all q sufficiently large, every square-free polynomial in $\mathbb{F}_q[X]$ of degree less than d represents a primitive root in $\mathbb{F}_q$. This is equivalent to saying that the inverse-image size of P under a square-free $h \in \mathcal{R}_d$ is *not zero*. It is known (see [9]) that the proportion of square-free polynomials is $1 - q^{-1}$ (*i.e.* almost all of them) and so our result is somewhat consistent with Madden's, although not as specific.

3.5 Direct Image Size and Classical Occupancy

We now move on to the more complicated direct-image problem and begin by calculating the moments of η^*:

Lemma 3.5.1.

$$\mathbb{E}\left(\binom{\eta^*}{k}\right) = \binom{N}{k}\left(1-\frac{k}{N}\right)^n \qquad \forall k \geq 0$$

Proof.

$$\begin{aligned}
\mathbb{E}\left(\binom{\eta^*}{k}\right) &= \mathbb{E}(\sharp\{\text{k-subsets of } B\backslash\phi(A)\}) \\
&= \frac{1}{\sharp B^A}\sum_{\phi\in B^A} \sharp\{\text{k-subsets of } B\backslash\phi(A)\} \\
&= \frac{1}{\sharp B^A}\sum_{\substack{K\subseteq B \\ \sharp K=k}} \sharp\{\phi\in B^A \ : \ \phi(A)\cap K=\emptyset\}
\end{aligned}$$

Let us write $A=\{a_1,\dots,a_n\}$ and define $E_i := \{\phi\in B^A \ : \ \phi(A)\in K\}$, where K is fixed for the moment.

Then $\{\phi\in B^A \ : \ \phi(A)\cap K=\emptyset\} = B^A\backslash\bigcup_{i=1}^n E_i$. By the inclusion-exclusion principle we have

$$\begin{aligned}
N-\sharp\{\phi\in B^A \ : \ \phi(A)\cap K=\emptyset\} &= \sum_{l=1}^{n}(-1)^{l-1}\sum_{\substack{L\subseteq\{1,\dots,n\} \\ \sharp L=l}} \sharp\bigcap_{i\in L} E_i \\
&= \sum_{l=1}^{n}(-1)^{l-1}\binom{n}{l}k^l N^{n-l} \\
&= N-(N-k)^n
\end{aligned}$$

Thus,

$$\begin{aligned}
\mathbb{E}\left(\binom{\eta^*}{k}\right) &= \frac{1}{N^n}\sum_{\substack{K\subseteq B \\ \sharp K=k}}(N-k)^n \\
&= \frac{1}{N^n}\binom{N}{k}(N-k)^n \\
&= \binom{N}{k}\left(1-\frac{k}{N}\right)^n
\end{aligned}$$

as required. □

Remark The above proof is long-winded in that the inclusion-exclusion principle is not needed. Indeed, there is a much more direct method of obtaining the result $\sharp\{\phi\in B^A \ : \ \phi(A)\cap K=\emptyset\}=(N-k)^n$. However, as we shall soon see, the use of the inclusion-exclusion principle is crucial to the estimation of the moments of η, which is why we have used it here.

Corollary 3.5.2. *The variable η^* has mean and variance given (respectively) by*

$$\mu = N\left(1-\tfrac{1}{N}\right)^n \qquad \textit{and} \qquad \sigma^2 = N(N-1)\left(1-\tfrac{2}{N}\right)^n + N\left(1-\tfrac{1}{N}\right)^n - N^2\left(1-\tfrac{1}{N}\right)^{2n}$$

Proof. Put $k=1$ and $k=2$ above (and use the formula from §1.9). □

The study of the asymptotic behaviour of η^* is complicated, splitting into several cases, depending on how n and N tend to infinity relative to each other. Heuristically, if n is much larger than N then the number of 'empty boxes' will be very close to zero, *i.e.* ϕ is surjective with probability approaching 1. This would lead us to limit ourselves to looking at paths of $n, N \to \infty$ in which n is 'not too large' compared with N. We make this idea precise with the following lemma:

Lemma 3.5.3. *Suppose $n, N \to \infty$ in such a way that $N\log N = o(n)$. Then η^* converges to zero in probability.*

Proof. Write $n = tN\log N$; then by Cor.3.5.2, we have

$$\begin{aligned} \log\mu &= \log N + n\log\left(1-\frac{1}{N}\right) \\ &= \log N\left\{1 + t\log\left(1-\frac{1}{N}\right)\right\} \\ &\sim (1-t)\log N \\ &\to -\infty \qquad \text{since } t\to\infty \\ \therefore \quad \mu &\to 0 \end{aligned}$$

Also, by Cor.3.5.2, $0 \le \sigma^2 \le \mu$ so that $\sigma \to 0$ also. The asserted result now follows from 1.7.4. □

From now on, we assume that $n = O(N\log N)$. Chapter 1 of [27] gives a comprehensive account of the behaviour of η^* under this condition (it calls this variable $\mu_0(n,N)$), and we summarise the results here.

Def 3.5.4. The behaviour of η^* splits into the following cases:

(a) $\frac{n}{N} \to 0,\ \ \frac{n^2}{2N} \to \lambda < \infty,\ \lambda \ge 0$ (Left-hand domain)
(b) $\frac{n}{N} \to 0,\ \ \frac{n^2}{2N} \to \infty$ (Left-hand intermediate domain)
(c) $0 < c_1 < \frac{n}{N} < c_2 < \infty$ (Central domain)
(d) $\frac{n}{N} \to \infty,\ \ \mathbb{E}(\eta^*) \to \infty$ (Right-hand intermediate domain)
(e) $\frac{n}{N} \to \infty,\ \ \mathbb{E}(\eta^*) \to \lambda < \infty$ (Right-hand domain)

Theorem 3.5.5. *In the left-hand domain, the variable $\eta^* - (N-n)$ converges in distribution to a Poisson variable with parameter λ.*

Proof. See chapter 1, sections 2 & 4 of [27]. □

Theorem 3.5.6. *In the left-hand intermediate, right-hand intermediate and central domains, $\frac{\eta^*-\mu}{\sigma}$ converges in distribution to a Gaussian variable.*

Proof. See Chapter 1, section 3 of [27]. □

Theorem 3.5.7. *In the right-hand domain, η^* converges in distribution to a Poisson variable with parameter λ.*

Proof. See Chapter 1, section 1 of [27]. □

Remarks

1. In the case $n = q^r$, $N = q^s$, the right-hand intermediate domain (d) never occurs, since

$$\frac{n}{N} \to \infty \quad \Rightarrow \quad r > s \quad \Rightarrow \quad q^s\left(1 - \frac{1}{q^s}\right)^{q^r} \to 0 \quad \Rightarrow \quad \mathbb{E}(\eta^*) \to 0$$

2. In (a) and (e) the case $\lambda = 0$ is allowed, the corresponding random variable then tending to 0 in probability.

3.6 The Moments of η and η^*

Now that we know the limiting behaviour of η^* we attempt to match this variable with η, our random polynomial equivalent. From now on we assume that $A = \mathbb{F}_q^r$ and $B = \mathbb{F}_q^s$ and let us attempt to calculate the moments for η:

$$\begin{aligned}
\mathbb{E}\left(\binom{\eta}{k}\right) &= \frac{1}{\sharp\mathcal{F}} \sum_{\mathbf{h}\in\mathcal{F}} \sharp\{\text{k-subsets of } (\mathbb{F}_q^s \backslash \mathbf{h}(\mathbb{F}_q^r))\} && (3.4)\\
&= \frac{1}{\sharp\mathcal{F}} \sum_{\substack{K\subseteq\mathbb{F}_q^s\\ \sharp K=k}} \sharp\{\mathbf{h}\in\mathcal{F} \;:\; \mathbf{h}(\mathbb{F}_q^r) \subseteq \mathbb{F}_q^s\backslash K\} && (3.5)\\
&= \frac{1}{\sharp\mathcal{F}} \sum_{\substack{K\subseteq\mathbb{F}_q^s\\ \sharp K=k}} \left(\frac{q^s-k}{q^s}\right)^{q^r} \sharp\mathcal{F} \qquad \text{provided } d \geq q^r \text{ (by 2.3.3)} && (3.6)\\
&= \binom{q^s}{k}\left(1 - \frac{k}{q^s}\right)^{q^r} \qquad \text{provided } d \geq q^r && (3.7)
\end{aligned}$$

In [12], Cohen obtained the following expression for the average number of values $v(t)$ taken by a polynomial in $\mathbb{F}_q[X]$ of degree t:

$$v(t) = q\left[1 - \left(1 - q^{-1}\right)^q\right] \qquad \text{for } t \geq q \tag{3.8}$$

From this it was deduced that

$$v(t) \sim q\left(1 - e^{-1}\right) \qquad \text{as } q \to \infty \tag{3.9}$$

We can generalise 3.8 to polynomial vectors satisfying $d \geq q^r$ by putting $k = 1$ in 3.7:

$$\mathbb{E}(q^s - \eta) = q^s\left[1 - \left(1 - q^{-s}\right)^{q^r}\right] \tag{3.10}$$

If $r = s$, then we get a generalisation of 3.9:

$$\mathbb{E}(q^s - \eta) \sim q^r\left(1 - e^{-1}\right) \qquad \text{as } q \to \infty \tag{3.11}$$

One can also find asymptotic formulae for other values of r and s but we do not wish to digress here. The main consequence of 3.7 is the following result:

Lemma 3.6.1.

$$\mathbb{E}\left(\binom{\eta}{k}\right) = \mathbb{E}\left(\binom{\eta^*}{k}\right) \qquad \forall k \geq 0$$

provided $d \geq q^r$

From 1.7.5 it follows that the variables η and η^* will have the same asymptotic behaviour as $q, d \to \infty$, provided $d \geq q^r$ always. This last condition however is unsatisfactory, since the degree is too large compared with the size of the field. (Recall from 2.4.1 that polynomials behave like maps anyway when the degree is large enough). The problem is, however, that for lower d we cannot calculate exactly the moments of η, so we revert to some estimation in the next lemma. From now on in this chapter, we assume that $d \leq q^r$.

Lemma 3.6.2.

$$\mathbb{E}\left(\binom{\eta}{k}\right) = \binom{q^s}{k}\sum_{l=0}^{d-1}(-1)^l\binom{q^r}{l}\left(\frac{k}{q^s}\right)^l \quad + \; R(q,d,k)$$

where

$$|R(q,d,k)| \leq \binom{q^s}{k}\binom{q^r}{d}\left(\frac{k}{q^s}\right)^d$$

Proof. As in equation 3.5, we have

$$\mathbb{E}\left(\binom{\eta}{k}\right) = \frac{1}{\sharp\mathcal{F}}\sum_{\substack{K \subseteq \mathbb{F}_q^s \\ \sharp K = k}} \sharp\{\mathbf{h} \in \mathcal{F} \;:\; \mathbf{h}(\mathbb{F}_q^r) \subseteq \mathbb{F}_q^s \backslash K\}$$

Let K be a fixed k-subset of $\mathbb{F}_q^s$ and write

$$\mathbb{F}_q^r = \{\mathbf{a}_1, \ldots, \mathbf{a}_{q^r}\} \quad \text{and} \quad E_i = \{\mathbf{h} \in \mathcal{F} \;:\; \mathbf{h}(\mathbf{a}) \in K\} \quad (1 \leq i \leq q^r)$$

Then, by the inclusion-exclusion principle,

$$\sharp\{\mathbf{h}\in\mathcal{F}\ :\ \mathbf{h}(\mathbb{F}_q^r)\subseteq\mathbb{F}_q^s\backslash K\}=\sharp\mathcal{F}-\sum_{l=1}^{q^r}(-1)^{l-1}\alpha_l \tag{3.12}$$

where the non-negative integers α_l satisfy

$$\alpha_l=\sum_{\substack{L\subseteq\{1,\dots,m\}\\ \sharp L=l}}\sharp\bigcap_{i\in L}E_i\ =\ \binom{q^r}{l}\left(\frac{k}{q^s}\right)^l\sharp\mathcal{F}\quad\text{for } l\leq d\quad\text{(by 2.3.3)} \tag{3.13}$$

Hence

$$\begin{aligned}\sharp\{\mathbf{h}\in\mathcal{F}\ :\ \mathbf{h}(\mathbb{F}_q^r)\subseteq\mathbb{F}_q^s\backslash K\}\ &=\ \sharp\mathcal{F}-\sum_{l=1}^{d-1}(-1)^{l-1}\alpha_l+\sum_{l=d}^{q^r}(-1)^l\alpha_l\\ &=\ \sharp\mathcal{F}\sum_{l=0}^{d-1}(-1)^l\binom{q^r}{l}\left(\frac{k}{q^s}\right)^l\ +\ R_1(q,d,K)\end{aligned}$$

where

$$|R_1(q,d,K)|\leq\alpha_l\qquad\text{(by note 2.3.4)} \tag{3.14}$$

Applying the Brun-Waring inequality 5.7 gives

$$|R_1(q,d,K)|\leq\binom{q^r}{d}\left(\frac{k}{q^s}\right)^d\sharp\mathcal{F} \tag{3.15}$$

so that, finally, we have the equation

$$\mathbb{E}\left(\binom{\eta}{k}\right)\ =\ \binom{q^s}{k}\sum_{l=0}^{d-1}(-1)^l\binom{q^r}{l}\left(\frac{k}{q^s}\right)^l\ +\ R(q,d,k)$$

and, as required,

$$|R(q,d,k)|\leq\binom{q^s}{k}\binom{q^r}{d}\left(\frac{k}{q^s}\right)^d$$

□

Corollary 3.6.3. *Let $A=\mathbb{F}_q^r$, $B=\mathbb{F}_q^s$ so that $n=q^r$, $N=q^s$. Then, for each $k\geq 0$*

$$\left|\mathbb{E}\left(\binom{\eta}{k}\right)-\mathbb{E}\left(\binom{\eta^*}{k}\right)\right|\leq 2\binom{q^s}{k}\binom{q^r}{d}\left(\frac{k}{q^s}\right)^d$$

Proof. Recall that, in the proof of 3.5.1, we used the inclusion-exclusion principle (and, at the time, it seemed unnecessary). Instead of using the exact value of $\mathbb{E}\left(\binom{\eta^*}{k}\right)$, let us truncate the alternating sum in the proof of 3.5.1 at $d-1$, as we did in the last lemma. This gives

$$\mathbb{E}\left(\binom{\eta^*}{k}\right)=\binom{q^s}{k}\sum_{l=0}^{d-1}(-1)^l\binom{q^r}{l}\left(\frac{k}{q^s}\right)^l\ +\ R^*(q,d,k) \tag{3.16}$$

where $|R^*(q,d,k)| \leq \binom{q^s}{k}\binom{q^r}{d}\left(\frac{k}{q^s}\right)^d$ by the Brun-Waring principle (5.7).

Hence,

$$\mathbb{E}\left(\binom{\eta}{k}\right) - \mathbb{E}\left(\binom{\eta^*}{k}\right) = R(q,d,k) - R^*(q,d,k) \tag{3.17}$$

and result follows from the triangle inequality. □

3.7 The Results for a Polynomial Vector

We are now in a position to state and prove the results for a typical polynomial vector which correspond to 3.5.5, 3.5.6, and 3.5.7.

