Chapter Twelve

$\{n^k\alpha\}$ and Poissonian Behavior

For $x \in \mathbb{R}$, let $\{x\}$ denote the fractional part of x and $[x]$ denote the greatest integer less than or equal to x, so $x = [x]+\{x\}$. Given a sequence z_n, let $x_n = \{z_n\}$. Thus $\{x_n\}_{n=1}^{\infty}$ is a sequence of numbers in $[0,1]$, and we can investigate its properties. We ask three questions of increasing difficulty about $x_n = \{n^k\alpha\}$ (where k and α are fixed): is the sequence dense, is it equidistributed, and what are the spacing statistics between ordered adjacent spacings. In many cases it is conjectured that the answers for these sequences should be the same as what we would observe if we chose the x_n uniformly in $[0,1]$. This is known as **Poissonian behavior**, and arises in many mathematical and physical systems. For example, numerical investigations of spacings between adjacent primes support the conjecture that the primes exhibit Poissonian behavior; see for example [Sch, Weir]; we encounter another example in Project 17.4.3. We assume the reader is familiar with probability theory and elementary Fourier analysis at the level of Chapters 8 and 11. See [Py] for a general survey on spacing results.

12.1 DEFINITIONS AND PROBLEMS

In this chapter we fix a positive integer k and an $\alpha \in \mathbb{R}$, and investigate the sequence $\{x_n\}_{n=1}^{\infty}$, where $x_n = \{n^k\alpha\}$. The first natural question (and the easiest) to ask about such sequences is whether or not the sequence gets arbitrarily close to every point:

Definition 12.1.1 (Dense)**.** *A sequence $\{x_n\}_{n=1}^{\infty}$, $x_n \in [0,1]$, is dense in $[0,1]$ if for all $x \in [0,1]$ there is a subsequence $\{x_{n_k}\}_{k=1}^{\infty}$ such that $x_{n_k} \to x$.*

Question 12.1.2. *Are the fractional parts $\{n^k\alpha\}$ dense in $[0,1]$? How does the answer depend on k and α?*

We show in §12.2 that if $\alpha \notin \mathbb{Q}$ then the fractional parts $\{n^k\alpha\}$ are dense. We prove this only for $k = 1$, and sketch the arguments for larger k. For a dense sequence, the next natural question to ask concerns how often the sequence is near a given point:

Definition 12.1.3 (Characteristic Function)**.**

$$\chi_{(a,b)}(x) \;=\; \begin{cases} 1 & \text{if } x \in [a,b] \\ 0 & \text{otherwise.} \end{cases} \tag{12.1}$$

We call $\chi_{(a,b)}$ the characteristic (or indicator) function of the interval $[a,b]$.

Definition 12.1.4 (Equidistributed)**.** *A sequence* $\{x_n\}_{n=-\infty}^{\infty}$, $x_n \in [0,1]$, *is equidistributed in* $[0,1]$ *if*

$$\lim_{N\to\infty} \frac{\#\{n : |n| \le N, x_n \in [a,b]\}}{2N+1} \;=\; \lim_{N\to\infty} \frac{\sum_{n=-N}^{N} \chi_{(a,b)}(x_n)}{2N+1} \;=\; b-a \tag{12.2}$$

for all $(a,b) \subset [0,1]$. *A similar definition holds for* $\{x_n\}_{n=0}^{\infty}$.

Question 12.1.5. *Assume the fractional parts* $\{n^k\alpha\}$ *are dense; are they equidistributed? How does the answer depend on* k *and* α*?*

We show in Theorem 12.3.2 that if $\alpha \notin \mathbb{Q}$ then the fractional parts $\{n\alpha\}$ are equidistributed. Equivalently, $n\alpha \bmod 1$ is equidistributed. For $k > 1$, $\{n^k\alpha\}$ is also equidistributed, and we sketch the proof in §12.3.2.

We have satisfactory answers to the first two questions. The last natural question is still very much open. Given a sequence of numbers $\{x_n\}_{n=1}^{N}$ in $[0,1)$, we arrange the terms in increasing order, say

$$0 \;\le\; y_1(=x_{n_1}) \;\le\; y_2(=x_{n_2}) \;\cdots\; \le\; y_N(=x_{n_N}) \;<\; 1. \tag{12.3}$$

The y_i's are called order statistics; see [DN] for more information.

Definition 12.1.6 (Wrapped unit interval)**.** *We call* $[0,1)$, *when all arithmetic operations are done mod* 1, *the wrapped unit interval. The distance between* $x, y \in \mathbb{R}$ *is given by*

$$||x-y|| \;=\; \min_{n\in\mathbb{Z}} |x-y-n|. \tag{12.4}$$

For example, $||8.45 - .41|| = .04$, and $||.999 - .001|| = .002$.

Exercise 12.1.7. *Show* $||x-y|| \le \frac{1}{2}$ *and* $||x-y|| \le |x-y|$*. Is* $||x-z|| \le ||x-y|| + ||y-z||$*?*

If the N elements $y_n \in [0,1)$ are distinct, every element has a unique element to the left and to the right, except for y_1 (no element to the left) and y_N (no element to the right). If we consider the wrapped unit interval, this technicality vanishes, as y_1 and y_N are neighbors. We have N spacings between neighbors: $||y_2 - y_1||, \dots, ||y_N - y_{N-1}||, ||y_1 - y_N||$.

Exercise 12.1.8. *Consider the spacings* $||y_2 - y_1||, \dots, ||y_N - y_{N-1}||, ||y_1 - y_N||$ *where all the* y_n*'s are distinct. Show the average spacing is* $\frac{1}{N}$.

How are the y_n spaced? How likely is it that two adjacent y_n's are very close or far apart? Since the average spacing is about $\frac{1}{N}$, it becomes very unlikely that two adjacent terms are far apart on an absolute scale. Explicitly, as $N \to \infty$, it is unlikely that two adjacent y_n's are separated by $\frac{1}{2}$ or more. This is not the natural question; the natural questions concern the normalized spacings; for example:

1. How often are two y_n's less than half their average spacing apart?

2. How often are two y_n's more than twice their average spacing apart?

3. How often do two adjacent y_n's differ by $\frac{c}{N}$? How does this depend on c?

Question 12.1.9. *Let $x_n = \{n^k\alpha\}$ for $1 \leq n \leq N$. Let $\{y_n\}_{n=1}^N$ be the x_n's arranged in increasing order. What rules govern the spacings between the y_n's? How does this depend on k and α and N?*

We answer such questions below for some choices of sequences x_n, and describe interesting conjectural results for $\{n^k\alpha\}$, $\alpha \notin \mathbb{Q}$. Of course, it is possible to study sequences other than $n^k\alpha$, and in fact much is known about the fractional parts of $g(n)\alpha$ for certain $g(n)$.

For example, let $g(n)$ be a **lacunary** sequence of integers. This means that $\liminf \frac{g(n+1)}{g(n)} > 1$, which implies that there are large gaps between adjacent values. A typical example is to take $g(n) = b^n$ for any integer $b \geq 2$. In [RZ2] it is shown that for any lacunary sequence $g(n)$ and almost all α, the fractional parts of $g(n)\alpha$ are equidistributed and exhibit Poissonian behavior. We refer the reader to [RZ1, RZ2] for complete details; we have chosen to concentrate on $n^k\alpha$ for historical reasons as well as the number of open problems.

Exercise 12.1.10. *Show $g(n) = n!$ is lacunary but $g(n) = n^k$ is not. Is $g(n) = \binom{2n}{1}$ or $g(n) = \binom{2n}{n}$ lacunary?*

Exercise 12.1.11. *For any integer b, show α is normal base b if and only if the fractional parts $b^n\alpha$ are equidistributed (see §10.5.3 for a review of normal numbers). As it can be shown that for almost all α the fractional parts $b^n\alpha$ are equidistributed, this implies almost all numbers are normal base b.*

12.2 DENSENESS OF $\{n^k\alpha\}$

We tackle Question 12.1.2, namely, when is $\{n^k\alpha\}$ dense in $[0,1]$?

Exercise 12.2.1. *For $\alpha \in \mathbb{Q}$, prove $\{n^k\alpha\}$ is never dense.*

Theorem 12.2.2 (Kronecker)**.** *For $k = 1$ and $\alpha \notin \mathbb{Q}$, $x_n = \{n\alpha\}$ is dense in $[0,1]$.*

The idea of the proof is as follows: by Dirichlet's Pigeon-Hole Principle, we can find a multiple of α that is "close" to zero. By taking sufficiently many copies of this multiple, we can move near any x.

Proof. We must show that, for any $x \in [0,1]$, there is a subsequence $x_{n_j} \to x$. It suffices to show that for each $\epsilon > 0$ we can find an x_{n_j} such that $||x_{n_j} - x|| < \epsilon$ (remember, .001 is close to .999). By Dirichlet's Pigeon-Hole Principle (see Theorem 5.5.4 or 7.9.4), as α is irrational there are *infinitely many relatively prime* p *and* q such that $|\alpha - \frac{p}{q}| < \frac{1}{q^2}$. Thus, $||q\alpha|| \leq |q\alpha - p| < \frac{1}{q}$. Choose q large enough so that $\frac{1}{q} < \epsilon$.

