DIMACS Series in Discrete Mathematics
and Theoretical Computer Science
Volume 4, 1991

Applied Geometry
and Discrete Mathematics
THE VICTOR KLEE FESTSCHRIFT

Recognizing Properties of Periodic Graphs

EDITH COHEN AND NIMROD MEGIDDO

ABSTRACT. A periodic (dynamic) graph is an infinite graph with a repetitive structure and a compact representation. A periodic graph is represented by a finite directed graph, called the dependence or the static graph, with d-dimensional integer vector weights associated with its edges. For every vertex in the dependence graph there corresponds a d-dimensional lattice in the periodic graph. For every edge (u, v) in the dependence graph with vector weight $\mathbf{a}$, there are infinitely many edges in the periodic graph, namely, from every point on the lattice corresponding to u to the point shifted by $\mathbf{a}$ on the lattice corresponding to v. Periodic graphs are used, for example, to model VLSI circuits and systems of uniform recurrence relations. In this paper we give algorithms to compute weakly connected components, to test bipartiteness, and to compute a minimum average cost spanning tree for d-dimensional periodic graphs.

1. Introduction

Periodic graphs are infinite graphs with a repetitive structure. They have a finite description (the "period") given by a directed graph with integer vector weights associated with the edges. A more formal definition follows.

DEFINITION 1.1. Given $G = (V, E, \mathbf{f})$, where $\mathbf{f}: E \to Z^d$ is a weight function on the edges of G, the *periodic* (*dynamic*) graph $G^* = (V^*, E^*)$ defined by G has:

$$V^* = \{(\mathbf{z}, v) | \mathbf{z} \in Z^d, \ v \in V\},$$
$$E^* = \{((\mathbf{z}, u); (\mathbf{z} + \mathbf{f}(u, v), v)) | (u, v) \in E\}.$$

The *dimension* of G^* is d. The graph G is called the *dependence* or the *static* graph. For an edge $(u, v) \in E$, the edges *generated* by (u, v) are $((\mathbf{z}, u); (\mathbf{z} + \mathbf{f}(u, v), v))$ $(\mathbf{z} \in Z^d)$.

See Figures 1 and 2 for examples of one- and two-dimensional periodic graphs, respectively. Note that for our purposes the periodic graph as defined

1980 *Mathematics Subject Classification* (1985 *Revision*). Primary 05CXX; Secondary 68R10, 68Q35.

Research partially supported by NSF PYI Grant CCR-8858097.

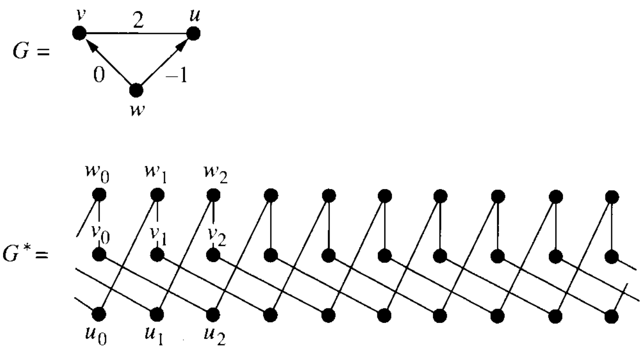

FIGURE 1. Example of a one-dimensional periodic graph.

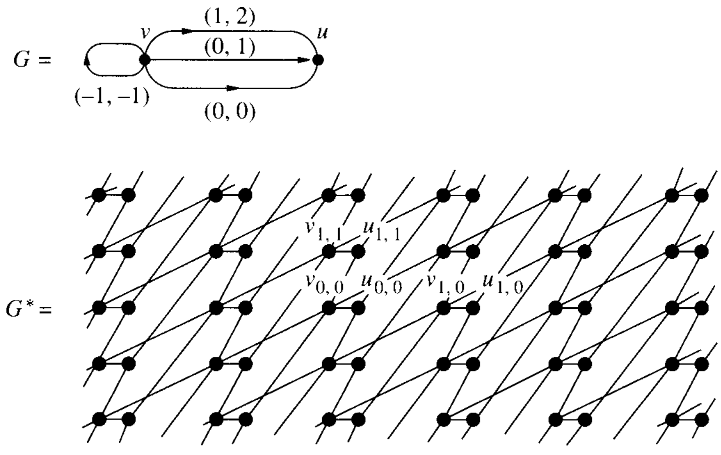

FIGURE 2. Example of a two-dimensional periodic graph.

above may be viewed as undirected. Therefore, a directed edge $(u, v) \in E$ in the dependence graph with $\mathbf{f}(u, v) = \mathbf{z}$ is equivalent to a directed edge (v, u) with weight $\mathbf{f}(v, u) = -\mathbf{z}$. The undirected infinite periodic graph defined by G does not change if (u, v) is replaced by (v, u). Periodic graphs can be used to model a system of uniform recurrence equations [**12**] and VLSI circuits [**8, 9**]. We are interested in testing properties of a periodic graph by working on the defining finite dependence graph. Previous work on periodic graphs was concerned with scheduling [**2, 12, 15**], planarity testing [**10**], cycle detection [**1, 2, 9, 12, 13**], and finding strongly connected components [2] of directed periodic graphs. Orlin [**14**] defined the concept of a one-dimensional dynamic/periodic graph and discussed the complexity of recognizing some properties of it. In particular, he considered finding the connected and strongly connected components, recognizing bipartiteness, and computing a minimum average cost spanning tree. Iwano [7] studied two-dimensional periodic graphs and gave algorithms that test bipartiteness, and compute connected components.