Theorem 3.7.1. *Let*

$$\mu = q^s(1-q^{-s})^{q^r} \quad \textit{and}$$
$$\sigma^2 = q^{2s}(1-q^{-s})(1-2q^{-s})^{q^r} + q^s(1-q^{-s})^{q^r} - q^{2s}(1-q^{-s})^{2q^r}$$

Then, for a random polynomial vector $\mathbf{f}$ *of type* (r,s), *the variable* η *has the following behaviour as* $q \to \infty$:

1. $\mathbf{s = 2r}$ $\eta - q^r(q^r-1)$ *will converge to a Poisson variable with parameter* $\frac{1}{2}$ *if* $d \to \infty$ *arbitrarily slowly with* q.
2. $\mathbf{r < s < 2r}$ $\frac{\eta-\mu}{\sigma}$ *will converge to a Gaussian distribution if* $d \to \infty$ *arbitrarily slowly with* q.
3. $\mathbf{s = r}$ $\frac{\eta-\mu}{\sigma}$ *will converge to a Gaussian distribution if* $d \to \infty$ *subject to the constraint* $\log q = o(d \log d)$.
4. $\mathbf{s < r}$ η *will converge in probability to zero if* $d \to \infty$ *arbitrarily slowly with* q.

Proof. Let us make the following definition:

$$2\Delta(q,d,k) := \frac{\left|\mathbb{E}\left(\binom{\eta}{k}\right) - \mathbb{E}\left(\binom{\eta^*}{k}\right)\right|}{\mathbb{E}\left(\binom{\eta^*}{k}\right)} \tag{3.18}$$

We shall show that, in each of the above cases, η and η^* have the same limiting behaviour under the stipulated conditions. For this it is sufficient to prove that $\Delta(d,q,k) \to 0$ (from 1.7.5). Note that, by 3.6.3, we have

$$\Delta(q,d,k) \leq \frac{\binom{q^r}{d}\left(\frac{k}{q^s}\right)^d}{\left(1-\frac{k}{q^s}\right)^{q^r}} \tag{3.19}$$

We claim the following:

$$\begin{aligned}\log RHS \;=\; & (r-s)d\log q \;-\; d\log d \;-\; \left\{\left(1-\frac{1}{2}\right)\frac{d^2}{q^r}+\left(\frac{1}{2}-\frac{1}{3}\right)\frac{d^3}{q^{2r}}+\dots\right\} \\ & +\; kq^{r-s} \;+\; k^2q^{r-2s} \;+\; O(\log q) \;+\; O(d) \;+\; O(q^{r-3s}) \quad (3.20)\end{aligned}$$

To see this, we use Stirling's formula to obtain:

$$\log\binom{q^r}{d} \;=\; rd\log q \;-\; d\log d \;+\; (q^r-d)\log\left(1-\frac{d}{q^r}\right) \;+\; O(\log q) \;+\; O(\log d) \qquad (3.21)$$

where

$$\begin{aligned}(q^r-d)\log\left(1-\frac{d}{q^r}\right) &= (q^r-d)\left(-\frac{d}{q^r}-\frac{d^2}{2q^{2r}}-\frac{d^3}{3q^{3r}}-\dots\right) && (3.22)\\ &= -d \;-\; \left\{\left(1-\frac{1}{2}\right)\frac{d^2}{q^r}+\left(\frac{1}{2}-\frac{1}{3}\right)\frac{d^3}{q^{2r}}+\dots\right\} && (3.23)\end{aligned}$$

We also have the following expansions:

$$\log\left(1-\frac{k}{q^s}\right)^{q^r} \;=\; q^r\left\{-\frac{k}{q^s}-\frac{k^2}{q^{2s}}+O\left(\frac{k^3}{q^{3s}}\right)\right\} \qquad (3.24)$$

$$\log\left(\frac{k}{q^s}\right)^d \;=\; d\log k \;-\; sd\log q \qquad (3.25)$$

Combining (3.21), (3.23), (3.24) and (3.25) proves the claim.

We now show that RHS of (3.20) tends to $-\infty$ in each of the cases 1-3. This will ensure that $\Delta(q,d,k)\to 0$ and these cases will be proved. Case 4 is dealt with separately.

Case 1 $\mathbf{s=2r}$ Equation (3.20) becomes

$$\begin{aligned}\log\Delta(q,d,k) \;\le\; & -rd\log q \;-\; d\log d \;-\; \left\{\left(1-\frac{1}{2}\right)\frac{d^2}{q^r}+\left(\frac{1}{2}-\frac{1}{3}\right)\frac{d^3}{q^{2r}}+\dots\right\} \\ & +\; O(q^{-r}) \;+\; O(\log q) \;+\; O(d) \quad (3.26)\end{aligned}$$

Clearly, $RHS\to-\infty$ as $q,d\to\infty$.

Case 2 $\mathbf{r<s<2r}$ Equation (3.20) becomes

$$\begin{aligned}\log\Delta(q,d,k) \;\le\; & -(s-r)d\log q \;-\; d\log d \;-\; \left\{\left(1-\frac{1}{2}\right)\frac{d^2}{q^r}+\left(\frac{1}{2}-\frac{1}{3}\right)\frac{d^3}{q^{2r}}+\dots\right\} \\ & +\; O(q^{-r}) \;+\; O(\log q) \;+\; O(d) \quad (3.27)\end{aligned}$$

Thus, $LHS \to -\infty$ as $q, d \to \infty$.

Case 3 $\mathbf{s} = \mathbf{r}$ Equation (3.20) becomes

$$\log\Delta(q,d,k) \;\le\; -\left[d\log d \;-\; \left\{\left(1-\frac{1}{2}\right)\frac{d^2}{q^r}+\left(\frac{1}{2}-\frac{1}{3}\right)\frac{d^3}{q^{2r}}+\dots\right\}\right] \;+\; k \\ +\; O(q^{-r}) \;+\; O(\log q) \;+\; O(d) \quad (3.28)$$

Hence, if $\log q = o(d\log d)$ we have that $LHS \to -\infty$ as $q, d \to \infty$.

Case 4 $\mathbf{s} < \mathbf{r}$ In this case we show that $\mu, \sigma \to 0$ as $q, d \to \infty$. This guarantees that η converges in probability to zero (see Lemma1.7.4).

Recall that, from Theorem 3.7.1, we have

$$\mu = q^s(1-q^{-s})^{q^r} \quad \text{and} \quad 0 \le \sigma^2 \le \mu \qquad (3.29)$$

Hence, since $r > s$,

$$\begin{aligned}\log\mu &= s\log q - \left(q^{r-s}+\frac{1}{2}q^{r-2s}+\dots\right)\\ &\to -\infty \qquad \text{as } q \to \infty\end{aligned}$$

It follows that $\mu, \sigma^2 \to 0$, and this completes the proof of the theorem. □

Remarks

The fact that the sizes of our sets are powers of a natural number q simplifies the random map scenario, causing anything 'right' of the central domain to be forced to converge to zero in probability. Also, going 'left' of centre means that, because of the difference in size between domain and codomain, the polynomial mimics the map no matter how slowly d grows with q.

The interesting case is the 'central domain', where the domain and codomain of our map/ polynomial actually have the same size. This leads to the added proviso that d be sufficiently large relative to q in order that the polynomial should mirror the map in its behaviour.

Chapter 4

Generalised Inverse-Image Variables

4.1 First Generalisation

Recall that in Chapter 3, the inverse-image problem was more straight forward than the direct-image problem. This was because, in the former, our moment-matching did not depend on q. Indeed, we had in Lemma 3.3.1:

$$\mathbb{E}\left(\binom{\zeta_C}{k}\right) = \mathbb{E}\left(\binom{\zeta_c^*}{k}\right) \qquad \forall k \leq d \tag{4.1}$$

We now go on to explore a whole class of random variables associated with a typical polynomial vector of type (r, s) which have properties analogous to (4.1).

With the same notation as before, let $A = \mathbb{F}_q^r$, $B = \mathbb{F}_q^s$ so that $n = q^r$, $N = q^s$. We first observe that the variable ζ_C of Chapter 3 can be written

$$\zeta_C = \sum_{\mathbf{a} \in A} \delta_C(\mathbf{f}(\mathbf{a})) \tag{4.2}$$

where

$$\delta_C(\mathbf{b}) = \begin{cases} 1, \ \mathbf{b} \in C \\ 0, \ \mathbf{b} \notin C \end{cases} \tag{4.3}$$

A natural way to generalise this is to define the complex-valued variable

$$S_1 = S_1(\mathbf{f}) = \sum_{\mathbf{a} \in A} \rho(\mathbf{f}(\mathbf{a})) \tag{4.4}$$

where $\rho : B \to \mathbb{C}$ satisfies $|\rho(\mathbf{b})| \leq 1 \quad \forall \mathbf{b} \in B$.
The map analogue of this variable would be

$$S_1^* = \sum_{\mathbf{a} \in A} \rho(\phi(\mathbf{a})) \tag{4.5}$$

As always, f and ϕ denote a random polynomial vector and a random map, respectively.

Our original motivation for studying this type of variable was not, as suggested by the above, to generalise ζ_C, but to gain insight into the 'average' behaviour of certain character sums over $\mathbb{F}_q$. Indeed, this will be the main application of these *'generalised inverse-image variables'*, as they will be called. However, throughout, we keep in mind that this variable is a generalisation of ζ_C and the moments behave accordingly:

Proposition 4.1.1. *For $k, l \geq 0$, the moments of the complex-valued variables S_1 and S_1^* satisfy:*

$$\mathbb{E}(S_1^k \overline{S_1}^l) = \mathbb{E}(S_1^{*k} \overline{S_1^*}^l) \qquad \forall k + l \leq d$$

Proof. We postpone the proof until Proposition 4.7.1 where a more general result is proved. □

4.2 A Simple Example

Let $r = s = 1$ so that $A = B = \mathbb{F}_q$ and let ρ be a non-principal additive character of $\mathbb{F}_q$. Then

$$|\rho(\mathbf{b})| = 1 \qquad \forall \mathbf{b} \in B$$

In fact, one can say more. If $\mathbf{b}$ is a random element of B (in the natural way), then $\rho(\mathbf{b})$ is a random variable in $\mathbb{C}$ which is uniformly distributed on some finite subgroup of the circle. (Note that this property will be used later.) Thus we have the trigonometric sum

$$S_1 = \sum_{\mathbf{a} \in \mathbb{F}_q} \rho(\mathbf{f}(\mathbf{a})) \tag{4.6}$$

Using the Riemann hypothesis for function fields over $\mathbb{F}_q$, Weil [41] proved the following estimate for the above sum:

$$|S_1| \leq (d-1)\sqrt{q} \tag{4.7}$$

where $d = \deg f$, as always.

Since there is no good uniform bound for $d >> \sqrt{q}$ (see [35] for examples), one might ask: "what is the asymptotic behaviour, as $q, d \to \infty$, of the real-valued random variable $|S_1(f)|$?" In fact, this problem is solved in [35], where the author uses a random walk to model $|S_1^*|$ and then matches the moments with those of $|S_1|$.

We shall go one further and ask: "what is the asymptotic behaviour, as $q, d \to \infty$, of the

complex-valued random variable $S_1(f)$?"
In fact, we shall ask this question for even more general variables which will turn out to have the property described in (4.1).

4.3 Second Generalisation

We now give an alternative generalisation of the variable ζ_C, where $C \subseteq B$ as in Chapter 3. Let $A_1, \ldots, A_m \subseteq A$ be m disjoint sets of equal size $\frac{\alpha}{m}$. Then

$$\sharp\,(A_1 \sqcup \ldots \sqcup A_m) = \alpha \leq q^r$$

To each A_j we assign a complex number h_j and we also define

$$\zeta_{C,j} = \sharp\left(\mathbf{f}^{-1}(C) \cap A_j\right) \;=\; \sharp\{\mathbf{a} \in A_j \;:\; \mathbf{f}(\mathbf{a}) \in C\}$$

This allows us to define the following complex-valued sum:

$$S_2 = S_2(\mathbf{A}, \mathbf{h}, \mathbf{f}) = \sum_{j=1}^{m} h_j \zeta_{C,j} \tag{4.8}$$

where $\mathbf{A}$, $\mathbf{h}$ denote the collections $\{A_1, \ldots, A_m\}$ and $(h_1, \ldots, h_m)$, respectively.

Note that, putting $\mathbf{A} = \{A\}$ (so that $m = 1$) and $h_1 = 1$ gives $S_2 = \zeta_C$. Therefore S_2 is indeed a generalisation of ζ_C.

Naturally, we define the random map analogue

$$S_2^* = S_2^*(\mathbf{A}, \mathbf{h}, \phi) = \sum_{j=1}^{m} h_j \zeta_{C,j}^* \tag{4.9}$$

where, by now, the definition of $\zeta_{C,j}^*$ should be obvious.

As hoped for, the moments of S_2 also turn out to have nice properties:

Proposition 4.3.1. *For $k, l \geq 0$, the moments of the complex-valued variables S_2 and S_2^* satisfy:*

$$\mathbb{E}(S_2^k \overline{S_2}^l) = \mathbb{E}(S_2^{*k} \overline{S_2^*}^l) \qquad \forall k + l \leq d$$

Proof. Again, we postpone the proof until Proposition 4.7.1, where a more general result is proved. □

4.4 A Simple Example

Let $r = s = 1$ so that $A = B = \mathbb{F}_q$.

Suppose that $m \mid q-1$, $m \neq 1$. If χ is a multiplicative character of order m on $\mathbb{F}_q$ then we have, for each $b_1, b_2 \in \mathbb{F}_q^*$,

$$\chi(b_1) = \chi(b_2) \iff b_2 = x^m b_1 \qquad \text{for some } x \in \mathbb{F}_q^*$$

Hence, if $t \in \mathbb{F}_q^*$ is an element of order m, then χ takes distinct complex values on the m cosets of t in $\mathbb{F}_q^*$. Let $A_1, \dots, A_m$ be these cosets and let $h_1, \dots, h_m \in \mathbb{C}$ be the distinct values of $\chi(b)$.

Finally, let $C = \{0\}$ so that

$$\zeta_{C,j} = \sharp\{\text{Zeros of } \mathbf{f} \text{ in } A_j\} \qquad 1 \leq j \leq m$$

Then we have

$$S_2 = \sum_{f(a)=0} \chi(a) \tag{4.10}$$

So this variable is a character sum over the zeros of a polynomial. Of course, we can also have the more general case of a character sum over the points of a random variety in $\mathbb{F}_q^r$, namely the variety defined by the s polynomials of a random polynomial vector. Such character sums have been studied by Bombieri, Adolphson and Sperber; see for example [7, 1].

4.5 Third Generalisation

Our third generalisation of the inverse-image variable will be a combination of the first and the second generalisations. It will be for this most general variable that we do the moment calculations, and ultimately prove our main convergence theorems (5.3.3 - 5.3.5).

First of all recall the definition of S_2:

$$S_2 = \sum_{j=1}^{m} h_j \zeta_{C,j} \tag{4.11}$$

Recall also, however, that we can write

$$\zeta_{C,j} = \sum_{\mathbf{a} \in A_j} \delta_C(\mathbf{f}(\mathbf{a})) \tag{4.12}$$

where δ_C is as in (4.3).

