

Geometric Approximation Algorithms

Sariel Har-Peled

DEPARTMENT OF COMPUTER SCIENCE, UNIVERSITY OF ILLINOIS IN URBANA-CHAMPAIGN, URBANA, ILLINOIS 61801

Current address: Department of Computer Science, University of Illinois in Urbana-Champaign, Urbana, Illinois 61801

E-mail address: sariel@illinois.edu

1991 *Mathematics Subject Classification*. Primary 68U05, 68W25;
Secondary 68P05, 52Cxx

Key words and phrases. Computational Geometry, Geometric approximation algorithms,
Discrete Geometry

Work on this book was partially supported by a NSF CAREER award CCR-0132901 and
NSF AF award CCF-0915984.

Contents

Preface	xi
Chapter 1. The Power of Grids – Closest Pair and Smallest Enclosing Disk	1
1.1. Preliminaries	1
1.2. Closest pair	1
1.3. A slow 2-approximation algorithm for the k -enclosing disk	5
1.4. A linear time 2-approximation for the k -enclosing disk	6
1.5. Bibliographical notes	10
1.6. Exercises	11
Chapter 2. Quadtrees – Hierarchical Grids	13
2.1. Quadtrees – a simple point-location data-structure	13
2.2. Compressed quadtrees	15
2.3. Dynamic quadtrees	20
2.4. Bibliographical notes	24
2.5. Exercises	26
Chapter 3. Well-Separated Pair Decomposition	29
3.1. Well-separated pair decomposition (WSPD)	29
3.2. Applications of WSPD	33
3.3. Semi-separated pair decomposition (SSPD)	40
3.4. Bibliographical notes	43
3.5. Exercises	44
Chapter 4. Clustering – Definitions and Basic Algorithms	47
4.1. Preliminaries	47
4.2. On k -center clustering	49
4.3. On k -median clustering	51
4.4. On k -means clustering	57
4.5. Bibliographical notes	57
4.6. Exercises	59
Chapter 5. On Complexity, Sampling, and ε -Nets and ε -Samples	61
5.1. VC dimension	61
5.2. Shattering dimension and the dual shattering dimension	64
5.3. On ε -nets and ε -sampling	70
5.4. Discrepancy	75
5.5. Proof of the ε -net theorem	80
5.6. A better bound on the growth function	82
5.7. Bibliographical notes	83
5.8. Exercises	84

Chapter 6. Approximation via Reweighting	87
6.1. Preliminaries	87
6.2. Computing a spanning tree with low crossing number	88
6.3. Geometric set cover	94
6.4. Geometric set cover via linear programming	98
6.5. Bibliographical notes	100
6.6. Exercises	100
Chapter 7. Yet Even More on Sampling	103
7.1. Introduction	103
7.2. Applications	106
7.3. Proof of Theorem 7.7	109
7.4. Bibliographical notes	119
7.5. Exercises	119
Chapter 8. Sampling and the Moments Technique	121
8.1. Vertical decomposition	121
8.2. General settings	125
8.3. Applications	128
8.4. Bounds on the probability of a region to be created	130
8.5. Bibliographical notes	131
8.6. Exercises	133
Chapter 9. Depth Estimation via Sampling	135
9.1. The at most k -levels	135
9.2. The crossing lemma	136
9.3. A general bound for the at most k -weight	140
9.4. Bibliographical notes	142
9.5. Exercises	143
Chapter 10. Approximating the Depth via Sampling and Emptiness	145
10.1. From emptiness to approximate range counting	145
10.2. Application: Halfplane and halfspace range counting	148
10.3. Relative approximation via sampling	149
10.4. Bibliographical notes	150
10.5. Exercises	150
Chapter 11. Random Partition via Shifting	151
11.1. Partition via shifting	151
11.2. Hierarchical representation of a point set	155
11.3. Low quality ANN search	158
11.4. Bibliographical notes	160
11.5. Exercises	160
Chapter 12. Good Triangulations and Meshing	163
12.1. Introduction – good triangulations	163
12.2. Triangulations and fat triangulations	164
12.3. Analysis	168
12.4. The result	175
12.5. Bibliographical notes	176