Either $0 < q\alpha - p < \frac{1}{q}$ or $-\frac{1}{q} < q\alpha - p < 0$. As the two cases are handled similarly, we only consider the first. Given $x \in [0,1]$, choose j so that $j(q\alpha - p)$ is within $\frac{1}{q}$ of x. This is always possible. Each time j increases by one, we increase

$j(q\alpha - p)$ by a fixed amount independent of j and at most $\frac{1}{q}$. As $||j(q\alpha - p)|| = ||jq\alpha||$, we have shown that $x_{n_j} = \{jq\alpha\}$ is within ϵ of x. □

Exercise 12.2.3. *In the above argument, show that* $q\alpha - p \neq 0$*.*

Exercise 12.2.4. *Handle the second case, and show the above argument does generate a sequence* $x_{n_j} \to x$ *for any* $x \in [0,1]$*.*

Remark 12.2.5 (Kronecker's Theorem)**.** Kronecker's Theorem is more general than what is stated above. Let $\alpha_1, \alpha_2, \dots, \alpha_k \in \mathbb{R}$ have the property that if

$$c_0 1 + c_1\alpha_1 + \cdots + c_n\alpha_k \;=\; 0 \tag{12.5}$$

with all $c_i \in \mathbb{Q}$, then $c_i = 0$ for all i. The standard terminology is to say the α_i are linearly independent (see Definition B.1.7) over $\mathbb{Q}$. Then Kronecker's theorem asserts that the sequence of points $(v_n)_{n\in\mathbb{N}}$ given by

$$v_n = (\{n\alpha_1\}, \dots, \{n\alpha_k\}) \tag{12.6}$$

is dense in $[0,1]^k$; we proved the case $k = 1$.

Exercise 12.2.6. *Show it is not enough to assume just* $c_1\alpha_1 + \cdots + c_k\alpha_k = 0$ *implies all the* $c_i = 0$*.*

Exercise 12.2.7. *Prove Kronecker's Theorem in full generality (a proof is given in [HW], Chapter XXIII).*

Theorem 12.2.8. *For* k *a positive integer greater than 1 and* $\alpha \notin \mathbb{Q}$*,* $x_n = \{n^k\alpha\}$ *is dense in* $[0,1]$*.*

If we can show $\{n^k\alpha\}$ is equidistributed, then the above theorem is an immediate corollary. As we sketch the proof of equidistribution later (see §12.3), we confine ourselves to proving the $k = 2$ case for very special α, which highlight how the algebraic structure of α *can* enter proofs as well as the different notions of sizes of sets. There are other methods to prove this theorem which can handle larger classes of α; we have chosen the method below because of the techniques it introduces (especially in showing how the approximation exponent can enter).

Proof. Assume α has approximation exponent $4 + \eta$ for *some* $\eta > 0$ (see §5.5). Then there are infinitely many solutions to $|\alpha - \frac{p}{q}| < \frac{1}{q^4}$ with p, q relatively prime. We must show that, given x and $\epsilon > 0$, there is an x_{n_j} within ϵ of x. Choose p, q so large that $\frac{1}{q} < \frac{\epsilon}{100}$ and $|\alpha - \frac{p}{q}| < \frac{1}{q^4}$. Thus there exists a δ such that

$$\alpha - \frac{p}{q} \;=\; \frac{\delta}{q^4}. \tag{12.7}$$

We assume $\delta > 0$ (the other case is handled similarly), and clearly $\delta < 1$. We have

$$q^2m^2\alpha - pqm^2 \;=\; (\delta m^2)\cdot\frac{1}{q^2}. \tag{12.8}$$

We claim we may choose m so that $\delta m^2 \in [1,4]$. If $\delta m_0^2 < 1$ then $m_0 < \frac{1}{\sqrt{\delta}}$ and

$$\delta(m_0+1)^2 - \delta m_0^2 \;=\; \delta(2m_0+1) \;<\; \frac{2\delta}{\sqrt{\delta}} + \delta \;<\; 3. \tag{12.9}$$

As we move at most 3 units each time we increment m_0 by 1, we can find an m such that $\delta m^2 \in [1,4]$. Later we will see why it is necessary to have upper bound for δm^2.

Given $x \in [0,1]$, as $\delta m^2 \geq 1$ we have $\frac{x}{\delta m^2} \in [0,1]$. We can find n such that $\frac{n}{q}$ is within $\frac{1}{q}$ of $\sqrt{\frac{x}{\delta m^2}}$ and $\frac{n}{q} < \sqrt{\frac{x}{\delta m^2}}$. Further we may take $\frac{n}{q} \in [0,1]$ so $n \leq q$. For some $\theta \in [0,1]$ we have

$$\frac{n}{q} + \frac{\theta}{q} \;=\; \sqrt{\frac{x}{\delta m^2}}, \qquad \frac{n^2}{q^2} + \frac{2n\theta + \theta^2}{q^2} \;=\; \frac{x}{\delta m^2}. \tag{12.10}$$

The point is that $\delta m^2 \frac{n^2}{q^2}$ is very close to x. We want to write x as almost a square times α (minus an integer), as we are working with numbers of the form $\{b^2\alpha\}$. Multiplying (12.8) by n^2 yields

$$\begin{aligned} q^2m^2n^2\alpha - pqm^2n^2 \;&=\; (\delta m^2)\cdot\frac{n^2}{q^2} \\ &=\; x - \frac{(\delta m^2)\cdot(2n\theta+\theta^2)}{q^2}. \end{aligned} \tag{12.11}$$

As $\delta m^2 \leq 4$, $n \leq q$, $\theta \leq 1 \leq q$ and $\frac{1}{q} < \frac{\epsilon}{100}$,

$$\frac{(\delta m^2)\cdot(2n\theta+\theta^2)}{q^2} \;<\; \frac{12}{q} \;<\epsilon, \tag{12.12}$$

which shows that $(qmn)^2\alpha$ minus an integer is within ϵ of x. □

Remark 12.2.9. It was essential that δm^2 was bounded from above, as the final error used $\delta m^2 \leq 4$. Of course, we could replace 4 with any finite number.

Remark 12.2.10. We do not need α to have order of approximation $4+\eta$; all we need is that there are infinitely many solutions to $|\alpha - \frac{p}{q}| < \frac{1}{q^4}$.

Remark 12.2.11. Similar to our investigations on the Cantor set (§A.5.2 and Remark A.5.10), we have a set of numbers that is “small” in the sense of measure, but “large” in the sense of cardinality. By Theorem 5.5.9 the measure of $\alpha \in [0,1]$ with infinitely many solutions to $|\alpha - \frac{p}{q}| < \frac{1}{q^4}$ is zero; on the other hand, every Liouville number (see §5.6.2) has infinitely many solutions to this equation, and in Exercise 5.6.5 we showed there are uncountably many Liouville numbers.

Exercise 12.2.12. *Prove that if α has order of approximation $4+\eta$ for some $\eta > 0$, then there are infinitely many solutions to $|\alpha - \frac{p}{q}| < \frac{1}{q^4}$ with p, q relatively prime.*

Exercise 12.2.13. *To show $\{n^k\alpha\}$ is dense (using the same arguments), what must we assume about the order of approximation of α?*

Remark 12.2.14. We can weaken the assumptions on the approximability of α by rationals and the argument follows similarly. Let $f(q)$ be any monotone increasing function tending to infinity. If $|\alpha - \frac{p}{q}| < \frac{1}{q^2 f(q)}$ has infinitely many solutions, then $\{n^2\alpha\}$ is dense. We argue as before to find m so that $\delta m^2 \in [1,4]$, and then find n so that $\frac{n^2}{f(q)}$ is close to $\frac{x}{\delta m^2}$. Does the set of such α still have measure zero (see Theorem 5.5.9)? Does the answer depend on f?

Exercise 12.2.15. *Prove the claim in Remark 12.2.14.*

Exercise 12.2.16. *Let $x_n = \sin n$; is this sequence dense in $[-1, 1]$? We are of course measuring angles in radians and not degrees.*

Exercise 12.2.17. *In Exercise 5.5.6 we showed $\sum_{n=1}^{\infty}(\cos n)^n$ diverge by showing $|(\cos n)^n|$ is close to 1 infinitely often. Is $x_n = (\cos n)^n$ dense in $[-1, 1]$?*

Exercise 12.2.18. *Let α be an irrational number, and as usual set $x_n = \{n\alpha\}$. Is the sequence $\{x_{n^2}\}_{n\in\mathbb{N}}$ dense? How about $\{x_{p_n}\}$? Here $\{p_1, p_2, p_3, \dots\}$ is the sequence of prime numbers.*

12.3 EQUIDISTRIBUTION OF $\{n^k\alpha\}$

We now turn to Question 12.1.5. We prove that $\{n^k\alpha\}$ is equidistributed for $k = 1$, and sketch the proof for $k > 1$. We will use the following functions in our proof:

Definition 12.3.1 ($e(x)$, $e_m(x)$)**.** *We set*

$$e(x) \;=\; e^{2\pi i x}, \qquad e_m(x) \;=\; e^{2\pi i m x}. \tag{12.13}$$

Theorem 12.3.2 (Weyl)**.** *Let α be an irrational number in $[0, 1]$, and let k be a fixed positive integer. Let $x_n = \{n^k\alpha\}$. Then $\{x_n\}_{n=1}^{\infty}$ is equidistributed.*

Contrast the above with

Exercise 12.3.3. *Let $\alpha \in \mathbb{Q}$. If $x_n = \{n^k\alpha\}$ prove $\{x_n\}_{n=1}^{\infty}$ is not equidistributed.*

We give complete details for the $k = 1$ case, and provide a sketch in the next subsection for general k. According to [HW], note on Chapter XXIII, the theorem for $k = 1$ was discovered independently by Bohl, Sierpiński and Weyl at about the same time. We follow Weyl's proof which, according to [HW], is undoubtedly "the best proof" of the theorem. There are other proofs, however; one can be found in [HW], Chapter XXIII. A more modern, useful reference is [KN], Chapter 1.

