美国数学会经典影印系列

AMS
AMERICAN MATHEMATICAL SOCIETY

Higher Order Fourier Analysis

高阶傅里叶分析

Terence Tao

高等教育出版社

美国数学会经典影印系列
AMS
AMERICAN MATHEMATICAL SOCIETY

Higher Order Fourier Analysis

高阶傅里叶分析

Terence Tao

高等教育出版社·北京

图字：01-2016-2502 号

高阶傅里叶分析
Gaojie Fuliye Fenxi

图书在版编目 (CIP) 数据

高阶傅里叶分析 = Higher Order Fourier Analysis：英文 / (澳) 陶哲轩 (Terence Tao) 著 . —影印本. —北京：高等教育出版社，2017.1
ISBN 978-7-04-046909-7
Ⅰ. ① 高… Ⅱ. ①陶… Ⅲ. ①傅里叶分析—英文
Ⅳ. ①O174.2
中国版本图书馆 CIP 数据核字 (2016) 第 281867 号

策划编辑 赵天夫　　责任编辑 赵天夫
封面设计 张申申　　责任印制 毛斯璐

出版发行 高等教育出版社
社址 北京市西城区德外大街 4 号
邮政编码 100120
购书热线 010-58581118
咨询电话 400-810-0598
网址 http://www.hep.edu.cn
http://www.hep.com.cn
网上订购 http://www.hepmall.com.cn
http://www.hepmall.com
http://www.hepmall.cn
印刷 北京中科印刷有限公司

开本 787mm × 1092mm 1/16
印张 12.75
字数 263 千字
版次 2017 年 1 月第 1 版
印次 2017 年 1 月第 1 次印刷
定价 99.00 元

[物 料 号 46909-00]

美国数学会经典影印系列

出版者的话

近年来，我国的科学技术取得了长足进步，特别是在数学等自然科学基础领域不断涌现出一流的研究成果。与此同时，国内的科研队伍与国外的交流合作也越来越密切，越来越多的科研工作者可以熟练地阅读英文文献，并在国际顶级期刊发表英文学术文章，在国外出版社出版英文学术著作。

然而，在国内阅读海外原版英文图书仍不是非常便捷。一方面，这些原版图书主要集中在科技、教育比较发达的大中城市的大型综合图书馆以及科研院所的资料室中，普通读者借阅不甚容易；另一方面，原版书价格昂贵，动辄上百美元，购买也很不方便。这极大地限制了科技工作者对于国外先进科学技术知识的获取，间接阻碍了我国科技的发展。

高等教育出版社本着植根教育、弘扬学术的宗旨服务我国广大科技和教育工作者，同美国数学会（American Mathematical Society）合作，在征求海内外众多专家学者意见的基础上，精选该学会近年出版的数十种专业著作，组织出版了“美国数学会经典影印系列”丛书。美国数学会创建于 1888 年，是国际上极具影响力的专业学术组织，目前拥有近 30000 会员和 580 余个机构成员，出版图书 3500 多种，冯·诺依曼、莱夫谢茨、陶哲轩等世界级数学大家都是其作者。本影印系列涵盖了代数、几何、分析、方程、拓扑、概率、动力系统等所有主要数学分支以及新近发展的数学主题。

我们希望这套书的出版，能够对国内的科研工作者、教育工作者以及青年学生起到重要的学术引领作用，也希望今后能有更多的海外优秀英文著作被介绍到中国。

高 等 教 育 出 版 社

2016 年 12 月

To Garth Gaudry, who set me on the road;
To my family, for their constant support;
And to the readers of my blog, for their feedback and contributions.

Preface

Traditionally, Fourier analysis has been focused on the analysis of functions in terms of linear phase functions such as the sequence $n \mapsto e(\alpha n) := e^{2\pi i \alpha n}$. In recent years, though, applications have arisen—particularly in connection with problems involving linear patterns such as arithmetic progressions—in which it has been necessary to go beyond the linear phases, replacing them to higher order functions such as quadratic phases $n \mapsto e(\alpha n^2)$. This has given rise to the subject of *quadratic Fourier analysis* and, more generally, to *higher order Fourier analysis*.

