Graph Embeddings

D. V. Chudnovsky, G. V. Chudnovsky, T. Morgan
IMAS
NYU Tandon School of Engineering
6 MetroTech Center
Brooklyn, NY 11201

April 21, 2016

Starting With Additive Number Theory

The relevant problem here that seems to be a 1-dimensional problem is known as the optimal Golomb Ruler. Here we are looking for a given (order) N at integers $x_1 < \cdots < x_N$ with a minimal **length** $x_N - x_1$ such that all pairwise differences $x_i - x_j$ for $i > j$ are distinct.

This problem was considered important in radio frequency sampling and antenna placement. This problem is **very** closely related to the problem of Sidon's sets (with B_2 or B_2^* notations, whichever is correct).

This problem is not (yet) proved to be NP-complete, but the general belief is that it is; definitely efforts for determine the optimal length for a given order N take exponential time in N.
For example, the largest order N for which one can prove the optimality of the length is $N = 27$ (for the length of 553). The distributed net run that established that took almost 5 years. The search for $N = 28$ is on for the last 2 years.

Integer Linear Programming Formulation

The relationship between this problem and the set of problems we will talk today.
This problem, just like those bothering us, can be properly reformulated, and reduced to the solution of Integer Linear Programming (ILP) problems, with unknowns/equations of a greatly similar two-dimensional form.

The variables here are $d_{ij} = x_j - x_i$, and the linear problem one needs to solve is

$$\min \sum_{k=1}^{N-1} d_{kk+1}; \;\; d_{ij} = \sum_{k=i}^{j-1} d_{kk+1},$$

with the constraint $\mathrm{AllDifferent}(\{d_{ij}\})$.

With a proper introduction of bipartite matching integer variables this problem reformulates as an ILP Problem in $O(N^3)$ variables, very similar to problems we will introduce below.

Grid Graph Embedding

We consider graphs that are all planar, and are subgraphs of the gird (square lattice), with the underlying l_1 (Manhattan) distance. The problem here is that of embedding one such graph (we call it X) into another (we call it G) bounding the "distortion/dilation" D of the embedding $g : X \to G$:

$$\mathrm{dist}(g(x), g(x')) \leq D$$

for all $x, x', \mathrm{dist}(x, x') = 1$.

In all interesting cases either of two graphs X or G is a full rectangle (typically a square or a "near square").

We make a distinction between the "distortion" D and "dilation" δ depending on which metrics is used. The dilation δ means that we use the graph distance, while "distortion" D means that we measure the distance as l_1. For simply-connected (convex) sets G these are the same.

Just like in studies of the Dimer/Domino tiling and percolation problems on square lattices, the main focus here is on the parameterized "rectilinear domains".

Rectilinear Discrete Domains

One starts with a connected (not necessarily simply connected) domain $\mathcal{D}$ in the square lattice, and replaces every "dot" with an $N \times N$ gird, creating from "fat dots" a domain $\mathcal{D}(N)$ with a cardinality $N^2 \cdot |\mathcal{D}|$.
These and rectangles of arbitrary aspect ratio are typical objects of study.

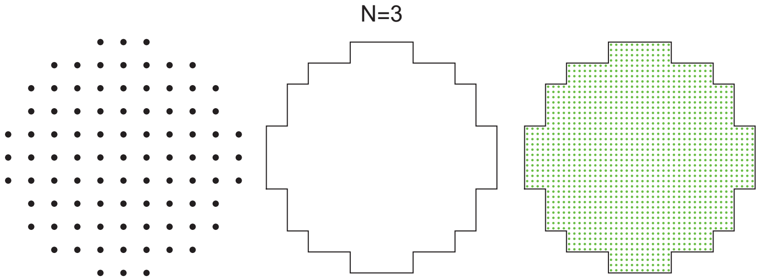


Figure: "circular" domain

Bounded dilation embedding of such graphs (typically to or from a square or near square) initially became very popular through 1980s-1990s in connection with two independent sets of acute problems:

A) The Placement Problem. How to create/move complex shapes of gate/transistor designs in particular area of a bigger VLSI design (core) without much distortion of the wiring. Here typically the mapping is into a square. This problem is still of a vital importance, very often needed in PD work.