We replace δ_C with the more generic ρ of (4.4) and then combine the ideas of (4.11) & (4.12) to get the following variable:

$$\Sigma = \Sigma(\mathbf{A}, \mathbf{h}, \mathbf{f}) = \sum_{j=1}^{m} h_j \sum_{\mathbf{a} \in A_j} \rho(\mathbf{f}(\mathbf{a})) \tag{4.13}$$

Putting $\mathbf{A} = A$ (so that $m = 1$) and $h_1 = 1$ gives $\Sigma = S_1$. On the other hand, putting $\rho = \delta_C$ gives $\Sigma = S_2$. Therefore Σ is a generalisation of both S_1 and S_2.

4.6 A Simple Example

Let $r = s = 1$ and let $A_1, \ldots, A_m$ & $h_1, \ldots, h_m$ be as in §4.4. Further, let ρ be an additive character of $\mathbb{F}_q$. Then we have

$$S_3 = \sum_{a \in \mathbb{F}_q} \chi(a)\rho(f(a)) \tag{4.14}$$

where χ is a multiplicative character on $\mathbb{F}_q$, so that S_3 is a type of Gauss sum (see [20]).

4.7 Moment Calculations

We have now set the scene for a collection of random variables associated with the probability space $\mathcal{F}$ which have applications to, among other things, several different types of character sum. In order to be able to determine the asymptotic distributions of these variables it is first necessary to match their moment sequences to those of the corresponding random map variables. Of course, we stated such results in Lemmas 4.1.1 & 4.3.1, and now we go on to prove the most general case, namely that of Σ.

Proposition 4.7.1. *For $k, l \geq 0$, the moments of the complex-valued variables Σ and Σ^* satisfy:*

$$\mathbb{E}(\Sigma^k \overline{\Sigma}^l) = \mathbb{E}(\Sigma^{*k} \overline{\Sigma^*}^l) \qquad \forall k + l \leq d$$

Before we can prove the proposition, we need the following preliminary lemma.

Lemma 4.7.2. *Let $\mathbf{a} \in A$ and let $i, j \in \mathbb{Z}_{\geq 0}$. Then*

$$\mathbb{E}\left(\rho(\mathbf{f}(\mathbf{a}))^i \overline{\rho(\mathbf{f}(\mathbf{a}))}^j\right) = \mathbb{E}\left(\rho(\phi(\mathbf{a}))^i \overline{\rho(\phi(\mathbf{a}))}^j\right) \qquad \textit{provided } d \geq 1$$

Proof.

$$\begin{aligned}
L.H.S. &= \frac{1}{\sharp\mathcal{F}}\sum_{\mathbf{h}\in\mathcal{F}}\rho(\mathbf{h}(\mathbf{a}))^i\overline{\rho(\mathbf{ha}))}^j \\
&= \frac{1}{\sharp\mathcal{F}}\sum_{\mathbf{b}\in B}\rho(\mathbf{b})^i\overline{\rho(\mathbf{b}))}^j\ \sharp\{\mathbf{h}\in\mathcal{F}:\mathbf{h}(\mathbf{a})=\mathbf{b}\} \\
&= \frac{1}{\sharp\mathcal{F}}\sum_{\mathbf{b}\in B}\rho(\mathbf{b})^i\overline{\rho(\mathbf{b}))}^j\ q^{-s}\sharp\mathcal{F} \qquad \text{by Cor. 2.3.1} \\
&= q^{-s}\sum_{\mathbf{b}\in B}\rho(\mathbf{b})^i\overline{\rho(\mathbf{b}))}^j
\end{aligned}$$

On the other hand, we have

$$\begin{aligned}
R.H.S. &= \frac{1}{\sharp B^A}\sum_{\phi\in B^A}\rho(\phi(\mathbf{a}))^i\overline{\rho(\phi(\mathbf{a}))}^j \\
&= \frac{1}{N^n}\sum_{\mathbf{b}\in B}\rho(\mathbf{b})^i\overline{\rho(\mathbf{b}))}^j\ \sharp\{\phi\in B^A:\phi(\mathbf{a})=\mathbf{b}\} \\
&= \frac{1}{N^n}\sum_{\mathbf{b}\in B}\rho(\mathbf{b})^i\overline{\rho(\mathbf{b}))}^j\ N^{n-1} \\
&= N^{-1}\sum_{\mathbf{b}\in B}\rho(\mathbf{b})^i\overline{\rho(\mathbf{b}))}^j
\end{aligned}$$

Since $B=\mathbb{F}_q^s$ in our case, we have $N=q^s$ and the result follows. □

We are now in a position to prove Proposition 4.7.1, although we refer the reader to Appendix B for some technical points which occur in the proof.

Proof. Let us rewrite the expression for Σ in a slightly different way:

$$\Sigma=\sum_{\mathbf{a}\in A'}h_{\mathbf{a}}\,\rho(\mathbf{f}(\mathbf{a})) \tag{4.15}$$

where A' denotes $A_1\sqcup\ldots\sqcup A_m$.
$H_{\mathbf{a}}$ is defined to be h_j (of (4.13)), where j is the unique integer such that $\mathbf{a}\in A_j$.

We can therefore write

$$\Sigma^k\overline{\Sigma}^l=\sum_{t=1}^{k+l}\ \sum_{\substack{T\subseteq A'\\ T=\{\mathbf{a}_1,\ldots,\mathbf{a}_t\}}}\ \sum_{\mathbf{i},\mathbf{j}}\binom{k}{\mathbf{i}}\binom{l}{\mathbf{j}}\prod_{w=1}^{t}h_{\mathbf{a}_w}^{i_w}\overline{h_{\mathbf{a}_w}}^{j_w}\ \rho(\mathbf{f}(\mathbf{a}_w))^{i_w}\overline{\rho(\mathbf{f}(\mathbf{a}_w))}^{j_w} \tag{4.16}$$

The innermost sum is over all multi-indices $\mathbf{i}=(i_1,\ldots,i_t)$ and $\mathbf{j}=(j_1,\ldots,j_t)$ satisfying $||\mathbf{i}||=k$, $||\mathbf{j}||=l$ and $(i_w,j_w)\neq\mathbf{0}$ $(\forall w)$. See Appendix B for a full explanation of this.
Similarly, we have

$$\Sigma^{*k}\overline{\Sigma}^{*l}=\sum_{t=1}^{k+l}\ \sum_{\substack{T\subseteq A'\\ T=\{\mathbf{a}_1,\ldots,\mathbf{a}_t\}}}\ \sum_{\mathbf{i},\mathbf{j}}\binom{k}{\mathbf{i}}\binom{l}{\mathbf{j}}\prod_{w=1}^{t}h_{\mathbf{a}_w}^{i_w}\overline{h_{\mathbf{a}_w}}^{j_w}\ \rho(\phi(\mathbf{a}_w))^{i_w}\overline{\rho(\phi(\mathbf{a}_w))}^{j_w} \tag{4.17}$$

Applying the expectation operator to (4.16) & (4.17), and tidying up the constants, we have

$$\mathbb{E}\left(\Sigma^k\overline{\Sigma}^l\right) = \sum_{t=1}^{k+l} \sum_{\substack{T\subseteq A' \\ T=\{\mathbf{a}_1,\dots,\mathbf{a}_t\}}} \sum_{\mathbf{i},\mathbf{j}} H(t,T,\mathbf{i},\mathbf{j}) \,.\, \mathbb{E}\left(\prod_{w=1}^{t} \rho(\mathbf{f}(\mathbf{a}_w))^{i_w}\overline{\rho(\mathbf{f}(\mathbf{a}_w))}^{j_w}\right) \tag{4.18}$$

$$\mathbb{E}\left(\Sigma^{*k}\overline{\Sigma}^{*l}\right) = \sum_{t=1}^{k+l} \sum_{\substack{T\subseteq A' \\ T=\{\mathbf{a}_1,\dots,\mathbf{a}_t\}}} \sum_{\mathbf{i},\mathbf{j}} H(t,T,\mathbf{i},\mathbf{j}) \,.\, \mathbb{E}\left(\prod_{w=1}^{t} \rho(\phi(\mathbf{a}_w))^{i_w}\overline{\rho(\phi(\mathbf{a}_w))}^{j_w}\right) \tag{4.19}$$

where

$$H(t,T,\mathbf{i},\mathbf{j}) = \binom{k}{\mathbf{i}}\binom{l}{\mathbf{j}} \prod_{w=1}^{t} h_{\mathbf{a}_w}^{i_w}\overline{h_{\mathbf{a}_w}}^{j_w} \tag{4.20}$$

By Proposition 2.5.1, the variables $\rho(\phi(\mathbf{a}_w))^{i_w}\overline{\rho(\phi(\mathbf{a}_w))}^{j_w}$, where $1 \leq w \leq t$, are independent and so the expectation operator can be taken inside the product sign in (4.19).

By the same result, the variables $\rho(\mathbf{f}(\mathbf{a}_w))^{i_w}\overline{\rho(\mathbf{f}(\mathbf{a}_w))}^{j_w}$ $(1 \leq w \leq t)$ will be independent *provided that* $d \geq t$. Thus, for $k + l \leq d$, we can take the expectation operator inside the product sign in (4.18) also.

In view of Lemma 4.7.2, we obtain the required result. □

Note 4.7.3. This also proves Propositions 4.1.1 and 4.3.1 also, since S_1 and S_2 are both special cases of Σ.

4.8 Summary

We have defined the very general variable

$$\Sigma = \Sigma(\mathbf{A},\mathbf{h},\mathbf{f}) = \sum_{j=1}^{m} h_j \sum_{\mathbf{a}\in A_j} \rho(\mathbf{f}(\mathbf{a}))$$

This is a complex-valued random variable associated with a random polynomial vector of type (r, s). It has applications to, most notably, character sums over $\mathbb{F}_q$, and it also generalises the variable ζ_C of chapter 2.

Σ has a natural random map analogue Σ^* - a complex-valued random variable associated with a random map of type (q^r, q^s).

The moments of Σ and Σ^* match up to an order which depends only on the maximum degree d that the random polynomial vector $\mathbf{f}$ can have. (Recall that $d := \min\{d_1, \dots, d_s\}$.)

As a result, if we know the asymptotic behaviour of any particular specialisation of Σ^*, we can deduce that Σ has the same asymptotic behaviour, provided d tends to infinity *arbitrarily slowly* with q.

Chapter 5

Random Character Sums

5.1 Characters on $\mathbb{F}_q$

We can now turn our attention to character sums over finite fields which involve one or more random polynomials. We saw in chapter 4 that, by construction, the generalised inverse-image variable Σ has applications to such character sums. If we can determine the asymptotic behaviour of the map analogue Σ^*, then the moment matching of §4.7 will infer that Σ has (in general) the same behaviour.

First, however, we must endow the function ρ (defined in §4.1) with certain specific properties which identify a product of complex characters on $\mathbb{F}_q$.

Let us define:

$$G_k = \{z \in \mathbb{C} : z^k = 1\} \qquad (k \in \mathbb{N})$$

$$G_\infty = \{z \in \mathbb{C} : |z| = 1\}$$

These are precisely the compact subgroups of $\mathbb{C}^*$, the former being finite for all $k \in \mathbb{N}$.

As an example of what we require of ρ, let us consider the 'mixed character' $\chi : B \to \mathbb{C}$ on $B = \mathbb{F}_q^s$ given by

$$\chi(\mathbf{b}) = \chi_1(b_1)\dots\chi_u(b_u)\psi_1(b_{u+1})\dots\psi_v(b_{u+v}) \qquad (u, v \in \mathbb{Z}_{\geq 0},\ u+v=s) \tag{5.1}$$

where χ_i is a multiplicative character on $\mathbb{F}_q$ $(1 \leq i \leq u)$
and ψ_i is an additive character on $\mathbb{F}_q$ $(1 \leq i \leq v)$.
Then the image of χ will be a finite group of roots of unity, possibly with zero adjoined (if $u \geq 1$). That is,

$$\chi(B) = G_t \quad \text{or} \quad \chi(B) = G_t \cup \{0\} \quad (\text{some } t \in \mathbb{N})$$

Hence, if we restrict χ to the domain $C = (\mathbb{F}_q^*)^u \times \mathbb{F}_q^{\,v}$ we have

$$\chi(B) = G_t \quad (\text{some } t \in \mathbb{N})$$

Further, if $\mathbf{b}$ is a random element of C (in the natural way), then $\chi(\mathbf{b})$ is uniformly distributed on G_t, so that, in particular, t divides $c = \sharp C$. This leads us to make the following definition:

Def 5.1.1. We will say that the function $\rho : B \to \mathbb{C}$ is of *arithmetic type* if

$$\rho(\mathbf{b}) = \delta_C(\mathbf{b})\hat{\rho}(\mathbf{b}) \quad \forall \mathbf{b} \in B \qquad (\text{some } C \subseteq B)$$

where δ_C is as in (4.3) and $\hat{\rho} : B \to \mathbb{C}$ is such that $\hat{\rho}(C) = G_t$ (some $t \in \mathbb{N}$) and that $\hat{\rho}(\mathbf{b})$ is *uniformly distributed* in G_t ($\mathbf{b}$ a random element of C).

5.2 Restrictions on ρ as $q \to \infty$

A function which is of arithmetic type exhibits the properties of a typical product of characters on $\mathbb{F}_q$. Next, consider the sum Σ as given in (4.15):

$$\Sigma = \sum_{\mathbf{a} \in A'} h_{\mathbf{a}}\, \rho(\mathbf{f}(\mathbf{a}))$$

where $A' = A_1 \sqcup \ldots \sqcup A_m$. We assume that $\frac{\alpha}{n} \to \gamma \leq 1$ as $n, N \to \infty$.

In our convergence theorems of this chapter, we will, as always, be letting $q \to \infty$ through any sequence of prime powers. So our sets A, B will be growing in size and our function ρ will be changing constantly. We must therefore first specify the behaviour of ρ as $q \to \infty$. Let us write $\hat{\rho}(C) = G_{t(q)}$. We shall consider two cases:

1. $t(q) \to \tau < \infty$ as $q \to \infty$.
   This is the same as saying that, for all q sufficiently large, $\hat{\rho}(C) = G_\tau$ for some fixed $\tau \in \mathbb{N}$.

2. $t(q) \to \infty$ as $q \to \infty$.
   This situation is slightly more complicated.

Case 1 is straight forward; we have that $\hat{\rho}(\mathbf{b})$ is eventually uniformly distributed on the finite group G_τ. For case 2, however, we need to know a result about the convergence of uniform distributions on the circle G_∞.

Theorem 5.2.1. *Let $\{z_k\}_{k \in \mathbb{N}}$ be a sequence of discrete complex-valued random variables such that, for each $k \in \mathbb{N}$, z_k is uniformly distributed on $G_{t(k)}$. If $t(k) \to \infty$ as $k \to \infty$, then z_k converges in distribution to a uniform distribution on G_∞.*

Proof. See Appendix C □

5.3 The Asymptotic Behaviour of Σ^* (and hence of Σ)

We wish to know the asymptotic behaviour of Σ^* $(q, d \to \infty)$ which, as always, will depend on how certain parameters vary as q tends to infinity. We shall see that the behaviour splits into five cases. Finding this limiting behaviour in each case amounts to calculating the limit of the characteristic function of a normalised version of Σ^*, and then using the Continuity Theorem (1.7.8). *i.e.* we work with the characteristic function of a variable of the form

$$\frac{\Sigma^* - \mu}{\sigma}$$

where μ, σ^2 are (resp.) the asymptotic mean and variance of Σ^* in that particular case.

This is the standard technique used to prove the Central Limit Theorem and in fact, our theorems will be very much of central-limit-type. Before we dive straight into the theorems we must first make several definitions in preparation for our analysis of Σ^*.