The current paper generalizes some of the results of Orlin and of Iwano to higher dimensions. We consider here the problems of recognizing connectivity, bipartiteness, and computing a minimum average cost spanning tree. The periodic graph of Figure 1 is bipartite and has three connected components. The periodic graph of Figure 2 is connected but is not bipartite.

§2 gives an algorithm for the connectivity problem. This algorithm is used in the succeeding sections. It is a generalization of the algorithms of [**14**, 7]. The connectivity algorithm computes dependence graphs whose corresponding periodic graphs are isomorphic to the connected components of the original periodic graph. §3 gives an algorithm that checks whether the periodic graph is bipartite. §4 gives an algorithm for finding a tree that spans a given periodic graph and has an asymptotically minimal average cost. The problems of recognizing bipartiteness and of computing a minimum average cost spanning tree in one-dimensional periodic graphs were studied by Orlin [**14**]. The problems of finding connected components and testing bipartiteness in two-dimensional periodic graphs were studied by Iwano [7]. In this paper we generalize these results to higher dimensions.

2. Connectivity

PROBLEM 2.1. Given $G = (V, E, \mathbf{f})$, recognize whether or not the periodic graph G^* is connected. If G^* is not connected, find $G_i = (V_i, E_i, \mathbf{f}_i)$, $i = 1, \ldots, r$, such that the corresponding periodic graphs G_i^* are isomorphic to the connected components of G^*.

An algorithm that solves Problem 2.1 for one-dimensional periodic graphs was given by Orlin [**14**]. An algorithm for two-dimensional periodic graphs was given by Iwano [7]. Orlin also proved that if the dependence graph G is connected, then all the connected components of G^* are isomorphic, and therefore only one dependence graph suffices for describing the connected components of G^*. We show that the same is true in higher dimensions.

The following proposition is a restatement of Lemma 4 of [**14**]. The proof given there works for any dimension.

DEFINITION 2.2. Let $\mathbf{f}: E \to R^d$ be a weight function on the edges of a graph. A function $\mathbf{d}: V \to R^d$ is called a *potential*. Denote by $\mathbf{f}^{(\mathbf{d})}: E \to R^d$ the function $\mathbf{f}^{(\mathbf{d})}(i, j) = \mathbf{f}(i, j) + \mathbf{d}(i) - \mathbf{d}(j)$.

PROPOSITION 2.3. *For any potential function* $\mathbf{d}: V \to Z^d$, *if* $G = (V, E, \mathbf{f})$ *and* $G' = (V, E, \mathbf{f}^{(\mathbf{d})})$, *then* G^* *is isomorphic to* $(G')^*$ *and the isomorphism* $\mathscr{H}: V^* \to (V')^*$ *is given by* $\mathscr{H}(\mathbf{z}, i) = (\mathbf{z} + \mathbf{d}(i), i)$.

DEFINITION 2.4. If $G = (V, E, \mathbf{f})$ contains a spanning tree $T \subset E$ such that for all the edges $e \in T$, $\mathbf{f}(e) = \mathbf{0}$, then G is in a *basic form*.

The following corollary of Lemma 4 of [**14**] generalizes to vector weights.

COROLLARY 2.5. *Given $G = (V, E, \mathbf{f})$ and a spanning tree $T \subset E$ of G, a potential function $\mathbf{d}$ can be computed such that for the dependence graph with new weights $G' = (V, E, \mathbf{f}^{(\mathbf{d})})$, the periodic graph $(G')^*$ is isomorphic to G^* and $\mathbf{f}^{(\mathbf{d})}(e) = \mathbf{0}$ for all $e \in T$.*

PROOF. This is done by computing the distances on the tree T from one vertex to all the other vertices, and choosing $\mathbf{d}\colon V \to Z^d$ to be the distance function. $\square$

We refer to the above procedure as a transformation of G into a basic form.

DEFINITION 2.6. Given an integer matrix $\mathbf{A} \in Z^{d\times m}$, we define an equivalence relation $\overset{\mathbf{A}}{\sim}$ on Z^d as follows. Two vectors $\mathbf{a}, \mathbf{b} \in Z^d$ are in the same equivalence class ($\mathbf{a} \overset{\mathbf{A}}{\sim} \mathbf{b}$) if and only if there exists an integer vector $\mathbf{z} \in Z^m$ such that $\mathbf{a} - \mathbf{b} = \mathbf{Az}$.

