Higher-order Fourier Analysis and Applications
Other titles in Foundations and Trends® in Theoretical Computer Science

Coding for Interactive Communication: A Survey
Ran Gelles
ISBN: 978-1-68083-346-1

Hashing, Load Balancing and Multiple Choice
Udi Wieder

Scalable Algorithms for Data and Network Analysis
Shang-Hua Teng
ISBN: 978-1-68083-130-6

Communication Complexity (for Algorithm Designers)
Tim Roughgarden
ISBN: 978-1-68083-114-6

Quantum Proofs
Thomas Vidick and John Watrous
Higher-order Fourier Analysis and Applications

Hamed Hatami
McGill University
hatami@cs.mcgill.ca

Pooya Hatami
Ohio State University
pooyahat@gmail.com

Shachar Lovett
UC San Diego
slovett@cs.ucsd.edu
Editorial Scope

Topics
Foundations and Trends® in Theoretical Computer Science publishes survey and tutorial articles in the following topics:

- Algorithmic game theory
- Computational algebra
- Computational aspects of combinatorics and graph theory
- Computational aspects of communication
- Computational biology
- Computational complexity
- Computational geometry
- Computational learning
- Computational Models and Complexity
- Computational Number Theory
- Cryptography and information security
- Data structures
- Database theory
- Design and analysis of algorithms
- Distributed computing
- Information retrieval
- Operations Research
- Parallel algorithms
- Quantum Computation
- Randomness in Computation

Information for Librarians
Foundations and Trends® in Theoretical Computer Science, 2019, Volume 13, 4 issues. ISSN paper version 1551-305X. ISSN online version 1551-3068. Also available as a combined paper and online subscription.
Contents

1 Introduction 3

1 Low Degree Testing 6

2 Fourier Analytic Property Testing 11
   2.1 Linearity testing ................................. 13
   2.2 Testing for affine linearity ....................... 15
   2.3 Limitations of Fourier analysis ................... 17

3 Low-degree Tests, the 99% Regime 19
   3.1 Basic properties of low-degree polynomials ........ 19
   3.2 Low-degree testing ................................ 22
   3.3 Analysis of the AKKLR test ....................... 23
   3.4 Implications for the Gowers norms ................. 30

4 Low-degree Tests, the 1% Regime 33
   4.1 Completeness ....................................... 35
   4.2 Soundness for $d = 1$ .............................. 36
   4.3 Soundness for $d = 2$ .............................. 37
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>Gowers Norms, the Inverse Gowers Conjecture and its Failure</td>
<td>51</td>
</tr>
<tr>
<td>5.1</td>
<td>Gowers norms</td>
<td>51</td>
</tr>
<tr>
<td>5.2</td>
<td>The counter-example</td>
<td>53</td>
</tr>
<tr>
<td>II</td>
<td>Higher Order Fourier Analysis</td>
<td>54</td>
</tr>
<tr>
<td>6</td>
<td>Nonclassical Polynomials, and the Inverse Gowers Theorem</td>
<td>65</td>
</tr>
<tr>
<td>6.1</td>
<td>Nonclassical polynomials</td>
<td>67</td>
</tr>
<tr>
<td>6.2</td>
<td>The inverse theorem for Gowers norms</td>
<td>69</td>
</tr>
<tr>
<td>7</td>
<td>Rank, Regularity, and Other Notions of Uniformity</td>
<td>73</td>
</tr>
<tr>
<td>7.1</td>
<td>Polynomial factors</td>
<td>75</td>
</tr>
<tr>
<td>7.2</td>
<td>Analytic measures of uniformity</td>
<td>77</td>
</tr>
<tr>
<td>7.3</td>
<td>The derivative polynomial</td>
<td>81</td>
</tr>
<tr>
<td>7.4</td>
<td>Equidistribution of regular factors</td>
<td>83</td>
</tr>
<tr>
<td>7.5</td>
<td>Regularization of factors</td>
<td>84</td>
</tr>
<tr>
<td>7.6</td>
<td>Strong equidistribution of regular factors</td>
<td>88</td>
</tr>
<tr>
<td>7.7</td>
<td>Joint distribution of high-rank polynomials over linear forms</td>
<td>92</td>
</tr>
<tr>
<td>8</td>
<td>Bias vs Low Rank in Large Fields</td>
<td>95</td>
</tr>
<tr>
<td>8.1</td>
<td>Bias implies low rank approximation</td>
<td>96</td>
</tr>
<tr>
<td>8.2</td>
<td>Bias implies low rank exact computation</td>
<td>99</td>
</tr>
<tr>
<td>9</td>
<td>Decomposition Theorems</td>
<td>105</td>
</tr>
<tr>
<td>9.1</td>
<td>Basic decomposition theorem</td>
<td>105</td>
</tr>
<tr>
<td>9.2</td>
<td>Higher-order Fourier expansion</td>
<td>107</td>
</tr>
<tr>
<td>9.3</td>
<td>Strong decomposition theorems</td>
<td>108</td>
</tr>
<tr>
<td>9.4</td>
<td>Sub-atom selection</td>
<td>110</td>
</tr>
<tr>
<td>10</td>
<td>Homogeneous Nonclassical Polynomials</td>
<td>115</td>
</tr>
<tr>
<td>10.1</td>
<td>A homogeneous basis for nonclassical polynomials</td>
<td>118</td>
</tr>
<tr>
<td>11</td>
<td>Complexity of Systems of Linear Forms</td>
<td>123</td>
</tr>
<tr>
<td>11.1</td>
<td>Cauchy-Schwarz complexity</td>
<td>124</td>
</tr>
<tr>
<td>11.2</td>
<td>The True Complexity</td>
<td>126</td>
</tr>
</tbody>
</table>
12 Deferred Technical Proofs