Def 5.3.1. Let us define

$$C(\mathbf{t}) = \mathbb{E}(\exp\ i\mathbf{t}.\Sigma^*) \tag{5.2}$$

where

$$\mathbf{t} = (t_1, t_2) \quad \text{and} \quad \mathbf{t}.\Sigma = t_1\Re(\Sigma^*) + t_2\Im(\Sigma^*)$$

$C(\mathbf{t})$ will be used to determine the behaviour of Σ^* in the first two cases, Theorems 5.3.3 and 5.3.4. In the latter, the following variable, P, will occur:

Let $Z_1, \dots, Z_m$ be random variables, uniformly distributed on the sets $h_1G_T, \dots, h_mG_T$, respectively. We define P_j to be the compound Poisson variable with parameters $\frac{\gamma\lambda}{m}$ and Z_j $(1 \leq j \leq m)$ (see §1.6). Finally, let P be the sum of the (independent) variables P_j, that is

$$P = \sum_{j=1}^{m} P_j$$

In the final three cases, we shall be looking for convergence to Gaussian distributions. These will require use of the following characteristic functions:

1.

$$C_1(\mathbf{t}) = \mathbb{E}\left(\exp i\mathbf{t}.\left(\frac{\Sigma^* - \mu_1}{\sigma_1}\right)\right) \tag{5.3}$$

where

$$\mu_1 = \frac{cn}{mN}\gamma\sum_{j=1}^{m} h_j \quad \text{and} \quad \sigma_1^2 = \frac{cn}{mN}\gamma\left(1 - \frac{c}{N}\right)$$

2.

$$C_2(\mathbf{t}) = \mathbb{E}\left(\exp it.\left(\frac{\Sigma^*}{\sigma_2}\right)\right) \tag{5.4}$$

where

$$\sigma_2^2 = \frac{cn}{mN}\gamma$$

3.

$$C_3(\mathbf{t}) = \mathbb{E}\left(\exp it.\left(\frac{\Sigma^*}{\sigma_3}\right)\right) \tag{5.5}$$

where

$$\sigma_3^2 = \frac{cn}{2mN}\gamma\sum_{j=1}^{m}|h_j|^2$$

Finally, we define the real 2 × 2 matrix **H** by

$$\mathbf{H} = \begin{bmatrix} \sum_{j=1}^m \Re(h_j)^2 & \sum_{j=1}^m \Re(h_j)\Im(h_j) \\ \sum_{j=1}^m \Re(h_j)\Im(h_j) & \sum_{j=1}^m \Im(h_j)^2 \end{bmatrix} \tag{5.6}$$

The following lemma tells us what $C(\mathbf{t})$ looks like:

Lemma 5.3.2.

$$\log C(\mathbf{t}) = \frac{1}{m}\sum_{j=1}^{m}\left\{\frac{\alpha}{n}.\frac{cn}{N}E_j(\mathbf{t}) - \frac{\alpha}{2n^2}\left(\frac{cn}{N}\right)^2 E_j(\mathbf{t})^2 + \dots\right\} \tag{5.7}$$

where

$$E_j(\mathbf{t}) = \mathbb{E}(\exp\left[i\mathbf{t}.h_j\rho(b)\right]) - 1 \tag{5.8}$$

for a random variable $b \in C$, uniformly distributed on C.

Proof.

$$\begin{aligned}
\mathbb{E}(\exp\ i\mathbf{t}.\Sigma^*) &= \mathbb{E}\left(\exp\left[i\mathbf{t}.\sum_{j=1}^{m} h_j \sum_{a\in A_j}\rho(\phi(a))\right]\right) \\
&= \mathbb{E}\left(\prod_{j=1}^{m}\prod_{a\in A_j}\exp\left[i\mathbf{t}.h_j\rho(\phi(a))\right]\right) \\
&= \prod_{j=1}^{m}\prod_{a\in A_j}\mathbb{E}(\exp\left[i\mathbf{t}.h_j\rho(\phi(a))\right]) \quad \text{(by Prop. 2.5.1)} \\
&= \prod_{j=1}^{m}\prod_{a\in A_j}\mathbb{E}\left(\exp\left[i\mathbf{t}.h_j\rho(b')\right]\right) \quad \text{(b' uniformly distributed on B)} \\
&= \prod_{j=1}^{m}\mathbb{E}\left(\exp\left[i\mathbf{t}.h_j\rho(b')\right]\right)^{\frac{\alpha}{m}}
\end{aligned}$$

But

$$\begin{aligned}
\mathbb{E}\left(\exp\left[it.h_j\rho(b')\right]\right) &= \frac{1}{N}\left(\sum_{x\in C}\exp\left[it.h_j\rho(x)\right]+\sum_{x\notin C}1\right)\\
&= \frac{1}{N}\sum_{x\in C}\exp\left[it.h_j\rho(x)\right]+\frac{N-c}{N}\\
&= \frac{c}{N}.\frac{1}{c}\sum_{x\in C}\exp\left[it.h_j\rho(x)\right] \quad +1-\frac{c}{N}\\
&= \frac{c}{N}\left(\mathbb{E}(\exp\left[it.h_j\rho(b)\right)\right]-1) \quad +1\\
&= \frac{c}{N}E_j(\mathbf{t})+1
\end{aligned}$$

Thus

$$C(\mathbf{t})=\prod_{j=1}^{m}\left(\frac{c}{N}E_j(\mathbf{t})+1\right)^{\frac{\alpha}{m}}$$

and so

$$\begin{aligned}
\log C(\mathbf{t}) &= \sum_{j=1}^{m}\frac{\alpha}{m}\log\left(1+\frac{c}{N}E_j(\mathbf{t})\right) && (5.9)\\
&= \sum_{j=1}^{m}\frac{\alpha}{m}\left\{\frac{c}{N}E_j(\mathbf{t})-\frac{c^2}{2N^2}E_j(\mathbf{t})^2+\ldots\right\} && (5.10)
\end{aligned}$$

and the result follows. □

We can now state and prove the five main results which govern the asymptotic behaviour of Σ^*. The proofs all rely on the characteristic functions defined above and the Continuity Theorem (Thm. 1.7.8). Because of our moment-matching of §4.7, the results all hold for Σ, provided $d \to \infty$ arbitrarily slowly with q.

Theorem 5.3.3. *Suppose that $c\frac{n}{N} = cq^{r-s} \to 0$. Then Σ^* (and hence Σ) converges to zero in probability.*

Proof. Under these conditions, $\log C(\mathbf{t}) \to 0$. That is, $C(\mathbf{t}) \to 1$ and the result follows from the Continuity Theorem. □

Theorem 5.3.4. *Suppose that $c\frac{n}{N} = cq^{r-s} \to \lambda > 0$. Then Σ^* (and hence Σ) converges in distribution to the variable P, defined in Def. 5.3.1.*

Proof. Under the above conditions, we have

$$\begin{aligned}
\log C(\mathbf{t}) &\to \sum_{j=1}^{m}\frac{\gamma\lambda}{m}E_j(\mathbf{t})\\
\therefore\ C(\mathbf{t}) &\to \prod_{j=1}^{m}\exp\left[\frac{\gamma\lambda}{m}\mathbb{E}\left(e^{i\mathbf{t}.h_j\rho(\mathbf{b})}\right)-1\right]
\end{aligned}$$

By Proposition 1.5.2, the RHS is the characteristic function of a sum of independent random variables having characteristic functions

$$\exp\left[\frac{\gamma\lambda}{m}(\mathbb{E}\left(e^{i\mathbf{t}.h_j\rho(\mathbf{b})}\right)-1\right] \qquad (1 \le j \le m)$$

By Appendix F, these variables are the P_j, and the result follows. □

Theorem 5.3.5. *Suppose that* $c\frac{n}{N} = cq^{\tau-s} \to \infty$ *and* $\tau = 1$. *Then* $\frac{\Sigma^*-\mu_1}{\sigma_1}$ *converges in distribution to a 2-dimensional Gaussian distribution with mean* 0 *and covariance matrix* $\mathbf{H}$.

Proof. Let $\mathbf{t}' = \frac{\mathbf{t}}{\sigma_1}$ and define $\mathbf{b}', \mathbf{b}$ to be random elements of B, C (respectively) in the natural way. Then we have

$$\begin{aligned} C_1(\mathbf{t}) &= \mathbb{E}\left(\exp i\mathbf{t}.\left(\frac{\Sigma^*-\mu_1}{\sigma_1}\right)\right) \\ &= \mathbb{E}\left(\exp i\mathbf{t}'.(\Sigma^*-\mu_1)\right) \\ &= \mathbb{E}\left(\exp i\mathbf{t}'.\left(\sum_{j=1}^{m}\sum_{\mathbf{a}\in A_j}\left[\rho(\phi(\mathbf{a}))-\frac{c}{N}\right]\right)\right) \\ &= \prod_{j=1}^{m}\mathbb{E}\left(\exp i\mathbf{t}'.\left[h_j\left(\rho(\mathbf{b}')-\frac{c}{n}\right)\right]\right)^{\frac{\alpha}{m}} \qquad \text{by Prop. 2.5.1} \end{aligned}$$

Now, for $1 \le j \le m$,

$$\begin{aligned} \mathbb{E}\left(\exp i\mathbf{t}'.\left[h_j\left(\rho(\mathbf{b}')-\frac{c}{N}\right)\right]\right) &= \frac{1}{N}\sum_{x\in C}\exp i\mathbf{t}'.\left[h_j\left(\rho(x)-\frac{c}{N}\right)\right]+\frac{1}{N}\sum_{x\notin C}\exp i\mathbf{t}'.\left[h_j\left(\rho(x)-\frac{c}{N}\right)\right] \\ &= \frac{c}{N}.\frac{1}{c}\sum_{x\in C}\exp i\mathbf{t}'.\left[h_j\left(\hat{\rho}(x)-\frac{c}{N}\right)\right]+\left(1-\frac{c}{N}\right)\exp\left[-i\mathbf{t}'.\frac{c}{N}h_j\right] \\ &= \frac{c}{N}\mathbb{E}\left(\exp i\mathbf{t}'.\left[h_j\left(\hat{\rho}(\mathbf{b})-\frac{c}{N}\right)\right]\right)+\left(1-\frac{c}{N}\right)\exp\left[-i\mathbf{t}'.\frac{c}{N}h_j\right] \\ &= \frac{c}{N}\mathbb{E}\left(1+\mathbf{t}'.h_j\left(\hat{\rho}(\mathbf{b})-\frac{c}{N}\right)-\frac{1}{2}\left(\mathbf{t}'.h_j\left(\hat{\rho}(\mathbf{b})-\frac{c}{N}\right)\right)^2\right) \\ &\quad+\left(1-\frac{c}{N}\right)\left(1-i\mathbf{t}'.\frac{c}{N}h_j-\frac{1}{2}\left[\mathbf{t}'.\frac{c}{N}h_j\right]^2+O(|\mathbf{t}'|^3)\right) \end{aligned}$$

But $\tau = 1$, which means that for n sufficiently large, $\hat{\rho}(\mathbf{b})$ is identically 1. This leads to

$$\begin{aligned} \mathbb{E}\left(\exp i\mathbf{t}'.\left[h_j\left(\rho(\mathbf{b}')-\frac{c}{N}\right)\right]\right) &= 1-\frac{c}{2N}\left(1-\frac{c}{N}\right)(\mathbf{t}'.h_j)^2+O(|\mathbf{t}'|^3) \\ &= 1-\frac{1}{2\left(\frac{\alpha}{m}\right)}(\mathbf{t}.h_j)^2+o\left(\frac{|\mathbf{t}|^2}{\alpha}\right) \qquad (n \text{ suff. large}) \end{aligned}$$

Since $\alpha \sim n\gamma \to \infty$, as $n \to \infty$, we get

$$\begin{aligned} C_1(\mathbf{t}) &\sim \prod_{j=1}^{m} \left(1 - \frac{1}{2\left(\frac{\alpha}{m}\right)}(\mathbf{t}.h_j)^2\right)^{\frac{\alpha}{m}} \\ &\to \prod_{j=1}^{m} \exp\left(-\frac{1}{2}(\mathbf{t}.h_j)^2\right) \\ &= \exp\left(-\frac{1}{2}\sum_{j=1}^{m}(\mathbf{t}.h_j)^2\right) \end{aligned}$$

By inspection, the matrix of the quadratic form

$$\sum_{j=1}^{m}(\mathbf{t}.h_j)^2$$

is $\mathbf{H}$ and the result follows from Appendix F and the Continuity Theorem.

□

Theorem 5.3.6. *Suppose that* $c\frac{n}{N} = cq^{\tau-s} \to \infty$ *and* $\tau = 2$. *Then* $\frac{\Sigma^*}{\sigma_2}$ *converges in distribution to a 2-dimensional Gaussian distribution with mean* 0 *and covariance matrix* $\mathbf{H}$.

Proof. Let $\mathbf{t}' = \frac{\mathbf{t}}{\sigma_2}$ and let $\mathbf{b}', \mathbf{b}$ be as before. Then we have

$$\begin{aligned} C_2(\mathbf{t}) &= \mathbb{E}\left(\exp i\mathbf{t}.\left(\frac{\Sigma^*}{\sigma_2}\right)\right) \\ &= \mathbb{E}\left(\exp i\mathbf{t}'.\Sigma^*\right) \\ &= \mathbb{E}\left(\exp i\mathbf{t}'.\left(\sum_{j=1}^{m}\sum_{\mathbf{a}\in A_j}\rho(\phi(\mathbf{a}))\right)\right) \\ &= \prod_{j=1}^{m}\mathbb{E}\left(\exp i\mathbf{t}'.\left[h_j\rho(\mathbf{b}')\right]\right)^{\frac{\alpha}{m}} \end{aligned}$$

Now, for $1 \le j \le m$,

$$\begin{aligned} \mathbb{E}\left(\exp i\mathbf{t}'.\left[h_j\rho(\mathbf{b}')\right]\right) &= \frac{1}{N}\sum_{x\in C}\exp i\mathbf{t}'.[h_j\rho(x)] + \frac{1}{N}\sum_{x\notin C}\exp i\mathbf{t}'.[h_j\rho(x)] \\ &= \frac{c}{N}.\frac{1}{c}\sum_{x\in C}\exp i\mathbf{t}'.[h_j\hat{\rho}(x)] + \left(1 - \frac{c}{N}\right) \\ &= \frac{c}{N}.\mathbb{E}\left(\exp i\mathbf{t}'.[h_j\hat{\rho}(\mathbf{b})]\right) + \left(1 - \frac{c}{N}\right) \\ &= \frac{c}{N}\mathbb{E}\left(1 + \mathbf{t}'.h_j\hat{\rho}(\mathbf{b}) - \frac{1}{2}\left(\mathbf{t}'.h_j\hat{\rho}(\mathbf{b})\right)^2 + O\left(|\mathbf{t}'|^3\right)\right) + \left(1 - \frac{c}{N}\right) \end{aligned}$$

But $\tau = 2$, which means that for n sufficiently large, $\hat{\rho}(\mathbf{b})$ is uniformly distributed on $\{-1, 1\}$. In particular,

$$\mathbb{E}\left(\mathbf{t}'.h_j\hat{\rho}(\mathbf{b})\right) = 0 \quad \textit{and} \quad \mathbb{E}\left(\left(\mathbf{t}'.h_j\hat{\rho}(\mathbf{b})\right)^2\right) = 2(\mathbf{t}.h_j)^2$$

We deduce that

$$\begin{aligned}\mathbb{E}\left(\exp i\mathbf{t}'.\left[h_j\rho(\mathbf{b}')\right]\right) &= 1 - \frac{c}{2N}\left(1 - \frac{c}{N}\right)(\mathbf{t}'.h_j)^2 + o(|\mathbf{t}'|^2) \\ &= 1 - \frac{1}{2\left(\frac{\alpha}{m}\right)}(\mathbf{t}.h_j)^2 + o\left(\frac{|\mathbf{t}|^2}{\alpha}\right) \qquad (n \text{ suff. large})\end{aligned}$$

Since $\alpha \sim n\gamma \to \infty$, as $n \to \infty$, we get

$$\begin{aligned}C_2(\mathbf{t}) &\to \prod_{j=1}^{m}\left(1 - \frac{1}{2\left(\frac{\alpha}{m}\right)}(\mathbf{t}.h_j)^2\right)^{\frac{\alpha}{m}} \\ &\to \prod_{j=1}^{m}\exp\left(-\frac{1}{2}(\mathbf{t}.h_j)^2\right) \\ &= \exp\left(-\frac{1}{2}\sum_{j=1}^{m}(\mathbf{t}.h_j)^2\right)\end{aligned}$$

and the result follows from Appendix F and the Continuity Theorem.