B) The Mesh of Computers. How to align complex placement and connection of computers into conventional two-dimensional grids. Particular attention was devoted to n-cube virtual configurations of computer clusters, with possibly failing computer nodes, remapped as rectangles in $2D$ grids.

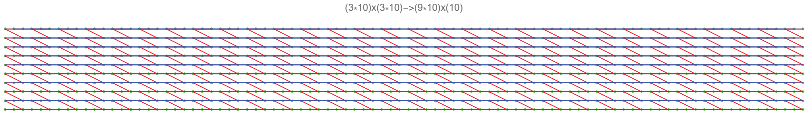
Mappings Between Rectangles and Squares

It may be counterintuitive (unlike a toothpaste): it is easier (lower dilation) to map a skinny rectangle (say, a line $N \times 1$) into a square ($N \times N$) than vice versa.

Use local modular transformations (unlike global affine transformation) to create mappings between rectangles. These transformations are of 3 categories:

stretching, squeezing, folding (but not "shaking").

An example of the optimal mapping of the $(3 \cdot 10)^2$ square onto the $(9 \cdot 10) \times (10)$ rectangle.



Boolean Linear Programming Framework

To get the optimal solution of the graph embedding problem with a given Distortion (D) or a Dilation (δ) one formulates it as an Integer Programming problem, similar to what maximal graph matching problem or TSP/Hamiltonian Path problem looks like.

To study the 1-1 embedding $g = g(x)$ of the graph X into G one introduces auxiliary Boolean: 0/1 variables $F[g, x]$ with a meaning that $F[g, x] = 1$ iff $g = g(x)$.
Then the basic equations that describing that this is a 1-1 embedding of all X into G are:

$$\sum_{g \in G} F[g, x] \; = \; 1 \qquad (1)$$

for every vertex $x \in X$, and

$$\sum_{x \in X} F[g, x] \; \leq \; 1 \qquad (2)$$

for every vertex $g \in G$.
One can also use alternative formulation focusing on edges of X, G instead.

Discrepancy/Dilation part

The equations (1), (2) are really just a graph matching problem, solvable in a polynomial (or often in nearly linear) time. The important part here is the control of mapping "dilation" D. Various equivalent formulations are possible:

$$F[g,x] - \sum_{g' \in G, \mathrm{dist}(g,g') \le D} F[g',x'] \ \le \ 0 \tag{3}$$

for all $g, g' \in G$, $x, x' \in X$, $\mathrm{dist}(x,x') = 1$, or instead:

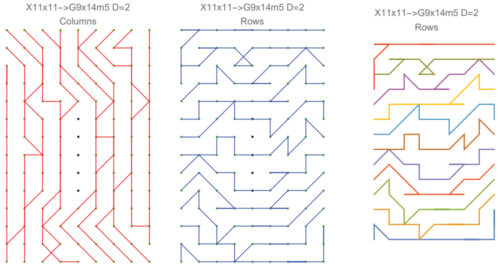
$$F[g,x] + F[g',x'] \ \le \ 1 \tag{4}$$

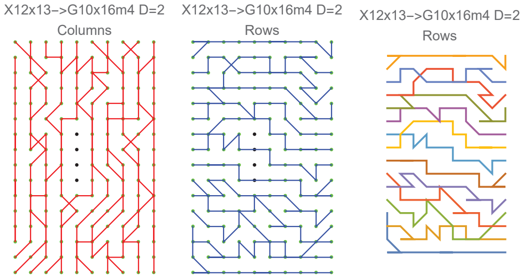
for all $g, g' \in G$, $\mathrm{dist}(g,g') > D$, $x, x' \in X$, $\mathrm{dist}(x,x') = 1$.

Here (4) provides more equations than (3), but the system (2),(4) has a huge advantage. It forms a totally unimodular matrix. The ILP problem for a totally unimodular matrix is simply reducible to just an LP problem (those can be solved in a polynomial time).