The following argument is very common in analysis. Often we want to prove a limit involving characteristic functions of intervals (and such functions are not continuous); here the characteristic function is the indicator function for the interval $[a, b]$. We first show that to each such characteristic function there is a continuous function "close" to it, reducing the original problem to a related one for continuous functions. We then show that, given any continuous function, there is a trigonometric polynomial that is "close" to the continuous function, reducing the problem to one involving trigonometric polynomials (see §11.3.4).

Why is it advantageous to recast the problem in terms of trigonometric polynomials rather than characteristic functions of intervals? Since $x_n = \{n\alpha\} = n\alpha - [n\alpha]$ and $e_m(x) = e_m(x + h)$ for every integer h,

$$e_m(x_n) \;=\; e^{2\pi i m\{n\alpha\}} \;=\; e^{2\pi i m n\alpha}. \tag{12.14}$$

The complications of looking at $n\alpha \bmod 1$ vanish, and it suffices to evaluate exponential functions at the simpler sequence $n\alpha$.

Remark 12.3.4. Here our "data" is discrete, namely we have a sequence of points $\{a_n\}$. See [MN] for another technique which is useful in investigating equidistribution of continuously distributed data (where one studies values $f(t)$ for $t \leq T$, which corresponds to taking terms a_n with $n \leq N$).

12.3.1 Weyl's Theorem when $k = 1$

Theorem 12.3.5 (Weyl Theorem, $k = 1$)**.** *Let α be an irrational number in $[0, 1]$. Let $x_n = \{n\alpha\}$. Then $\{x_n\}_{n=1}^{\infty}$ is equidistributed.*

Proof. For $\chi_{a,b}$ as in Definition 12.1.3 we must show

$$\lim_{N\to\infty} \frac{1}{2N+1} \sum_{n=-N}^{N} \chi_{(a,b)}(x_n) \ = \ b - a. \tag{12.15}$$

We have

$$\begin{aligned} \frac{1}{2N+1} \sum_{n=-N}^{N} e_m(x_n) \ &= \ \frac{1}{2N+1} \sum_{n=-N}^{N} e_m(n\alpha) \\ &= \ \frac{1}{2N+1} \sum_{n=-N}^{N} (e^{2\pi i m\alpha})^n \\ &= \ \begin{cases} 1 & \text{if } m = 0 \\ \frac{1}{2N+1} \frac{e_m(-N\alpha) - e_m((N+1)\alpha)}{1 - e_m(\alpha)} & \text{if } m > 0, \end{cases} \end{aligned} \tag{12.16}$$

where the last follows from the geometric series formula. For a fixed irrational α, $|1 - e_m(\alpha)| > 0$; *this is where we use $\alpha \notin \mathbb{Q}$*. Therefore if $m \neq 0$,

$$\lim_{N\to\infty} \frac{1}{2N+1} \frac{e_m(-N\alpha) - e_m((N+1)\alpha)}{1 - e_m(\alpha)} \ = \ 0. \tag{12.17}$$

Let $P(x) = \sum_m a_m e_m(x)$ be a finite sum (i.e., $P(x)$ is a trigonometric polynomial). By possibly adding some zero coefficients, we can write $P(x)$ as a sum over a symmetric range: $P(x) = \sum_{m=-M}^{M} a_m e_m(x)$.

Exercise 12.3.6. *Show* $\int_0^1 P(x)dx = a_0$.

We have shown that for any finite trigonometric polynomial $P(x)$:

$$\lim_{N\to\infty} \frac{1}{2N+1} \sum_{n=-N}^{N} P(x_n) \ \longrightarrow \ a_0 \ = \ \int_0^1 P(x)dx. \tag{12.18}$$

Consider two continuous approximations to the characteristic function $\chi_{(a,b)}$:

1. A_{1j}: $A_{1j}(x) = 1$ if $a + \frac{1}{j} \leq x \leq b - \frac{1}{j}$, drops linearly to 0 at a and b, and is zero elsewhere (see Figure 12.1).

2. A_{2j}: $A_{1j}(x) = 1$ if $a \leq x \leq b$, drops linearly to 0 at $a - \frac{1}{j}$ and $b + \frac{1}{j}$, and is zero elsewhere.

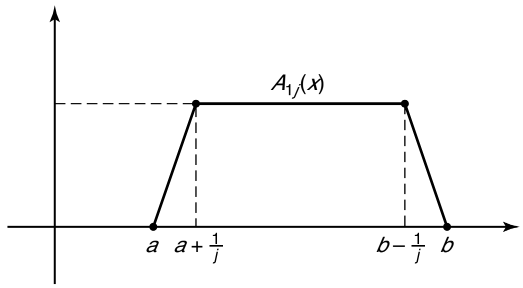

Figure 12.1 Plot of $A_{1j}(x)$

Note there are trivial modifications if $a = 0$ or $b = 1$. Clearly

$$A_{1j}(x) \ \leq\ \chi_{(a,b)}(x) \ \leq\ A_{2j}(x). \tag{12.19}$$

Therefore

$$\frac{1}{2N+1}\sum_{n=-N}^{N} A_{1j}(x_n) \ \leq\ \frac{1}{2N+1}\sum_{n=-N}^{N} \chi_{(a,b)}(x_n) \ \leq\ \frac{1}{2N+1}\sum_{n=-N}^{N} A_{2j}(x_n). \tag{12.20}$$

By Theorem 11.3.1, for each j, given $\epsilon > 0$ we can find symmetric trigonometric polynomials $P_{1j}(x)$ and $P_{2j}(x)$ such that $|P_{1j}(x) - A_{1j}(x)| < \epsilon$ and $|P_{2j}(x) - A_{2j}(x)| < \epsilon$. As A_{1j} and A_{2j} are continuous functions, we can replace

$$\frac{1}{2N+1}\sum_{n=-N}^{N} A_{ij}(x_n) \quad \text{with} \quad \frac{1}{2N+1}\sum_{n=-N}^{N} P_{ij}(x_n) \tag{12.21}$$

at a cost of at most ϵ. As $N \to \infty$,

$$\frac{1}{2N+1}\sum_{n=-N}^{N} P_{ij}(x_n) \ \longrightarrow\ \int_0^1 P_{ij}(x)dx. \tag{12.22}$$

But $\int_0^1 P_{1j}(x)dx = (b-a) - \frac{1}{j}$ and $\int_0^1 P_{2j}(x)dx = (b-a) + \frac{1}{j}$. Therefore, given j and ϵ, we can choose N large enough so that

$$(b-a) - \frac{1}{j} - \epsilon \ \leq\ \frac{1}{2N+1}\sum_{n=-N}^{N} \chi_{(a,b)}(x_n) \ \leq\ (b-a) + \frac{1}{j} + \epsilon. \tag{12.23}$$

Letting j tend to ∞ and ϵ tend to 0, we see $\frac{1}{2N+1}\sum_{n=-N}^{N} \chi_{(a,b)}(x_n) \to b-a$, completing the proof. □

Exercise 12.3.7. *Rigorously do the necessary book-keeping to prove the previous theorem.*

Exercise 12.3.8. *Prove for k a positive integer, if $\alpha \in \mathbb{Q}$ then $\{n^k\alpha\}$ is periodic while if $\alpha \notin \mathbb{Q}$ then no two $\{n^k\alpha\}$ are equal.*

12.3.2 Weyl's Theorem for $k > 1$

We sketch the proof of the equidistribution of $\{n^k\alpha\}$. Recall that by Theorem 12.3.17 a sequence $\{\xi_n\}_{n=1}^{\infty}$ is equidistributed mod 1 if and only if for every integer $m \neq 0$,

$$\lim_{N\to\infty} \frac{1}{2N+1} \sum_{n=-N}^{N} e_m(z_n) \ = \ 0. \tag{12.24}$$

We follow the presentation in [Ca]. We need the following technical lemma:

Lemma 12.3.9. *Let $u_1, u_2, \dots, u_N \in \mathbb{C}$ and let $1 \leq H \leq N$. Then*

$$H^2 \left| \sum_{1\leq n\leq N} u_n \right|^2 \ \leq \ H(H+N-1) \sum_{1\leq n\leq N} |u_n|^2 + 2(H+N-1) \sum_{0<h<H} (H-h) \sum_{1\leq n\leq N-h} \overline{u}_n u_{n+h}. \tag{12.25}$$

For a proof, see [Ca], page 71. From the above lemma we can conclude

Corollary 12.3.10. *For a sequence $\{z_n\}_{n=1}^{\infty}$, suppose that for each $h > 0$ and integer $m \neq 0$*

$$\lim_{N\to\infty} \frac{1}{2N+1} \sum_{n=-N}^{N} e_m(z_{n+h} - z_n) \ = \ 0. \tag{12.26}$$

Then

$$\lim_{N\to\infty} \frac{1}{2N+1} \sum_{n=-N}^{N} e_m(z_n) \ = \ 0. \tag{12.27}$$

In particular if $\{z_{n+h} - z_n\}_{n=1}^{\infty}$ is equidistributed for all $h \in \mathbb{N}$ then $\{z_n\}_{n=1}^{\infty}$ is equidistributed.