Chapter 13. Approximating the Euclidean Traveling Salesman Problem (TSP)	177
13.1. The TSP problem – introduction	177
13.2. When the optimal solution is friendly	178
13.3. TSP approximation via portals and sliding quadtrees	182
13.4. Bibliographical notes	190
13.5. Exercises	190
Chapter 14. Approximating the Euclidean TSP Using Bridges	191
14.1. Overview	191
14.2. Cuts and bridges	192
14.3. The dynamic programming	198
14.4. The result	202
14.5. Bibliographical notes	202
Chapter 15. Linear Programming in Low Dimensions	203
15.1. Linear programming	203
15.2. Low-dimensional linear programming	205
15.3. Linear programming with violations	208
15.4. Approximate linear programming with violations	209
15.5. LP-type problems	210
15.6. Bibliographical notes	213
15.7. Exercises	215
Chapter 16. Polyhedrons, Polytopes, and Linear Programming	217
16.1. Preliminaries	217
16.2. Properties of polyhedrons	219
16.3. Vertices of a polytope	227
16.4. Linear programming correctness	230
16.5. Bibliographical notes	232
16.6. Exercises	232
Chapter 17. Approximate Nearest Neighbor Search in Low Dimension	233
17.1. Introduction	233
17.2. The bounded spread case	233
17.3. ANN – the unbounded general case	236
17.4. Low quality ANN search via the ring separator tree	238
17.5. Bibliographical notes	240
17.6. Exercises	242
Chapter 18. Approximate Nearest Neighbor via Point-Location	243
18.1. ANN using point-location among balls	243
18.2. ANN using point-location among approximate balls	250
18.3. ANN using point-location among balls in low dimensions	252
18.4. Approximate Voronoi diagrams	253
18.5. Bibliographical notes	255
18.6. Exercises	256
Chapter 19. Dimension Reduction – The Johnson-Lindenstrauss (JL) Lemma	257
19.1. The Brunn-Minkowski inequality	257
19.2. Measure concentration on the sphere	260

19.3.	Concentration of Lipschitz functions	263
19.4.	The Johnson-Lindenstrauss lemma	263
19.5.	Bibliographical notes	266
19.6.	Exercises	267
Chapter 20.	Approximate Nearest Neighbor (ANN) Search in High Dimensions	269
20.1.	ANN on the hypercube	269
20.2.	LSH and ANN in Euclidean space	274
20.3.	Bibliographical notes	276
Chapter 21.	Approximating a Convex Body by an Ellipsoid	279
21.1.	Ellipsoids	279
21.2.	Bibliographical notes	282
Chapter 22.	Approximating the Minimum Volume Bounding Box of a Point Set	283
22.1.	Some geometry	283
22.2.	Approximating the minimum volume bounding box	284
22.3.	Exact algorithm for the minimum volume bounding box	286
22.4.	Approximating the minimum volume bounding box in three dimensions	288
22.5.	Bibliographical notes	289
22.6.	Exercises	289
Chapter 23.	Coresets	291
23.1.	Coreset for directional width	291
23.2.	Approximating the extent of lines and other creatures	297
23.3.	Extent of polynomials	301
23.4.	Roots of polynomials	302
23.5.	Bibliographical notes	306
23.6.	Exercises	306
Chapter 24.	Approximation Using Shell Sets	307
24.1.	Covering problems, expansion, and shell sets	307
24.2.	Covering by cylinders	310
24.3.	Bibliographical notes	313
24.4.	Exercises	313
Chapter 25.	Duality	315
25.1.	Duality of lines and points	315
25.2.	Higher dimensions	318
25.3.	Bibliographical notes	319
25.4.	Exercises	321
Chapter 26.	Finite Metric Spaces and Partitions	323
26.1.	Finite metric spaces	323
26.2.	Random partitions	325
26.3.	Probabilistic embedding into trees	327
26.4.	Embedding any metric space into Euclidean space	329
26.5.	Bibliographical notes	332
26.6.	Exercises	333
Chapter 27.	Some Probability and Tail Inequalities	335

27.1. Some probability background	335
27.2. Tail inequalities	338
27.3. Hoeffding's inequality	342
27.4. Bibliographical notes	344
27.5. Exercises	344
Chapter 28. Miscellaneous Prerequisite	347
28.1. Geometry and linear algebra	347
28.2. Calculus	348
Bibliography	349
Index	357

Preface

Finally: It was stated at the outset, that this system would not be here, and at once, perfected. You cannot but plainly see that I have kept my word. But I now leave my cetological system standing thus unfinished, even as the great Cathedral of Cologne was left, with the crane still standing upon the top of the uncompleted tower. For small erections may be finished by their first architects; grand ones, true ones, ever leave the copestone to posterity. God keep me from ever completing anything. This whole book is but a draft – nay, but the draft of a draft. Oh, Time, Strength, Cash, and Patience!