The classical results of Weyl on the equidistribution of polynomials (and their generalisations to other orbits on homogeneous spaces) can be interpreted through this perspective as foundational results in this subject. However, the modern theory of higher order Fourier analysis is very recent indeed (and still incomplete to some extent), beginning with the breakthrough work of Gowers [**Go1998**], [**Go2001**] and also heavily influenced by parallel work in ergodic theory, in particular, the seminal work of Host and Kra [**HoKr2005**]. This area was also quickly seen to have much in common with areas of theoretical computer science related to polynomiality testing, and in joint work with Ben Green and Tamar Ziegler [**GrTa2010**], [**GrTa2008c**], [**GrTaZi2010b**], applications of this theory were given to asymptotics for various linear patterns in the prime numbers.

There are already several surveys or texts in the literature (e.g. [**Gr2007**], [**Kr2006**], [**Kr2007**], [**Ho2006**], [**Ta2007**], [**TaVu2006**]) that seek to cover some aspects of these developments. In this text (based on a topics graduate course I taught in the spring of 2010), I attempt to give a broad tour of this nascent field. This text is not intended to directly substitute for the core papers on the subject (many of which are quite technical and lengthy), but focuses instead on basic foundational and preparatory material, and on the simplest illustrative examples of key results, and should thus hopefully serve as a companion to the existing literature on the subject. In accordance with this complementary intention of this text, we also present certain approaches to the material that is not explicitly present in the literature, such as the abstract approach to Gowers-type norms (Section 2.2) or the ultrafilter approach to equidistribution (Section 1.1.3).

There is, however, one important omission in this text that should be pointed out. In order to keep the material here focused, self-contained, and of a reasonable length (in particular, of a length that can be mostly covered in a single graduate course), I have focused on the combinatorial aspects of higher order Fourier analysis, and only very briefly touched upon the equally significant ergodic theory side of the subject. In particular, the breakthrough work of Host and Kra [**HoKr2005**], establishing an ergodic-theoretic precursor to the inverse conjecture for the Gowers norms, is not discussed in detail here; nor is the very recent work of Szegedy [**Sz2009**], [**Sz2009b**], [**Sz2010**], [**Sz2010b**] and Camarena-Szegedy [**CaSz2010**] in which the Host-Kra machinery is adapted to the combinatorial setting. However, some of the foundational material for these papers, such as the ultralimit approach to equidistribution and structural decomposition, or the analysis of parallelopipeds on nilmanifolds, is covered in this text.

This text presumes a graduate-level familiarity with basic real analysis and measure theory, such as is covered in [**Ta2011**], [**Ta2010**], particularly with regard to the "soft" or "qualitative" side of the subject.

The core of the text is Chapter 1, which comprises the main lecture material. The material in Chapter 2 is optional to these lectures, except for the ultrafilter material in Section 2.1 which would be needed to some extent in order to facilitate the ultralimit analysis in Chapter 1. However, it is possible to omit the portions of the text involving ultrafilters and still be able to cover most of the material (though from a narrower set of perspectives).

Acknowledgments

I am greatly indebted to my students of the course on which this text was based, as well as many further commenters on my blog, including Sungjin Kim, William Meyerson, Joel Moreira, Thomas Sauvaget, Siming Tu, and Mads Sørensen. These comments, as well as the original lecture notes for this course, can be viewed online at

`terrytao.wordpress.com/category/teaching/254b-higher-order-fourier-analysis/`

Thanks also to Ben Green for suggestions. The author is supported by a grant from the MacArthur Foundation, by NSF grant DMS-0649473, and by the NSF Waterman award.

Contents

Chapter 1

Higher order Fourier analysis

1.1. Equidistribution of polynomial sequences in tori

(Linear) *Fourier analysis* can be viewed as a tool to study an arbitrary function f on (say) the integers $\mathbf{Z}$, by looking at how such a function correlates with *linear phases* such as $n \mapsto e(\xi n)$, where $e(x) := e^{2\pi i x}$ is the fundamental character, and $\xi \in \mathbf{R}$ is a frequency. These correlations control a number of expressions relating to f, such as the expected behaviour of f on arithmetic progressions $n, n+r, n+2r$ of length three.