Exercise 12.3.11. *Prove the corollary follows from the lemma.*

Exercise^(h) 12.3.12. *Prove $\{n\alpha\}$ is equidistributed using the results of this subsection.*

Exercise^(h) 12.3.13. *Prove $\{n^2\alpha\}$ is equidistributed.*

Exercise^(h) 12.3.14. *Prove $\{n^k\alpha\}$ is equidistributed for $k \geq 1$.*

Exercise 12.3.15. *Let $\theta = 0.123456789101112\dots$ be the transcendental number from Exercise 10.5.6. By Theorem 12.3.5, $\{n\theta\}$ is equidistributed as $\theta \notin \mathbb{Q}$. Show directly that $\{10^n\theta\}$ is also equidistributed.*

Exercise 12.3.16. *Show that the sequence $\{n!e\}$ is not equidistributed. In fact, the only limit point of this sequence is 0.*

12.3.3 Weyl's Criterion

We generalize the methods of §12.3.1 to provide a useful test for equidistribution.

Theorem 12.3.17 (Weyl's Criterion)**.** *A sequence $\{\xi_n\}$ is equidistributed (modulo 1) in $[0, 1)$ if and only if*

$$\forall h \in \mathbb{Z} - \{0\}, \quad \lim_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} e_h(\xi_n) \ = \ 0. \tag{12.28}$$

Proof. We sketch the proof of the sufficiency of (12.28); for necessity see Exercise 12.3.21. Suppose that we are given that (12.28) is true, and we wish to prove that $\{\xi_n\}$, considered modulo 1, is equidistributed in $[0, 1)$.

Claim 12.3.18. *Let f be any continuous periodic function with period 1. If* (12.28) *holds, then*

$$\lim_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} f(\xi_n) \ = \ \int_0^1 f(x)dx. \tag{12.29}$$

Proof. First note that (12.29) holds for $f(x) = e(hx)$ for all $h \in \mathbb{Z}$ (for $h = 0$ trivially and for $h \neq 0$ by assumption). Thus (12.29) holds for all trigonometric polynomials. Now suppose f is any continuous periodic function with period 1. As in §12.3.1, fix an $\epsilon > 0$ and choose a trigonometric polynomial P satisfying

$$\sup_{x\in\mathbb{R}} |f(x) - P(x)| \ < \ \epsilon. \tag{12.30}$$

Choose N large enough so that

$$\left| \frac{1}{N}\sum_{n=1}^{N} P(\xi_n) - \int_0^1 P(x)dx \right| \ < \ \epsilon. \tag{12.31}$$

By the triangle inequality (Exercise 12.3.19) we have

$$\begin{aligned} \left| \frac{1}{N}\sum_{n=1}^{N} f(\xi_n) - \int_0^1 f(x)dx \right| \ \le \ & \frac{1}{N}\sum_{n=1}^{N} |f(\xi_n) - P(\xi_n)| \\ & + \left| \frac{1}{N}\sum_{n=1}^{N} P(\xi_n) - \int_0^1 P(x)dx \right| \\ & + \int_0^1 |P(x) - f(x)|dx \ < \ 3\epsilon, \end{aligned} \tag{12.32}$$

which proves the claim. □

The above argument uses several standard techniques: we *approximate our function with a trigonometric polynomial*, *add zero* and then do a *three epsilon proof*.

The rest of the proof of Theorem 12.3.17 mirrors the argument in §12.3.1. We choose functions $A_{1j}(x)$ and $A_{2k}(x)$ as before. We leave the rest of the proof to the reader. For a proof of the necessity of (12.28), see Exercise 12.3.21. □

Exercise 12.3.19 (Triangle Inequality)**.** *Prove for $a, b \in \mathbb{C}$ that $|a + b| \le |a| + |b|$.*

Exercise 12.3.20. *Complete the proof of Theorem 12.3.17.*

Exercise 12.3.21. *Prove the following statements:*

1. *If* $0 \le x - y < \frac{\pi}{2}$, *then* $|e^{ix} - e^{iy}| \le |x - y|$. Hint: *If* $0 < x < \frac{\pi}{2}$, *then* $\sin x \le x$.

2. *Show that if the sequence* $\{\xi_n\}$ *is equidistributed then* (12.28) *holds for all* h. Hint: *One can proceed as follows:*

   (a) *Define* $I_j = [\frac{j}{L}, \frac{j+1}{L})$ *for each* $j \in \{0, 1, \dots, L-1\}$ *and let* $A_{j,N} = \{n : 1 \le n \le N, \xi_n \in I_j\}$. *Show that there exists* N_0 *such that* $\forall N \ge N_0$,
   $$\frac{1}{L} - \frac{1}{L^2} \le \frac{|A_{j,N}|}{N} \le \frac{1}{L} + \frac{1}{L^2} \qquad \forall j \in \{0, 1, \dots, L-1\}.$$

   (b) *Using the triangle inequality and part (a) of this exercise show that*
   $$\left| \frac{1}{N} \sum_{n=1}^{N} e_h(\xi_n) \right| \le \frac{2\pi h}{L} + \left| \frac{1}{N} \sum_{j=0}^{L-1} \sum_{n \in A_{j,N}} e_h\left(\frac{j}{L}\right) \right|. \tag{12.33}$$

   (c) *Show that for* $N \ge N_0$ *we have*
   $$\left| \frac{1}{N} \sum_{j=0}^{L-1} \sum_{n \in A_{j,N}} e_h\left(\frac{j}{L}\right) \right| \le \frac{1}{L}. \tag{12.34}$$

   (d) *Use equation* (12.33) *and equation* (12.34) *to prove the desired statement.*

Exercise(h) 12.3.22. *Is the sequence* $\{\log_b n\}_{n=1}^{\infty}$ *equidistributed for* $b = e$ *or* $b = 10$*?*

12.4 SPACING PRELIMINARIES

For many α and k we have shown $\{n^k\alpha\}$ is equidistributed in $[0, 1)$. We now ask finer questions about the sequences. Consider the uniform distribution on $[0, 1)$. If we were to choose N points from this distribution, we expect for N large that the points should look equidistributed. This *suggests* the possibility that these two processes (looking at the fractional parts $\{n^k\alpha\}$ and choosing N points uniformly in $[0, 1)$) could share other behavior as well. We see this is the case for some triples (N, k, α) and violently false for others. The answer often depends on the structure of α and its approximations by continued fractions.

We call the behavior of points chosen uniformly in $[0, 1)$ **Poissonian behavior**, and conjecture that often $\{x_n\}_{n=1}^{N}$ ($x_n = \{n^k\alpha\}$) exhibits Poissonian behavior. One must be very careful, of course, in making such conjectures. The Prime Number Theorem implies that for primes of size x, the average spacing between primes is like $\frac{1}{\log x}$. One natural model for the distribution of primes (the **Cramér model**) is that a number near x is prime with probability $\frac{1}{\log x}$, and each number's "primality" is determined independent of its neighbors. For many statistics (for example the number of primes, number of twin primes, spacings between primes; see for example [Sch, Weir]) this gives answers that are reasonably fit by the actual data; however, in [MS] a statistic was investigated where the actual data disagrees with the Cramér model. It agrees with another model, Random Matrix Theory; see Chapter 15 for an introduction to Random Matrix Theory.

12.5 POINT MASSES AND INDUCED PROBABILITY MEASURES

In §11.2.2 we introduced the Dirac delta functional $\delta(x-a)$, which we interpreted as a unit point mass at a. Recall $\int f(x)\delta(x-a)dx = f(a)$. Given N point masses located at $x_1, x_2, \dots, x_N$, we can form a probability measure

$$\mu_N(x)dx \ = \ \frac{1}{N}\sum_{n=1}^{N}\delta(x-x_n)dx. \tag{12.35}$$

We say probability measure and not probability density as we do not have a nice, continuous density – our density "function" involves the Dirac functional. Note $\int \mu_N(x)dx = 1$, and if $f(x)$ is continuous

$$\int f(x)\mu_N(x)dx \ = \ \frac{1}{N}\sum_{n=1}^{N} f(x_n). \tag{12.36}$$

Exercise 12.5.1. *Prove* (12.36) *for continuous* $f(x)$.

Note the right hand side of (12.36) looks like a Riemann sum. If the x_n's are equidistributed, we expect this to converge to $\int f(x)dx$. In general the x_n's might not be equidistributed; for example, they may be drawn from a fixed probability distribution $p(x)$. We *sketch* what should happen.