□

Theorem 5.3.7. *Suppose that* $c\frac{n}{N} = cq^{r-s} \to \infty$ *and* $\tau \geq 3$. *Then* $\frac{\Sigma^*}{\sigma_3}$ *converges in distribution to a 2-dimensional Gaussian distribution with mean* 0 *and covariance matrix* $\mathbf{I}$.

Proof. Let $\mathbf{t}' = \frac{\mathbf{t}}{\sigma_2}$ and let $\mathbf{b}', \mathbf{b}$ be as before. A similar argument to that in the previous proof produces

$$C_3(\mathbf{t}) = \mathbb{E}\left(\exp i\mathbf{t}.\left(\frac{\Sigma^*}{\sigma_3}\right)\right) \tag{5.11}$$

$$= \prod_{j=1}^{m}\mathbb{E}\left(\exp i\mathbf{t}'.\left[h_j\rho(\mathbf{b}')\right]\right)^{\frac{\alpha}{m}} \tag{5.12}$$

Again, mimicking the same previous argument gives

$$\mathbb{E}\left(\exp i\mathbf{t}'.h_j\rho(\mathbf{b}')\right) = \frac{c}{N}\mathbb{E}\left(1 + \mathbf{t}'.h_j\hat{\rho}(\mathbf{b}) - \frac{1}{2}\left(\mathbf{t}'.h_j\hat{\rho}(\mathbf{b})\right)^2 + O\left(|\mathbf{t}'|^3\right)\right) + \left(1 - \frac{c}{N}\right) \tag{5.13}$$

$$= 1 - \frac{c}{2N}\mathbb{E}\left(\left(\mathbf{t}'.h_j\hat{\rho}(\mathbf{b})\right)^2\right) + o\left(\frac{|\mathbf{t}|^2}{\alpha}\right) \qquad (n \text{ suff. large}) \tag{5.14}$$

Since $\tau \geq 3$, we have that for n sufficiently large,

$$\mathbb{E}\left(\mathbf{t}'.h_j\hat{\rho}(\mathbf{b})\right) = 0$$

and we claim that

$$\sum_{j=1}^{m} \mathbb{E}\left((\mathbf{t}'.h_j\hat{\rho}(\mathbf{b}))^2\right) = \frac{1}{2}\left(t_1'^2 + t_2'^2\right) \sum_{j=1}^{m} |h_j|^2 \qquad (n \text{ suff. large}) \tag{5.15}$$

For the moment, let us assume (5.15). Then 5.12 and 5.14 yield

$$\begin{aligned} C_3(\mathbf{t}) &\sim \prod_{j=1}^{m}\left(1 - \frac{c}{2N}\mathbb{E}\left(\left[\mathbf{t}'.h_j\hat{\rho}(\mathbf{b})\right]^2\right)\right)^{\frac{\alpha}{m}} \\ &\to \prod_{j=1}^{m}\exp\left(-\frac{\alpha c}{2mN}\mathbb{E}\left(\left[\mathbf{t}'.h_j\hat{\rho}(\mathbf{b})\right]^2\right)\right) \\ &\to \exp\left(-\frac{c\alpha}{2mN}\sum_{j=1}^{m}\mathbb{E}\left(\left[\mathbf{t}'.h_j\hat{\rho}(\mathbf{b})\right]^2\right)\right) \\ &= \exp\left(-\frac{c\alpha}{2mN}.\frac{1}{2}\left(t_1'^2 + t_2'^2\right)\sum_{j=1}^{m}|h_j|^2\right) \qquad \text{(by 5.15)} \\ &= \exp\left(-\frac{1}{2}\sigma_3^2\left(t_1'^2 + t_2'^2\right)\right) \\ &= \exp\left(-\frac{1}{2}\left(t_1^2 + t_2^2\right)\right) \end{aligned}$$

and the result follows from Appendix F and the Continuity Theorem.

It remains to prove 5.15. This will be an elementary calculation. For $1 \leq j \leq m$, let z_j denote the random variable $h_j\hat{\rho}(\mathbf{b})$. By definition,

$$\begin{aligned} \mathbb{E}\left((\mathbf{t}.h_j\hat{\rho}(\mathbf{b}))^2\right) &= \mathbb{E}([t_1\Re(z_j) + t_2\Im(z_j)])^2 & (5.16) \\ &= t_1^2\mathbb{E}\left(\Re(z_j)^2\right) + 2t_1t_2\mathbb{E}(\Re(z_j)\Im(z_j)) + t_2^2\mathbb{E}\left(\Im(z_j)^2\right) & (5.17) \end{aligned}$$

where $k = \tau$ if $\tau < \infty$ and $k \to \infty$ if $\tau = \infty$. Now,

$$\Re(z_j)^2 = \left[\frac{1}{2}(z_j + \overline{z_j})\right]^2 = \frac{1}{4}\left[z_j^2 + \overline{z_j}^2 + 2|z_j|^2\right] \tag{5.18}$$

By the same token, we have

$$\begin{aligned} \Im(z_j)^2 &= -\frac{1}{4}\left[z_j^2 + \overline{z_j}^2 - 2|z_j|^2\right] & (5.19) \\ \text{and} \quad \Re(z_j)\Im(z_j) &= -\frac{1}{4}i\left[z_j^2 - \overline{z_j}^2\right] & (5.20) \end{aligned}$$

Substituting 5.18, 5.19 and 5.20 into 5.17 gives

$$4.\mathbb{E}\left((\mathbf{t}.h_j\hat{\rho}(\mathbf{b}))^2\right) = \mathbb{E}\left(z_j^2\right)[t_1 - it_2]^2 + 2\mathbb{E}\left(|z_j|^2\right)\left[t_1^2 + t_2^2\right] + \mathbb{E}\left(\overline{z_j}^2\right)[t_1 + it_2]^2 \tag{5.21}$$

It is easy to see that

$$\mathbb{E}\left(z_j^2\right) = \mathbb{E}\left(\overline{z_j}^2\right) = 0 \quad \text{and} \quad \mathbb{E}\left(|z_j|^2\right) = \mathbb{E}\left(|h_j|^2\right)$$

whence

$$\mathbb{E}\left((\mathbf{t}.h_j\hat{\rho}(\mathbf{b}))^2\right) = \frac{1}{2}\mathbb{E}\left(|h_j|^2\right)\left[t_1^2 + t_2^2\right] \tag{5.22}$$

Summing over j gives 5.15 and hence the theorem is proved. □

5.4 First Application

Let $\rho : \mathbb{F}_q^s \to \mathbb{C}$ be a product of additive and multiplicative characters on $\mathbb{F}_q$. *i.e.*

$$\rho = \chi_1 \cdots \chi_s$$

Then the natural domain of definition C for ρ is the subset $C_1 \times \ldots \times C_s \subseteq \mathbb{F}_q^s$ where

$$C_v = \begin{cases} \mathbb{F}_q, \ \chi_v \text{ additive} \\ \mathbb{F}_q^*, \ \chi_v \text{ multiplicative} \end{cases}$$

Clearly $cq^{-s} \to 1$ as $q \to \infty$.

We put $m = 1$ so that $A' = A$ (i.e. the trivial partition of $A = \mathbb{F}_q^r$) and we also put $h_1 = 1$. Then we have, as in § 4.1, the sum

$$S_1 = \sum_{\mathbf{a} \in \mathbb{F}_q^r} \chi_1(f_1(\mathbf{a})) \ldots \chi_s(f_s(\mathbf{a})) \tag{5.23}$$

where f is a random polynomial vector in $\mathcal{F}$.

As in §5.2 we assume that the order of the character product approaches a limit $\tau \leq \infty$ as $q \to \infty$. We can then apply Theorems 5.3.3-5.3.7 to S_1 to get the following corollary:-

Corollary 5.4.1. *Let $q \to \infty$. Then the character sum S_1 has the following behaviour as $d \to \infty$ arbitrarily slowly with q:*

1. *If $\tau = 1$ then $q^r - S_1$ converges to zero in probability.*
2. *If $\tau = 2$ then $\frac{S_1}{q^{\frac{r}{2}}}$ converges in distribution to a Gaussian variable.*
3. *If $\tau \geq 3$ then $2q^{-\frac{r}{2}}S_1$ converges in distribution to an isotropic $\mathbb{R}^2$-Gaussian variable.*

Proof. Since $cq^{r-s} \sim q^r \to \infty$, Theorem 5.3.5 applies. We note that $\gamma = 1$ and $cq^{-s} \to 1$.

For $\tau = 1$ we have

$$\mathbb{E}(q^r - S_1) = q^r - \mu_1 = q^r - cq^{r-s}\gamma \to 0$$

Also,

$$\sigma_1^2 = cq^{r-s}\gamma(1 - cq^{-s})$$

Hence, by Lemma 1.7.4, S_1 converges in probability to zero.

For $\tau = 2$ we note that

$$\mathbf{H} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$$

That is, we are dealing with a distribution in $\mathbb{R}$. The result comes from applying part (2) of Theorem 5.3.5.

The case $\tau \geq 3$ is proved in a similar way. □

From this we can deduce the limiting distribution of the real-valued random variable, $|S_1|$, in the case that $\tau \geq 3$:

Corollary 5.4.2. *If $\tau \geq 3$, then $2q^{-\frac{r}{2}}|S_1|$ converges in distribution to a Weibull distribution with parameters $-\frac{1}{2}, 2$. That is,*

$$Prob\left(|S_1| \leq \frac{1}{2}q^{\frac{r}{2}}x\right) \to 1 - \exp\left(-\frac{1}{2}x^2\right) \qquad \forall x \geq 0$$

Proof. We apply the result in Appendix D which says that the modulus of an isotropic $\mathbb{R}^2$-Gaussian variable has a Weibull distribution with the required parameters. □

A very similar result to the above was proved in [35], but using a slightly different probability space and a completely different method.

As mentioned in §4.2, bounds for character sums in one variable are known but only when the degree of the polynomial is sufficiently small compared with the size of the field. This result gives the average behaviour over all degrees of polynomial and includes the multivariate case, which is much more difficult to find estimates for using number-theoretic and geometric techniques.

5.5 Second Application

Let $\chi_1, \dots, \chi_r$ be multiplicative characters of $\mathbb{F}_q^*$ whose product χ has order m. This means that χ takes m distinct values over $A' = (\mathbb{F}_q^*)^s$.

We partition A' into $A_1 \sqcup \dots \sqcup A_m$ according to these m values and define $h_1, \dots, h_m$ to be these m values in $\mathbb{C}$. That is,

$$h_j = \chi(\mathbf{a}_j)$$

where $\mathbf{a}_j \in A_j$ is any representative.

Putting $C = \{\mathbf{0}\} \subseteq \mathbb{F}_q^s$ and $\rho = \delta_C$ gives us the following sum which is like that of §4.4:

$$S_2 = \sum_{\mathbf{f}(\mathbf{a})=\mathbf{0}} \chi_1(a_1)\dots\chi_r(a_r) \tag{5.24}$$

Applying Theorems 5.3.3-5.3.7 to this sum gives us the following:

Corollary 5.5.1. *Let $q \to \infty$. Then the character sum S_2 has the following behaviour as $d \to \infty$ arbitrarily slowly with q:*

1. *If $r < s$ then S_2 converges to zero in probability.*
2. *If $r = s$ then S_2 converges in distribution to the variable*
$$Y = \sum_{j=1}^{m} h_j Y_j$$
where the Y_j are independent Poisson variables with parameter $\frac{1}{m}$.
3. *If $r > s$ then $\frac{S_2}{\sigma}$ converges in distribution to an isotropic $\mathbb{R}^2$-Gaussian variable; here*
$$\sigma^2 = \frac{1}{2}q^{r-s}$$

Proof. The result for $r < s$ follows from Theorem 5.3.3. If $r = s$, we have $cq^{r-s} \to 1$, whence our mixed Poisson variables have parameter $\frac{1}{m}$, from Theorem 5.3.4.

Finally, suppose $r > s$. Note that $\gamma = c = 1$, so that in the case $\tau = 2$, we apply Theorem 5.3.6 with

$$\sigma_2^2 = \frac{n}{mN}$$

Also, it can be shown, using the orthogonality relations for characters, that the matrix $\mathbf{H}$ simplifies to

$$\mathbf{H} = \begin{bmatrix} \frac{m}{2} & 0 \\ 0 & \frac{m}{2} \end{bmatrix}$$

It follows that the covariance matrix of $\frac{\Sigma}{\sigma}$ is $\mathbf{I}$ in this case.

For $r \geq 3$ we have

$$\sigma_3^2 = \frac{cn}{2mN}\gamma\sum_{j=1}^{m}|h_j|^2 = \frac{n}{2N}$$

and the result follows from Theorem 5.3.7. □

We now discuss an interesting example. Consider the character sum

$$S_2' = \sum_{\substack{a\in\mathbb{F}_p \\ f(a)=0}} \left(\frac{a}{p}\right) \tag{5.25}$$

over the prime finite field $\mathbb{F}_p$. Here $\left(\frac{a}{p}\right)$ denotes the Legendre symbol, that is, the quadratic character on $\mathbb{F}_p$.

To get this sum, we have put $q = p$, $m = 2$, $h_1 = 1$, $h_2 = -1$ and $r = s = 1$ in the definition of S_2. By the above corollary, S_2 will converge to

$$Y = Y_1 - Y_2$$

where Y_1 and Y_2 are independent Poisson variables with parameter $\frac{1}{2}$. However, we can actually say more about this variable.