12.1 Near-orthogonality: Proof of Theorem 7.32 .................................. 129
12.2 Proof of Theorem 11.8 ............................................................... 138

13 Algorithmic Regularity

13.1 A lemma of Bogdanov and Viola ............................................. 145
13.2 Algorithmic regularity lemmas ................................................. 147
13.3 Algorithmic inverse theorem for polynomials .......................... 151
13.4 Derandomization via PRGs for polynomials ........................... 154
13.5 Algorithmic Decomposition Theorems ...................................... 156

III Algebraic Property Testing

15 One-Sided Algebraic Property Testing

15.1 Proof overview ................................................................. 161
15.2 Big picture functions .......................................................... 163
15.3 Proof of testability ............................................................. 164

16 Degree Structural Properties

16.1 Proof of Theorem 16.3 .......................................................... 176

17 Estimating the Distance from Algebraic Properties

17.1 Proof sketch of Theorem 17.4 ............................................... 183

IV Open Problems

18 Open Problems

18.1 Testability of hereditary properties ....................................... 190
18.2 Testing correlation with classical polynomials ....................... 191
18.3 Quantitative bounds for inverse theorems ............................ 191
18.4 Complexity of linear forms .................................................. 192
18.5 Norms defined by linear forms. ............................................. 193

References ................................................................. 194
Higher-order Fourier Analysis and Applications

Hamed Hatami$^1$, Pooya Hatami$^2$ and Shachar Lovett$^3$

$^1$McGill University; hatami@cs.mcgill.ca
$^2$Ohio State University; pooyahat@gmail.com
$^3$University of California, San Diego; slovett@cs.ucsd.edu

ABSTRACT

Fourier analysis has been extremely useful in many areas of mathematics. In the last several decades, it has been used extensively in theoretical computer science. Higher-order Fourier analysis is an extension of the classical Fourier analysis, where one allows to generalize the “linear phases” to higher degree polynomials. It has emerged from the seminal proof of Gowers of Szemerédi’s theorem with improved quantitative bounds, and has been developed since, chiefly by the number theory community. In parallel, it has found applications also in theoretical computer science, mostly in algebraic property testing, coding theory and complexity theory.

The purpose of this book is to lay the foundations of higher-order Fourier analysis, aimed towards applications in theoretical computer science with a focus on algebraic property testing.
The purpose of this text is to provide an introduction to the field of higher-order Fourier analysis with an emphasis on its applications to theoretical computer science. Higher-order Fourier analysis is an extension of the classical Fourier analysis. It was initiated by a seminal paper of Gowers [37] on a new proof for Szemerédi’s theorem, and has been developed by several mathematicians over the past few decades in order to study problems in an area of mathematics called additive combinatorics, which is primarily concerned with linear patterns such as arithmetic progressions in subsets of integers. While most of the developments in additive combinatorics were focused on the group $\mathbb{Z}$, it was quickly noticed that the analogous questions and results for the group $\mathbb{F}_2^n$ are of great importance to theoretical computer scientists as they are related to basic concepts in areas such as property testing and coding theory.

