Konrad-Zuse-Zentrum
für Informationstechnik Berlin

Takustraße 7
D-14195 Berlin-Dahlem
Germany

AGNES DITTEL[1]
ARMIN FÜGENSCHUH[2]
ALEXANDER MARTIN[3]

Polyhedral Aspects of Self-Avoiding Walks

[1]Fraunhofer Institut für Techno- und Wirtschaftsmathematik, Kaiserslautern, Germany
[2]Konrad Zuse Zentrum für Informationstechnik Berlin, Abteilung Optimierung, Berlin, Germany
[3]Friedrich-Alexander Universität Erlangen-Nürnberg, Fachbereich Mathematik, Erlangen, Germany

ZIB-Report 11-11 (April 2011)

POLYHEDRAL ASPECTS OF SELF-AVOIDING WALKS

AGNES DITTEL, ARMIN FÜGENSCHUH, AND ALEXANDER MARTIN

Abstract. In this paper, we study self-avoiding walks of a given length on a graph. We consider a formulation of this problem as a binary linear program. We analyze the polyhedral structure of the underlying polytope and describe valid inequalities. Proofs for their facial properties for certain special cases are given. In a variation of this problem one is interested in optimal configurations, where an energy function measures the benefit if certain path elements are placed on adjacent vertices of the graph. The most prominent application of this problem is the protein folding problem in biochemistry. On a set of selected instances, we demonstrate the computational merits of our approach.

1. Introduction

A path in a graph is a sequence of adjacent vertices. A path is called simple if multiple occurrence of vertices is prohibited. Paths in graphs give rise to various optimization questions. One of the most prominent is the shortest-path problem, where one is interested in an optimal connection between two given distinct vertices of the graph with respect to certain edge weights. If the edge weights are all positive or, more general, if there is no cycle with negative total weight, then an optimal path is automatically a simple path. Moreover, in this case, the solution of the shortest-path problem can be obtained in polynomial time complexity.

Our work is motivated by a field of applications in physical chemistry, where linear polymer molecules are modeled as simple paths in graphs featuring a certain regularity. These graphs are then referred to as *lattices*, and a simple path in this context is called a *self-avoiding walk* on the lattice. A prominent example is the protein folding problem which refers to the assembly ("*folding*") of a three-dimensional structure of a polypeptide molecule, which is a linear polymer consisting of amino acids, in an aqueous solvent. Formulations of this problem as binary linear programs were given in [10, 17]. Our work contributes to a deeper understanding of the respective underlying polytopes. In particular, we are interested in the convex hull of the incidence vectors of self-avoiding walks. To the best of our knowledge, a polyhedral analysis of families of valid inequalities was not done so far. Under certain conditions we are able to prove facet-defining criteria for some substructures of interest.

The outline of the remainder of this article is the following. In Section 2 we introduce the necessary mathematical description of the problem. In Section 3, we state a complete outer description for $P^{(2)}$ by facet-defining inequalities. In the general case, one technical difficulty arising with the description of the facial structure of $P^{(n)}$ is the lack of dimensionality of these polytopes. We therefore consider their *down-monotonization* as a full-dimensional relaxation. The down-monotonization of $P^{(n)}$ yields the submonotone SAW-n polytope $P^{(n;\leq)}$ which we study in Section 4. We describe the structure of valid inequalities, and we provide a facet characterization for two special cases. In Section 5, we demonstrate the application of cutting planes derived from the polyhedral structures of $P^{(2)}$ and $P^{(n;\leq)}$, respectively. Section 6 contains a conclusion and an outlook to further research opportunities.

Date: April 11, 2011.