In this text we will be studying higher-order correlations, such as the correlation of f with quadratic phases such as $n \mapsto e(\xi n^2)$, as these will control the expected behaviour of f on more complex patterns, such as arithmetic progressions $n, n+r, n+2r, n+3r$ of length four. In order to do this, we must first understand the behaviour of *exponential sums* such as

$$\sum_{n=1}^{N} e(\alpha n^2).$$

Such sums are closely related to the *distribution* of expressions such as $\alpha n^2 \bmod 1$ in the unit circle $\mathbf{T} := \mathbf{R}/\mathbf{Z}$, as n varies from 1 to N. More generally, one is interested in the distribution of polynomials $P\colon \mathbf{Z}^d \to \mathbf{T}$ of one or more variables taking values in a torus $\mathbf{T}$; for instance, one might be interested in the distribution of the quadruplet $(\alpha n^2, \alpha(n+r)^2, \alpha(n+2r)^2, \alpha(n+3r)^2)$ as n, r both vary from 1 to N. Roughly speaking, once we understand these types of distributions, then the general machinery of quadratic Fourier analysis will then allow us to understand the distribution of the quadruplet $(f(n), f(n+r), f(n+2r), f(n+3r))$ for more general classes of functions f; this can lead for instance to an understanding of the distribution of arithmetic progressions of length 4 in the primes, if f is somehow related to the primes.

More generally, to find arithmetic progressions such as $n, n+r, n+2r, n+3r$ in a set A, it would suffice to understand the equidistribution of the quadruplet[1] $(1_A(n), 1_A(n+r), 1_A(n+2r), 1_A(n+3r))$ in $\{0,1\}^4$ as n and r vary. This is the starting point for the fundamental connection between *combinatorics* (and more specifically, the task of finding patterns inside sets) and *dynamics* (and more specifically, the theory of equidistribution and recurrence in measure-preserving dynamical systems, which is a subfield of *ergodic theory*). This connection was explored in the previous monograph [**Ta2009**]; it will also be important in this text (particularly as a source of motivation), but the primary focus will be on finitary, and Fourier-based, methods.

[1] Here 1_A is the *indicator function* of A, defined by setting $1_A(n)$ equal to 1 when $n \in A$ and equal to zero otherwise.

The theory of equidistribution of polynomial orbits was developed in the linear case by Dirichlet and Kronecker, and in the polynomial case by Weyl. There are two regimes of interest; the (qualitative) *asymptotic regime* in which the scale parameter N is sent to infinity, and the (quantitative) *single-scale regime* in which N is kept fixed (but large). Traditionally, it is the asymptotic regime which is studied, which connects the subject to other asymptotic fields of mathematics, such as dynamical systems and ergodic theory. However, for many applications (such as the study of the primes), it is the single-scale regime which is of greater importance. The two regimes are not directly equivalent, but are closely related: the single-scale theory can be usually used to derive analogous results in the asymptotic regime, and conversely the arguments in the asymptotic regime can serve as a simplified model to show the way to proceed in the single-scale regime. The analogy between the two can be made tighter by introducing the (qualitative) *ultralimit regime*, which is formally equivalent to the single-scale regime (except for the fact that explicitly quantitative bounds are abandoned in the ultralimit), but resembles the asymptotic regime quite closely.

For the finitary portion of the text, we will be using *asymptotic notation*: $X \ll Y$, $Y \gg X$, or $X = O(Y)$ denotes the bound $|X| \leq CY$ for some absolute constant C, and if we need C to depend on additional parameters, then we will indicate this by subscripts, e.g., $X \ll_d Y$ means that $|X| \leq C_d Y$ for some C_d depending only on d. In the ultralimit theory we will use an analogue of asymptotic notation, which we will review later in this section.

1.1.1. Asymptotic equidistribution theory. Before we look at the single-scale equidistribution theory (both in its finitary form, and its ultralimit form), we will first study the slightly simpler, and much more classical, *asymptotic* equidistribution theory.