REMARK 2.7. Observe that if $d = m = 1$, then the relation $a \overset{\mathbf{A}}{\sim} b$ simply means $a \equiv b \pmod{A_{11}}$. In general, an equivalence class of the relation $\overset{\mathbf{A}}{\sim}$ is simply a translate (by a vector in Z^d) of the lattice spanned by the columns of $\mathbf{A}$. When $m < d$, there are infinitely many equivalence classes. When $m = d$ and $\mathbf{A}$ is nonsingular, there are finitely many equivalence classes. When $m > d$, there exists a $d \times d$ matrix $\mathbf{B}$ that defines the same equivalence relations as $\mathbf{A}$ (see Problem 2.8).

PROBLEM 2.8. Let

$$\mathscr{L} = \left\{ \sum_{j=1}^{m} z_j \mathbf{a}^j \,|\, z_j \in Z \right\} \subset Z^d$$

be the lattice spanned by m vectors $\mathbf{a}^1, \dots, \mathbf{a}^m \in Z^d$. Find vectors $\mathbf{b}^1, \dots, \mathbf{b}^l \in Z^d$, a permutation $(i_1, \dots, i_d)$ of $\{1, \dots, d\}$, and vectors $\mathbf{z}^1, \dots, \mathbf{z}^m \in Z^l$ with the following properties:

(i) $\mathscr{L} = \{\sum_{j=1}^{l} z_j \mathbf{b}^j \,|\, z_j \in Z\}$.

(ii) The vectors $\mathbf{b}^j$ are linearly independent.

(iii) The permutation $(i_1, \dots, i_d)$ is such that for $j = 1, \dots, l$, $\mathbf{b}^j_{i_k} = 0$ for $k < j$.

(iv) For all $j \in \{1, \dots, m\}$, $\mathbf{a}^j = \mathbf{Bz}^j$, where $\mathbf{B}$ is the matrix whose columns are the $\mathbf{b}^j$'s.

PROPOSITION 2.9. *Problem 2.8 can be solved in $O(d^2 mB(d\log(d\|\mathbf{A}\|)))$ time, where $\|\mathbf{A}\| = \max_{ij} |\mathbf{a}^j_i|$ and $B(t)$ bounds the time required to perform the extended Euclidean algorithm on two t bit integers.*

PROOF. Denote by $\mathbf{A} \in Z^{d\times m}$ the matrix whose jth column is $\mathbf{a}^j$ ($j = 1, \dots, m$). Note that the rows of $\mathbf{A}$ are not necessarily independent. We begin by identifying a maximal set of linearly independent rows of $\mathbf{A}$. This can be done in polynomial time by "careful" Gaussian elimination [4]. Without loss of generality, suppose that the first l rows comprise such a set (otherwise, the rows of $\mathbf{A}$ can be permuted accordingly), and denote the submatrix by $\mathbf{A}' \in Z^{l\times m}$. Let $\mathbf{A}'' \in Z^{(d-l)\times m}$ be the submatrix consisting of the last $d-l$ rows of $\mathbf{A}$. Moreover, during the process of identifying the submatrices, we also find a $\mathbf{Y} \in Q^{(d-l)\times l}$ such that $\mathbf{A}'' = \mathbf{Y}\mathbf{A}'$, and the size of each entry of $\mathbf{Y}$ is bounded by the determinants of the square submatrices of $\mathbf{A}$ (that is, $\|\mathbf{Y}\| = O(\|\mathbf{A}\|^l)$). This computation can be done in $O(ldm)$ arithmetic operations on numbers of size $O(\|\mathbf{A}\|^l)$. Let $\mathscr{L}' \subset Z^l$ denote the lattice spanned by the columns of $\mathbf{A}'$. It is easy to verify that $(\mathbf{b}', \mathbf{b}'')$ $(\mathbf{b}' \in R^l, \mathbf{b}'' \in R^{d-l})$ is in $\mathscr{L}$ if and only if $\mathbf{b}' \in \mathscr{L}'$ and $\mathbf{b}'' = \mathbf{Y}\mathbf{b}'$. Moreover, if $\mathbf{u}^1, \dots, \mathbf{u}^l \in \mathscr{L}'$ span the lattice $\mathscr{L}'$, then the vectors $(\mathbf{u}^1, \mathbf{Y}\mathbf{u}^1), \dots, (\mathbf{u}^l, \mathbf{Y}\mathbf{u}^l)$ span the lattice $\mathscr{L}$.