2. Problem Description

Let $G = (V, E)$ be a graph. A path in G is a sequence

$$\omega = (\omega_0, \dots, \omega_m)$$

with $\omega_i \in V$ for all $i \in \{0, \dots, m\}$ and $\{\omega_{i-1}, \omega_i\} \in E$ for all $i \in \{1, \dots, m\}$. We alternatively call ω an m-step path (which refers to the number of its edges) or a path of length $m+1$ (which refers to the number of its vertices). A self-avoiding walk (SAW) on G is a path in G without repetition of vertices, i.e., $\omega_i \neq \omega_j$ for $i \neq j$. At this point, we remark that the term "self-avoiding walk" is typically associated with a *lattice*, which can be seen as an infinite graph featuring a certain regular structure (usually originating from a regular tiling of the plane or space). Although the results of this paper are valid for general graphs, we put the focus on finite subgraphs of regular lattices, which we denote as $x \times y (\times z)$ lattices. Examples of regular grid graphs include $x \times y$ square ($Q_{x \times y}$) and triangular ($T_{x \times y}$) lattices in two dimensions, as well as $x \times y \times z$ cubic ($Q_{x \times y \times z}$) and tetrahedral ($T_{x \times y \times z}$) lattices in three dimensions. As an example, the 3×3 square lattice $Q_{3\times 3}$ is shown in Figure 1. As a convention, we number the vertices consecutively, starting with zero, as indicated in Figure 1.

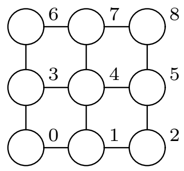

Figure 1. 3×3 square lattice

In order to emphasize the connection to lattice graphs, we use the term "self-avoiding walk" instead of "path". In the following, we consider SAWs in the context of their vertices and therefore denote the set of all SAWs of length n on G (or $(n-1)$-step SAWs, respectively) by $\Omega_G^{(n)}$. We set $S^{(n)} = \{0, \dots, n-1\}$ and define the incidence vector for an SAW $\omega \in \Omega_G^{(n)}$ as

$$x(\omega) = (x(\omega)_v^s)_{v \in V, s \in S^{(n)}} \quad \text{with} \quad x(\omega)_v^s = \begin{cases} 1, & \text{if } \omega_s = v, \\ 0, & \text{otherwise.} \end{cases}$$

The aim of this paper is the investigation of the convex hull $P^{(n)}$ of the incidence vectors for all SAWs of length n on a given graph G, i.e., we are going to study the structure of the polytope

$$P^{(n)} = \operatorname{conv}\{x(\omega) \mid \omega \in \Omega_G^{(n)}\}$$

which we call the SAW-n polytope.

We start with the introduction of a terminology which enables a set-based representation for SAWs of length n on a graph $G = (V, E)$. For the remainder of this paper, we assume that G is connected and consists of at least two vertices, and we assume $n \geq 2$ to be fixed. Next, we introduce the SAW-n graph associated to a graph G as the expansion

$$G^{(n)} = \left(V^{(n)}, E^{(n)}\right)$$

where $V^{(n)} = V \times S^{(n)}$ and $E^{(n)} = \bigcup_{j \in S^{(n)} \setminus \{0\}} \{\{(v, j-1), (w, j)\} \mid \{v, w\} \in E\}$.
In the following, we mention two properties of the SAW-n graph. The proofs for the respective statements can be found in [16].

Property 2.1. *The SAW-n graph $G^{(n)}$ is bipartite with the partition*

$$V^{(n)} = V \times S_0^{(n)} \,\dot\cup\, V \times S_1^{(n)},$$

where $S_0^{(n)}$ contains all even and $S_1^{(n)}$ all odd elements of $S^{(n)}$, i.e.,

$$S_k^{(n)} = \left\{ j \in S^{(n)} \mid j \equiv k_{(\mathrm{mod}\, 2)} \right\} \quad \textit{for} \quad k = 0, 1.$$

Property 2.2.

a) *The number* $\kappa\left(G^{(n)}\right)$ *of connected components of* $G^{(n)}$ *is at most two.*
b) $G^{(n)}$ *is connected if and only if* G *is non-bipartite.*

Example 2.3. *Consider the* 3×3 *square lattice* $G = Q_{3\times 3}$. *Since* G *is bipartite, the corresponding SAW-n graph* $G^{(n)}$ *consists of two connected components for each* $n \geq 2$. *Figure 2 shows the corresponding SAW-4 graph and the decomposition into its two connected components.*