Suppose we have a sequence of points $x(1), x(2), x(3), \ldots$ in a compact metric space X. For any finite $N > 0$, we can define the probability measure

$$\mu_N := \mathbf{E}_{n \in [N]} \delta_{x(n)}$$

which is the average of the *Dirac point masses* on each of the points $x(1), \ldots, x(N)$, where we use $\mathbf{E}_{n\in[N]}$ as shorthand for $\frac{1}{N}\sum_{n=1}^{N}$ (with $[N] := \{1, \ldots, N\}$). *Asymptotic equidistribution theory* is concerned with the limiting behaviour of these probability measures μ_N in the limit $N \to \infty$, for various sequences $x(1), x(2), \ldots$ of interest. In particular, we say that the sequence $x \colon \mathbf{N} \to X$ is *asymptotically equidistributed* on $\mathbf{N}$ with respect to a reference *Borel probability measure* μ on X if the μ_N converge in the vague topology to μ or, in other words, that

$$\mathbf{E}_{n\in[N]} f(x(n)) = \int_X f \, d\mu_N \to \int_X f \, d\mu \tag{1.1}$$

for all continuous scalar-valued functions $f \in C(X)$. Note (from the *Riesz representation theorem*) that any sequence is asymptotically equidistributed with respect to at most one Borel probability measure μ.

It is also useful to have a slightly stronger notion of equidistribution: we say that a sequence $x \colon \mathbf{N} \to X$ is *totally asymptotically equidistributed* if it is asymptotically equidistributed on every infinite arithmetic progression, i.e. that the sequence $n \mapsto x(qn+r)$ is asymptotically equidistributed for all integers $q \geq 1$ and $r \geq 0$.

A doubly infinite sequence $(x(n))_{n \in \mathbf{Z}}$, indexed by the integers rather than the natural numbers, is said to be asymptotically equidistributed relative to μ if both halves[2] of the sequence $x(1), x(2), x(3), \dots$ and $x(-1)$, $x(-2), x(-3), \dots$ are asymptotically equidistributed relative to μ. Similarly, one can define the notion of a doubly infinite sequence being totally asymptotically equidistributed relative to μ.

Example 1.1.1. If $X = \{0,1\}$, and $x(n) := 1$ whenever $2^{2j} \leq n < 2^{2j+1}$ for some natural number j and $x(n) := 0$ otherwise, show that the sequence x is not asymptotically equidistributed with respect to any measure. Thus we see that asymptotic equidistribution requires all scales to behave "the same" in the limit.

Exercise 1.1.1. If $x \colon \mathbf{N} \to X$ is a sequence into a compact metric space X, and μ is a probability measure on X, show that x is asymptotically equidistributed with respect to μ if and only if one has

$$\lim_{N \to \infty} \frac{1}{N} |\{1 \leq n \leq N : x(n) \in U\}| = \mu(U)$$

for all open sets U in X whose boundary ∂U has measure zero. (*Hint:* For the "only if" part, use *Urysohn's lemma*. For the "if" part, reduce (1.1) to functions f taking values between 0 and 1, and observe that almost all of the level sets $\{y \in X : f(y) < t\}$ have a boundary of measure zero.) What happens if the requirement that ∂U have measure zero is omitted?

Exercise 1.1.2. Let x be a sequence in a compact metric space X which is equidistributed relative to some probability measure μ. Show that for any open set U in X with $\mu(U) > 0$, the set $\{n \in \mathbf{N} : x(n) \in U\}$ is infinite, and furthermore has positive lower density in the sense that

$$\liminf_{N \to \infty} \frac{1}{N} |\{1 \leq n \leq N : x(n) \in U\}| > 0.$$

In particular, if the support of μ is equal to X, show that the set $\{x(n) : n \in \mathbf{N}\}$ is dense in X.

[2]This omits $x(0)$ entirely, but it is easy to see that any individual element of the sequence has no impact on the asymptotic equidistribution.