Corollary 5.5.2. *Let $q \to \infty$ and let $d \to \infty$ arbitrarily slowly with q. Then the character sum S_2' converges in distribution to a Bessel distribution with parameter* 1. *That is,*

$$Prob(S_2' = v) \to e^{-1}I_v(1) \qquad (\forall v \in \mathbb{Z})$$

Proof. From the proof of 5.3.4, we have

$$C(\mathbf{t}) \to \prod_{j=1}^{m}\exp\left[\frac{\gamma\lambda}{m}\mathbb{E}\left(e^{i\mathbf{t}.h_j\rho(\mathbf{b})}\right) - 1\right]$$

In this particular case, it reduces to

$$\begin{aligned} C(\mathbf{t}) &\to \exp\left[\frac{1}{2}\mathbb{E}\left(e^{i\mathbf{t}.\rho(\mathbf{b})}\right) - 1\right]\exp\left[\frac{1}{2}\mathbb{E}\left(e^{-i\mathbf{t}.\rho(\mathbf{b})}\right) - 1\right] \\ &= \exp\left[\frac{1}{2}\left(e^{it_1} - 1\right)\right]\exp\left[\frac{1}{2}\left(e^{-it_1} - 1\right)\right] \\ &= \exp\left[\cos t_1 - 1\right] \end{aligned}$$

By Appendix F, this is the characteristic function of a Bessel distribution with parameter 1 and the result follows from Theorem 1.7.8. □

We also refer the reader to [16], p59 for a discussion on the 'randomised random walk in one dimension' and its relation to Bessel functions.

Another example along the same lines is the character sum

$$S_2'' = \sum_{\substack{a\in\mathbb{F}_p \\ f(a)=0}} \chi(a) \tag{5.26}$$

where χ is a multiplicative character of order 4 on $\mathbb{F}_p$, $p \equiv 1 \bmod 4$. Then S_2'' is a discrete random variable taking values in $\mathbb{Z}[i]$, the Gaussian integers. Our parameters this time are: $q = p$, $r = s = 1$, $m = 4$, $h_j = i^j$ $(1 \leq j \leq 4)$. According to Corollary 5.5.1, the sum S_2'' converges in distribution to the variable

$$Y = (Y_4 - Y_2) + i(Y_1 - Y_3)$$

where the Y_j are independent Poisson variables with parameter $\frac{1}{4}$. From the last example, however, we infer that $(Y_4 - Y_2)$ and $(Y_1 - Y_3)$ must each be Bessel variables with parameter $\frac{1}{2}$. Hence we have, for all $u, v \in \mathbb{Z}$,

$$Prob\left(S_2'' = u + iv\right) \to e^{-1} I_u\left(\frac{1}{2}\right) I_v\left(\frac{1}{2}\right) \tag{5.27}$$

as $d \to \infty$ arbitrarily slowly with p, $p \equiv 1$ mod 4.

5.6 Third Application

Let $\chi_1, \ldots, \chi_r$ and hence $\mathbf{A}, \mathbf{h}$ be as in §5.5. This time, let ρ be a product of additive characters $\chi_1 \cdots \chi_s$ on $\mathbb{F}_q$ so that this particular version of Σ is a kind of 'Gauss sum':

$$S_3 = \sum_{\mathbf{a}\in\mathbb{F}_q^r} \psi_1(a_1)\cdots\psi_r(a_r)\chi_1(f_1(\mathbf{a}))\cdots\chi_s(f_s(\mathbf{a}))$$

Of course, this is just a more general version of the example given in §4.6. We assume that the products of the χ_j and of the ψ_j are non-trivial characters on $\mathbb{F}_q^r$ and $\mathbb{F}_q^s$, respectively. Again we apply Theorems 5.3.3-5.3.7:

Corollary 5.6.1. *Let $q \to \infty$. Then $\frac{S_3}{\sigma}$ converges in distribution to an isotropic $\mathbb{R}^2$-Gaussian variable; here*

$$\sigma^2 = \frac{1}{2}q^r$$

Proof. $\frac{c}{n}N \sim n = q^r \to \infty$ and $\chi_1 \cdots \chi_s$ is non-trivial, so only Theorems 5.3.6 & 5.3.7 apply. In the case $\tau = 2$ we have, as in the proof of Corollary 5.5.1, that $\mathbf{H} = \frac{m}{2}\mathbf{I}$ and so the result is clear in this case. Also, if $\tau \geq 3$ then Theorem 5.3.7 applies directly to give the result (as in proof of Corollary 5.5.1). □

We note that the resulting distribution in this case is the same as that of Corollary 5.4.1, part 3; that is, S_1 and S_3 have the same asymptotic behaviour has $q, d \to \infty$, at least in the case $\tau \geq 3$. It follows that $|S_3|$ follows the same behaviour as $|S_1|$, namely that described in Corollary 5.4.2.

5.7 Fourth Application

Consider the character sum defined by

$$S_4 = \sum_{f_1(a)=0} \psi(a)\chi(f_2(a)) \tag{5.28}$$

where f_1, f_2 are random polynomials in $\mathbb{F}_q[X]$, and each of χ, ψ can be an additive or a multiplicative character. Then this sum can be obtained from Σ by putting:

1. $r = 1$ and $s = 2$.
2. $C = \{0\} \times \mathbb{F}_q$ or $C = \{0\} \times \mathbb{F}_q^*$, so that $cq^{r-s} \to 1$.
3. $A_1, \dots, A_m$ to be the cosets in $\mathbb{F}_q$ (or $\mathbb{F}_q^*$) of the distinct values of χ, and $h_1, \dots, h_m$ to be these values (so $\gamma = 1$).
4. $\rho = \delta_C \chi$

Further, if we assume that $\tau = \infty$, *i.e* the order of χ tends to infinity with q. Then Theorem 5.2.1 applies and Theorems 5.3.3-5.3.7 give:

Corollary 5.7.1. *The character sum S_4 converges to the compound Poisson distribution with parameters* 1 *and Z, Z being a complex random variable uniformly distributed on the circle G_∞.*

Proof. Since $cq^{r-s} \to 1$, only Theorem 5.3.4 applies in this case. Since $\tau = \infty$, $h_j G_\tau = G_\tau$ and hence S_4 converges in distribution to a sum of m identically distributed compound Poisson-$(\frac{1}{m}, Z)$ variables. This is clearly equivalent to a Poisson-$(1, Z)$ variable, hence the result. □

This is an interesting result, as the compound variable obtained is a random walk in two-dimensions with a random number of steps. Each step has length 1 and can be equally likely in any direction, and the total number of steps is a Poisson variable with parameter 1. In particular, the expected number of steps is 1. A good question might be, "what is the expected length of the resulting vector?" or even, "what is the distribution of this length?".

Let Z_n denote the result of a random walk consisting of precisely n steps in $\mathbb{R}^2$, each step being uniformly distributed on G_∞. The distribution of Z_n was studied by Kluyver [23] and Rayleigh [37] who obtained

$$F_n(x) = Prob(|Z_n| \le x) = x \int_0^\infty J_1(xt) J_0(t)^n dt \quad (n \ge 1) \tag{5.29}$$

where J_v denotes the standard Bessel function of order v (see Appendix E). A proof of (5.29) can be found in [40], p420.

From this we deduce the density of $|Z_n|$:

$$f_n(x) = \frac{dF_n}{dx} = \int_0^\infty J_0(xt)tJ_0(t)^n dt \quad (n \geq 1) \tag{5.30}$$

This formula was used by Odoni in [35] to study a random character sum.

Now, let S denote the compound Poisson distribution described in Corollary 5.7.1. We have

$$\begin{aligned}
F(x) = Prob(|S| \leq x) &= e^{-1}\frac{1}{0!}F_0(x) + \sum_{k=1}^{\infty} e^{-1}\frac{1}{k!}F_k(x) \\
&= e^{-1} + \sum_{k=1}^{\infty} e^{-1}\frac{1}{k!}\int_0^\infty xJ_1(xt)j_0(t)^k dt \\
&= e^{-1} + \int_0^\infty e^{-1}xJ_1(xt)\left(\sum_{k=1}^{\infty}\frac{J_0(t)^k}{k!}\right)dt \\
&= e^{-1} + \int_0^\infty xJ_1(xt)\left(\exp(J_0(t)-1)-1\right)dt
\end{aligned}$$

Unfortunately, this integral does not appear to be an elementary function. By a similar calculation, we can obtain the density $f(x)$ of S:

$$\begin{aligned}
f(x) &= \sum_{k=0}^{\infty} e^{-1}\frac{1}{k!}f_k(x) \\
&= \int_0^\infty tJ_0(xt)\left(\exp(J_0(t)-1)-1\right)dt
\end{aligned}$$

This allows us to get an expression for the expectation of $|S|$, using

$$\mathbb{E}(|S|) = \int_0^\infty xf(x)dx$$

If required to, one could estimate this integral using numerical methods to get an approximate value for the expectation.

Concluding Remarks

Up to now we have exploited the similarities between mappings and polynomials over finite fields in order to prove some convergence results for certain variables associated with random polynomials. These results had applications to zeros, primitive roots, image sizes and character sums. We shall finish off by suggesting more ways in which the theory could be developed, hopefully giving motivation for future research.

Classes of Polynomials

In our work, we have used the probability space $\mathcal{F}$, essentially consisting of all polynomials up to degree d. Instead, one could restrict one's attention to a certain class of polynomials and attempt to gain information about them. A typical example might be *linearized polynomials*, defined thus:

Def 5.7.2. Let F be a finite extension of $\mathbb{F}_q$. A polynomial $f \in F[X]$ is q-linearized if

$$f(X) = c_d X^{q^d} + \ldots + c_1 X^q + c_0$$

It is easy to see that

$$\begin{aligned} f(a+b) &= f(a) + f(b) \quad \forall a, b \in F \\ f(\lambda a) &= \lambda f(a) \quad \forall a \in F, \lambda \in \mathbb{F}_q \end{aligned}$$

so that f induces an $\mathbb{F}_q$-linear map from F to F. Linearized polynomials exhibit some very nice properties (see [30]) and would be amenable to probabilistic techniques.

Def 5.7.3. Another important class of polynomials are the Dickson polynomials (of the first kind), defined by:

$$g_d(X, a) = \sum_{j=0}^{[\frac{d}{2}]} \frac{d}{d-j} \binom{d-j}{j} (-a)^j \; X^{d-2j}$$

where $d \in \mathbb{N}$ and $a \in \mathbb{F}_q$. We also have the Dickson polynomials of the second kind:

$$f_d(X, a) = \sum_{j=0}^{[\frac{d}{2}]} \binom{d-j}{j} (-a)^j \; X^{d-2j}$$

There also exist generalisations of these to several variables.

Dickson polynomials were introduced by L.E.Dickson about a hundred years ago. Their interesting algebraic properties mean that they have applications in the pure and applied theory of finite fields, including an application to cryptosystems (see [34]). We also remark that the direct-image size of Dickson polynomials has been studied in [11]. For a full account of the theory of Dickson polynomials, see [29].

More on Zeros of Random Polynomials

The results on inverse-image size in Chapter 3 had applications to zeros of random polynomials. As in [36], we obtained the asymptotic distribution of the number of zeros of f *in the ground field*, $\mathbb{F}_q$. We could ask the same question, but this time allow the zeros to lie in some extension field $\mathbb{F}_{q^u}$.

Def 5.7.4. Let $s = 1$ so that our random polynomial vector f consists of a single polynomial in $\mathbb{F}_q[X_1, \dots, X_r]$. We define

$$\zeta_{0,u}(f) = \sharp\{\mathbf{a} \in \mathbb{F}_{q^u}^r \ : \ f(a) = 0\}$$

When we try to calculate the moments of $\zeta_{0,u}$ we obtain

$$\begin{aligned} \mathbb{E}\left(\binom{\zeta_{0,u}}{k}\right) &= \frac{1}{\sharp\mathcal{F}} \sum_{h\in\mathcal{F}} \sharp\{k\text{-subsets of zeros of } h\} \\ &= \frac{1}{\sharp\mathcal{F}} \sum_{\substack{K\subset F \\ \sharp K=k}} \sharp\{h \in F \ : \ h(K) = 0\} \end{aligned}$$

Clearly, one would need some generalisation of Lemma 2.2.1 or one of its corollaries in order to obtain the full moment sequence for $\zeta_{0,u}$.

More Character Sums

Our aim in Chapters 4&5 was to generalise the inverse-image variable of Chapter 3 and to get the most general variable possible. Instead, if one had a particular character sum in mind which did not fit into the Chapter 5 picture, one could go back and try to construct a specific variable which behaved accordingly. For example, consider the sum

$$S = \sum_{\substack{\mathbf{a}\in\mathbb{F}_q^r \\ f_2(\mathbf{a})\neq 0}} \psi\left(\frac{f_1(\mathbf{a})}{f_2(\mathbf{a})}\right)$$

where ψ is an additive character on $\mathbb{F}_q$. This cannot be dealt with by our variable Σ, but one could modify the probability space $\mathcal{F}$ to include rational functions and proceed in the usual way.

Another interesting problem would be to study incomplete character sums. That is, character sums over a proper subset of $\mathbb{F}_q$ which has size asymptotically 'smaller' than q *e.g.* $O\left(q^{\frac{1}{2}}\right)$. Such sums have important applications in coding theory (see [19]).

Random Polynomials over Galois Rings

A Galois ring is basically a generalisation of a finite field. It is a finite commutative ring with identity, having characteristic p^m (p prime, $m \in \mathbb{N}$). The idea behind Galois rings is to have a theory which covers both finite fields $\mathbb{F}_q$ and the residue class rings $\frac{\mathbb{Z}}{p^m\mathbb{Z}}$. This is outlined by B.McDonald in [33]:

"It is classically accepted that in certain classes of problems in combinatorial mathematics the researcher handles separately the finite field $GF(p^n)$ and the prime ring $\frac{\mathbb{Z}}{p^m\mathbb{Z}}$. It is our belief that both cases should be treated simultaneously in the setting of a Galois ring."

Unfortunately, as far as random polynomials are concerned, this statement is perhaps a little ambitious, as we shall soon see. We can however say a few things about random polynomials over Galois rings. Let us begin with a definition.

Def 5.7.5. let p be a prime and $m, e \in \mathbb{N}$. By the Galois ring $GR(p^m, e)$ we mean the unique separable extension of $\frac{\mathbb{Z}}{p^m\mathbb{Z}}$ of degree e. We have that

$$GR(p^m, e) \cong \frac{\frac{\mathbb{Z}}{p^n\mathbb{Z}}[X]}{\langle f \rangle}$$

where f has degree e and is a "basic" monic irreducible polynomial in $\frac{\mathbb{Z}}{p^m\mathbb{Z}[X]}$, that is, a monic polynomial which is irreducible modulo p.

Properties of Galois Rings

1. $\#GR(p^m, e) = p^{me}$
2. $GR(p^m, e)$ has characteristic p^m
3. $GR(p^m, e)$ is a commutative local ring, having residue field $\mathbb{F}_{p^e}$
4. The maximal ideal of $GR(p^m, e)$ is principal, generated by p *i.e.* $\mathfrak{m} = \langle p \rangle$
5. Every ideal of $GR(p^m, e)$ is of the form $\langle p^k \rangle$ for some $k \in \mathbb{N}$

For a fuller account of Galois rings, we refer the reader to [33].

We define random polynomials over Galois rings in an analagous way to those over finite fields. However, we shall stick to the simplest case, namely polynomials in one variable. Let $R = GR(p^m, e)$ and define $\mathcal{R}_d = \{f \in R[X] \ : \ \deg f \leq d\}$. This is a finite subset of $R[\mathbf{X}]$ and is made into a probability space in the obvious way.
For a random polynomial $f \in \mathcal{R}_d$ we define the random variable ζ_0 to be the number of zeros of f in R.