For simplicity, assume $x_n \in [0,1]$ and assume for any interval $[a,b] \subset [0,1]$, as $N \to \infty$ the fraction of x_n's ($1 \le n \le N$) in $[a,b]$ tends to $\int_a^b p(x)dx$ for some continuous function $p(x)$:

$$\lim_{N\to\infty} \frac{\#\{n : 1 \le n \le N \text{ and } x_n \in [a,b]\}}{N} \ = \ \int_a^b p(x)dx. \tag{12.37}$$

Assume f', p' are bounded. We want to compare $\int f(x)p(x)dx$ with $\int f(x)\mu_N(x)dx$. Then

$$\begin{aligned}
\int_0^1 f(x)\mu_N(x)dx \ &= \ \frac{1}{N}\sum_{n=1}^{N} f(x_n) \\
&\approx \ \sum_{k=0}^{M-1} f\left(\frac{k}{M}\right) \frac{\#\{n : 1 \le n \le N \text{ and } x_n \in \left[\frac{k}{M}, \frac{k+1}{M}\right]\}}{N} \\
&\approx \ \sum_{k=0}^{M-1} f\left(\frac{k}{M}\right) \int_{\frac{k}{M}}^{\frac{k+1}{M}} p(x)dx \\
&\approx \ \int_0^1 f(x)p(x)dx.
\end{aligned} \tag{12.38}$$

Exercise 12.5.2 (Monte Carlo Method)**.** *Make the above argument rigorous when f', p' are bounded on $[0, 1]$. Explicitly, show given an $\epsilon > 0$ there exists an N_0 such that for all $N > N_0$, $|\int f(x)\mu_N(x)dx - \int f(x)p(x)dx| < \epsilon$. This gives a numerical method to evaluate integrals, known as the **Monte Carlo Method**. See [Met, MU] for more details and references.*

Monte Carlo methods are extremely popular and useful for determining multidimensional integrals with nice boundary conditions. For example, say we wish to determine the n-dimensional volume of a set $S \subset \mathbb{R}^n$ that is contained in some finite n-dimensional box B. If for each point $x \in B$ we can easily determine whether or not $x \in S$, we generate points uniformly in B and see what fraction lie in S. For "nice" regions S the percent of points in B that are also in S will converge to $\frac{\text{vol}(S)}{\text{vol}(B)}$. Monte Carlo integration is one of the most frequently used methods to numerically approximate such solutions.

Definition 12.5.3 (Convergence to $p(x)$)**.** *If the sequence of points x_n satisfies* (12.37) *for some nice function $p(x)$, we say the probability measures $\mu_N(x)dx$ converge to $p(x)dx$.*

12.6 NEIGHBOR SPACINGS

Let $\{\alpha_n\}_{n=1}^N$ be a collection of points in $[0, 1)$. We arrange them in increasing order:

$$0 \le \alpha_{n_1} \le \alpha_{n_2} \le \cdots \le \alpha_{n_N}. \tag{12.39}$$

For notational convenience, let $\beta_j = \alpha_{n_j}$. We investigate how the differences $\beta_{j+1} - \beta_j$ are distributed. Remember as we are working on the wrapped unit interval, the distance is $||\beta_{j+1} - \beta_j||$ (see §12.1). In looking at spacings between the β_j's, we have $N - 1$ pairs of neighbors:

$$(\beta_2, \beta_1),\ (\beta_3, \beta_2),\ \ldots,\ (\beta_N, \beta_{N-1}). \tag{12.40}$$

These pairs give rise to spacings $\beta_{j+1} - \beta_j \in [0, 1)$. We can also consider the pair (β_1, β_N). This gives rise to the spacing $\beta_1 - \beta_N \in [-1, 0)$; however, as we are studying this sequence mod 1, this is equivalent to $\beta_1 - \beta_N + 1 \in [0, 1)$.

Definition 12.6.1 (Neighbor Spacings)**.** *Given a sequence of numbers α_n in $[0, 1)$, fix an N and arrange the numbers α_n ($n \le N$) in increasing order. Label the new sequence β_j; note the ordering will depend on N. It is convenient to periodically extend our sequence, letting $\beta_{N+j} = \beta_j$ for all j.*

1. *The nearest neighbor spacings are the numbers $||\beta_{j+1} - \beta_j||$, $j = 1$ to N.*
2. *The m^{th} neighbor spacings are the numbers $||\beta_{j+m} - \beta_j||$, $j = 1$ to N.*

Exercise 12.6.2 (Surprising)**.** *Let $\alpha = \sqrt{2}$, and let $\alpha_n = \{n\alpha\}$ or $\{n^2\alpha\}$. Calculate the nearest neighbor and the next nearest neighbor spacings in each case for $N = 10$ and $N = 20$. Note the different behavior for $n\alpha$ and $n^2\alpha$. See also Exercise 12.6.3.*

Exercise 12.6.3. *Prove that if $\alpha \notin \mathbb{Q}$ then for each N there are at most three different values for the nearest neighbor spacings. Bound the number of possible values for the m^{th} neighbor spacings. Is the bound sharp? For more on such spacings, see [Bl1, Bl2].*

Remark 12.6.4. While we have concentrated on the distribution of the differences of independent uniformly distributed random variables, there are many other interesting questions we could ask. For example, how are the leading digits of these differences distributed? See [MN] for an interesting application of Poisson summation to this problem, which shows the distribution of the leading digits of these differences almost obey Benford's Law.

12.7 POISSONIAN BEHAVIOR

Before investigating Question 12.1.9 (concerning the spacings of $\{n^k\alpha\}$), we consider a simpler case. Fix N and consider N independent random variables x_n. Each random variable is chosen from the uniform distribution on $[0, 1)$; thus the probability that $x_n \in [a, b]$ is $b - a$. Let $\{y_n\}_{n=1}^N$ be the x_n's arranged in increasing order. How do the neighbor spacings behave?

We first need to decide what is the correct scale to use for our investigations. As we have N objects on the wrapped unit interval, we have N nearest neighbor spacings and we expect the average spacing to be $\frac{1}{N}$.

Definition 12.7.1 (Unfolding)**.** *Let $z_n = Ny_n$. The numbers $z_n = Ny_n$ have unit mean spacing. While we expect the average spacing between adjacent y_n's to be $\frac{1}{N}$ units for typical sequences, we expect the average spacing between adjacent z_n's to be 1 unit.*

The probability of observing a spacing as large as $\frac{1}{2}$ between adjacent y_n's becomes negligible as $N \to \infty$. What we should ask is what is the probability of observing a nearest neighbor spacing of adjacent y_n's that is *half* the average spacing. In terms of the z_n's, this corresponds to a spacing between adjacent z_n's of $\frac{1}{2}$ a unit.

For the rest of the section, x_n, y_n and z_n are as defined above.

12.7.1 Nearest Neighbor Spacings

Let the x_n's in increasing order be labeled $y_1 \leq y_2 \leq \cdots \leq y_N$, $y_n = x_{n_j}$. As the average distance between adjacent y_n's is $\frac{1}{N}$, it is natural to look at nearest neighbor spacings of size $\frac{t}{N}$. In particular, we study the probability of observing a nearest neighbor spacing between $\frac{t}{N}$ and $\frac{t+\Delta t}{N}$. By symmetry, on the wrapped unit interval the expected nearest neighbor spacing is independent of j. Explicitly, we expect $y_{l+1} - y_l$ to have the same distribution as $y_{i+1} - y_i$. *Remember, the y_n's are ordered and the x_n's are unordered.* We assume familiarity with elementary properties of binomial coefficients (see §1.2.4 and §A.1.3).

We give a simple argument first which highlights the ideas, although it is slightly wrong (though only in lower order terms). As the x_n's are chosen independently, there are $\binom{N-1}{1}$ choices of subscript n such that x_n is the *first neighbor to the right* of x_1. This can also be seen by symmetry, as each x_n is equally likely to be the first to the *right* of x_1 (where, of course, $.001$ is just a little to the right of $.999$), and we have $N-1$ choices left for x_n. As x_n was chosen uniformly in $[0,1)$, the probability that $x_n \in \left[\frac{t}{N}, \frac{t+\Delta t}{N}\right]$ is $\frac{\Delta t}{N}$.

For the remaining $N-2$ of the x_n's, each must be further than $\frac{t+\Delta t}{N}$ to the right x_n. They must *all* lie in an interval (or possibly two intervals if we wrap around) of length $1 - \frac{t+\Delta t}{N}$. The probability that they all lie in this region is $\left(1 - \frac{t+\Delta t}{N}\right)^{N-2}$.

Thus, if $x_1 = y_l$, we want to calculate the probability that $||y_{l+1} - y_l|| \in \left[\frac{t}{N}, \frac{t+\Delta t}{N}\right]$. This is

$$\begin{aligned}\text{Prob}\left(||y_{l+1} - y_l|| \in \left[\frac{t}{N}, \frac{t+\Delta t}{N}\right]\right) &= \binom{N-1}{1} \cdot \frac{\Delta t}{N} \cdot \left(1 - \frac{t+\Delta t}{N}\right)^{N-2} \\ &= \left(1 - \frac{1}{N}\right) \cdot \left(1 - \frac{t+\Delta t}{N}\right)^{N-2} \Delta t.\end{aligned} \tag{12.41}$$

For N enormous and Δt small,

$$\begin{aligned}\left(1 - \frac{1}{N}\right) &\approx 1 \\ \left(1 - \frac{t+\Delta t}{N}\right)^{N-2} &\approx e^{-(t+\Delta t)} \approx e^{-t}(1 + O(\Delta t));\end{aligned} \tag{12.42}$$

See §5.4 for a review of the needed properties of e. Thus

$$\text{Prob}\left(||y_{l+1} - y_l|| \in \left[\frac{t}{N}, \frac{t+\Delta t}{N}\right]\right) \xrightarrow[N \to \infty]{} e^{-t}\Delta t. \tag{12.43}$$

In terms of the z_n's, which have mean spacing 1, this yields

$$\text{Prob}\left(||z_{l+1} - z_l|| \in [t, t+\Delta t]\right) \xrightarrow[N \to \infty]{} e^{-t}\Delta t. \tag{12.44}$$

Remark 12.7.2. The above argument is infinitesimally wrong. Once we have located y_{l+1}, the remaining x_n's do not need to be more than $\frac{t+\Delta t}{N}$ units to the right of $x_1 = y_l$; they only need to be further to the right than y_{l+1}. As the incremental gain in probabilities for the locations of the remaining x_n's is of order Δt, these contributions will not influence the large N, small Δt limits, and we may safely ignore these effects.