Denote the columns of $\mathbf{A}'$ by $\mathbf{a}'^j$. Observe that the lattice $\mathscr{L}'$ spanned by the $\mathbf{a}'^j$'s does not change if any vector $\mathbf{a}'^j$ is replaced by $-\mathbf{a}'^j$, or by $\mathbf{a}'^j + \mathbf{a}'^k$ or $\mathbf{a}'^j - \mathbf{a}'^k$ for some $k \neq j$. By a classical theorem of Hermite [6] (note that $\mathbf{A}'$ has a full row rank), the matrix $\mathbf{A}'$ can be triangulated by elementary (unimodular) column operations, i.e., the only operations used in this transformation are additions of integral multiples of one column to another, and exchanges of columns. Neither of these operations alters the lattice spanned by the columns of $\mathbf{A}'$. The matrix $\mathbf{A}'$ is transformed so that the last $m-l$ $(l \le d)$ columns become zero, and the first l columns comprise an integer lower triangular matrix with nonzero diagonal entries. If the diagonal entries are maximal in their rows, then the triangular matrix is unique and is called the Hermite normal form of $\mathbf{A}'$. It follows from Hermite's proof that this can be accomplished in at most lm applications of the extended Euclidean algorithm for the g.c.d. (of two numbers) plus $O(l^2m)$ arithmetic operations on integers. Kannan and Bachem [11] proposed an improvement of this algorithm, where all the operations are done on integers of polynomial size. Hafner and McCurley [5] proposed an asymptotically faster algorithm where all integers encountered do not exceed $l^l\|\mathbf{A}'\|^{2l}$. Their algorithm runs in $O(l^2mB(l\log(l\|\mathbf{A}\|)))$ bit operations. Moreover, for the simpler task of triangulating $\mathbf{A}'$, they used fast matrix multiplication to get even better time bounds. Using $O(l^{\theta-1}m\log(2m/l)B(l\log(l\|\mathbf{A}'\|)))$ bit operations (where $\theta < 2.376$, following Coppersmith and Winograd [3]), the algorithm computes a triangular matrix whose entries have absolute values bounded by $l^{l/2}\|\mathbf{A}'\|^l$.

Assume $\mathbf{A}'$ has been transformed into triangular form, replace $\mathbf{A}''$ by $\mathbf{Y}\mathbf{A}'$. Note that the new $\mathbf{A}''$ has integer entries of absolute values $O(l^{l/2+1}\|\mathbf{A}\|^{2l})$. Thus, the new matrix $\mathbf{A}$ has for every k, $k = 1, \dots, l$, $A_{kk} \neq 0$ and $A_{kj} = 0$ for $j > k$. Also, for $k = l+1, \dots, d$, $A_{kj} = 0$ for all $j > l$. Finally, let the vector $\mathbf{b}^j$ be equal to the jth column of the matrix $\mathbf{A}$ $(j = 1, \dots, l)$. Note that properties (i) and (ii) of Problem 2.8 imply that $l \leq d$ and that the permutation $(i_1, \dots, i_d)$ and the vectors $\mathbf{z}^1, \dots, \mathbf{z}^m$ exist. Given the $\mathbf{b}^j$'s, the vectors $\mathbf{z}^j$ can be computed easily. □

COROLLARY 2.10. (i) *If* $\mathbf{A}$ *has full row rank, then Problem* 2.8 *can be solved in*

$$O(l^{\theta-1} m \log(2m/l) B(l \log(l \|\mathbf{A}'\|)))$$

bit operations (*where* $\theta < 2.376$).

(ii) *When* d *is fixed, Problem* 2.8 *can be solved in* $O(m \log(\|\mathbf{A}\|))$ *arithmetic operations.*

(iii) *If an integer multiple* h *of* $\det \mathbf{B}$ *is given, then it follows from Theorem* 1 *of* [5] *that Problem* 2.8 *can be solved in* $O(d^2 m B(h))$ *arithmetic operations.*

Algorithm 2.11 solves Problem 2.1. It first transforms the weights of G to a basic form. The algorithm continues and solves Problem 2.8 with respect to the transformed edge weights. The number of connected components of G^* is $\det \mathbf{B}$.

ALGORITHM 2.11. (Connectivity). (i) *Compute a spanning tree* $T \subset E$ *of* G *and the appropriate potential function* $\mathbf{d}$, *and transform* $\mathbf{f}$ *to a basic form, i.e., replace* $\mathbf{f}$ *by* $\mathbf{f}^{(\mathbf{d})}$ (*see Corollary* 2.5).

(ii) *Denote by* $\mathbf{a}^1, \dots, \mathbf{a}^m$ *the vector weights in the range of the* (*transformed*) $\mathbf{f}$. *Solve Problem* 2.8 *with respect to the* $\mathbf{a}^j$*'s, i.e., compute the vectors* $\mathbf{b}^j, \mathbf{z}^j$.

(iii) *If* $l = d$, *then the number of connected components of* G^* *is* $\prod_{i=1}^d B_{ii}$. *Otherwise, there are infinitely many connected components, each of which is an* l*-dimensional periodic graph. Consider the new edge weights* $\tilde{\mathbf{f}}\colon E \to Z^l$ *given by* $\tilde{\mathbf{f}}(e) = \mathbf{z}^j$, *where* $\mathbf{f}(e) = \mathbf{a}^j$. *Construct the dependence graph* $\tilde{G} = (V, E, \tilde{\mathbf{f}})$. *The periodic graph* $(\tilde{G})^*$ *is isomorphic to each of the connected components of* G^*. *The isomorphism* $\mathscr{H}\colon \tilde{V}^* \to V^*$ *is given by* $\mathscr{H}(\mathbf{a}, u) = (\mathbf{Ba} + \mathbf{z}, u)$, *so that different connected components have different values of* $\mathbf{z} \in Z^d$ *associated with them.*

The correctness of the algorithm is proved in the following proposition.