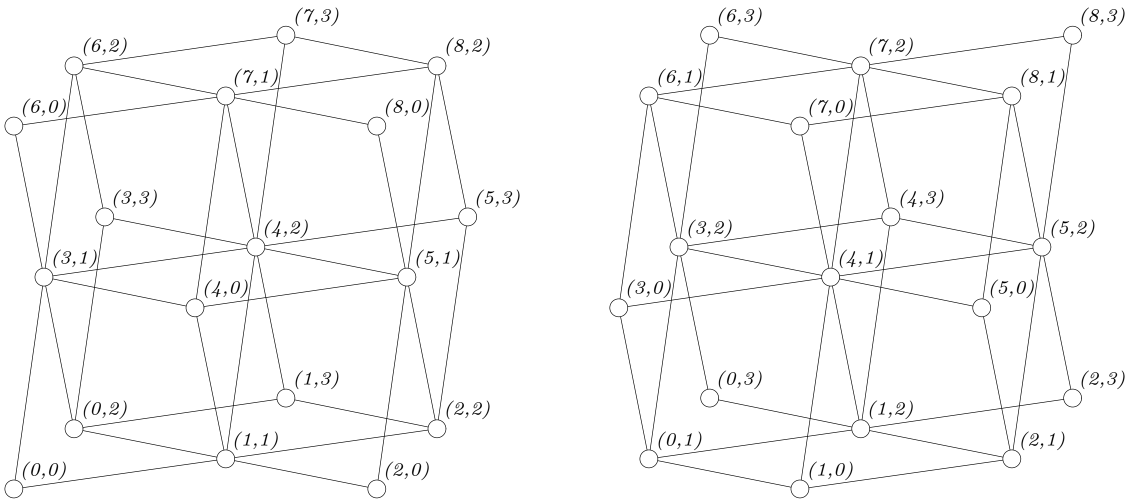

FIGURE 2. Connected components of the SAW-4 graph corresponding to the 3×3 square lattice

The SAW-n graph enables the representation of an SAW of length n on G as a subset of its vertices. For this, we introduce the following terminology.

Definition 2.4. *Let* $G = (V, E)$ *be a graph and* $G^{(n)}$ *the corresponding SAW-n graph.*

a) *An* SAW-n conformation *in* $G^{(n)}$ *is a set*

$$\psi = \left\{(v, s) \in V^{(n)} \mid \omega_s = v \;\; \textit{for some} \;\; \omega \in \Omega_G^{(n)}\right\}.$$

b) *We denote the set of all SAW-n conformations in* $G^{(n)}$ *with* $\Psi_G^{(n)}$.
c) *For an SAW-n conformation* $\psi \in \Psi_G^{(n)}$ *in* $G^{(n)}$, *the corresponding* SAW-n vector *is given by the incidence vector* $\chi^{\psi} \in \{0,1\}^{|V^{(n)}|}$.
d) *We denote the set of SAW-n vectors in* $G^{(n)}$ *with* $X_G^{(n)} = \{\chi^{\psi} \mid \psi \in \Psi_G^{(n)}\}$.

Thus the SAW-n polytope $P^{(n)}$ is given by $P^{(n)} = \operatorname{conv}(X_)^{(n)}$. In the sequel, we present a straightforward description of $P^{(n)}$ by linear constraints and integrality conditions which has already been stated in [10, 17]. Throughout this article we will refer to this as the classical 0/1 model.
For each vertex $v \in V$ and for each element $s \in S^{(n)}$, we introduce a binary variable x_v^s with

$$x_v^s = \begin{cases} 1, & \text{if } \omega_s = v, \\ 0, & \text{otherwise.} \end{cases} \tag{1}$$

In order to guarantee the correct representation of a SAW of length n by the x-variables, the set of possible assignments of these variables has to be restricted by the following constraints:

- **Total Deployment**
  Each element $s \in S^{(n)}$ must occupy exactly one vertex $v \in V$:

(2)
$$\sum_{v\in V} x_v^s = 1, \quad \forall\ s \in S^{(n)}.$$

- **Self-Avoidance**
  Each vertex $v \in V$ can be occupied by at most one element $s \in S^{(n)}$:

(3)
$$\sum_{s\in S^{(n)}} x_v^s \leq 1, \quad \forall\ v \in V.$$

- **Contiguity**
  Successive elements $s, s+1 \in S^{(n)}$ have to occupy adjacent vertices of G:

(4)
$$x_v^s \leq \sum_{w\in \delta_G(v)} x_w^{s+1}, \quad \forall\ v \in V,\ s \in S^{(n)}\backslash\{n-1\},$$

  where $\delta_G(v)$ denotes the set of all vertices adjacent to v in G.

In the sections below, we are going to study certain polytopes related to $P^{(n)}$. Of particular interest in this context are the inequalities defining their facets. Having regard to this issue, we introduce the following notation which will be used throughout this article.

Definition 2.5. *Let P be a polyhedron and $a^T x \leq \alpha$ an inequality.*

a) *An element $\overline{x} \in P$ is called a* root *of $a^T x \leq \alpha$ if $a^T\overline{x} = \alpha$. The set of all such roots is denoted as* $\mathrm{eq}(P; a^T x \leq \alpha)$.
b) *We call (a subset of) an SAW-n conformation ψ an* SAW-n root (sub-) conformation *of $a^T x \leq \alpha$ for the associated polytope P if its incidence vector χ^ψ constitutes a root of $a^T x \leq \alpha$.*

3. The SAW-2 Polytope

The aim of this section is the investigation of a class of polytopes for the representation of one-step paths (which are naturally self-avoiding). For these polytopes, we provide a complete outer description by facet-defining inequalities. We observe that the set of 1-step SAWs on a graph G can be bijectively mapped to the set of edges of the SAW-2 graph $G^{(2)}$ in the sense that an edge $e = \{(v,s), (w, 1-s)\} \in E^{(2)}$ is assigned the 1-step SAW $\omega \in \Omega_G^{(2)}$ given by $\omega_s = v$, $\omega_{1-s} = w$. Thus an SAW-2 conformation in $G^{(2)}$ is equivalent to an edge of $G^{(2)}$.

3.1. **Dimension.** We provide an upper bound for the dimension of $P^{(n)}$ which is given by a linear independent set of valid equations.

Theorem 3.1. *For a given graph G, an upper bound for the dimension of the corresponding SAW-n polytope $P^{(n)}$ is given by*

$$\dim P^{(n)} \leq \begin{cases} n\cdot|V| - n - (n-1), & \text{if } G \text{ is bipartite,} \\ n\cdot|V| - n, & \text{if } G \text{ is not bipartite.} \end{cases}$$

Proof. Since $P^{(n)} \subseteq [0,1]^{|V^{(n)}|}$, its dimension is at most $|V^{(n)}| = n\cdot|V|$. Each vertex of $P^{(n)}$ satisfies the total deployment conditions (2). For $S^{(n)} = \{0, ..., n-1\}$, we denote the deployment condition for the element $s \in S^{(n)}$ with $\mathrm{TD}(s)$. The set $\{\mathrm{TD}(s) \mid s \in S^{(n)}\}$ of all deployment equations is linearly independent. Consequently, each of these n equations reduces the upper bound of the dimension of $P^{(n)}$ by one. If moreover $G = (V, E)$ is bipartite with vertex partition $V = V_I \,\dot\cup\, V_{II}$, there are $n-1$ additional parity equations

(5)
$$\sum_{v\in V_I} \left(x_v^{s-1} + x_v^s\right) = 1 \quad \forall\, s \in S^{(n)}\backslash\{0\}.$$

Then this set $\{\mathrm{PAR}_I(s) \mid s \in S^{(n)}\backslash\{0\}\}$ of equations, joined with the set $\{\mathrm{TD}(s) \mid s \in S^{(n)}\}$ of deployment equations, is linearly independent, yielding the theorem. □