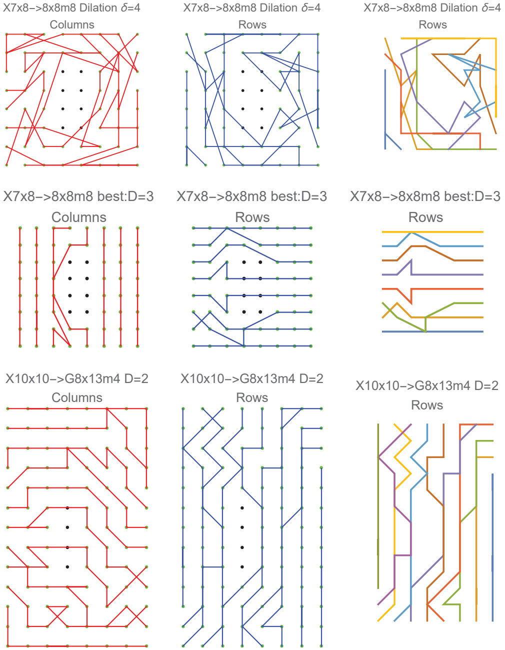
While (1) spoils the unimodularity (otherwise NP=P), the (2), (4) parts help with solving rather giant ILP problems.

Small Examples with $g = 1$ Topology but Computed Dilation I





Small Examples with $g = 1$ Topology but Computed Dilation II



Lower Bounds for Distortion/Dilation

The simplest low bound for the dilation of the mapping of X into G is the ratio of the diameters of G and X.

For better bounds one need vertex-isoperimetric inequalities for the rectilinear domain, particularly for rectangles, determined by Bollobas and Leader as nested sets.

These inequalities provide the best scan order for rectangles. In the scan order two points $\bar{x}$, $\bar{y}$ in a square $[k]^2$ are ordered as

$$\bar{x} <_S \bar{y} \text{ iff } \parallel \bar{x} \parallel_{l_1} < \parallel \bar{y} \parallel_{l_1}$$

or when

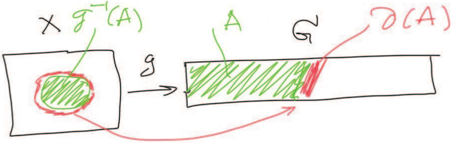
$$\parallel \bar{x} \parallel_{l_1} = \parallel \bar{y} \parallel_{l_1} \text{ and } x_1 > y_1$$

Such order finds the "smallest" boundary of the set $A \subset [k]^2$ for a given cardinality $\mid A \mid$ just by looking at the first $\mid A \mid$ elements of $[k]^2$ in the scan order.

Sketches

To get lower bounds we look at "front propagation" of the scan order in the source graph X (typically a square) and watch the boundary of its image expanding in the target G. One can use the vertex isoperimetric order moving it back to the preimage $g^{-1}(G)$ in X. This shows the extent of a "front propagation" of g.
When X is a square of a linear size H, and G is a rectangle of (nearly) the same size but with the height $h < H$, we can follow a "diagonal" of X with $\| \bar{x} \|_{l_1} = a$ with $a \sim H$. The size of the largest boundary is then (as follows from the nested scan order) H.

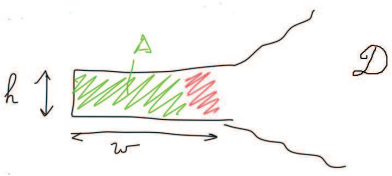
Thus the dilation of the map $g : X \to G$ becomes bounded by $\frac{H}{h}$.



The low bound is actually met in the squeezing map.

General Domains

This argument allows to get lower bounds for rectilinear domains with rectangular $w \times h$ "protrusions" with $w \gg h$, where we are determining the lower bound of the boundary of a "filled" set A.



Using the vertex-isoperimetric inequality in the square source X on $g^{-1}(A)$ we can see that for a set $C \subset X$ of the cardinality $\mid C \mid$ the smallest boundary is $\sqrt{2 \mid C \mid}$. Then somewhere at the end of the $w \times h$ protrusion, the set of $h \cdot w$ vertices in G has a boundary of the size $\sqrt{2hw}$, and for the height of h we get the width of the boundary as

$$\sqrt{2w/h}$$

This becomes a low bound on the dilation of the mapping from the square onto the domain $\mathcal{D}$.

Percolation

Typically G (in $X \to G$) is not a solid domain. In computer grids or a standard cell design, there are locations that are either:
deliberately left unused (for redundancy or openings in packaging/routing), or are getting defective over time.