Classical Fourier analysis is a powerful tool that studies functions by expanding them in terms of the Fourier characters, which are “linear phase functions” such as $n \mapsto e^{-\frac{2\pi}{N}n}$ for the group $\mathbb{Z}_N$, or $(x_1, \ldots, x_n) \mapsto (-1)\sum a_jx_j$ for the group $\mathbb{F}_2^n$. Note that $n$ and $\sum a_jx_j$ are both linear functions. Fourier analysis has been extremely successful in the study
of certain linear patterns such as three-term arithmetic progressions. For example, if the number of three-term arithmetic progressions in a subset $A \subseteq \mathbb{Z}_N$ deviates from the expected number of them in a random subset of $\mathbb{Z}_N$ with the same cardinality as $A$, then $A$ must have significant correlation with a linear phase function. In other words, the characteristic function of $A$ must have a large non-principal Fourier coefficient. Roth [66] used these ideas to show that every subset of integers of positive upper density contains an arithmetic progression of length 3. However, classical Fourier analysis seems to be inadequate in detecting more complex linear patterns such as four-term or longer arithmetic progressions. Indeed, one can easily construct dense sets $A \subseteq \mathbb{Z}_N$ that do not have significant correlation with any linear phase function, and nevertheless do not contain the number of four-term arithmetic progressions that one expects by considering random subsets of the same cardinality. Hence in order to generalize Roth’s theorem to arithmetic progressions of arbitrary length, Szemerédi [76, 77] departed from the Fourier analytic approach and appealed to purely combinatorial ideas. However, his proof of this major result, originally conjectured by Erdös and Turán [27], provided poor quantitative bounds on the minimal density that guarantees the existence of the arithmetic progressions of the desired length. Later Furstenberg [31] developed an ergodic-theoretic framework and gave a new proof for Szemerédi’s theorem, but his proof was still qualitative. His theory is further developed by - to name a few - Host, Kra, Ziegler, Bergelson, Tao (See e.g. [51], [88], and [10, 82]), and there are important parallels between this theory and higher-order Fourier analysis. Indeed some of the terms that are commonly used in higher-order Fourier analysis such as “phase functions” or “factors” are ergodic theoretic terms.

Generalizing Roth’s original proof and obtaining good quantitative bounds for Szemerédi’s theorem remained a challenge until finally Gowers [37] discovered that the essential idea to overcome the obstacles described above is to consider higher-order phase functions. His proof laid the foundation for the area of higher-order Fourier analysis, where one studies a function by approximating it by a linear combination of few higher-order phase functions. Although the idea of using higher-order phase functions already appears in Gowers’s work [37], it was not
Introduction

until more than fifteen years later that some of the major technical difficulties in achieving a satisfactory theory of higher-order Fourier analysis have been resolved. By now, due to great contributions by prominent mathematicians such as Gowers, Green, Tao, Szegedy, Host, Kra and Ziegler (See [75] and [80] and the references there), there is a deep understanding of qualitative aspects of this theory. However, despite these major breakthroughs, still very little is known from a quantitative perspective as many of the proofs are based on soft analytic techniques, and obtaining efficient bounds is one of the major challenges in this area.

This survey will emphasize the applications of the theory of higher-order Fourier analysis to theoretical computer science, and to this end, we will present the foundations of this theory through such applications, in particular to the area of property testing. In the early nineties, it was noticed by Blum et al. [20] and Babai et al. [6] that Fourier analysis can be used to design a very efficient algorithm that distinguishes linear functions \( f : \mathbb{F}_2^n \rightarrow \mathbb{F}_2 \) from functions that are far from being linear. This initiated the area of property testing, the study of algorithms that query their input a very small number of times and with high probability decide correctly whether their input satisfies a given property or is “far” from satisfying that property. It was soon noticed that generalizing the linearity test of Blum et al. [20] and Babai et al. [6] to other properties such as the property of being a quadratic polynomial requires overcoming the same obstacles that one faces in an attempt to generalize Fourier analytic study of three-term arithmetic progressions to four-term arithmetic progressions. Hence in parallel to additive combinatorics, theoretical computer scientists have also been working on developing tools in higher-order Fourier analysis to tackle such problems. In fact some of the most basic results, such as the inverse theorem for the Gowers \( U^3 \) norm for the group \( \mathbb{F}_2^n \), were first proved by Samorodnitsky [70] in the context of property testing for quadratic polynomials.

In Part I we discuss the linearity test due to Blum et al. [20] and its generalization to higher degree polynomials. We will see how this naturally necessitates the development of a theory of higher-order Fourier analysis. In Part II we present the fundamental results of the theory of higher-order Fourier analysis. Since we are interested in the
applications to theoretical computer science, we will only consider the group $\mathbb{F}_p^n$ where $p$ is a fixed prime, and asymptotics are as $n$ tends to infinity. Higher-order Fourier analysis for the group $\mathbb{Z}_N$, which is of more interest for number theoretic applications, shares the same basic ideas but differs on some technical aspects. For this group, the higher order phase functions, rather than being exponentials of polynomials, are the so called nilsequences. We refer the interested reader to Tao [80] for more details. In Part III we use the tools developed in Part II to prove some general results about property testing for algebraic properties.

Throughout most of the text, we will consider fields of constant prime order, namely $\mathbb{F} = \mathbb{F}_p$ where $p$ is a constant, and study functions from $\mathbb{F}_p^n$ to $\mathbb{R}$, $\mathbb{C}$, or $\mathbb{F}_p$ when $n$ is growing. Our choice is mainly for simplicity of exposition, as there have been recent research that extend several of the tools from higher-order Fourier analysis to large or non-prime fields. We refer the interested reader to a paper by Bhattacharyya et al. [12] for treatment of non-prime fields. In Chapter 8 we will discuss a paper by Bhowmick and Lovett [19] considering the case $\mathbb{F}_p^n$ when $p$ is allowed to grow as a function of $n$. 
References


References