Exercise 1.1.3. Let $x\colon \mathbf{N} \to X$ be a sequence into a compact metric space X which is equidistributed relative to some probability measure μ. Let $\varphi\colon \mathbf{R} \to \mathbf{R}$ be a compactly supported, piecewise continuous function with only finitely many pieces. Show that for any $f \in C(X)$ one has

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n\in\mathbf{N}} \varphi(n/N) f(x(n)) = \left(\int_X f\, d\mu\right)\left(\int_0^\infty \varphi(t)\, dt\right)$$

and for any open U whose boundary has measure zero, one has

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n\in\mathbf{N}: x(n)\in U} \varphi(n/N) = \mu(U)\left(\int_0^\infty \varphi(t)\, dt\right).$$

In this section, X will be a torus (i.e., a compact connected abelian Lie group), which from the theory of Lie groups is isomorphic to the standard torus $\mathbf{T}^d$, where d is the dimension of the torus. This torus is then equipped with *Haar measure*, which is the unique Borel probability measure on the torus which is translation-invariant. One can identify the standard torus $\mathbf{T}^d$ with the standard fundamental domain $[0,1)^d$, in which case the Haar measure is equated with the usual Lebesgue measure. We shall call a sequence $x_1, x_2, \ldots$ in $\mathbf{T}^d$ (asymptotically) *equidistributed* if it is (asymptotically) equidistributed with respect to Haar measure.

We have a simple criterion for when a sequence is asymptotically equidistributed, that reduces the problem to that of estimating exponential sums:

Proposition 1.1.2 (Weyl equidistribution criterion). *Let $x\colon \mathbf{N} \to \mathbf{T}^d$. Then x is asymptotically equidistributed if and only if*

$$\lim_{N\to\infty} \mathbf{E}_{n\in[N]} e(k \cdot x(n)) = 0 \tag{1.2}$$

for all $k \in \mathbf{Z}^d \backslash \{0\}$, where $e(y) := e^{2\pi i y}$. Here we use the dot product

$$(k_1, \ldots, k_d) \cdot (x_1, \ldots, x_d) := k_1 x_1 + \cdots + k_d x_d$$

which maps $\mathbf{Z}^d \times \mathbf{T}^d$ to $\mathbf{T}$.

Proof. The "only if" part is immediate from (1.1). For the "if" part, we see from (1.2) that (1.1) holds whenever f is a plane wave $f(y) := e(k \cdot y)$ for some $k \in \mathbf{Z}^d$ (checking the $k = 0$ case separately), and thus by linearity whenever f is a trigonometric polynomial. But by Fourier analysis (or from the *Stone-Weierstrass theorem*), the trigonometric polynomials are dense in $C(\mathbf{T}^d)$ in the uniform topology. The claim now follows from a standard limiting argument. □

As one consequence of this proposition, one can reduce multidimensional equidistribution to single-dimensional equidistribution:

Corollary 1.1.3. *Let $x\colon \mathbf{N} \to \mathbf{T}^d$. Then x is asymptotically equidistributed in $\mathbf{T}^d$ if and only if, for each $k \in \mathbf{Z}^d\backslash\{0\}$, the sequence $n \mapsto k \cdot x(n)$ is asymptotically equidistributed in $\mathbf{T}$.*

Exercise 1.1.4. Show that a sequence $x : \mathbf{N} \to \mathbf{T}^d$ is totally asymptotically equidistributed if and only if one has

$$\lim_{N\to\infty} \mathbf{E}_{n\in[N]} e(k \cdot x(n))e(\alpha n) = 0 \tag{1.3}$$

for all $k \in \mathbf{Z}^d\backslash\{0\}$ and all rational α.

This quickly gives a test for equidistribution for linear sequences, sometimes known as the *equidistribution theorem*:

Exercise 1.1.5. Let $\alpha, \beta \in \mathbf{T}^d$. By using the geometric series formula, show that the following are equivalent:

(i) The sequence $n \mapsto n\alpha + \beta$ is asymptotically equidistributed on $\mathbf{N}$.

(ii) The sequence $n \mapsto n\alpha + \beta$ is totally asymptotically equidistributed on $\mathbf{N}$.

(iii) The sequence $n \mapsto n\alpha + \beta$ is totally asymptotically equidistributed on $\mathbf{Z}$.