Deliberate destruction of vertices in G can simply render G unconnected. Instead a realistic scenario is that of "percolation model". One treats missing vertices as occurring randomly with a probability p, $0 < p < 0.42$ (below the critical threshold where the connectivity of G is statistically insured). If one uses p as a probability of a site missing (even if the distribution is not Poisson) then all lower bounds on the dilation acquire an extra factor of $(1 + p)$.

E.g. protrusions dilation lower bound becomes

$$\sqrt{\frac{2w}{h}} \cdot (1 + p)$$

At high percolation rate, in order to reduce the dilation often it makes sense to remove some of the "protrusions".

General Upper Bound for Rectilinear Domains

For a given (connected) rectilinear domain $\mathcal{D}$ and its $N \times N$ "dots" blow-up domain $\mathcal{D}(N)$ one can prove a reasonable upper bound:

Rectilinear Distortion/Dilation Upper Bound. For all sufficiently large N ($N \geq C(\mathcal{D})$), there is an embedding of the "nearly same size square" $X(N)$ into the domain $\mathcal{D}(N)$ with a constant discrepancy.

Here "nearly same size square" $X(N)$ has a linear dimension of

$$\sqrt{|\mathcal{D}|} \cdot N - o(N)$$

for sufficiently large N.

In the opposite direction – embedding of general connected grid graphs into its nearly ideal square – there is no constant upper bound for a domain with holes, but there is a constant upper bound of $D = 5$ for arbitrary rectangles.

Known and Still Unknown

Fundamental dilation problems of mappings between rectangles are still open. There is a difference between embedding rectangles into rectangles with a larger aspect ratio –

there the lower bound (from isoperimetric inequalities) and upper bounds (from modular matrices) meet,

and mappings into rectangles with a smaller aspect ratio –
there the best upper bound is still unknown.

The lower bound is 2, but the best upper bound of embedding the $h \times w$ rectangle to a lower aspect ratio rectangle $H \times W$ with

$$h < H \leq W < w$$

of nearly the same size (i.e. $H \cdot (W - 1) < h \cdot w \leq H \cdot W$) is still **unknown**.

There are no counter-examples for the dilation $D = 2$, and the best upper bound is $D = 5$.

Complexity

Because the bandwidth problem for grid subgraphs is *NP*-complete, even simplest distortion/dilation problems are *NP*-complete as well.

1. The problem of embedding of an n-vertex grid subgraph into its "ideal square" (i.e. of the linear size of $\lceil N \rceil$) with the dilation/distortion of at most k is *NP*-complete.

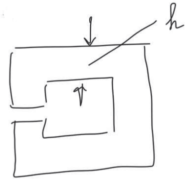
Even seemingly simpler problem is very hard, in general.
2. Let G be a subgraph, with n vertices, of a Grid graph $G(h, w)$. The problem of embedding of G into a Grid $h \times \frac{N}{h}$ with the dilation 1 is *NP*-complete.

Also, the problem of $1-1$ embedding of the line onto the grid graph with dilation 1 (i.e., a problem of finding a Hamiltonian path in such a graph) is *NP*-complete.

While the embedding of the minimal dilation is very hard to find, for many classes of grid subgraphs nearly optimal dilation embeddings are possible to find using ILP methodology.

Conformal Mappings...

Relations to Koebe(-Andreev-Thurston) coin packing theorem, and its variation.
However, in the digital/discrete case we can have embedding from multiply-connected grid domains to squares with bounded dilations.



The embeddings we need are not actually distance-preserving or angle-preserving. However one can use the conformal mappings as an initial approximation to dilation bounded mappings in the ultimate ILP solution problem (including the multiply-connected cases).

Some References

R. Aleliunas, A. Rosenberg, On embedding rectangular grids in square grids, IEEE Trans. Comput. C-31 (9) (1982) 907-913.

B. Bollobas, I. Leader, Compressions and isoperimetric inequalities, J. Combin. Theory A-56 (1991), 47-62.

Ellis, John, "Embedding Graphs in Lines, Trees and Grids", Ph.D. Thesis, Northwestern Univ., Evanston, IL, 1984.

S.-H. Huang, H.L. Liu, R.M. Verma: On Embedding Rectangular Meshes into Rectangular Meshes of Smaller Aspect Ratio. Inf. Proc. Let. 63, (1997), 123129.