To rigorously derive the limiting behavior of the nearest neighbor spacings using the above arguments, one would integrate over x_m ranging from $\frac{t}{N}$ to $\frac{t+\Delta t}{N}$, and the remaining events x_n would be in the a segment of length $1 - x_m$. As

$$\left|(1 - x_m) - \left(1 - \frac{t+\Delta t}{N}\right)\right| \leq \frac{\Delta t}{N}, \tag{12.45}$$

this will lead to corrections of higher order in Δt, hence negligible.

We can rigorously avoid this complication by instead considering the following:

1. Calculate the probability that *all* $N-1$ of the other x_n's are at least $\frac{t}{N}$ units to the right of x_1. This is
$$p_N(t) \;=\; \left(1-\frac{t}{N}\right)^{N-1}, \quad \lim_{N\to\infty} p_N(t) \;=\; e^{-t}. \tag{12.46}$$

2. Calculate the probability that *all* $N-1$ of the other x_n's are at least $\frac{t+\Delta t}{N}$ units to the right of x_1. This is
$$p_N(t+\Delta t) \;=\; \left(1-\frac{t+\Delta t}{N}\right)^{N-1}, \quad \lim_{N\to\infty} p_N(t+\Delta t) \;=\; e^{-(t+\Delta t)}. \tag{12.47}$$

3. The probability that no x_n's are within $\frac{t}{N}$ units to the right of x_1 but at least one x_n is between $\frac{t}{N}$ and $\frac{t+\Delta t}{N}$ units to the right is $p_N(t) - p_N(t+\Delta t)$, and
$$\begin{aligned}\lim_{N\to\infty}\left(p_N(t)-p_N(t+\Delta t)\right) \;&=\; e^{-t}-e^{-(t+\Delta t)}\\ &=\; e^{-t}\left(1-e^{-\Delta t}\right)\\ &=\; e^{-t}\left(1-1+\Delta t+O\left((\Delta t)^2\right)\right)\\ &=\; e^{-t}\Delta t+O\left((\Delta t)^2\right).\end{aligned} \tag{12.48}$$

We have shown

Theorem 12.7.3 (Nearest Neighbor Spacings)**.** *As $N\to\infty$ and then $\Delta t\to 0$,*
$$\begin{aligned}\text{Prob}\left(||y_{l+1}-y_l|| \in \left[\frac{t}{N},\frac{t+\Delta t}{N}\right]\right) \;&\longrightarrow\; e^{-t}\Delta t\\ \text{Prob}\left(||z_{l+1}-z_l|| \in [t,t+\Delta t]\right) \;&\longrightarrow\; e^{-t}\Delta t.\end{aligned} \tag{12.49}$$

Exercise(h) 12.7.4. *Generalize the above arguments to analyze the nearest neighbor spacings when $x_1,\dots,x_N$ are independently drawn from a "nice" probability distribution. Explicitly, for any $\delta\in(0,1)$, show that as $N\to\infty$ the normalized nearest neighbor spacings for any N^δ consecutive x_i's tend to independent standard exponentials. See Appendix A of [MN] for a proof.*

12.7.2 m^{th} Neighbor Spacings

Similarly, one can easily analyze the distribution of the m^{th} nearest neighbor spacings (for fixed m) when each x_n is chosen independently from the uniform distribution on $[0,1)$.

1. We first calculate the probability that exactly $m-1$ of the other x_n's are at most $\frac{t}{N}$ units to the right of x_1, and the remaining $(N-1)-(m-1)$ of the x_n's are at least $\frac{t}{N}$ units to the right of x_1. There are $\binom{N-1}{m-1}$ ways to choose $m-1$ of the x_n's to be at most $\frac{t}{N}$ units to the right of x_1. As m is fixed, for N large, m is significantly smaller than N. The probability is
$$\begin{aligned}p_N(t) \;&=\; \binom{N-1}{m-1}\left(\frac{t}{N}\right)^{m-1}\left(1-\frac{t}{N}\right)^{(N-1)-(m-1)}\\ \lim_{N\to\infty} p_N(t) \;&=\; \frac{t^{m-1}}{(m-1)!}e^{-t},\end{aligned} \tag{12.50}$$

because

$$\lim_{N\to\infty} \binom{N-1}{m-1}\frac{1}{N!} \;=\; \frac{1}{(m-1)!}. \tag{12.51}$$

It is not surprising that the limiting arguments break down for m of comparable size to N; for example, if $m = N$, the N^{th} nearest neighbor spacing is always 1.

2. We calculate the probability that exactly $m-1$ of the other x_n's are at most $\frac{t}{N}$ units to the right of x_1, and the remaining $(N-1)-(m-1)$ of the x_n's are at least $\frac{t+\Delta t}{N}$ units to the right of x_1. Similar to the above, this gives

$$\begin{aligned} p_N(t+\Delta t) \;&=\; \binom{N-1}{m-1}\left(\frac{t}{N}\right)^{m-1}\left(1-\frac{t+\Delta t}{N}\right)^{(N-1)-(m-1)} \\ \lim_{N\to\infty} p_N(t+\Delta t) \;&=\; \frac{t^{m-1}}{(m-1)!}e^{-(t+\Delta t)}. \end{aligned} \tag{12.52}$$

3. The probability that exactly $m-1$ of the x_n's are within $\frac{t}{N}$ units to the right of x_1 and at least one x_n is between $\frac{t}{N}$ and $\frac{t+\Delta t}{N}$ units to the right is $p_N(t) - p_N(t+\Delta t)$, and

$$\begin{aligned} \lim_{N\to\infty}\left(p_N(t)-p_N(t+\Delta t)\right) \;&=\; \frac{t^{m-1}}{(m-1)!}e^{-t} - \frac{t^{m-1}}{(m-1)!}e^{-(t+\Delta t)} \\ &=\; \frac{t^{m-1}}{(m-1)!}e^{-t}\left(1-e^{-\Delta t}\right) \\ &=\; \frac{t^{m-1}}{(m-1)!}e^{-t}\Delta t + O((\Delta t)^2). \end{aligned} \tag{12.53}$$

Note that when $m = 1$, we recover the nearest neighbor spacings. We have shown

Theorem 12.7.5 (Neighbor Spacings)**.** *For fixed m, as $N \to \infty$ and then $\Delta t \to 0$,*

$$\begin{aligned} \text{Prob}\left(||y_{l+m} - y_l|| \in \left[\frac{t}{N}, \frac{t+\Delta t}{N}\right]\right) \;&\longrightarrow\; \frac{t^{m-1}}{(m-1)!}e^{-t}\Delta t \\ \text{Prob}\left(||z_{l+1} - z_l|| \in [t, t+\Delta t]\right) \;&\longrightarrow\; \frac{t^{m-1}}{(m-1)!}e^{-t}\Delta t. \end{aligned} \tag{12.54}$$

Exercise 12.7.6. *Keep track of the lower order terms in N and Δt in* (12.50) *and* (12.52) *to prove Theorem 12.7.5.*

Exercise 12.7.7 (Median)**.** *We give another application of order statistics. Let $X_1, \dots, X_n$ be independent random variables chosen from a probability distribution p. The median of p is defined as the number $\widetilde{\mu}$ such that $\int_{-\infty}^{\widetilde{\mu}} p(x)dx = \frac{1}{2}$. If $n = 2m+1$, let the random variable $\widetilde{X}$ be the median of $X_1, \dots, X_n$. Show the density function of $\widetilde{X}$ is*

$$f(\widetilde{x}) \;=\; \frac{(2m+1)!}{m!m!}\left[\int_{-\infty}^{\widetilde{x}} p(x)dx\right]^m \cdot p(\widetilde{x}) \cdot \left[\int_{\widetilde{x}}^{\infty} p(x)dx\right]^m. \tag{12.55}$$

It can be shown that if p is continuous and $p(\widetilde{\mu}) \neq 0$ then for large m, $f(\widetilde{x})$ is approximately a normal distribution with mean $\widetilde{\mu}$ and variance $\frac{1}{8p(\widetilde{\mu})^2 m}$. For distributions that are symmetric about the mean μ, the mean μ equals the median $\widetilde{\mu}$ and the above is another way to estimate the mean. This is useful in situations where the distribution has infinite variance and the Central Limit Theorem is not applicable, for example, if p is a translated Cauchy distribution:

$$p(x) \ = \ \frac{1}{\pi}\frac{1}{1+(x-\widetilde{\mu})^2}. \tag{12.56}$$

In fact, while the above is symmetric about μ, the expected value does not exist yet the Median Theorem can still be used to estimate the parameter $\widetilde{\mu}$. Another advantage of the median over the mean is that changing one or two data points will not change the median much but could greatly change the mean. Thus, in situations where there is a good chance of recording the wrong value of an observation, the median is often a better statistic to study.