One would like to be able to calculate the moments $\binom{\zeta_0}{k}$. However, for $k \geq 3$ this appears to be combinatorially infeasible, and so we shall calculate only the mean and variance of ζ_0. We first need a preliminary lemma and corollary:

Lemma 5.7.6. *Let $a, b \in R$, $a \neq b$, and put $v = ord_p(a - b)$. Then, for $g \in R[X]$,*

1. $g(a) = 0 \quad \Leftrightarrow \quad g \in \langle X - a \rangle$
2. $g(a) = g(b) = 0 \quad \Leftrightarrow \quad g \in \langle\, (X - a)(X - b),\ p^{m-v}(X - a)\,\rangle$

Proof. 1. The first assertion follows from the factor theorem, which holds in S[X] for any ring S.

2. ($\Leftarrow$) is easily verified.

($\Rightarrow$) Suppose that $f(a) = f(b) = 0$
Since $f(a) = 0$, we have $f = (X - a)g_1$ (some $g_1 \in R[X]$)
Then $0 = f(b) = (b - a)g_1(b)$
i.e. $up^v g_1(b) = 0$ (for some unit $u \in R$)
i.e. $g_1(b) \equiv 0 \ (\text{mod} p^{m-v})$
i.e. $g_1 = (X - b)g_2 + p^{m-v}g_3$
from which the result follows.

□

Def 5.7.7. Let $K \subseteq R$. For convenience, we define

$$\mathfrak{I}(K) = \{f \in R[X] \ : \ f(c) = 0\, \forall c \in K\}$$

As common sense would suggest, we abbreviate $\mathfrak{I}(\{a\})$ to $\mathfrak{I}(a)$ and $\mathfrak{I}(\{a, b\})$ to $\mathfrak{I}(a, b)$.

Corollary 5.7.8. *With a and b as above, we have*

1. $\#\,(\mathcal{R}_d \cap \mathfrak{I}(a)) = p^{med}$ *for* $d \geq 1$
2. $\#\,(\mathcal{R}_d \cap \mathfrak{I}(a, b)) = p^{e(m(d-1)+v)}$ *for* $d \geq 2$

Proof. 1. Each element g of $\mathcal{R}_d \cap \mathfrak{I}(a)$ is uniquely expressible in the form $g = (X - a)g_1$ where g_1 has degree $d - 1$. Since there are p^{me} possibilites for each coefficient of g_1, the total number is p^{med}.

2. We claim that in this case, each element of our ideal $\mathfrak{I}(a,b)$ is uniquely expressible in the form $g = (X-a)(X-b)g_1 + (X-a)\alpha$ where α is chosen modulo p^v. The result will then follow.

   To prove the claim, first note that if g is of the above form then it is obviously an element of the ideal.

   Next, suppose that $g \in \mathfrak{I}(a,b)$. Then

$$g = (X-a)(X-b)g_1 + p^{m-v}(X-a)g_2$$

for some $g_1, g_2 \in R[X]$.
By the factor theorem,

$$g_2 = (X-b)g_3 + \alpha$$

for some $g \in R[X]$, $\alpha \in R$. Therefore

$$\begin{aligned} f &= (X-a)(X-b)(g_1 + p^{m-v}g_3) + p^{m-v}(X-a)\alpha \\ &= (X-a)(X-b)g_1' + p^{m-v}(X-a)\alpha \end{aligned}$$

which is of the required form. It now remains to prove the uniqueness.

Suppose that

$$(X-a)(X-b)g_1 + p^{m-v}(X-a)\alpha_1 = (X-a)(X-b)g_2 + p^{m-v}(X-a)\alpha_2$$

Then

$$(X-b)g_1 + p^{m-v}\alpha_1 = (X-b)g_2 + p^{m-v}\alpha_2$$

Taking the degree of both sides we see that $\deg g_1 = \deg g_2 = \delta$, say. We proceed by induction on δ. Without loss, let us assume that $b = 0$.
When $\delta = 0$ we have $g_1, g_2 \in R$ and

$$Xg_1 + p^{m-v}\alpha_1 = Xg_2 + p^{m-v}\alpha_2$$

from which it follows that $g_1 = g_2$ and $\alpha_1 \equiv \alpha_2 \, (\mathrm{mod} p^v)$, as required. Assume now that the uniqueness is true for $\delta = k-1$, $(k \in \mathbb{N})$ and let $\delta = k$. Then

$$\begin{aligned} & Xg_1 + p^{m-v}\alpha_1 = Xg_2 + p^{m-v}\alpha_2 \\ \Rightarrow \quad & (\text{Coeff. of } X^k in\ g_1) = (\text{Coeff. of } X^k in\ g_2) \\ \Rightarrow \quad & g_1 = g_2 \quad \text{and} \quad \alpha_1 \equiv \alpha_2 \, (\mathrm{mod} p^v) \quad \text{by induction} \end{aligned}$$

and the claim is proved

□

We are now in a position to calculate the mean and variance of ζ_0.

Theorem 5.7.9. *For $d \geq 2$, the random variable ζ_0 has mean* 1 *and variance $m\left(1 - \frac{1}{p^e}\right)$.*

Proof. The mean of ζ_0 is given by

$$\begin{aligned}
\mathbb{E}(\zeta_0) &= \frac{1}{\sharp\mathcal{R}_d} \sum_{f \in \mathcal{R}_d} \sharp\{a \in R \,:\, f(a) = 0\} \\
&= \frac{1}{p^{me(d+1)}} \sum_{a \in R} \sharp\{f \in \mathcal{R}_d \,:\, f(a) = 0\} \\
&= \frac{1}{p^{me(d+1)}} \sum_{a \in R} p^{med} \\
&= 1
\end{aligned}$$

The second moment is given by

$$\begin{aligned}
\mathbb{E}\left(\binom{\zeta_0}{2}\right) &= \frac{1}{\sharp\mathcal{R}_d} \sum_{f \in \mathcal{R}_d} \sharp\{\{a,b\} \subseteq R \,:\, f(a) = f(b) = 0\} \\
&= \frac{1}{p^{me(d+1)}} \sum_{\{a,b\} \subseteq R} \sharp\{f \in \mathcal{R}_d \,:\, f(a) = f(b) = 0\} \\
&= \frac{1}{p^{me(d+1)}} \frac{1}{2} \sum_{a \in R} \sum_{v=0}^{m-1} \sum_{\substack{u \bmod p^{m-v} \\ p \nmid u}} \sharp\{f \in \mathcal{R}_d \,:\, f(a) = f(a + up^v) = 0\} \\
&= \frac{1}{p^{me(d+1)}} \frac{1}{2} \sum_{a \in R} \sum_{v=0}^{m-1} \sum_{\substack{u \bmod p^{m-v} \\ p \nmid u}} p^{e(m(d-1)+v)} \qquad \text{by Corollary 5.7.8} \\
&= \frac{1}{p^{me(d+1)}} \frac{1}{2} p^{me(d-1)} \sum_{a \in R} \sum_{v=0}^{m-1} \left(p^{e(m-v)} - p^{e(m-v-1)}\right) p^{ev} \\
&= \frac{1}{2} p^{-2me} \sum_{a \in R} p^{me} \sum_{v=0}^{m-1} \left(1 - \frac{1}{p^e}\right) \\
&= \frac{1}{2} m \left(1 - \frac{1}{p^e}\right)
\end{aligned}$$

Hence, by 1.1, we have

$$V(\zeta_0) = m\left(1 - \frac{1}{p^e}\right)$$

□

Note that, putting $m = 1$, we have that R is the finite field $\mathbb{F}_q$ where $q = p^e$ and the above theorem agrees with our results in Chapter 2.

For Galois rings, it would be nice to be able to prove convergence results analogous to Theorems 3.3.2 and 3.7.1, that is, results about inverse and direct image sizes. Unfortunately, the techniques we have used in studying random polynomials over finite fields seem to fail in general for Galois rings. This is because we have been exploiting the fact that every map on a finite field is a polynomial (Prop. 2.4.1), while this is far from being true in a general Galois ring. We refer the reader to [10, 8] for a study of maps and polynomials over Galois rings. It is likely that a different approach is required to study random polynomials over such rings.

Appendices

A. A Combinatorial Sieve

In Chapter 3, we require an estimate of the size of a set which is calculated using the *inclusion-exclusion principle* in which not all the details are known. We have the following notation and result: Let Ω be a finite non-empty set and let $E_1, \ldots, E_m \subseteq \Omega$. For $1 \leq l \leq m$ put

$$\alpha_l = \sum_{\substack{L \subseteq \{1,\ldots,m\} \\ \sharp L = l}} \sharp \bigcap_{i \in L} E_i$$

and define $\alpha_0 = 1$ so that $\alpha_l \geq 0 \quad (\forall l)$. Then, by the inclusion-exclusion principle,

$$\sharp \bigcup_{i=1}^{m} E_i = \sum_{l=1}^{m} (-1)^{l-1} \alpha_l$$

Lemma. *(Brun-Waring principle)*
For each $1 \leq j \leq m-1$,

$$\left| \sharp \bigcup_{i=1}^{m} E_i - \sum_{l=1}^{j} (-1)^{l-1} \alpha_l \right| \leq \alpha_{j+1}$$

Proof. For $1 \leq i \leq m$, let us define

$$\delta_i(x) = \begin{cases} 1, & x \in E_i \\ 0, & x \notin E_i \end{cases} \qquad (\forall x \in \Omega)$$

Then we have that

$$\sharp \left(\Omega \backslash \bigcup_{i=1}^{m} E_i \right) = \sum_{x \in \Omega} \prod_{i=1}^{m} (1 - \delta_i(x)) \tag{5.31}$$

For the moment, fix $x \in \Omega$ and define $f_x : \mathbb{R} \to \mathbb{R}$ by

$$f_x(t) = \prod_{i=1}^{m} (1 - t\delta_i(x)) \tag{5.32}$$

Expanding this out gives

$$f_x(t) \;=\; \sum_{l=0}^{m}(-1)^l\, t^l \left(\sum_{\substack{L\subseteq\{1,\dots,m\}\\ \sharp L=l}} \; \prod_{i\in L} \delta_i(x) \right) \tag{5.33}$$

$$\;=\; \sum_{l=0}^{m}(-1)^l\, t^l \quad \alpha_l(x) \tag{5.34}$$

where

$$\sum_{x\in\Omega} \alpha_l(x) = \alpha_l \tag{5.35}$$

From 5.34 we have that

$$\frac{f_x^{(l)}(0)}{l!} = (-1)^l\; \alpha_l(x) \tag{5.36}$$

Also, differentiating 5.32 l times gives

$$(-1)^l\; f_x^{(l)}(t) = \sum_{\substack{L\subseteq\{1,\dots,m\}\\ \sharp L=l}} \; \prod_{i\notin L} (1 - t\delta_i(x)) \tag{5.37}$$

from which we see that $(-1)^l f_x^{(l)}(t)$ is monotone decreasing in $[0,1]$.

Applying MacLaurin's theorem (Lagrange remainder) to f_x, we have, for $0 \leq l < m$

$$f_x(t) = \sum_{l=0}^{j}(-1)^l\; \alpha_l(x).t^l \;+\; R_{j+1}(x) \tag{5.38}$$

where

$$R_{j+1}(x) = \frac{f_x^{(j+1)}}{(j+1)!} \tag{5.39}$$

Since $(-1)^{j+1} f_x^{(j+1)}(t)$ is monotone decreasing on $[0,1]$ we have

$$(-1)^{j+1} R_{j+1}(x) \;\leq\; (-1)^{j+1}\, \frac{f_x^{(j+1)}(0)}{(j+1)!} \;=\; \alpha_{j+1}(x) \qquad \text{(from 5.36)} \tag{5.40}$$

Combining 5.38 and 5.40 gives

$$(-1)^{j+1} \left(f_x(t) \;-\; \sum_{l=0}^{j}(-1)^l\; \alpha_l(x).t^l \right) \;\leq\; \alpha_{j+1}(x) \tag{5.41}$$

Summing over all $x \in \Omega$, putting $t = 1$ and using 5.31 & 5.35 gives the result. □

B. A Multinomial Expansion

Consider the following expression:

$$z = (x_1 + \ldots + x_v)^k (y_1 + \ldots + y_v)^l \tag{5.42}$$

where $v, k, l \in \mathbb{Z}_{\geq 0}$ and $v \geq k + l$.

To expand this expression, normally one would use the multinomial theorem twice, and combine the answer. That is,

$$\begin{aligned} z &= \left(\sum_{\substack{\mathbf{i} \\ i_1 + \ldots + i_v = k}} \binom{k}{\mathbf{i}} x_1^{i_1} \ldots x_v^{i_v} \right) \left(\sum_{\substack{\mathbf{j} \\ j_1 + \ldots + j_v = k}} \binom{l}{\mathbf{j}} y_1^{j_1} \ldots y_v^{j_v} \right) && (5.43) \\ &= \sum_{\mathbf{i}, \mathbf{j}} \binom{k}{\mathbf{i}} \binom{l}{\mathbf{j}} \prod_{u=1}^{v} x_u^{i_u} y_u^{j_u} && (5.44) \end{aligned}$$

In the the above, $\mathbf{i}$ denotes the multi-index $(i_1, \ldots, i_v)$ where $i_u \in \mathbb{Z}_{\geq 0}$ $(1 \leq u \leq v)$ (and likewise for $\mathbf{j}$) . The multinomial coefficient is defined by

$$\binom{k}{\mathbf{i}} = \frac{k!}{i_1! \ldots i_v!}$$

Also, for convenience we shall define

$$||\mathbf{i}|| = i_1 + \ldots + i_v$$

Since $v \geq k + l$, each summand in (5.44) has at least one of the pairs (i_u, j_u) equal to zero. This is inefficient as far as our moment calculations of Chapter 4 are concerned, and we therefore require a better way of writing (5.44), namely one in which none of the (i_u, j_u) are zero.

Consider a typical, term of the summation:

$$\prod_{u=1}^{v} x_u^{i_u} y_u^{j_u}$$

Suppose that, in this summand, the total number of non-zero (i_u, j_u) is t. Then, certainly t can lie anywhere in the range of 1 to $k + l$. We then have a subset $T \in \{1, \ldots, v\}$ of size t, where $u \in T$ if and only if $(i_u, j_u) \neq \mathbf{0}$.

Finally, for each subset T, the set of possible products of the x_w and y_w will now be indexed by $\mathbf{i} = (i_1, \ldots, i_t)$ & $\mathbf{j} = (j_1, \ldots, j_t)$ where $(i_w, j_w) \neq \mathbf{0}$ for each $1 \leq w \leq t$.

Hence, we have

$$z = \sum_{t=1}^{k+l} \sum_{\substack{T \in \{1, \ldots, v\} \\ T = \{s(1), \ldots, s(t)\}}} \sum_{\mathbf{i}, \mathbf{j}} \binom{k}{\mathbf{i}} \binom{l}{\mathbf{j}} \prod_{w=1}^{t} x_{s(w)}^{i_w} y_{s(w)}^{j_w} \tag{5.45}$$

Note that the innermost sum is over all $\mathbf{i}, \mathbf{j} \in (\mathbb{Z}_{\geq 0})^t$ such that $(i_w, j_w) \neq \mathbf{0}$ $(1 \leq w \leq t)$.