PROPOSITION 2.12. *Suppose* $G = (V, E, \mathbf{f})$ *is connected and* $\mathbf{f}$ *is in a basic form. Denote by* $\mathbf{B} \in Z^{d \times l}$ *the matrix sought in Problem* 2.8 *with respect to the vectors* $\mathbf{f}(e)$ $(e \in E)$. *Under these conditions:*

(i) *If* $l = d$, *then the number of connected components of* G^* *is* $\prod_{i=1}^d B_{ii}$.
(ii) *The connected components of* G^* *correspond to the equivalence classes of* $\overset{\mathbf{B}}{\sim}$, *i.e., two vertices* $(\mathbf{a}, u), (\mathbf{b}, v) \in V^*$ *are in the same connected component if and only if* $\mathbf{a} \overset{\mathbf{B}}{\sim} \mathbf{b}$.
(iii) *The graph* $(\tilde{G})^*$ *is connected and isomorphic to each of the connected components of* G^*.

PROOF. For the proof of part (i), note that $\det \mathbf{B}$ is the volume of a "cell" of the lattice, which is the same as the number of distinct lattices that can be obtained by an integral translation of the given lattice. To prove part (ii) observe that $(\mathbf{a}, u)$ and $(\mathbf{b}, v)$ are in the same connected component if and only if $\mathbf{a} - \mathbf{b}$ equals an integer combination of the vectors $\mathbf{f}(E)$. Part (iii) follows from part (ii). The isomorphism $\mathscr{H}: \tilde{V}^* \to V^*$ is given by $\mathscr{H}(\mathbf{a}, u) = (\mathbf{Ba} + z, u)$. □

REMARK 2.13. The connected components of G^* are l-dimensional periodic graphs. Observe that when the weights $\mathbf{f}$ are in the basic form, then l is the dimension of the vector space spanned by $\mathbf{f}(E)$.

Denote $m = |E|$ and $\|\mathbf{f}\| = \max_{e\in E} \|\mathbf{f}(e)\|_\infty$. Denote by $\operatorname{conn}(m, \|\mathbf{f}\|, d)$ the number of operations required by the connectivity algorithm.

PROPOSITION 2.14. *Let $B(t)$ be as in Proposition* 2.9. *We have*

$$\operatorname{conn}(m, \|\mathbf{f}\|, d) = O(d^2 m B(d \log(d\|\mathbf{f}\|))).$$

PROOF. The computation done by Algorithm 2.11 amounts to transforming G to a basic form, and to a solution of an instance of Problem 2.8. Recall that the transformation to a basic form involves a computation of a spanning tree, and computing single source shortest paths on the tree. Therefore, the transformation is done in $O(m)$ operations. The proof follows from Proposition 2.9. □

COROLLARY 2.15. *The following are consequences of Corollary* 2.10:

(i) *If the connected components of G^* are l-dimensional periodic graphs, then Algorithm* 2.11 *requires $O(l^{\theta-1} m \log(2m/l) B(l \log(l\|\mathbf{f}\|)))$ operations $(\theta < 2.376)$.*

(ii) *If the dimension d is fixed, the algorithm requires $O(m \log\|\mathbf{f}\|)$ operations.*

(iii) *Consider the problem of recognizing whether the number of connected components of G^* equals k, where k is some fixed integer. It follows from Corollary* 2.10(iii) *that this task can be done in $O(d^2 m)$ time.*

3. Bipartiteness

In this section we consider the following problem.

PROBLEM 3.1. Given a dependence graph $G = (V, E, \mathbf{f})$, decide whether or not G^* is bipartite.

DEFINITION 3.2. Given a dependence graph G, define $\tilde{G} = (V, \tilde{E}, \tilde{\mathbf{f}})$ as follows. Let $\tilde{E}$ be the set of pairs $(u, v) \in V \times V$ such that there exists $w \in V$ with $\{(u, w), (w, v)\} \subseteq E$. For such pairs let $\tilde{\mathbf{f}}(u, v) = \mathbf{f}(u, w) + \mathbf{f}(w, v)$.

REMARK 3.3. Note that the graph $\tilde{G}$ is well defined even when G does not have edge weights. In this case the edges of $\tilde{G}$ correspond to paths of length 2 in G. If G is connected, then it is bipartite if and only if $\tilde{G}$ has two connected components. If G is a nonbipartite connected graph, then $\tilde{G}$ is connected.

We assume in this section that G^* is connected. If this is not true, then the dependence graphs that define the connected components of G^* should be considered independently.

ALGORITHM 3.4. (Bipartiteness). (i) *Construct* $\tilde{G}$ *as described in Definition* 3.2.

(ii) *If* $\tilde{G}$ *is not connected, then* G^* *is bipartite. Stop.*

(iii) *Otherwise, the graph* G^* *is bipartite if and only if* $(\tilde{G})^*$ *is not connected.*

Correctness. Observe that $(\tilde{G})^* = \widetilde{G^*}$, that is, (u, v) is an edge of $(\tilde{G})^*$ if and only if there is a path of length two between u and v in G^*. Therefore, it follows from Remark 3.3 that G^* is bipartite if and only if the graph $(\tilde{G})^*$ is not connected. This condition is checked in steps (ii) and (iii) of the algorithm.