(iv) α is *irrational*, in the sense that $k \cdot \alpha \neq 0$ for any non-zero $k \in \mathbf{Z}^d$.

Remark 1.1.4. One can view Exercise 1.1.5 as an assertion that a linear sequence x_n will equidistribute itself unless there is an "obvious" algebraic obstruction to it doing so, such as $k \cdot x_n$ being constant for some non-zero k. This theme of algebraic obstructions being the "only" obstructions to uniform distribution will be present throughout the text.

Exercise 1.1.5 shows that linear sequences with irrational shift α are equidistributed. At the other extreme, if α is *rational* in the sense that $m\alpha = 0$ for some positive integer m, then the sequence $n \mapsto n\alpha + \beta$ is clearly periodic of period m, and definitely not equidistributed.

In the one-dimensional case $d = 1$, these are the only two possibilities. But in higher dimensions, one can have a mixture of the two extremes, that exhibits irrational behaviour in some directions and periodic behaviour in others. Consider for instance the two-dimensional sequence $n \mapsto (\sqrt{2}n, \frac{1}{2}n) \bmod \mathbf{Z}^2$. The first coordinate is totally asymptotically equidistributed in $\mathbf{T}$, while the second coordinate is periodic; the shift $(\sqrt{2}, \frac{1}{2})$ is neither irrational nor rational, but is a mixture of both. As such, we see that the two-dimensional sequence is equidistributed with respect to Haar measure on the group $\mathbf{T} \times (\frac{1}{2}\mathbf{Z}/\mathbf{Z})$.

This phenomenon generalises:

Proposition 1.1.5 (Equidistribution for abelian linear sequences). *Let T be a torus, and let $x(n) := n\alpha + \beta$ for some $\alpha, \beta \in T$. Then there exists a decomposition $x = x' + x''$, where $x'(n) := n\alpha'$ is totally asymptotically equidistributed on $\mathbf{Z}$ in a subtorus T' of T (with $\alpha' \in T'$, of course), and $x''(n) = n\alpha'' + \beta$ is periodic (or equivalently, that $\alpha'' \in T$ is rational).*

Proof. We induct on the dimension d of the torus T. The claim is vacuous for $d = 0$, so suppose that $d \geq 1$ and that the claim has already been proven for tori of smaller dimension. Without loss of generality we may identify T with $\mathbf{T}^d$.

If α is irrational, then we are done by Exercise 1.1.5, so we may assume that α is not irrational; thus $k \cdot \alpha = 0$ for some non-zero $k \in \mathbf{Z}^d$. We then write $k = mk'$, where m is a positive integer and $k' \in \mathbf{Z}^d$ is *irreducible* (i.e., k' is not a proper multiple of any other element of $\mathbf{Z}^d$); thus $k' \cdot \alpha$ is rational. We may thus write $\alpha = \alpha_1 + \alpha_2$, where α_2 is rational, and $k' \cdot \alpha_1 = 0$. Thus, we can split $x = x_1 + x_2$, where $x_1(n) := n\alpha_1$ and $x_2(n) := n\alpha_2 + \beta$. Clearly x_2 is periodic, while x_1 takes values in the subtorus $T_1 := \{y \in T \colon k' \cdot y = 0\}$ of T. The claim now follows by applying the induction hypothesis to T_1 (and noting that the sum of two periodic sequences is again periodic). □

As a corollary of the above proposition, we see that any linear sequence $n \mapsto n\alpha + \beta$ in a torus T is equidistributed in some union of finite cosets of a subtorus T'. It is easy to see that this torus T is uniquely determined by α, although there is a slight ambiguity in the decomposition $x = x' + x''$ because one can add or subtract a periodic linear sequence taking values in T from x' and add it to x'' (or vice versa).