C. Uniform Distributions on the Circle

In this section we prove Theorem 5.2.1

Theorem. *Let $\{z_k\}_{k\in\mathbb{N}}$ be a sequence of discrete complex-valued random variables such that, for each $k \in \mathbb{N}$, z_k is uniformly distributed on $G_{t(k)}$. If $t(k) \to \infty$ as $k \to \infty$, then z_k converges in distribution to a uniform distribution on G_∞.*

Proof. There is a 1-1 correspondence between the circle G_∞ and the interval $[0, 1)$ arising from

$$e^{2\pi i\theta} \mapsto \theta \bmod 1 \qquad \theta \in \mathbb{R}$$

We can therefore think of our probability distributions as being on $[0, 1)$ (see [16], p61 for details).

Suppose that X_m is uniformly distributed on $\{0, \frac{1}{m}, \ldots, \frac{m-1}{m}\}$, the set corresponding to G_m $(m \in \mathbb{N})$. Also, let X be uniformly distributed on $[0, 1)$. Then it suffices to show that the X_m converge in distribution to X as $m \to \infty$.

Now

$$\begin{aligned} \mathbb{E}(e^{itX}) &= \int_0^1 e^{itx}dx \\ &= \frac{1}{it}(e^{it} - 1) \end{aligned}$$

while

$$\begin{aligned} \mathbb{E}(e^{itX_m}) &= \frac{1}{m}\sum_{j=0}^{m-1} e^{itj} \\ &= \frac{e^{it} - 1}{m(e^{i\frac{t}{m}} - 1)} \end{aligned}$$

But $m(e^{i\frac{t}{m}} - 1) = m\frac{it}{m} + o(1)$ and so $\mathbb{E}(e^{itX_m}) \to \mathbb{E}(e^{itX})$ as $m \to \infty$. Hence, by the Continuity Theorem (1.7.8), the X_m converge in distribution to X, and the result follows. □

D. The Modulus of an Isotropic $\mathbb{R}^2$-Gaussian Variable

Suppose that Z is a Gaussian variable in $\mathbb{R}^2$ with covariance matrix $\sigma\mathbf{I}$. Then $|Z|$ is a random variable in $\mathbb{R}$ with the following distribution:

$$\begin{aligned}
Prob(|Z| \leq x) &= \frac{1}{2\pi\sigma^2}\int\int_{s^2+t^2\leq x^2} \exp\left(-\frac{1}{2\sigma^2}(s^2+t^2)\right)\,ds\,dt \\
&= \frac{1}{2\pi\sigma^2}\int_0^{2\pi}\int_0^{x} r.\exp\left(-\frac{1}{2\sigma^2}r^2\right)\,dr\,d\theta \\
&= \frac{1}{2\pi}\int_0^{2\pi}\left[-\exp\left(-\frac{1}{2}r^2\right)\right]_0^x\,d\theta \\
&= 1-\exp\left(-\frac{1}{2\sigma^2}x^2\right)
\end{aligned}$$

That is to say, $|Z|$ has a Weibull distribution with parameters $-\frac{1}{2\sigma^2}$, 2.

E. A Note on Bessel Functions

The Bessel function $J_v(\lambda)$ of order v (where $v \in \mathbb{Z},\ \lambda \in \mathbb{R}$) is defined by

$$\exp \lambda \left(T - \frac{1}{T}\right) = \sum_{v=-\infty}^{\infty} T^v J_v(\lambda) \tag{5.46}$$

which leads to the expressions

$$J_v(\lambda) = \sum_{l=0}^{\infty} \frac{(-1)^l \left(\frac{\lambda}{2}\right)^{2l+v}}{l!(l+v)!} \qquad (v \geq 0) \tag{5.47}$$

$$J_v(\lambda) = (-1)^v J_{-v}(\lambda) \qquad (v < 0) \tag{5.48}$$

However, the study of random walks in one dimension (see [16]) leads to distributions involving a modified version of the Bessel function:

$$I_v(\lambda) = \sum_{l=0}^{\infty} \frac{\left(\frac{\lambda}{2}\right)^{2l+v}}{l!(l+v)!} \qquad (v \geq 0) \tag{5.49}$$

$$I_v(\lambda) = I_{-v}(\lambda) \qquad (v < 0) \tag{5.50}$$

It is clear that I_v is related to J_v via

$$J_v(i\lambda) = i^v I_v(\lambda) \tag{5.51}$$

and from this we get the relation

$$\exp i\lambda \left(T - \frac{1}{T}\right) = \sum_{v=-\infty}^{\infty} (iT)^v I_v(\lambda) \tag{5.52}$$

Putting $T = -i$ in 5.52, gives

$$e^{-\lambda} \sum_{v=-\infty}^{\infty} I_v(\lambda) = 1 \tag{5.53}$$

from which it follows that, for a fixed $\lambda \in \mathbb{R}$, $f(v) = e^{-\lambda} I_v(\lambda)$ is a valid probability density function. Also, putting $T = -ie^{it}$ in 5.52 gives

$$\exp \lambda(\cos t - 1) = e^{-\lambda} \sum_{v=-\infty}^{\infty} e^{itv} I_v(\lambda) \tag{5.54}$$

Hence if X is a $\mathbb{Z}$-valued random variable with density function $f(v) = e^{-\lambda} I_v(\lambda)$, the characteristic function of X is

$$C_X(t) = \mathbb{E}(\exp itX) = \exp \lambda(\cos t - 1) \tag{5.55}$$

We refer the reader to [40], ch.2 for an introduction to Bessel functions.

F. Table of Characteristic Functions

DISTRIBUTION	PARAMETERS	$C(\mathbf{t})$	REFERENCE
Bernoulli	p	$1-p+pe^{it}$	[17], §5.8
Binomial	n, p	$(1-p+pe^{it})^n$	[17], §5.8
Poisson	λ	$\exp\lambda(e^{it}-1)$	Below
Gaussian in $\mathbb{R}$		$\exp\left(-\frac{1}{2}t^2\right)$	[17], §5.8
Gaussian in $\mathbb{R}^2$	$\mathbf{H}$	$\exp\left(-\frac{1}{2}\mathbf{t}^t\mathbf{H}\mathbf{t}\right)$	[17], §5.8
Compound Poisson	λ, F	$\exp\lambda(\mathbb{E}(e^{itX_1})-1)$	Below
Bessel	λ	$\exp\lambda(\cos t-1)$	Appendix E

Compound Poisson Distribution

Let X have a compound Poisson distribution with parameters F and λ. So

$$X = X_1 + \ldots + X_Y$$

where the X_j have distribution F, and Y is a Poisson-λ variable. Then

$$\begin{aligned}
\mathbb{E}\left(e^{itX}\right) &= \sum_{k=0}^{\infty}\mathbb{E}\left(\exp it\sum_{j=1}^{k}X_j\right) P(Y=k) \\
&= e^{-\lambda}\sum_{k=0}^{\infty}\mathbb{E}\left(e^{itX_1}\right)^k\frac{\lambda^k}{k!} \qquad \text{(by 1.5.2)} \\
&= e^{-\lambda}.\exp\lambda\mathbb{E}\left(e^{itX_1}\right) \\
&= \exp\lambda\left(\mathbb{E}\left(e^{itX_1}\right)-1\right)
\end{aligned}$$

Note that putting $F \equiv 1$ gives the result for an ordinary Poisson-λ variable.

G. Index of Notation

NOTATION	SECTION	DESCRIPTION
A	2.4	the domain of a random map, usually $\mathbb{F}_q^r$
A'	5.2	a particular subset of A
$\mathbf{a}$	2.4	an element of A
$\mathcal{A}$	1.1	a σ-algebra
B	2.4	the codomain of a random map, usually $\mathbb{F}_q^s$
$\mathbf{b}$	2.4	an element of B
C	3.1	a subset of B
c	3.1	the size of C
d_j	2.3	the degree of $f_j \quad (1 \leq j \leq s)$
d	2.3	the minimum of the d_j
δ_U	4.1	indicator function of a set U
$\mathbb{E}$	1.3	expectation operator
$\mathbb{F}_q$	2.1	finite field with q elements
$\mathcal{F}$	2.3	probability space for random polynomial vector
$\mathbf{f}$	2.3	random polynomial vector
f_j	2.3	components of random polynomial vector
G_k	5.1	the group of kth roots of unity in $\mathbb{C} \quad (k \in \mathbb{N})$
G_∞	5.1	the unit circle in $\mathbb{C}$
g_j	2.3	fixed polynomials in construction of $\mathcal{F}$
h_j	4.3	complex constants varying with $\mathcal{F}$
$\mathbf{H}$	1.4	covariance matrix
$\Im$	5.3	imaginary part
$\mathbf{I}$	5.3	identity matrix
m	4.3	number of subsets in partition of A'
N	2.4	the size of B, usually q^s
n	2.4	the size of A, usually q^r
P	1.1	a probability measure
p	1.6	a probability
$Prob(E)$	1.1	the probability of an event E

NOTATION	SECTION	DESCRIPTION
q	2.1	a prime power
r	2.2	a fixed natural number
$\Re$	5.3	real part
s	2.3	a fixed natural number
S_1	4.1	a character sum
S_2	4.2	a character sum
S_2', S_2''	5.5	a character sum
S_3	5.6	a character sum
S_4	5.7	a character sum
α	4.3	the size of A'
γ	5.2	the limit of $\frac{\alpha}{n}$
ζ_C, ζ_C^*	3.1	inverse-image size
η, η^*	3.1	direct-image size
λ	3.2	limit of $\frac{cn}{N}$
μ	1.4	mean
ρ	4.1	a function from B to $\mathbb{C}$
Σ, Σ^*	4.5	generalised inverse-image variable
σ	4.1	standard deviation
τ	5.2	the limit of a particular parameter
ϕ	2.4	random map

Bibliography

[1] A. Adolphson and Sperber S. Character sums in finite fields. *Comp. Math.*, 52:325–354, 1984.

[2] M.F. Atiyah and I.G. MacDonald. *Introduction to Commutative Algebra.* Addison-Wesley, 1969.

[3] A. Békéssy. On classical occupancy problems, I. *Magyar. Tud. Akad. Mat. Kutató Int Közl.*, 8:59–71, 1963.

[4] A. Békéssy. On classical occupancy problems, II. *Magyar. Tud. Akad. Mat. Kutató Int Közl.*, 9:133–141, 1964.

[5] B. J. Birch and H. P. F. Swinnerton-Dyer. Note on a problem of Chowla. *Acta Arith.*, 5:417–423, 1959.

[6] B. Bollobás. *Random Graphs.* Academic Press, 1st edition, 1985.

[7] E. Bombieri and Sperber S. On the esitimation of certain exponential sums. *Acta Arith.*, 69(4):329–358, 1995.

[8] J.V. Brawley and G.L. Mullen. Functions and polynomials over Galois rings. *J.Number Theory*, 41(2):156–166, 1992.

[9] L. Carlitz. Some topics in the arithmetic of polynomials. *Bull. Amer. Math. Soc.*, 48:679–691, 1942.

[10] L. Carlitz. Functions and polynomials mod p^n. *Acta Arith.*, 9:67–78, 1964.

[11] W-S. Chou and J.G. Calderon. Value sets of Dickson polynomials over finite fields. *J.Number Theory*, 30(3):334–344, 1988.

[12] S.D. Cohen. The values of a polynomial over a finite field. *Glasgow Math. J.*, 14:205–208, 1973.

[13] P. Deligne. Les intersections complètes de niveau de Hodge un. *Invent. Math.*, 15:237–250, 1972.

[14] P. Erdös and A. Rényi. On a classical problem of probability theory. *Magyar. Tud. Akad. Mat. Kutató Int Közl.*, 6(1-2):215–220, 1967.

[15] W. Feller. *An Introduction to Probability Theory and its Applications, vol 1.* J. Wiley and sons, 1957.

[16] W. Feller. *An Introduction to Probability Theory and its Applications, vol 2.* J. Wiley and sons, 1971.

[17] G.R. Grimmett and D.R. Stirzaker. *Probability and Random Processes.* Oxford, 2nd edition, 1992.

[18] G.H. Hardy and E.M. Wright. *An Introduction to the Theory of Numbers.* Oxford Science Publications, 5th edition, 1979.

[19] N.E. Hurt. Exponential sums and coding theory. *Acta Applicandae Mathematicae*, 46:49–91, 1997.

[20] K. Ireland and Rosen M. *A Classical Introduction to Modern Number Theory.* Springer-Verlag, 2nd edition, 1990.

[21] G.I. Ivchenko and Ju.I. Medvedev. Asymptotic behaviour of the number of aggregates of particles in the classical occupancy problem. *Theory Prob. Appl.*, 11, 1966.

[22] N. Katz. An overview of Deligne's proof of the Riemann Hypothesis for varieties over finite fields. *Proc. of Symposia in Pure Math.*, 28:275–305, 1976.

[23] J.C. Kluyver. A local probability problem. *Proc. Section of Sci., K. Akad. van Wet. te Amsterdam*, 8:749–755, 1906.

[24] A. Knopfmacher and J. Knopfmacher. Counting polynomials with a given number of zeros in a finite field. *Linear and Multilinear Algebra*, 26(4):287–292, 1990.

[25] A. Knopfmacher and J. Knopfmacher. The distribution of values of polynomials over a finite field. *Linear Algebra Appl.*, 134:145–151, 1990.

[26] V.F. Kolchin. The speed of convergence to limiting distributions in the classical ball problem. *Theory Prob. Appl.*, 11:128–140, 1966.

[27] V.F. Kolchin, B.A. Sevastianov, and V.P. Christyakov. *Random Allocations.* V.H.Winston, 1978.

[28] S. Kotz and N.L. Johnson. *Urn Models and their Applications.* J.Wiley and sons, 1977.

[29] R. Lidl. *Topics in Polynomials of One and Several Variables and their Applications.* World Science Publications, 1993.

[30] R. Lidl and H. Niederreiter. *Introduction to finite fields and their applications.* Cambridge University Press, 1986.

[31] J.E. Littlewood and A.C. Offord. On the number of real roots of a random algebraic equation. *J. Lond. Math. Soc.*, 13:288–295, 1938.

[32] D.J. Madden. Polynomials and primitive roots in finite fields. *J. Number Theory*, 13:499–514, 1981.

[33] B.R. McDonald. *Finite Rings with identity.* Marcel Dekker, 1st edition, 1974.

[34] R. Nöbauer. Cryptanalysis of a public-key cryptosystem based on Dickson-polynomials. *Math. Slovaca*, 38:309–323, 1988.

[35] R.W.K. Odoni. The statistics of Weil's trigonometric sums. *Proc. Camb. Phil. Soc.*, 74:467–471, 1973.

[36] R.W.K. Odoni. Zeros of random polynomials over finite fields. *Math. Proc. Camb. Phil. Soc.*, 111(2):193–197, 1992.

[37] Lord Rayleigh. On the problem of random vibrations and random flights. *Phil. Mag.*, 6:321–347, 1919.

[38] A. Rényi. Three new proofs and a generalization of a theorem by Irving Weiss. *Magyar. Tud. Akad. Mat. Kutató Int Közl.*, 7:203–214, 1962.

[39] B.A. Sevastyanov. Convergence of the distribution of the number of empty boxes to Gaussian and Poisson processes in the classical ball problem. *Theory Prob. Appl.*, 12:126–134, 1967.

[40] G.N. Watson. *A Treatise on the Theory of Bessel Functions.* Cambridge University Press, 2nd edition, 1944.

[41] A. Weil. Sur les courbes algébriques et les variétés qui s'en déduisent. Herman, Paris, 1948.

[42] I. Weiss. Limiting distributions in some occupancy problems. *Ann. Math. Statist.*, 29:878–884, 1958.