Complexity. The algorithm amounts to constructing the graph $\tilde{G}$, and then testing if the periodic graphs corresponding to $\tilde{G}$ (or to its connected components) have one or two connected components. The construction of $\tilde{G}$ requires $O(m^2)$ time. Observe that $|\tilde{E}| = O(|E|^2)$. Thus, the time complexity of testing if $\widetilde{G^*}$ has one or two connected components is $O(d^2m^2)$ (see Corollary 2.15(iii)). It follows that the time complexity of the algorithm is $O(d^2m^2)$.

4. Minimum average cost spanning tree

Orlin [14] defined a minimum average cost spanning tree of a one-dimensional periodic graph and gave an algorithm that computes one. We extend his definition to higher dimensions as follows.

DEFINITION 4.1. Let $G = (V, E, \mathbf{f}, c)$ be a dependence graph together with edge *costs* given by $c: E \to R^+$.

(i) Extend c to the edges of G^* in the obvious way: $c((\mathbf{z}, u); (\mathbf{z} + \mathbf{f}(u, v), v)) = c(u, v)$.

(ii) For any positive integer n, denote by $G^n = (V^n, E^n)$ the finite subgraph of G^* induced by the set of vertices $(\mathbf{z}, v)$ $(v \in V)$ with $|z_i| \le n$ $(i = 1, \dots, d)$.

(iii) A *minimum average cost spanning tree* (MACST) $T^* \subset E^*$ of G^* is a spanning tree for which the limit

$$\lim_{n\to\infty} (2n)^{-d} \sum_{e\in E^n \cap T^*} c(e)$$

exists and is minimal among all trees for which such a limit exists.

(iv) A spanning tree T^* is said to be *optimal* if each finite subtree $T_1^* \subset T^*$ is a minimum cost spanning tree of the subgraph of G^* induced by the nodes of T_1^*.

DEFINITION 4.2. Let $\mathbf{B} \in Z^{d\times d}$ and let $\overset{\mathbf{B}}{\sim}$ be the equivalence relation on Z^d defined by $\mathbf{B}$ in Definition 2.6. Let $\mathscr{L}$ be the lattice spanned by the columns of $\mathbf{B}$. For any $\mathbf{a} \in Z^d$, denote by $\mathscr{L}^{(\mathbf{a})}$ the lattice spanned by the columns of $\mathbf{B}$ together with $\mathbf{a}$. Denote by $\mathbf{B}^{(\mathbf{a})} \in Z^{d\times d}$ a matrix whose columns span the lattice $\mathscr{L}^{(\mathbf{a})}$. Denote by $R(\mathbf{B}) \subset Z^d$ any set of representatives of the equivalence classes of the relation $\overset{\mathbf{B}}{\sim}$.

REMARK 4.3. If $\mathbf{B}$ is full rank and triangular, then a triangular matrix $\mathbf{B}^{(\mathbf{a})}$ can be computed within $O(d^2 B(\log\det\mathbf{B}))$ time, which amounts to $O(d^2\log\det\mathbf{B})$ arithmetic operations (see [5]).

Algorithm 4.4 below computes an optimal MACST of G^*. The algorithm maintains a set of isomorphic trees and a matrix $\mathbf{B}$, such that the equivalence classes of $\overset{\mathbf{B}}{\sim}$ correspond to these trees. The infinite tree T^* computed by the algorithm is represented by a collection of pairs $(e, \mathbf{B})$, where $e \in E$ and $\mathbf{B} \in Z^{d\times i}$ is an integer matrix. Each such pair represents a set of edges of G^*:

(i) When $i < d$, each such pair represents infinitely many edges: for $(u, v) \in E$, $((\mathbf{z}, u); (\mathbf{z} + \mathbf{f}(u, v), v)) \in T^*$ if and only if $\mathbf{z} \in R(\mathbf{B})$. The algorithm generates one such pair for each $i < d$.

(ii) When $i = d$, the matrix $\mathbf{B}$ is nonsingular. The algorithm may produce as many as $m - d$ such pairs. Each pair represents finitely many edges: for $(u, v) \in E$, $((\mathbf{z}, u); (\mathbf{z} + \mathbf{f}(u, v), v)) \in T^*$ if and only if $\mathbf{z} \in R(\mathbf{B})$ and $\mathbf{z} \overset{\mathbf{B}}{\not\sim} \mathbf{y}$ for all $\mathbf{y} \in R(\mathbf{B}^{(\mathbf{f}(e))})$. Observe that the number of edges that correspond to such pairs is $|R(\mathbf{B})| - |R(\mathbf{B}^{(\mathbf{f}(e))})|$.

Each edge $e \in E$ can occur in at most one pair $(e, \mathbf{B})$. Algorithm 4.4 below consists of two phases. During the first phase (steps (iii), (iv)), pairs with $i < d$ are selected. During the second phase (steps (v), (vi)), pairs with $i = d$ are selected. Within each phase, the algorithm is greedy and selects the edges (with a certain desired property) according to increasing cost. Without loss of generality, assume G^* is connected.