Exercise(hr) 12.7.8. *Prove the Median Theorem from the previous problem if p is sufficiently nice.*

12.7.3 Induced Probability Measures

We reinterpret our results in terms of probability measures.

Theorem 12.7.9. *Consider N independent random variables x_n chosen from the uniform distribution on the wrapped unit interval $[0,1)$. For fixed N, arrange the x_n's in increasing order, labeled $y_1 \le y_2 \le \cdots \le y_N$. Fix m and form the probability measure from the m^{th} nearest neighbor spacings. Then as $N \to \infty$*

$$\mu_{N,m;y}(t)dt \ = \ \frac{1}{N}\sum_{n=1}^{N}\delta\left(t - N(y_n - y_{n-m})\right)dt \ \longrightarrow \ \frac{t^{m-1}}{(m-1)!}e^{-t}dt. \tag{12.57}$$

Equivalently, using $z_n = Ny_n$ gives

$$\mu_{N,m;z}(t)dt \ = \ \frac{1}{N}\sum_{n=1}^{N}\delta\left(t - (z_n - z_{n-m})\right)dt \ \longrightarrow \ \frac{t^{m-1}}{(m-1)!}e^{-t}dt. \tag{12.58}$$

Definition 12.7.10 (Poissonian Behavior)**.** *Let $\{x_n\}_{n=1}^{\infty}$ be a sequence of points in $[0,1)$. For each N, let $\{y_n\}_{n=1}^{N}$ be $x_1,\dots,x_N$ arranged in increasing order. We say x_n has* Poissonian behavior *if in the limit as $N\to\infty$, for each m the induced probability measure $\mu_{N,m;y}(t)dt$ converges to $\frac{t^{m-1}}{(m-1)!}e^{-t}dt$.*

Exercise 12.7.11. *Let $\alpha \in \mathbb{Q}$ and define $\alpha_n = \{n^m\alpha\}$ for some positive integer m. Show the sequence of points α_n does not have Poissonian Behavior.*

Exercise(h) 12.7.12. *Let $\alpha \notin \mathbb{Q}$ and define $\alpha_n = \{n\alpha\}$. Show the sequence of points α_n does not have Poissonian behavior.*

12.8 NEIGHBOR SPACINGS OF $\{n^k\alpha\}$

We now come to Question 12.1.9. Note there are three pieces of data: k, α and N.

Conjecture 12.8.1. *For any integer $k \geq 2$, for most irrational α (in the sense of measure), if $\alpha_n = \{n^k\alpha\}$ then the sequence $\{\alpha_n\}_{n=1}^N$ is Poissonian as $N \to \infty$.*

Many results towards this conjecture are proved in n [RS2, RSZ]. We merely state some of them; see [RS2, RSZ] for complete statements and proofs. We concentrate on the case $k = 2$.

Let $\alpha \notin \mathbb{Q}$ be such that there are infinitely many rationals b_j and q_j (with q_j prime) such that $|\alpha - \frac{b_j}{q_j}| < \frac{1}{q_j^3}$. Then there is a sequence $\{N_j\}_{j=1}^\infty$ such that $\{n^2\alpha\}$ exhibits Poissonian behavior along this subsequence. Explicitly, $\{\{n^2\alpha\}\}_{n=1}^N$ is Poissonian as $N \to \infty$ *if* we restrict N to the sequence $\{N_j\}$. If α has slightly better approximations by rationals, one can construct a subsequence where the behavior is *not* Poissonian. For $k = 2$ and α that are well approximated by rationals, we show there is a sequence of N_j's such that the triples (k, α, N_j) do not exhibit Poissonian behavior (see Theorem 12.8.5).

Thus, depending on how N tends to infinity, wildly different behavior can be seen. This leads to many open problems, which we summarize at the end of the chapter. For more information, see [Li, Mi1].

12.8.1 Poissonian Behavior

Consider first the case when $\alpha \in \mathbb{Q}$. If $\alpha = \frac{p}{q}$, the fractional parts $\{n^2\frac{p}{q}\}$ are periodic with period N; contrast this with $\alpha \notin \mathbb{Q}$, where the fractional parts are distinct. *Assume q is prime.* Davenport [Da3, Da4] investigated the neighbor spacings of $\{\{n^2\frac{p}{q}\}\}_{n=1}^N$ for $N = q$ as both tend to infinity; [RSZ] investigate the above for smaller N. As the sequence is periodic with period q, it suffices to study $1 \leq N \leq q$. For simplicity, let $p = 1$. If $N \leq \sqrt{q}$, the sequence is already in increasing order, and the adjacent spacings are given by $\frac{2n+1}{q}$. Clearly, for such small N, one does not expect the sequence $\{n^2\alpha\}$ to behave like N points chosen uniformly in $[0, 1)$. Once N is larger than $\sqrt{q}$, say $N > q^{\frac{1}{2}+\epsilon}$, we have q^ϵ blocks of $\sqrt{q}$ terms. These will now be intermingled in $[0, 1)$, and there is now the possibility of seeing "randomness." Similarly, we must avoid N being too close to q, as if $N = q$ then the normalized spacings are integers. In [RSZ], evidence towards Poissonian behavior is proved for $N \in [q^{\frac{1}{2}+\epsilon}, \frac{q}{\log q}]$.

For $\alpha \notin \mathbb{Q}$, [RSZ] note the presence of large square factors in the convergents to α (the q_n in the continued fraction expansion) prevent Poissonian behavior. While it is possible to construct numbers that always have large square factors in the convergents (if q_{m-1} and q_m are squares, letting $a_{m+1} = q_m + 2\sqrt{q_{m-1}}$ yields $q_{m+1} = a_{m+1}q_m + q_{m-1} = (q_m + \sqrt{q_{m-1}})^2$), for most α in the sense of measure this will not occur, and one obtains subsequences N_j where the behavior is Poissonian. In particular, their techniques can prove

Theorem 12.8.2 (RSZ). *Let $\alpha \notin \mathbb{Q}$ have infinitely many rational approximations satisfying $|\alpha - \frac{p_j}{q_j}| < \frac{1}{q_j^3}$ and $\lim_{j\to\infty} \frac{\log \widetilde{q}_j}{\log q_j} = 1$, where $\widetilde{q}_j$ is the square-free part of*

q_j. *Then there is a subsequence* $N_j \to \infty$ *with* $\frac{\log N_j}{\log q_j} \to 1$ *such that the fractional parts* $\{n^2\alpha\}$ *are Poissonian along this subsequence.*

For data on rates of convergence to Poissonian behavior, see [Li, Mi1].

Exercise 12.8.3. *Construct an* $\alpha \notin \mathbb{Q}$ *such that the denominators in the convergents are perfect cubes (or, more generally, perfect* k^{th} *powers).*

One useful statistic to investigate whether or not a sequence exhibits Poissonian behavior is the n**-level correlation**. We content ourselves with discussing the **2-level** or **pair correlation** (see [RS2, RSZ, RZ1] for more details). Given a sequence of increasing numbers x_n in $[0,1)$, let

$$R_2([-s,s],x_n,N) \ = \ \frac{1}{N}\#\{1 \le j \ne k \le N : ||x_j - x_k|| \le s/N\}. \quad (12.59)$$

As the expected difference between such x's is $\frac{1}{N}$, this statistic measures the distribution of spacings between x's on the order of the average spacing. The adjacent neighbor spacing distribution looks at $||x_{j+1} - x_j||$; here we allow x's to be between x_k and x_j. One can show that knowing *all* the n-level correlations allows one to determine all the neighbor spacings; see for example [Meh2].

Exercise 12.8.4 (Poissonian Behavior)**.** *For* $j \le N$ *let the* θ_j *be chosen independently and uniformly in* $[0,1)$. *Prove as* $N \to \infty$ *that* $R_2([-s,s],\theta_j,N) \to 2s$.

In [RS2] it is shown that the 2-level correlation for almost all α (in the sense of measure) converges to $2s$, providing support for the conjectured Poissonian behavior. The proof proceeds in three steps. First, they show the mean over α is $2s$. This is not enough to conclude that most α have 2-level correlations close to $2s$. For example, consider a sequence that half the time is 0 and the other half of the time is 2. The mean is 1, but clearly no elements are close to the mean. If, however, we knew the variance were small, *then* we could conclude that most elements are close to the mean. The second step in [RS2] is to show the variance is small, and then standard bounds on exponential sums and basic probability arguments allow them to pass to almost all α satisfy the claim.

Numerous systems exhibit Poissonian behavior. We see another example in Project 17.4.3, where we study the spacings between eigenvalues of real symmetric Toeplitz matrices (these matrices are constant along diagonals). See [GK, KR] for additional systems.

12.8.2 Non-Poissonian Behavior

While we give the details of a sequence that is non-Poissonian, remember that the *typical* α is expected to exhibit Poissonian behavior. See also [RZ1, RZ2] for Poissonian behavior of other (lacunary) sequences.