Having discussed the linear case, we now consider the more general situation of *polynomial* sequences in tori. To get from the linear case to the polynomial case, the fundamental tool is

Lemma 1.1.6 (van der Corput inequality). *Let $a_1, a_2, \ldots$ be a sequence of complex numbers of magnitude at most 1. Then for every $1 \leq H \leq N$, we have*

$$|\mathbf{E}_{n\in[N]} a_n| \ll \left(\mathbf{E}_{h\in[H]} |\mathbf{E}_{n\in[N]} a_{n+h}\overline{a_n}|\right)^{1/2} + \frac{1}{H^{1/2}} + \frac{H^{1/2}}{N^{1/2}}.$$

Proof. For each $h \in [H]$, we have

$$\mathbf{E}_{n\in[N]} a_n = \mathbf{E}_{n\in[N]} a_{n+h} + O\left(\frac{H}{N}\right)$$

and hence, on averaging,

$$\mathbf{E}_{n\in[N]} a_n = \mathbf{E}_{n\in[N]} \mathbf{E}_{h\in[H]} a_{n+h} + O\left(\frac{H}{N}\right).$$

Applying Cauchy-Schwarz, we conclude

$$\mathbf{E}_{n\in[N]}a_n \ll (\mathbf{E}_{n\in[N]}|\mathbf{E}_{h\in[H]}a_{n+h}|^2)^{1/2} + \frac{H}{N}.$$

We expand out the left-hand side as

$$\mathbf{E}_{n\in[N]}a_n \ll (\mathbf{E}_{h,h'\in[H]}\mathbf{E}_{n\in[N]}a_{n+h}\overline{a_{n+h'}})^{1/2} + \frac{H}{N}.$$

The diagonal contribution $h = h'$ is $O(1/H)$. By symmetry, the off-diagonal contribution can be dominated by the contribution when $h > h'$. Making the change of variables $n \mapsto n - h'$, $h \mapsto h + h'$ (accepting a further error of $O(H^{1/2}/N^{1/2})$), we obtain the claim. □

Corollary 1.1.7 (van der Corput lemma). *Let $x\colon \mathbf{N} \to \mathbf{T}^d$ be such that the derivative sequence $\partial_h x\colon n \mapsto x(n+h) - x(n)$ is asymptotically equidistributed on $\mathbf{N}$ for all positive integers h. Then x_n is asymptotically equidistributed on $\mathbf{N}$. Similarly with $\mathbf{N}$ replaced by $\mathbf{Z}$.*

Proof. We just prove the claim for $\mathbf{N}$, as the claim for $\mathbf{Z}$ is analogous (and can in any case be deduced from the $\mathbf{N}$ case).

By Proposition 1.1.2, we need to show that for each non-zero $k \in \mathbf{Z}^d$, the exponential sum

$$|\mathbf{E}_{n\in[N]}e(k\cdot x(n))|$$

goes to zero as $N \to \infty$. Fix an $H > 0$. By Lemma 1.1.6, this expression is bounded by

$$\ll (\mathbf{E}_{h\in[H]}|\mathbf{E}_{n\in[N]}e(k\cdot(x(n+h) - x(n)))|)^{1/2} + \frac{1}{H^{1/2}} + \frac{H^{1/2}}{N^{1/2}}.$$

On the other hand, for each fixed positive integer h, we have from hypothesis and Proposition 1.1.2 that $|\mathbf{E}_{n\in[N]}e(k\cdot(x(n+h) - x(n)))|$ goes to zero as $N \to \infty$. Taking limit superior as $N \to \infty$, we conclude that

$$\limsup_{N\to\infty}|\mathbf{E}_{n\in[N]}e(k\cdot x(n))| \ll \frac{1}{H^{1/2}}.$$

Since H is arbitrary, the claim follows. □

Remark 1.1.8. There is another famous lemma by van der Corput concerning oscillatory integrals, but it is not directly related to the material discussed here.

Corollary 1.1.7 has the following immediate corollary:

Corollary 1.1.9 (Weyl equidistribution theorem for polynomials). *Let $s \geq 1$ be an integer, and let $P(n) = \alpha_s n^s + \dots + \alpha_0$ be a polynomial of degree s with $\alpha_0, \dots, \alpha_s \in \mathbf{T}^d$. If α_s is irrational, then $n \mapsto P(n)$ is asymptotically equidistributed on $\mathbf{Z}$.*