ALGORITHM 4.4. (Optimal Minimum Average Cost Spanning Tree).

(i) *Find a minimum cost spanning tree T of G and transform G to the basic form corresponding to T (see Corollary* 2.5).

(ii) *Initialize* $T^* \leftarrow \{(t, \{\mathbf{0}\}) | t \in T\}$.

(iii) *Do step* (iv) *for* $i = 1, \dots, d$, *and then go to* (iv).

(iv) *Find an edge $e_i \in E$ of minimum cost, such that $\mathbf{f}(e_i)$ is linearly independent of the weights of previously selected edges: $\mathbf{f}(e_j)$ $(j < i)$. Let $\mathbf{B} \in Z^{d\times i}$ be the matrix whose columns are the vectors $\mathbf{f}(e_1), \dots, \mathbf{f}(e_i)$. Add the pair $(e_i, \mathbf{B})$ to the forest T^*.*

(v) *At this point $\mathbf{B} \in Z^{d\times d}$ is a square matrix, and T^* consists of* $\det\mathbf{B}$ *isomorphic trees. Transform $\mathbf{B}$ into a triangular matrix that defines the same lattice (see Proposition* 2.9). *Repeat step* (vi) *until* $\det\mathbf{B} = 1$.

(vi) *Select an edge* e *of minimal cost such that* $\mathbf{f}(e) \neq \mathbf{B}\mathbf{z}$ *for all* $\mathbf{z} \in Z^d$. *Add the pair* $(e, \mathbf{B})$ *to* T^*. *Update* $\mathbf{B}$ *to be* $\mathbf{B}^{(\mathbf{f}(\mathbf{e}))}$ (*see Definition* 4.2 *and Remark* 4.3).

Correctness. Denote by $e_1, \dots, e_s$ $(d \le s \le m)$ the edges selected by the algorithm. The edges $e_1, \dots, e_d$ are selected during step (iv). First we prove that T^* is indeed a tree.

PROPOSITION 4.5. *Throughout the execution of the algorithm, the connected components of* T^* *constitute a collection of isomorphic trees corresponding to the equivalence classes of the relation* $\overset{\mathbf{B}}{\sim}$.

PROOF. Consider any point during the execution of the algorithm and suppose $e_1, \dots, e_k$ are the edges that have been selected so far. Recall that two vectors are in the same equivalence class of $\overset{\mathbf{B}}{\sim}$ if and only if they differ by an integer linear combination of the vectors $\mathbf{f}(e_1), \dots, \mathbf{f}(e_k)$. First, observe that after step (ii) of the algorithm, T^* consists of a collection of finite trees. The vertex sets of these trees have the form $\{(i, v) | v \in V\}$. The proposition follows by induction on k. □

The graph G^* is connected. Therefore, if all edges are selected, we must have $\det \mathbf{B} = 1$. Thus the algorithm is guaranteed to terminate with $\det \mathbf{B} = 1$. It follows from Proposition 4.5 that when the algorithm terminates the forest T^* has one connected component.

PROPOSITION 4.6. *The tree* T^* *computed by the algorithm has a minimum average cost.*

PROOF. Observe that the set $\overline{T}$ of edges of G^* corresponding to the set $T \cup \{e_1\}$ is contained in T^* and does not contain a cycle. Thus, $\overline{T}$ is a forest of G^*. We will show that it can be converted into a tree by adding an asymptotically small number of edges from E^*, so that the limit of the average cost is the same as in $\overline{T}$.

Consider the intersection of $\overline{T}$ with G^n. Imagine running any greedy algorithm for a minimum spanning tree of G^n. We show that such an algorithm selects all the edges of $\overline{T} \cap G^n$. Obviously, the edges generated by T do not introduce any cycles, and they are of minimum cost. All the other edges of G, whose weights equal $\mathbf{0}$, generate edges of G^* that introduce cycles. However, the algorithm selects the next minimum cost edge $e_1 \in E$ such that $\mathbf{f}(e_1) \neq \mathbf{0}$, and adds to T^* all the edges generated by e_1. Thus, no cycle is introduced.

By now, we have a forest of isomorphic trees of size $\Omega(n/\|\mathbf{f}\|)$ that cover G^n. The asymptotic average cost of the forest is

$$\lim_{n\to\infty} \frac{(2n)^d c(T \cup \{e_1\})}{(2n)^d} = c(T \cup \{e_1\}).$$

This gives a lower bound on the asymptotic average cost of the spanning tree of G^*. Observe that it is an upper bound as well, since the number of additional edges needed to complete this forest to a tree is $O((2n)^{d-1}\|\mathbf{f}\|)$. It is easy to see that, regardless of the choice of these additional edges, the asymptotically minimum average cost of the tree is the same as that of the initial forest, i.e., $c(T \cup \{e_1\})$. $\square$

PROPOSITION 4.7. *The tree* T^* *is optimal* (*see Definition* 4.1(iv)).