Theorem 12.8.5 ([RSZ])**.** *Let* $\alpha \notin \mathbb{Q}$ *be such that* $\left|\alpha - \frac{p_j}{q_j}\right| < \frac{a_j}{q_j^3}$ *holds infinitely often, with* $a_j \to 0$. *Then there exist integers* $N_j \to \infty$ *such that* $\mu_{N_j,1}(t)$ *does not converge to* $e^{-t}dt$. *In particular, along this subsequence the behavior is non-Poissonian.*

The idea of the proof is as follows: since α is well approximated by rationals, when we write the $\{n^2\alpha\}$ in increasing order and look at the normalized nearest neighbor differences (the z_n's in Theorem 12.7.3), the differences will be arbitrarily close to integers. This cannot be Poissonian behavior, as the limiting distribution there is $e^{-t}dt$, which is clearly not concentrated near integers.

As $a_j \to 0$, eventually $a_j < \frac{1}{10}$ for all j large. Let $N_j = q_j$, where $\frac{p_j}{q_j}$ is a good rational approximation to α:

$$\left|\alpha - \frac{p_j}{q_j}\right| \ < \ \frac{a_j}{q_j^3}. \tag{12.60}$$

We look at $\alpha_n = \{n^2\alpha\}$, $1 \le n \le N_j = q_j$. Let the β_n's be the α_n's arranged in increasing order, and let the γ_n's be the numbers $\{n^2\frac{p_j}{q_j}\}$ arranged in increasing order:

$$\begin{aligned} &\beta_1 \ \le \ \beta_2 \ \le \ \cdots \ \le \ \beta_{N_j} \\ &\gamma_1 \ \le \ \gamma_2 \ \le \ \cdots \ \le \ \gamma_{N_j}. \end{aligned} \tag{12.61}$$

Lemma 12.8.6. *If $\beta_l = \alpha_n = \{n^2\alpha\}$, then $\gamma_l = \{n^2\frac{p_j}{q_j}\}$. Thus the same permutation orders both the α_n's and the γ_n's.*

Proof. Multiplying both sides of (12.60) by $n^2 \le q_j^2$ yields for j large

$$\left|n^2\alpha - n^2\frac{p_j}{q_j}\right| \ < \ n^2\frac{a_j}{q_j^3} \ \le \ \frac{a_j}{q_j} \ < \ \frac{1}{10q_j}, \tag{12.62}$$

and $n^2\alpha$ and $n^2\frac{p_j}{q_j}$ differ by at most $\frac{1}{10q_j}$. Therefore

$$\left|\left|\{n^2\alpha\} - \left\{n^2\frac{p_j}{q_j}\right\}\right|\right| \ < \ \frac{1}{10q_j}. \tag{12.63}$$

As p_j and q_j are relatively prime, and all numbers $\{m^2\frac{p_j}{q_j}\}$ have denominator at most q_j, two such numbers cannot be closer than $\frac{1}{q_j}$ unless they are equal. For example, if q_j is a perfect square, $m_1 = \sqrt{q_j}$ and $m_2 = 2\sqrt{q_j}$ give the same number; the presence of large square factors of q_j has important consequences (see [RSZ]).

Thus $\{n^2\frac{p_j}{q_j}\}$ is the closest (or tied for being the closest) of the $\{m^2\frac{p_j}{q_j}\}$ to $\{n^2\alpha\}$. This implies that if $\beta_l = \{n^2\alpha\}$, then $\gamma_l = \{n^2\frac{p_j}{q_j}\}$, completing the proof. □

Exercise 12.8.7. *Prove two of the $\{m^2\frac{p_j}{q_j}\}$ are either equal, or at least $\frac{1}{q_j}$ apart.*

Exercise 12.8.8. *Assume $||a-b||, ||c-d|| < \frac{1}{10}$. Show*

$$||(a-b)-(c-d)|| \ < \ ||a-b|| + ||c-d||. \tag{12.64}$$

We now prove Theorem 12.8.5: We have shown

$$||\beta_l - \gamma_l|| \ < \ \frac{a_j}{q_j}. \tag{12.65}$$

As $N_j = q_j$,

$$||N_j(\beta_l - \gamma_l)|| \ < \ a_j, \tag{12.66}$$

and the same result holds with l replaced by $l-1$. By Exercise 12.8.8,

$$||N_j(\beta_l - \gamma_l) - N_j(\beta_{l-1} - \gamma_{l-1})|| \; < \; 2a_j. \tag{12.67}$$

Rearranging gives

$$||N_j(\beta_l - \beta_{l-1}) - N_j(\gamma_l - \gamma_{l-1})|| \; < \; 2a_j. \tag{12.68}$$

As $a_j \to 0$, this implies that the difference between $||N_j(\beta_l - \beta_{l-1})||$ and $||N_j(\gamma_l - \gamma_{l-1})||$ tends to zero.

The above distance calculations were done modulo 1. The actual differences will differ by an integer. Therefore

$$\mu_{N_j,1;\alpha}(t)dt \;=\; \frac{1}{N_j}\sum_{l=1}^{N_j} \delta\left(t - N_j(\beta_l - \beta_{l-1})\right) \tag{12.69}$$

and

$$\mu_{N_j,1;\frac{p_j}{q_j}}(t)dt \;=\; \frac{1}{N_j}\sum_{l=1}^{N_j} \delta\left(t - N_j(\gamma_l - \gamma_{l-1})\right) \tag{12.70}$$

are in some sense extremely close to one another: each point mass from the difference between adjacent β_l's is within $k+a_j$ units of a point mass from the difference between adjacent γ_l's for some integer k, and a_j tends to zero. Note, however, that if $\gamma_l = \{n^2\frac{p_j}{q_j}\}$, then

$$N_j \cdot \gamma_l \;=\; q_j\left\{n^2\frac{p_j}{q_j}\right\} \;\in\; \mathbb{N}. \tag{12.71}$$

Thus the induced probability measure $\mu_{N_j,1;\frac{p_j}{q_j}}(t)dt$ formed from the γ_l's is supported on the integers! It is therefore impossible for $\mu_{N_j,1;\frac{p_j}{q_j}}(t)dt$ to converge to $e^{-t}dt$.

As $\mu_{N_j,1;\alpha}(t)dt$, modulo some possible integer shifts, is arbitrarily close to $\mu_{N_j,1;\frac{p_j}{q_j}}(t)dt$, the sequence $\{n^2\alpha\}$ is *not* Poissonian along the subsequence of Ns given by N_j, where $N_j = q_j$, q_j is a denominator in a good rational approximation to α.

12.9 RESEARCH PROJECTS

We have shown that the fractional parts $\{\{n^k\alpha\}\}_{n=1}^N$ (for $\alpha \notin \mathbb{Q}$) are dense and equidistributed for all positive integers k as $N \to \infty$; however, the finer question of neighbor spacings depends greatly on α and how N tends to infinity.

Research Project 12.9.1. Much is known about the fractional parts $\{n^2\alpha\}$ if $\alpha = \frac{p}{q}$ with q prime. What if q is the product of two primes (of comparable size or of wildly different size)? Is the same behavior true for $\{n^k\frac{p}{q}\}$?

Research Project 12.9.2. For certain α there is a subsequence N_j such that, along this subsequence, the behavior of $\{\{n^2\alpha\}\}_{n\le N_j}$ is non-Poissonian. What about the behavior of $\{n^2\alpha\}$ away from such bad N_j's? Is the behavior Poissonian between N_j and N_{j+1}, and if so, how far must one move? See [Mi1] for some observations and results.

Research Project 12.9.3. For $\alpha \notin \mathbb{Q}$, the fractional parts $x_n = \{n\alpha\}$ are equidistributed but not Poissonian. In particular, for each N there are only three possible values for the nearest neighbor differences (the values depend on N and α); for the m^{th} neighbor spacings, there are also only finitely many possible differences. How are the differences distributed among the possible values? Is each value equally likely? Does the answer depend on α, and if so, on what properties of α? For more details, see [Bl1, Bl2, Mar].

Research Project 12.9.4. Generalize the results of [RSZ] for the fractional parts $\{n^k\alpha\}$, $k \ge 3$. It is likely that the obstruction to non-Poissonian behavior will be the presence of convergents to α with q_n a large k^{th} power.

Research Project 12.9.5 (Primes)**.** Similar to investigating the Poissonian behavior of $\{n^k\alpha\}$, one can study primes. For $\{n^k\alpha\}$ we looked at $n \le N$. The average spacing was $\frac{1}{N}$ and we then took the limit as $N \to \infty$. For primes we can investigate $p \in [x, x+y]$. If we choose x large and y large on an absolute scale but small relative to x, then there will be a lot of primes and the average spacing will be approximately constant. For example, the Prime Number Theorem states the number of primes at most x is about $\frac{x}{\log x}$. If $x = 10^{14}$ and $y = 10^6$, then there should be a lot of primes in this interval and the average spacing should be fairly constant (about $\log 10^{14}$). Are the spacings between primes Poissonian? What about primes in arithmetic progression, or twin primes or generalized twin primes or prime tuples (see the introduction to Chapter 14 for more details about such primes). See [Sch, Weir] for numerical investigations along these lines. The Circle Method (Chapters 13 and 14) provides estimates as to the number of such primes. Another sequence to investigate are Carmichael numbers; see Project 1.4.25.