– Moby Dick, Herman Melville.

This book started as a collection of class notes on geometric approximation algorithms that was expanded to cover some additional topics. As the book title suggests, the target audience of this book is people interested in geometric approximation algorithms.

What is covered. The book describes some key techniques in geometric approximation algorithms. In addition, more traditional computational geometry techniques are described in detail (sampling, linear programming, etc.) as they are widely used in developing geometric approximation algorithms. The chapters are relatively independent and try to provide a concise introduction to their respective topics. In particular, certain topics are covered only to the extent needed to present specific results that are of interest. I also tried to provide detailed bibliographical notes at the end of each chapter.

Generally speaking, I tried to cover all the topics that I believe a person working on geometric approximation algorithms should at least know about. Naturally, the selection reflects my own personal taste and the topics I care about. While I tried to cover many of the basic techniques, the field of geometric approximation algorithms is too large (and grows too quickly) to be covered by a single book. For an exact list of the topics covered, see the table of contents.

Naturally, the field of geometric approximation algorithms is a subfield of both computational geometry and approximation algorithms. A more general treatment of approximation algorithms is provided by Williamson and Shmoys [WS11] and Vazirani [Vaz01]. As for computational geometry, a good introduction is provided by de Berg et al. [dBCvKO08].

What to cover? The material in this book is too much to cover in one semester. I have tried to explicitly point out in the text the parts that are more advanced and that can be skipped. In particular, the first 12 chapters of this book (skipping Chapter 7) provide (I hope) a reasonable introduction to modern techniques in computational geometry.

Intellectual merit. I have tried to do several things that I consider to be different than other texts on computational geometry:

- (A) Unified several data-structures to use compressed quadtrees as the basic building block and in the process provided a detailed description of compressed quadtrees.

- (B) Provided a more elaborate introduction to VC dimension, since I find this topic to be somewhat mysterious.
- (C) Covered some worthy topics that are not part of traditional computational geometry (for example, locality sensitive hashing and metric space partitions).
- (D) Embedded numerous color figures into the text to illustrate proofs and ideas.

Prerequisites. The text assumes minimal familiarity with some concepts in computational geometry including arrangements, Delaunay triangulations, Voronoi diagrams, point-location, etc. A reader unfamiliar with these concepts would probably benefit from skimming or reading the fine material available online on these topics (i.e., Wikipedia) as necessary. Tail inequalities (i.e., Chernoff's inequality) are described in detail in Chapter 27. Some specific prerequisites are discussed in Chapter 28.

Cross-references. For the convenience of the reader, text cross-references to theorems, lemmas, etc., often have a subscript giving the page location of the theorem, lemma, etc., being referenced. One would look like the following: Theorem 19.3_{p257}.

Acknowledgments. I had the benefit of interacting with numerous people during the work on this book. In particular, I would like to thank the students that took the class (based on earlier versions of this book) for their input, which helped in discovering numerous typos and errors in the manuscript. Furthermore, the content was greatly affected by numerous insightful discussions with Jeff Erickson and Edgar Ramos. Other people who provided comments and insights, or who answered nagging emails from me, for which I am thankful, include Bernard Chazelle, Chandra Chekuri, John Fischer, Samuel Hornus, Piotr Indyk, Mira Lee, Jirka Matoušek, and Manor Mendel.

I would especially like to thank Benjamin Raichel for painstakingly reading the text and pointing out numerous errors and typos and for giving guidance on what needed improvement. His work has significantly improved the quality of the text.

I am sure that there are other people who have contributed to this work, whom I have forgotten to mention – they have my thanks and apologies.

A significant portion of the work on this book was done during my sabbatical (taken during 2006/2007). I thank the people that hosted me during the sabbatical for their hospitality and help. Specifically, I would like to thank Lars Arge (Aarhus, Denmark), Sandeep Sen (IIT, New Delhi, India), and Otfried Cheong (KAIST, Daejeon, South Korea).

Work on this book was also partially supported by NSF (CAREER) award CCR-0132901, and AF award CCF-0915984.

Errors. There are without doubt errors and mistakes in the text and I would like to know about them. Please email me about any of them that you find.

Sariel Har-Peled
sariel@uiuc.edu
Urbana, IL
April 2011