PROOF. To prove optimality, consider some subtree T_1^* of T^*. Denote by H the subgraph of G^* induced by the nodes of T_1^*. We need to show that T_1^* is a minimum cost spanning tree of H, i.e., the edges of T are the same as the ones picked by some greedy algorithm. Algorithm 4.4 is greedy within each of its two phases. Edges are selected according to increasing cost. When an edge e is selected, the algorithm adds to T^* a maximal subset of the edges generated by e, so that T^* remains acyclic. We need to show that during the first phase of the algorithm, no two trees that are connected by an edge can be connected by an edge of a lower cost. Consider the state before choosing a new edge in the first phase. Observe that all vertices of a connected component always lie on a "flat" which is a translate of the subspace spanned by the columns of $\mathbf{B}$ (see Proposition 4.5). The weight of any edge selected by the algorithm during step (iv) is linearly independent of the columns of $\mathbf{B}$. Thus, the new edges must lie between two such "flats". All edges of lower cost are linearly dependent on the columns of $\mathbf{B}$. Therefore, they lie within such flats. It follows that the new edges added to T^* are minimum cost edges, linking connected components of the current T^*. Hence, they occur in some minimum cost spanning tree. $\square$

The algorithm amounts essentially to computing a minimum cost spanning tree in G, and then computing the matrix $\mathbf{B}$ and updating it at most m times (see Remark 4.3). Thus, the complexity is $O(n \log n + md^2 B(d \log(d\|\mathbf{f}\|)))$.

REFERENCES

1. E. Cohen and N. Megiddo, *Strongly polynomial and NC algorithms for detecting cycles in dynamic graphs*, Proc. 21st Annual ACM Sympos. on Theory of Computing, ACM, 1989, pp. 523–534.
2. ——, *Strongly polynomial time and NC algorithms for detecting cycles in periodic graphs*, IBM Research Report RJ 7587 (70764), IBM Almaden Research Center, San Jose, Ca., July 1990.
3. D. Coppersmith and S. Winograd, *Matrix multiplication via arithmetic progressions*, Proc. 19th Annual ACM Sympos. on Theory of Computing, ACM, 1987, pp. 1–6.
4. J. Edmonds, *Systems of distinct representatives and linear algebra*, J. Res. Nat. Bur. Standards **71B** (1967), 241–245.
5. J. L. Hafner and K. S. McCurley, *Asymptotically fast triangularization of matrices over rings*, IBM Research Report RJ 6921 (66027), IBM Almaden Research Center, San Jose, Ca., July 1989. Also in: Proc. 1st Annual ACM-SIAM Sympos. on Discrete Algorithms, ACM-SIAM, 1990, 194–200.
6. C. Hermite, *Sur l'introduction des variables continues dans la théorie des nombres*, J. Reine Angew. Math. **41** (1851), 431–451.
7. K. Iwano, *Some problems on doubly periodic infinite graphs*, Tech. Rep. CS-TR-078-87, Princeton University, 1987.

8. K. Iwano and K. Steiglitz, *Optimization of one-bit full adders embedded in regular structures*, IEEE Trans. Acoust. Speech Signal Process., IEEE, New York, 1986.
9. ——, *Testing for cycles in infinite graphs with periodic structure*, Proc. 19th Annual ACM Sympos. on Theory of Computing, ACM, 1987, pp. 46–53.
10. ——, *Planarity testing of doubly connected periodic infinite graphs*, Networks **18** (1988), 205–222.
11. R. Kannan and A. Bachem, *Polynomial algorithms for computing the Smith and Hermite normal forms of an integer matrix*, SIAM J. Comput. **8** (1979), 499–507.
12. R. M. Karp, R. E. Miller, and S. Winograd, *The organization of computations for uniform recurrence equations*, J. Assoc. Comput. Mach. **14** (1967), 563–590.
13. K. S. Kosaraju and G. F. Sullivan, *Detecting cycles in dynamic graphs in polynomial time*, Proc. 27th Annual IEEE Sympos. on Foundations of Computer Science, ACM, 1988, pp. 398–406.
14. J. B. Orlin, *Some problems in dynamic/periodic graphs*, in Progress in Combinatorial Optimization (W. R. Pullyblank, ed.), Academic Press, Orlando, Florida, 1984, pp. 273–293.
15. V. P. Roychowdhury and T. Kailath, *Study of parallelism in regular iterative algorithms*, Proc. 1990 ACM Sympos. on Parallel Algorithms and Architectures, ACM, 1990, pp. 367–376.

DEPARTMENT OF COMPUTER SCIENCE, STANFORD UNIVERSITY, STANFORD, CALIFORNIA 94305 AND IBM ALMADEN RESEARCH CENTER, SAN JOSE, CALIFORNIA 95120-6099
E-mail address: edith@cs.stanford.edu

IBM ALMADEN RESEARCH CENTER, SAN JOSE, CALIFORNIA 95120-6099 AND SCHOOL OF MATHEMATICAL SCIENCES, TEL AVIV UNIVERSITY, TEL AVIV, ISRAEL
E-mail address: megiddo@ibm.com